



OXFORD

# THE GRAMMAR OF INTERACTIVES

BERND HEINE

# The Grammar of Interactives



# The Grammar of Interactives

BERND HEINE

OXFORD  
UNIVERSITY PRESS



OXFORD  
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,  
United Kingdom

Oxford University Press is a department of the University of Oxford.  
It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide. Oxford is a registered trade mark of  
Oxford University Press in the UK and in certain other countries

© Bernd Heine 2023

The moral rights of the author have been asserted

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in  
a retrieval system, or transmitted, in any form or by any means, without the  
prior permission in writing of Oxford University Press, or as expressly permitted  
by law, by licence or under terms agreed with the appropriate reprographics  
rights organization. Enquiries concerning reproduction outside the scope of the  
above should be sent to the Rights Department, Oxford University Press, at the  
address above

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data  
Data available

Library of Congress Control Number: 2022940625

ISBN 978-0-19-287149-7

DOI: 10.1093/oso/9780192871497.001.0001

Printed and bound by  
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and  
for information only. Oxford disclaims any responsibility for the materials  
contained in any third party website referenced in this work.

# Contents

|                                                     |     |
|-----------------------------------------------------|-----|
| <i>Acknowledgments</i>                              | x   |
| <i>List of figures and tables</i>                   | xi  |
| <i>List of abbreviations</i>                        | xiv |
| <br>                                                |     |
| 1. Interactives                                     | 1   |
| 1.1 Introduction                                    | 1   |
| 1.2 Definition                                      | 7   |
| 1.3 Properties                                      | 11  |
| 1.3.1 Definitional properties                       | 11  |
| 1.3.2 Further features                              | 31  |
| 1.3.3 Conclusion                                    | 42  |
| 1.4 Properties shared by the two domains of grammar | 42  |
| 1.5 A classification of interactives                | 47  |
| 1.6 This volume                                     | 52  |
| <br>                                                |     |
| 2. Argument structure                               | 55  |
| 2.1 Meaning                                         | 55  |
| 2.1.1 Paraphrase                                    | 56  |
| 2.1.2 Expressive meaning                            | 63  |
| 2.1.3 Interactives as depictions                    | 68  |
| 2.1.4 Discussion                                    | 79  |
| 2.1.5 Conclusions                                   | 82  |
| 2.2 Identifying arguments                           | 83  |
| 2.2.1 Introduction                                  | 83  |
| 2.2.2 The arguments                                 | 87  |
| 2.2.3 More on T arguments                           | 94  |
| 2.2.4 More than one argument structure              | 97  |
| 2.2.5 Adjuncts                                      | 100 |
| 2.2.6 How to identify arguments                     | 102 |
| 2.3 Discussion                                      | 103 |
| 2.4 Contrasting ways of discourse coding            | 104 |
| 2.5 Conclusions                                     | 106 |
| <br>                                                |     |
| 3. Types of interactives                            | 108 |
| 3.1 Attention signals                               | 110 |
| 3.1.1 Introduction                                  | 110 |
| 3.1.2 Working definition                            | 111 |
| 3.1.3 Grammar                                       | 111 |

|       |                                               |     |
|-------|-----------------------------------------------|-----|
| 3.1.4 | Discussion                                    | 115 |
| 3.1.5 | The attention signal <i>*hey</i>              | 116 |
| 3.2   | Directives                                    | 119 |
| 3.2.1 | Introduction                                  | 120 |
| 3.2.2 | Working definition                            | 120 |
| 3.2.3 | Grammar                                       | 122 |
| 3.2.4 | Discussion                                    | 123 |
| 3.2.5 | Canonical imperatives                         | 125 |
| 3.2.6 | Special domains of usage                      | 126 |
| 3.3   | Discourse markers                             | 127 |
| 3.3.1 | Introduction                                  | 127 |
| 3.3.2 | Working definition                            | 128 |
| 3.3.3 | Grammar                                       | 129 |
| 3.3.4 | Fillers                                       | 133 |
| 3.3.5 | A note on clicks as English discourse markers | 136 |
| 3.3.6 | Discourse markers in language contact         | 137 |
| 3.3.7 | Grammaticalization                            | 139 |
| 3.3.8 | Discussion                                    | 140 |
| 3.4   | Evaluatives                                   | 141 |
| 3.4.1 | Introduction                                  | 141 |
| 3.4.2 | Working definition                            | 142 |
| 3.4.3 | Grammar                                       | 143 |
| 3.4.4 | Discussion                                    | 144 |
| 3.5   | Ideophones                                    | 145 |
| 3.5.1 | Introduction                                  | 146 |
| 3.5.2 | Working definition                            | 147 |
| 3.5.3 | Grammar                                       | 149 |
| 3.5.4 | Ideophones vs. interjections                  | 164 |
| 3.5.5 | Grammaticalization                            | 165 |
| 3.5.6 | A note on sound symbolism                     | 166 |
| 3.5.7 | Are ideophones interactives?                  | 168 |
| 3.5.8 | Ideophones as a potentially open-ended class  | 178 |
| 3.6   | Interjections                                 | 181 |
| 3.6.1 | Introduction                                  | 182 |
| 3.6.2 | Working definition                            | 183 |
| 3.6.3 | Grammar                                       | 188 |
| 3.6.4 | Sub-types                                     | 191 |
| 3.6.5 | Discussion                                    | 199 |
| 3.7   | Response elicitors                            | 199 |
| 3.7.1 | Introduction                                  | 199 |
| 3.7.2 | Working definition                            | 200 |
| 3.7.3 | Grammar                                       | 201 |
| 3.7.4 | Functional space                              | 203 |
| 3.7.5 | Other-initiated repair markers                | 206 |
| 3.8   | Response signals                              | 209 |

|        |                                               |     |
|--------|-----------------------------------------------|-----|
| 3.8.1  | Introduction                                  | 210 |
| 3.8.2  | Working definition                            | 211 |
| 3.8.3  | Grammar                                       | 213 |
| 3.8.4  | Response signals expressed by clicks          | 218 |
| 3.8.5  | Discussion                                    | 219 |
| 3.9    | Social formulae                               | 220 |
| 3.9.1  | Introduction                                  | 221 |
| 3.9.2  | Working definition                            | 222 |
| 3.9.3  | Grammar                                       | 223 |
| 3.9.4  | Grammaticalization                            | 226 |
| 3.9.5  | The English formula <i>please</i>             | 227 |
| 3.9.6  | How to use social formulae                    | 230 |
| 3.10   | Vocatives                                     | 231 |
| 3.10.1 | Introduction                                  | 233 |
| 3.10.2 | Working definition                            | 233 |
| 3.10.3 | Grammar                                       | 235 |
| 3.10.4 | Sub-types                                     | 238 |
| 3.10.5 | Typological variation                         | 240 |
| 3.10.6 | Special features of vocatives                 | 241 |
| 3.11   | Interaction with animals                      | 248 |
| 3.11.1 | Introduction                                  | 248 |
| 3.11.2 | Working definition                            | 249 |
| 3.11.3 | Grammar                                       | 250 |
| 3.11.4 | Usage and functions                           | 251 |
| 3.11.5 | Form                                          | 254 |
| 3.11.6 | Voiceless sibilants in animal dispersal calls | 258 |
| 3.11.7 | Discussion                                    | 259 |
| 3.11.8 | Conclusions                                   | 261 |
| 3.12   | An overview                                   | 262 |
| 3.13   | Conclusions                                   | 265 |
| 4.     | Development                                   | 266 |
| 4.1    | Cooptation: How interactives arise            | 267 |
| 4.1.1  | The mechanism                                 | 268 |
| 4.1.2  | Cooptation and ideophones                     | 271 |
| 4.1.3  | From propositional to expressive meaning      | 275 |
| 4.1.4  | A note on pragmaticalization                  | 276 |
| 4.2    | Grammaticalization                            | 277 |
| 4.2.1  | Parameters of grammaticalization              | 278 |
| 4.2.2  | Main pathways                                 | 280 |
| 4.2.3  | A network of grammaticalization               | 293 |
| 4.2.4  | Constructionalization                         | 295 |
| 4.2.5  | From interactive grammar to sentence grammar  | 297 |
| 4.3    | Other mechanisms                              | 300 |
| 4.3.1  | Expressive reinforcement                      | 300 |
| 4.3.2  | Camouflaging                                  | 303 |

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| 4.4 Grammaticalization of ideophones                               | 307 |
| 4.4.1 A case study                                                 | 309 |
| 4.4.2 A common pathway: from ideophone to lexical category         | 313 |
| 4.5 Conclusions                                                    | 319 |
| 5. Related approaches                                              | 322 |
| 5.1 Notions of interactive grammar                                 | 322 |
| 5.1.1 Grammar traditions                                           | 322 |
| 5.1.2 Inserts                                                      | 330 |
| 5.1.3 Expressives                                                  | 333 |
| 5.1.4 Discussion                                                   | 336 |
| 5.2 Dual-process frameworks                                        | 336 |
| 5.2.1 An overview                                                  | 337 |
| 5.2.2 Microgrammar vs. macrogrammar                                | 344 |
| 5.2.3 Sentence grammar vs. thetical grammar                        | 345 |
| 5.2.4 Propositional structure vs. interactional structure          | 348 |
| 5.3 Neurolinguistic correlates                                     | 352 |
| 5.4 Dualism in social psychology                                   | 356 |
| 5.5 Discussion                                                     | 360 |
| 5.6 Conclusions                                                    | 363 |
| 6. The status of interactives                                      | 364 |
| 6.1 Where are interactives located in grammar and discourse?       | 364 |
| 6.1.1 They are part of language structure                          | 365 |
| 6.1.2 They are located at the periphery of language structure      | 365 |
| 6.1.3 They are intermediate between language and something else    | 366 |
| 6.1.4 They are not part of language                                | 367 |
| 6.1.5 They form a domain of grammar separate from sentence grammar | 368 |
| 6.1.6 Conclusion                                                   | 369 |
| 6.2 Are interactives words?                                        | 369 |
| 6.2.1 On terminology                                               | 370 |
| 6.2.2 What is a word?                                              | 373 |
| 6.2.3 Discussion                                                   | 375 |
| 6.3 Do interactives form a system?                                 | 376 |
| 6.3.1 Is there a set of entities making up the system?             | 377 |
| 6.3.2 Do the entities interact with one another?                   | 377 |
| 6.3.3 If there is a system, which purpose does it serve?           | 379 |
| 6.3.4 How is the system set off from other systems or phenomena?   | 379 |
| 6.4 Are interactives linguistic marginalia?                        | 380 |
| 6.5 Interactives in language contact                               | 383 |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| 6.6 Interactives and language acquisition                       | 384 |
| 6.7 Are interactives fossils of language evolution?             | 385 |
| 6.8 Conclusions                                                 | 390 |
| 7. Deciding between two grammars                                | 391 |
| 7.1 Distinguishing features                                     | 392 |
| 7.1.1 Conceptual space                                          | 392 |
| 7.1.2 Meaning                                                   | 393 |
| 7.1.3 Mental representation                                     | 393 |
| 7.1.4 Form                                                      | 394 |
| 7.1.5 Speed of processing                                       | 394 |
| 7.1.6 Sequence of activation                                    | 395 |
| 7.1.7 The locus of semantic-pragmatic anchoring                 | 396 |
| 7.1.8 The role of context                                       | 396 |
| 7.1.9 Constraints on usage                                      | 396 |
| 7.1.10 Use of gesture                                           | 397 |
| 7.2 Types of discourse                                          | 397 |
| 7.2.1 Usage of interactive grammar is predominant               | 398 |
| 7.2.2 Both grammars are equally important                       | 400 |
| 7.2.3 Use of sentence grammar is predominant                    | 402 |
| 7.3 How do the two grammars contribute to discourse processing? | 403 |
| 7.3.1 General observations                                      | 403 |
| 7.3.2 Interactives in adjacency pairs                           | 405 |
| 7.4 Conclusion                                                  | 410 |
| 8. Conclusions                                                  | 411 |
| <i>References</i>                                               | 419 |
| <i>Index of Languages</i>                                       | 444 |
| <i>Index of Names</i>                                           | 447 |
| <i>Index of Subjects</i>                                        | 453 |

# Acknowledgments

The present book has benefitted greatly from cooperation by a wide range of colleagues and friends. We wish to thank Sasha Aikhenvald, Azeb Amha, Jim Bennett, Matthias Brenzinger, Kate Burridge, Herbert Clark, Ulrike Claudi, Bob Dixon, Nick Evans, Anne-Maria Fehn, Elke Gehweiler, Tom Givón, John Haiman, Christa König, Jim Matisoff, Zsahra Mirkhaef, Nico Nassenstein, Damara Nübling, Seongha Rhee, Barbara Sonnenhauser, Ulrike Stange-Hundsörfer, Tim Wharton, and Anna Wierzbicka for their valuable comments and all their support. We also wish to thank the participants of the Conference ‘Evidentiality and Modality at the Crossroads of Grammar and Lexicon,’ which took place from June 10 to 11, 2021, in Montpellier, especially to Sascha Diwersy, Hans Kronning, Eric Mélac, Aksu-Koc Ayhan, and Seongha Rhee, Giulio Scivoletto, for their stimulating comments on our presentation on interactive grammar.

Our special gratitude is due to Felix Ameka, Alexander Andrason, Doug Biber, Danna Chandra-Menendez, Mark Dingemanse, Franck Floricic, Gunther Kaltenböck, Tania Kuteva, Lachlan Mackenzie, and Neal Norrick for having struggled through an earlier version of this book, suggesting a wide range of changes and improvements. Especially Franck Floricic and Ad Foolen have enriched our knowledge with a wealth of inspirations, inducing us to look at the nature of interactives in a new light.

# List of figures and tables

## Figures

|                                                                           |     |
|---------------------------------------------------------------------------|-----|
| 4.1. A network of grammaticalization chains linking types of interactives | 294 |
| 5.1. Classification of the C-units and inserts of Biber et al. (1999)     | 333 |

## Tables

|                                                                                            |     |
|--------------------------------------------------------------------------------------------|-----|
| 1.1. The main types of interactives                                                        | 10  |
| 1.2. Examples of English interactives in the left periphery slot of an utterance           | 19  |
| 1.3. Tonal distinctions in the German interactives <i>ach</i> , <i>hm</i> , and <i>oje</i> | 24  |
| 1.4. English interactives and negation                                                     | 29  |
| 1.5. Unusual phonological features in English types of interactives                        | 32  |
| 1.6. The main types of interactives with English examples                                  | 49  |
| 2.1. Two kinds of meaning of interactives according to Kaplan (1999)                       | 66  |
| 2.2. Methods, functions, and displays used in communication                                | 69  |
| 2.3. Three basic modes of representing information for discourse processing                | 77  |
| 2.4. Types of English interactives                                                         | 88  |
| 2.5. Approximate paraphrases of the interactives in Table 2.4                              | 88  |
| 2.6. Prototypical argument structure of the interactives in Table 2.5                      | 89  |
| 3.1. Main functions of the ten types of interactives                                       | 109 |
| 3.2. Correspondences between inserts (Biber et al. 1999) and interactives                  | 110 |
| 3.3. Corresponding forms of the attention signal <i>*hey</i> in different languages        | 117 |
| 3.4. Functions expressed by directives in two Australian languages                         | 124 |
| 3.5. (Logudorese) Sardinian directives                                                     | 125 |
| 3.6. Suppletive imperative stems in !Xun (E1 dialect, Kx'a family)                         | 126 |
| 3.7. Partial suppletism in Swahili imperative forms                                        | 126 |
| 3.8. Aspectual features of Japanese adverbial ideophones ('mimetics')                      | 151 |
| 3.9. The main discourse constructions for presenting ideophones                            | 153 |
| 3.10. Ideophone constructions in Siwu                                                      | 157 |
| 3.11. Correlations in sound symbolism observed in ideophones                               | 166 |
| 3.12. Iconic sound–meaning distinctions in West African languages                          | 166 |



|                                                                                               |     |
|-----------------------------------------------------------------------------------------------|-----|
| 3.13. English forms that have been classified as primary and secondary interjections          | 193 |
| 3.14. Types of English interactives showing a primary vs. secondary form distinction          | 194 |
| 3.15. Structural schema imposed by other-initiated repair (OIR) markers in conversations      | 206 |
| 3.16. A sample of 'other-initiated repair markers'                                            | 208 |
| 3.17. Degrees of politeness in Korean response signals based on speech levels                 | 213 |
| 3.18. Tonological variants of the German response signal <i>hm</i>                            | 216 |
| 3.19. Types of English vocatives                                                              | 238 |
| 3.20. Reduction of some Italian first names                                                   | 242 |
| 3.21. Reduction of some Catalan first names                                                   | 243 |
| 3.22. Reduction of some Zulu nouns (Bantu, Niger-Congo)                                       | 243 |
| 3.23. Extensions of the term <i>îiyá</i> 'mother' in Datooga (Southern Nilotic, Nilo-Saharan) | 247 |
| 3.24. Animal directives for deictic motion in Zargulla (Omoti, Afroasiatic)                   | 252 |
| 3.25. Functions of animal directives for mules in Ayt Hadiddu (Berber, Afroasiatic)           | 253 |
| 3.26. Animal directives in Manambu (Ndu family)                                               | 255 |
| 3.27. Classes of animal directives in Arusa Maasai (Eastern Nilotic, Nilo-Saharan)            | 257 |
| 3.28. From animal directive to verb in Muna (Malayo-Polynesian, Austronesian)                 | 260 |
| 3.29. Nursery terms derived from animal directives in Ayt Hadiddu (Berber, Afroasiatic)       | 261 |
| 3.30. Distinguishing grammatical features of interactives                                     | 264 |
| 4.1. Cooptation of some ideophones from verbs in Swahili (Bantu, Niger-Congo)                 | 273 |
| 4.2. Cooptation of ideophones from verbs in Mwera (Bantu, Niger-Congo)                        | 274 |
| 4.3. The parameters of grammaticalization                                                     | 278 |
| 4.4. Hypothesis on the development of some interactives towards text anchoring                | 280 |
| 4.5. From vocative to discourse marker                                                        | 286 |
| 4.6. Unidirectionality in argument development                                                | 295 |
| 4.7. English examples of camouflaging                                                         | 306 |
| 4.8. German examples of camouflaging                                                          | 307 |
| 4.9. The initial and final stages to be expected in the grammaticalization of ideophones      | 315 |
| 4.10. From ideophone to verb, selected examples                                               | 316 |
| 4.11. From ideophone to noun in Southern Sotho (Bantu, Niger-Congo)                           | 316 |
| 5.1. 'Interjections' distinguished by Bloomfield ([1933] 1962)                                | 323 |
| 5.2. 'Interjections' distinguished by Leech et al. (1982)                                     | 324 |

|                                                                                                          |     |
|----------------------------------------------------------------------------------------------------------|-----|
| 5.3. ‘Discourse markers’ distinguished by Zwicky (1985)                                                  | 324 |
| 5.4. ‘Pragmatic markers’ distinguished by Fraser (1990; 1999) qualifying as interactives                 | 325 |
| 5.5. ‘Interjections’ distinguished by Norrick (2009)                                                     | 326 |
| 5.6. Classes of ‘interjections’ distinguished by Ameka and Wilkins (2006)                                | 326 |
| 5.7. Korean ‘discourse adverbials’ distinguished by Sohn (1999)                                          | 327 |
| 5.8. Wampis ‘interjections’ distinguished by Peña (2015)                                                 | 328 |
| 5.9. Kisi ‘interjections’ distinguished by Childs (1995)                                                 | 328 |
| 5.10. Zulu ‘interjectives’ distinguished by Doke ([1927] 1988)                                           | 329 |
| 5.11. Thompson ‘expletives’ distinguished by Thompson and Thompson (1992)                                | 329 |
| 5.12. English inserts distinguished by Biber et al. (1999)                                               | 332 |
| 5.13. English examples of formulaic theticals                                                            | 347 |
| 5.14. Sentence grammar vs. interactive grammar in a speech sample of an English-speaking aphasic patient | 356 |
| 5.15. Some distinguishing features associated with processes of reasoning and judgment                   | 358 |
| 5.16. Dual-process frameworks relating most strongly to interactive grammar                              | 360 |
| 6.1. Examples of English interactives in the left periphery slot of an utterance                         | 378 |
| 6.2. Examples of English primary interactives                                                            | 387 |
| 7.1. Features prototypically distinguishing sentence grammar from interactive grammar                    | 392 |
| 7.2. Example of an opening section in English telephone conversations                                    | 399 |

# List of abbreviations

|         |                                     |
|---------|-------------------------------------|
| [ ], () | contextual information              |
| ↑       | prosodic foregrounding              |
| 1, 2, 3 | first, second, third person         |
| A       | accusative                          |
| a.n.    | note added by the present author    |
| ABL     | ablative                            |
| ACAD    | academic text                       |
| ACC     | accusative                          |
| ADD     | additive                            |
| ADN     | adnominal                           |
| ADV     | adverb                              |
| ALL     | allative                            |
| AmE     | American English                    |
| ART     | article                             |
| ATT     | attention signal                    |
| ATTR    | attribute                           |
| AUX     | auxiliary                           |
| BEN     | benefactive                         |
| BrE     | British English                     |
| CNV     | converb                             |
| COMP    | complementizer                      |
| CONN    | connective                          |
| CONV    | conversation transcription; converb |
| COP     | copula, copular verb                |
| CPL     | completive                          |
| CRCM    | circumstance                        |
| CU      | coopted unit                        |
| DAT     | dative                              |
| DEF     | definite marker                     |
| DEM     | demonstrative                       |
| DET     | determiner                          |
| DIM     | diminutive                          |
| DIR     | directive                           |
| DM      | discourse marker                    |
| DO      | direct object                       |
| DUR     | durative                            |
| EM      | expressive morphology               |
| EMPH    | emphasizer                          |

|              |                              |
|--------------|------------------------------|
| END          | sentence ender               |
| EVID         | evidential                   |
| EXC          | exclamative                  |
| EXP          | expletive                    |
| F            | feminine                     |
| FICT         | fiction text                 |
| FOC          | focus                        |
| FP           | final particle               |
| FUT          | future                       |
| GEN          | genitive                     |
| GER          | gerundive form               |
| HAB          | habitual                     |
| HESI         | hesitation marker            |
| HON          | honorific                    |
| HOR          | hortative                    |
| IDE          | ideophone                    |
| IMP          | imperative                   |
| INT          | interjection                 |
| ITV          | interactive                  |
| JII          | juncture II                  |
| Lit.         | literal meaning              |
| LK           | linker, linking form of verb |
| LOC          | locative                     |
| MP           | modal particle               |
| MR           | multiple reference form      |
| N            | nominative                   |
| NC1, 2, etc. | noun class 1, 2, etc.        |
| NEG          | negative marker              |
| NEM          | new event marker             |
| NEWS         | newspaper text               |
| NF           | non-finite marker            |
| NML          | nominalizer                  |
| NOM          | nominative, nominalizer      |
| O            | object marker                |
| PAS          | passive                      |
| PAST         | past tense                   |
| PE           | polite ending                |
| PERF         | perfect                      |
| PFV          | perfective                   |
| PINT         | primary interjection         |
| PL           | plural                       |
| PM           | pragmatic marker             |
| POL          | politeness                   |
| POSS         | possessive                   |
| PREF         | prefix                       |

|      |                                      |
|------|--------------------------------------|
| PRES | present tense                        |
| PROB | probability                          |
| PROG | progressive                          |
| PROM | prominence                           |
| PRON | pronouns                             |
| PS   | primary suffix                       |
| PAST | past tense                           |
| PTCP | participle                           |
| Q    | interrogative marker                 |
| Q1   | interrogative marker                 |
| QP   | question particle                    |
| QT   | quotative marker                     |
| QUO  | quotative                            |
| REE  | response elicitor                    |
| REF  | referential demonstrative            |
| REFL | reflexive marker                     |
| REP  | repetitive                           |
| RES  | response signal                      |
| RESP | respectful                           |
| S    | subject marker                       |
| SB   | subject marker                       |
| SFP  | sentence final particle              |
| SG   | singular                             |
| SGU  | unit of sentence grammar             |
| SIM  | simultaneous adnominal               |
| SINT | secondary interjection               |
| SIT  | situative                            |
| SOF  | social formula                       |
| SUB  | subordinate                          |
| SUBJ | subjunctive                          |
| SUF  | suffix                               |
| T    | transitivizing suffix                |
| TAG  | question tag                         |
| TMA  | marker of tense, modality, or aspect |
| TOP  | topic marker                         |
| V    | verb                                 |
| VEN  | venitive                             |
| VOC  | vocative                             |

# 1

## Interactives

One short interjection may be more powerful, more to the point, more eloquent than a long speech. In fact, interjections, together with gestures, the movements of the muscles of the mouth, and the eye, would be quite sufficient for all purposes which language answers with the majority of mankind.

(Max Müller 1861: 410)

### 1.1 Introduction

In many monographs of linguistics a view is expressed or implied that is described succinctly by Givón (1984) in the following way:

Most languages display this mixed-bag category with expressions such as ‘yes’, ‘no’, ‘hey’, ‘oh’, ‘hi’, ‘wow’, ‘ouch’, etc. or their functional equivalents. It is not a unified category functionally, morphologically or syntactically and it is highly language specific.

(Givón 1984: 84)<sup>1</sup>

The goal of the present study is to take issue with this view.<sup>2</sup> It is argued that there is in fact such a category and rather than being a ‘mixed bag’, it can be described across languages as a grammatical category of its own. This category includes but is not restricted to the items mentioned by Givón, which are referred to here summarily as interactive forms or, in short, as *interactives*.

Consider the following conversation that took place in Tjwao, a language spoken in western Zimbabwe.

- (1) Tjwao (Khoë (‘Central Khoisan’); Andrason, Fehn, and Phiri 2020: 14, (17))  
A: **Yii!**  
B: **A-a!**  
A: **Yee!**  
B: **Ehe!**

<sup>1</sup> A more differentiated view is found in the revised edition of Givón’s introduction to syntax (Givón 2001: 102).

<sup>2</sup> Throughout this book, interactives are printed in bold.

The conversation consists entirely of interactives, that is, of the kind of expressions mentioned by Givón, called ‘interjections’ by the authors of (1). The conversation begins with person A meeting person B. A calls B, using the attention signal *yii*. Hearing this, B expresses his surprise by means of the emotive interjection *a-a*.<sup>3</sup> Subsequently, A produces the interjection *yee* to voice his excitement and happiness about meeting B. Speaker B experiences the same feeling and expresses this by means of the response signal *ehe*!.<sup>4</sup>

Examples of the kind illustrated in (1) are not isolated cases. Moving from Africa to Australia, cases such as the one in (2) can be found, once again consisting entirely of interactives. The excerpt of a conversation in (2) is taken from the Mayali language of Western Arnhem Land. It contains two directives (*boi!*, *njudj!*) and two instances of the interjection *njonj-njonj!*.

- (2) Mayali (Gunwinyguan, Arnhem; Evans 1992: 227, (3))

A: **Boi!**

Hey, come here!

[Old woman to youth:] ‘Hey, come here!’ [Old woman holds up baby]

B: **Njonj-njonj!**

**Njudj!**

**Njonj-njonj!**

What a little sweetheart! Blow your nose! What a little sweetheart!

[...]

‘Isn’t she a little sweetheart! Blow your nose! Isn’t she cute!’

Such examples are not restricted to ‘interjections’, they may as well involve other kinds of interactive forms. In the following example taken from Baka, the language of a traditional hunter-gatherer people of Southeast Cameroon, there is a ‘mini-narrative’ consisting of a string of six ideophones, ‘which evokes in the hearer the illusion of a direct participation’ (Kilian-Hatz 2001: 157).

- (3) Baka (Ubangi, Niger-Congo; Kilian-Hatz 2001: 157)<sup>5</sup>

**wòàwòàwòà**

**pèðð**

**kung**

**wóòò**

hunters.discuss

chimp.interrupts.eating

spear.strikes.chimp

falls down

**pao**

**tung.**

breaks.a.branch

falls.hard.on.ground

<sup>3</sup> In accordance with a distinction proposed by Caffi and Janney (1994: 328–9), interjections are treated here as belonging to emotive rather than emotional communication—pending further research. According to these authors, emotional communication is a ‘type of spontaneous, unintentional leakage or bursting out of emotion’, while emotive communication is ‘inherently strategic, persuasive, interactional and other-directed by its very nature’ (see Stange 2016: 29 for further discussion).

<sup>4</sup> Throughout this book, the terms ‘speaker’ and ‘hearer’ are used to refer to interlocutors in conversational exchanges. The terms are occasionally extended to writers and readers, respectively, but our focus is on spoken rather than written discourse. Note that in more than 90% of the languages of the world there is only spoken discourse, while there is essentially no ‘natural’ language in the world that disposes only of written discourse (ignoring special cases like the classical languages Latin or Sanskrit).

<sup>5</sup> The author does not provide a translation of this text piece but it seems that the meaning is recoverable from the glosses, apparently describing a chimpanzee hunt.

But interactives are not restricted to usage among humans—they are also found cross-linguistically in the interaction with animals.<sup>6</sup> Take the utterance in (4) from the Arusa Maasai language of northern Tanzania: A man calls for his donkey's attention (*kuk*), instructs it to move forward (*mape*), and expresses his surprise and annoyance (*fie*) as he did not expect the animal to slow down and stop moving. The utterance is composed entirely of interactives, where *kuk* and *mape* are directives dedicated specifically to interact with animals while *fie* is an interjection, signaling the emotive state of the speaker.

- (4) Arusa Maasai (Eastern Nilotic, Nilo-Saharan; [Andrason and Karani 2021b: 31](#))<sup>7</sup>

**Kuk mape, fie!**

DIR DIR INT

'*Kuk* [hey, said to a donkey] *mape* [let's go], ah!'

Moving on from somewhat 'exotic' languages and forms of communication, consider now the following conversational exchange in a better-known language—a ritual exchange that most of us are familiar with in some form or other. Like the preceding examples, (5) consists entirely of interactives, namely the greeting forms *hello* and *hi* and the interjection *oh*.<sup>8</sup>

- (5) English ([Levinson 1983: 311](#))

C: (C causes telephone to ring)

R: **Hello**

C: **Hi**

R: **Oh hi::**

The example in (5) presents the opening section of a telephone conversation as it can commonly be observed between speakers of English. Being part of the overall organization of a telephone call, it may somehow constitute a self-contained interaction, consisting of an exchange of greeting tokens between two persons, C and R. But the information conveyed appears to be more complex, as suggested by [Levinson \(1983: 311–12\)](#): While *hello* is a greeting token, speaker R does not seem to use it as such in this situation; rather, *hello* is simply a response to C's phone call and a display for recognition of R's identity. C's subsequent use of *hi* then functions in fact as a greeting token, and this token is returned by R with the greeting token *oh hi::*. The latter, however, is not restricted to greeting but in addition signals that R has recognized C.

<sup>6</sup> See Section 3.11 for a detailed discussion of animal directives.

<sup>7</sup> The gloss 'DIR' stands for animal directive forms. The authors [Andrason and Karani \(2021b: 3\)](#) use 'CAC' instead, which stands for 'conative animal call'.

<sup>8</sup> Both the example in (5) and its discussion are based on the analysis by [Levinson \(1983: 311–12\)](#). We will return to the example in Section 7.2.1.



The examples presented so far consist entirely of combinations or sequences of interactives but, more commonly, interactives are found co-occurring with other pieces of text, as in the following exchange, based on an actual experience of the author:<sup>9</sup>

- (6) English (Wharton 2003: 41, (4), (5))  
 Patient: Be careful with that needle!  
 Dentist: **Oops.**  
 Patient: **Ouch!**  
 Dentist: **Hell! I'm sorry.**  
 Patient: **Shit!** Get the bloody thing out of my cheek!

What all these examples suggest is, first, that interactives have, or may have rich social meaning content. Second, social exchanges like the ones in (1) to (6) can consist largely or entirely of interactives. Third, interactives can and frequently do constitute utterances or conversational turns of their own—if I say in English *Oops!*, *Ouch!*, *Hell!*, or *Bang!* then these are commonly accepted as self-contained utterances.<sup>10</sup> Fourth, interactives express conventionalized meaning contents that are stored and retrieved like other forms of a language (Section 1.4). And finally, the fact that interactives can occur as distinct utterances, and that they can be combined and arranged to function as self-contained pieces of discourse, would seem to suggest that they somehow have features of an independent domain of linguistic communication.

Such observations have not really overawed students of linguistics (Section 6.1). Way back in the nineteenth century, well-known linguist Müller (1861: 346, 352) observed that interjections ‘are playthings, not the tools of language’ and that ‘language begins where interjections end.’ Benfey (1869: 295) concluded in his history of linguistics that interjection is ‘the negation of language’, and for Sapir (1921: 6–7) interjections were at best a ‘decorative edging to the ample, complex fabric’ of language and ‘the nearest of all language sounds to instinctive utterance’.

The situation has not changed dramatically in the ensuing history of linguistics; suffice it to mention the influential study by Goffman (1978: 809–10), where the author maintains that ‘response cries’, that is, the interjections studied by him, are a variety of ‘non-words’ and as such, ‘can’t quite be called part of a language.’

In the course of the last decades, linguistics has been enriched by a plethora of reference grammars of languages, many of which are spoken in the remotest corners of the world. These grammars, mostly based on years of field work, describe

<sup>9</sup> Tim Wharton (p.c. of March 5, 2022).

<sup>10</sup> An ‘utterance’ is taken here to stand for a piece of self-contained text produced by a particular individual on a particular occasion for a particular purpose. Our concern here is mainly, though not exclusively, with spoken utterances. Such a piece frequently is, but need not be a sentence.

the grammatical structure of the language concerned in great detail and provide the linguist with goldmines of information. But somewhat surprisingly, a substantial part of these grammars contains no mention of interactives, and many other grammars deal with interactives at best in a few casual notes or as a kind of short appendage.

To mention another kind of example: The Korean language is especially rich in interactives, disposing of an inventory of over 5000 ideophones (Section 3.5.1)—the number in fact comes close to that of nouns or verbs, and Korean ideophones have a remarkable impact on the semantic and grammatical structure of the language. But the reader may be surprised to find no description of ideophones in the influential reference grammar of Korean by [Sohn \(1999\)](#). Ideophones are mentioned in a section on sound symbolism ([Sohn 1999](#): 96–102) but—again perhaps surprisingly—many of the ideophones have no conceivable relationship to sound symbolism.<sup>11</sup>

For someone who wants to learn a language, these grammars may therefore be somewhat disappointing. After having battled one's way through the 400 to 800 pages of these grammars, having internalized the sentences and other structures of the language, one may still have the feeling of not being 'communicatively competent', knowing how to greet, be polite, exhort, warn, beseech, surprise, call, persuade, cooperate, disapprove, challenge, or entertain others—in short, to do the kind of things one would feel obliged or like to do when being confronted with the community speaking the relevant language.

Such disregard for language as a tool of social interaction is by no means restricted to the tradition of grammar writing just alluded to; it is quite common in all kinds of linguistic publications and schools of linguistics. Suffice it to mention a paradigm example: One of the most comprehensive reference grammars of English, *The Cambridge Grammar of the English Language* ([Huddleston and Pullum 2002](#)) discusses most aspects of English grammar in great detail, but interactives are blatantly absent. While at least interjections are classified as one of the nine 'lexical categories' of English, only a dozen lines are devoted to them in a work comprising 1842 pages. The attitude of one of the two authors of this reference grammar is nicely reflected in the following remark:

Interjections are so unimportant to the fabric of the language that they are almost completely ignored in grammars. There is almost nothing to say. They have no syntactic properties at all, you pop one in when the spirit moves you. And their basic meaning is simply expressive of a transitory state.

(Geoffrey K. Pullum, January 2005 post. Quoted from [Ameka 2020](#): 57)

<sup>11</sup> In his earlier grammar of Korean though, [Sohn \(1994\)](#) had a whole chapter devoted to ideophones and interjections (Chapter 4).

This remark is reminiscent of some views found in nineteenth-century linguistics—for example, when Benfey (1869: 295) concluded in his history of linguistics that ‘interjections are employed when one either cannot or will not speak.’

Overall, a change in attitude can be observed in the history of grammar writing. Prior to the mid twentieth century, authors working on the languages of Africa, Asia, Australia, or the Americas were likely to at least mention some of the social functions of the languages studied by them. But with the rise of linguistic structuralism in all its many versions, an interest in language as a tool of social interaction declined and grammatical description became largely restricted to apprehending the structural properties to be observed in the categories of sentence grammar.

Nevertheless, in the 1990s, interactives were the subject of a range of seminal studies, revealing all the richness of the semantic space that interactives present (e.g., Ameka 1992a, 1992b; Evans 1992; Wierzbicka 1992; Wilkins 1992). These studies had some influence on ensuing discussions on meaning (see Section 2.1). However, the studies had no major impact on the tradition of grammar writing or of linguistic analysis in general, as witnessed, for example, in the remark by Geoffrey Pullum just quoted.

Underlying much of mainstream linguistics up to the present there appears to be a widespread assumption to the effect that there is no grammar to the language of social interaction or, even if there is, it is not worthy of scholarly attention. On this view, interactives are not clearly part of language, conceived somehow as ‘stowaways’ to be located somewhere between language and something else: They have been claimed to be peripheral to the language system, being located at the boundary between verbal and non-verbal behavior, or being paralinguistic and asystematic elements (see Section 6.1). Accordingly, they are called ‘response cries’, ‘non-words’, ‘semi-words’, ‘quasi-words’, or ‘partly natural, partly conventional semi-words’ (see Ameka 1992a: 112; 2006: 744–5; Nübling 2004; Ameka and Wilkins 2006: 5; Stange and Nübling 2014: 1986; Wharton 2016: 21), or as Ferguson (1981: 21) puts it, ‘the little snippets of ritual used in everyday encounters’.

Ideophones have been called ‘a step-child of modern linguistic science’ (Voeltz and Kilian-Hatz 2001: 2) and have been the subject of ‘a recurrent narrative of marginalisation’ (Dingemanse 2018: 2). In a similar way, O’Connell and Kowal (2008: 133) characterize interjections as ‘a phenomenon that has been historically a marginal, thoroughly neglected linguistic category’.

In sum, interactives, like interjections, ideophones, social formulae, discourse markers, etc., have been either neglected or ignored in most previous linguistic work—be that descriptive or prescriptive work. Understandably, this attitude is deeply deplored by students of interactives.

The goal of the present study is to argue that this traditional attitude is in need of reconsideration. Rather than being peripheral, or a marginal phenomenon, interactives are described as a category on its own, contrasting with what [Haiman \(2018\)](#) calls ‘prosaic’ or ‘propositional’ grammar and what is referred to here as ‘sentence grammar’. Whereas sentence grammar has been described as having denotational, informational, conceptual, descriptive, truth-conditional, or objective functions, interactives tend to be portrayed as serving connotative, or expressive functions. Perhaps most commonly they are referred to as expressives, that is, as verbal means having an expressive rather than a referential function ([Foolen 2016: 473](#)). The reason for calling them here ‘interactives’ rather than ‘expressives’ is that the latter term has frequently been applied also to a range of phenomena far beyond those that are covered by the definition of interactives to be provided in the next section.<sup>12</sup>

After defining interactives in Section 1.2, the properties contained in the definition are discussed in Section 1.3. Section 1.3 is meant to demonstrate that the grammar of interactives differs from that of sentence grammar in a principled way, while Section 1.4 shows that the two nevertheless share a common base. An inventory of the ten types of interactives distinguished is provided in Section 1.5 with English examples, and Section 1.6 then concludes the chapter with an outline of the content of the book.

## 1.2 Definition

What do vocative forms like *Mom!* or *Sir!* possibly have in common with interjections like *oops* or *yuck*, or discourse markers like *indeed* or *um*? One important answer is provided in this section: They all conform to the same definition.

Interactive forms or, in short, interactives provide insights into how speakers conceive themselves in the world of social communication. They are prefabricated routine forms and include but are not restricted to what [Ferguson \(1981\)](#) describes as social formulas and [Coulmas \(1981: 2–3\)](#) as conversational routines, that is, ‘highly conventionalized prepackaged expressions whose occurrence is tied to more or less standardized communication situations’. They are ‘interactive’ in the sense that they are grounded in social interaction and we propose the prototypical definition in (1) for them.

- (1) An interactive is an invariable deictic form that is in some way set off from the surrounding text semantically, syntactically and prosodically and can neither be negated nor questioned.

<sup>12</sup> For an overview of research on expressives see [Foolen \(1997, 2012, 2016\)](#).

It goes without saying that like other linguistic elements, interactives do not occur in isolation, sealed off from the rest of the text of which they are a part. Rather, they may show features of assimilation to their textual environment, such as traces of prosodic and other kinds of integration.

The definition in (1) makes no mention of the functions of interactives, for the following reason: As we will see in Table 1.1, interactives include a number of different types, and each type has its own functional focus. The reader is therefore referred to Chapter 3 for information on the functions of the various types of interactives (see Table 3.1).

The properties mentioned in (1) are discussed and looked at in more detail in Section 1.3.1. In addition, a number of other properties have been reported and are looked at in Section 1.3.2. The term ‘interactive’ is used here in a more general sense in that it includes both the ‘expressives’ and ‘interactives’ distinguished in the framework of Functional Discourse Grammar (Hengeveld and Mackenzie 2008: 77),<sup>13</sup> it largely corresponds to the term ‘insert’ of Biber et al. (1999; see Section 5.1). It includes but is not restricted to the ‘interactional structure’ of Wiltschko (2021; see Section 5.2.4). However, it does not include what Ogi (2017) calls ‘interactive markers’ in Japanese since the latter lack salient features of interactives in the sense of the term used here, such as distinct syntactic and prosodic status.

Following other authors dealing with specific categories of interactives, such as Evans (1992: 22), Gehweiler (2008), Andrason and Dlaki (2020: 164) on interjections or Childs (1994) on ideophones, we propose (1) to be a *prototypical* definition. Thus, rather than as a discrete category, interactives are assumed here to instantiate a prototype of linguistic forms. In doing so, we are relying on the notion of prototype theory as developed in the tradition of Rosch (1973), ignoring more specific problems and theoretical issues raised, for example, in work on family resemblance (cf. Wittgenstein 1953) or on radial categorization (Lakoff 1987). Prototypes can be characterized in the following way (e.g., Taylor 1989): (a) They are categories that are not defined in terms of a set of necessary and sufficient attributes. (b) Category membership is graded: Some members are better instances of the category than others. (c) The most central members function as cognitive reference points of the category, being more representative of the category than peripheral members.

In the framework proposed here, ‘members’ are instantiated by individual interactives. Interactives showing the whole set of properties are the most prototypical

<sup>13</sup> What distinguishes ‘expressives’ and ‘interactives’ in the framework of Functional Discourse Grammar (Hengeveld and Mackenzie 2008) is that unlike the former, the latter are ‘clearly directed to the Addressee’. As argued by Lachlan Mackenzie (p.c. of February 17, 2022), expressives ‘are like involuntary cries of pain, joy, etc., i.e. they are not inherently communicative’.

We will return to the distinction between presence and absence of addressee or, ‘hearer’ as we say here, in Section 2.2.

or ‘central members’ of the category, and they are the main concern of the paragraphs to follow.

The term ‘deictic form,’ or ‘indexical,’ refers to the fact that interactives *are immediately anchored to the situation of discourse*.<sup>14</sup> The term relates to the notion ‘discourse-deixis’ of Weinreich ([1966] 1989: 69), which is said to provide the cognitive frame used by interlocutors to design and interpret spoken or written texts; it concerns ‘the presuppositions about discourse context, or “discourse placedness conditions”’ as Evans (1992: 228) puts it with reference to the nature of his organizing interjections. Being semantically, syntactically, and prosodically separated, interactives can be ‘stand-alones.’ They can and quite commonly do form what has been referred to as ‘situation bound utterances’ (Felix Ameka, p.c. of July 25, 2021).

According to the definition in (1), a linguistic form that does not have the whole set of properties is not a prototypical member of the category, and the main concern of this study is with prototypical, or core members of the category—in other words, with forms that exhibit all the properties in (1). As we will see in Section 1.3.1, however, some properties are not entirely stable. Deviations from the prototype are most of all of two kinds. On the one hand, an interactive may have a set of variants, some of which are invariable while others are not. On the other hand, by being anchored to the situation of discourse rather than to the structure of a clause, interactives are highly context-sensitive and tend to exhibit a number of context-related usages where not all usages conform to the properties in (1).

We will not generally discard less prototypical instances as far as they concern problems with only one or two of the properties. If, however, central issues on the nature of interactives are at stake we will insist that all properties be honored in the discussions to follow.

Interactives, as they are defined in (1), include a range of different elements, and these elements can be classified into ten types, namely the ones listed in Table 1.1 (see Section 1.5 for more information). It is these types that will be the concern of the chapters to follow, and in Chapter 3 we will deal with each of them in greater detail.

As the classification proposed in Table 1.1 shows, the term ‘interjection’ is used here in a narrower sense than in many earlier classifications, where types such as attention signals, directives, response elicitors, and/or response signals were frequently subsumed under the heading ‘interjection.’ The reason for this discrepancy will be discussed in Section 3.6.

<sup>14</sup> The terms ‘deixis’ and ‘indexicality’ can be traced back each to different histories of research traditions and tend to be associated with slightly different references, the former being prevalent in linguistics and the latter in philosophy and pragmatics (see also Section 2.1.3.1). In the present study, ‘deictic form’ and ‘indexical’ are both treated as referring to linguistic elements whose meaning is immediately anchored to the situation of discourse and, hence, must be interpreted with reference to the ‘context’ in which the relevant linguistic discourse takes place.

**Table 1.1** The main types of interactives

| Type               | English example |
|--------------------|-----------------|
| Attention signals  | <i>Hey!</i>     |
| Directives         | <i>Shh!</i>     |
| Discourse markers  | <i>Indeed.</i>  |
| Evaluatives        | <i>Great!</i>   |
| Ideophones         | <i>Thud!</i>    |
| Interjections      | <i>Ouch!</i>    |
| Response elicitors | <i>Right?</i>   |
| Response signals   | <i>No.</i>      |
| Social formulae    | <i>Bye!</i>     |
| Vocatives          | <i>Mom!</i>     |

The ten types distinguished in Table 1.1 differ greatly from one another; suffice it here to draw attention to some salient features distinguishing types of interactives. First, there are differences in the role played by context in the use of interactives. For example, whereas discourse markers, like *indeed* in Table 1.1, relate to the surrounding text, this is not clearly the case with some other types, especially with interjections (e.g., *ouch*). Second, interactives like *shh*, *thud*, or *wow* are associated with a pronounced exclamational prosody; some other forms again are prosodically clearly less strongly marked. Third, response elicitors (e.g., *right?*) require an interrogative intonation while this is not the case with other types of interactives. Fourth, ideophones (e.g., *bang*, *plop*, *thud*) differ from other interactives in a number of ways, but we will see in Section 3.5.7 that there are reasons to maintain that they as well belong to the category. Fifth, all types can be used as stand-alones, that is, as utterances of their own, even if this is rarely the case with discourse markers.

Sixth, interactives are typically closed-class forms, but this does not apply to vocatives (e.g., *mom*, *darling*, *Sir*, *Madame Moneyppenny*), and in some languages also to ideophones (Section 3.5). Seventh, some forms are transparently related to corresponding forms of sentence grammar (e.g., *great*, *indeed*, *right?*) while others are not (e.g., *hey*, *shh*, *wow*) (see Section 3.6.4.1 for discussion). And finally, one might hesitate to call interjections ‘interactives’ considering that they can well be uttered in isolation without any other interlocutor being present. The main reason for extending the term to interjections is that in most of their occurrences in texts they involve or imply interactive situations.

One may also wonder whether the English examples in Table 1.1 are really appropriate instantiations of the ten types proposed. Take the form *indeed*: It has in fact been described as a discourse marker (see Traugott 1995; Brinton 2008: 246) but, like all the other forms listed, *indeed* has a wide range of usages, and

depending on the context in which it occurs its meaning may vary, having in some contexts more in common with a response signal like *yes* than with discourse markers.<sup>15</sup> Nevertheless, the examples in Table 1.1 have been found to show usages that justify treating them as representing one and the same category, namely that of interactives.

Most of the types listed in Table 1.1 are in many studies referred to summarily as ‘interjections’, in some other studies as ‘interjectives’, ‘discourse markers’, or ‘exclamations’, or something else (see Chapter 5.1). It would seem, however, that such a broad understanding of terms like ‘interjection’ or ‘discourse marker’ is not without problems. The various types differ from one another not only in their functions, their usage, etc. but also in the way they are used in structuring discourse as we will see in the chapters to follow.

To conclude, all the forms listed in Table 1.1 contrast with elements of sentence grammar in that they conform essentially to the definition in (1), forming a domain of their own—one that we propose to call ‘interactive grammar’. As we will see later in this chapter (Section 1.4), however, there are also features that the two grammars have in common.

## 1.3 Properties

In discussing the properties of interactives, our concern is with languages in different parts of the world. The overview to be presented in this section therefore is based on a comparative perspective, resting on the analysis of all the data that we were able to access.

### 1.3.1 Definitional properties

Some speakers of English may hesitate to include an item like *fuck* in their vocabulary, yet the item is fairly widely used—more than one might be inclined to appreciate or to believe. It has many different uses as a swearword; Mackenzie (2019) has devoted a study to them and he concludes that they are classified best into the five kinds of constructions illustrated in (1).

- (1) A taxonomy of uses of the English swearword *fuck* (Mackenzie 2019: 61)
  - a Literal Representational use: *I fucked her.*
  - b Single Discourse Act use: *Fuck!*
  - c Metaphorical Representational use: *They fucked me over.*
  - d Lexical Substitution use: *good as fuck*
  - e Expletive use: *the fucking towel*

<sup>15</sup> See Chapter 4.2 for an explanation of this fact.



With one exception, all these uses suggest that *fuck* is part of more or less pre-fabricated constructions that fall squarely within the domain of sentence grammar. Thus, the item is part of the clause in which it occurs—be that as a verb (cf. (1a), (1c)), a nominal (cf. (1d)), or an adjectival (cf. (1e))—in short, they do not conform to the definition of interactives in Section 1.2.<sup>16</sup>

This is different with the one exception in (1b), where *fuck*—in the wording of Mackenzie (2019: 64)—occurs ‘as the sole expression of a Discourse Act’. In this use, *fuck* is commonly classified as an interjection, and hence, as an interactive (e.g., Ameka 1992a; Huddleston and Pullum 2002: 1361; Norrick 2009). Accordingly, this use is in accordance with the definition of interactives proposed in Section 1.2: *Fuck* in (1b) has the following properties: (a) It is invariable, e.g., not taking inflectional or derivational affixes. (b) Its meaning is not a semantic part of any clause it may be associated with. (c) It is syntactically unattached, being commonly used as an utterance of its own or inserted within an utterance. (d) It is typically set off prosodically from surrounding text material. (e) Its use can only be interpreted meaningfully with reference to the here-and-now of the situation of discourse. (f) It cannot be negated—that is, if it is negated, it is no longer an interjection (e.g., *Don’t fuck!*). And (g), as an interjection it cannot normally be turned into a question.

To conclude, of all the uses of *fuck* in (1), it is only the one used as an interjection in (1b) that is within the scope of the present study: Only this use conforms to the definitional properties proposed in Section 1.2. We will now look at these properties in more detail.

### 1.3.1.1 Morphology

Dixon (2010b: 5) defines a grammatical word as an ‘inflected form of a lexeme’, and for Zwicky (1985: 288), words ‘frequently are morphologically complex, in the sense that they are to be analysed as being composed of two or more morphemes’ (see Section 6.2). Neither of these characterizations applies to interactives. Interactives have been characterized as *holophrases* (cf. Mackenzie 1998). While expressing complete, at times complex meaningful thoughts, they normally lack both external and internal morphological complexity, being unanalyzable and as a rule short grammatical expressions devoid of any inflectional or derivational elements. This generalization was already made by Biber et al. (1999: 1082) when they observed that their inserts, a category largely corresponding to that of interactives, are morphologically simple and unanalyzable.

<sup>16</sup> In Functional Discourse Grammar (FDG), the expletive *fuck* exemplified by (1e) is ‘Interpersonal’. Thus, ‘the distinction between interactive and sentence grammar is not isomorphic with the FDG distinction between Interpersonal and Representational meanings’ (Lachlan Mackenzie, p.c. of February 17, 2022).

For Wilkins (1992: 124), an ‘interjection’ is usually monomorphemic and ‘does not host inflectional or derivational morphemes’, and much the same observations have been made on other types of interactives. According to Childs (1995: 132), ideophones ‘have no morphology’, and Haiman (2018: 104, 265) says that ideophones are holistic (unparsable) signs that exhibit no internal syntax. And seemingly complex social formulae are usually unanalyzable forms, allowing no internal modification. For example, *How do you do?* is a paradigm English greeting formula, while neither *How will you do?* nor *How does he do?* are. And with reference to English interactives like *thank you*, *excuse me*, and *God almighty*, Biber et al. (1999) write:

They also cannot be easily varied: consider, for example, non-existent variations such as *?\*thank me*, *\*thank you a lot* (contrast *thanks a lot*), *?\*excuse me a little* (non-occurring as an apology), *?\*almighty God*, non-occurring as an expletive, as contrasted with the irregular *God almighty*).

(Biber et al. 1999: 1082)

Similar observations have been made in other languages. Based on findings in languages from different parts of the world, Andrason and Dlaki (2020: 166) note that ‘interjections’ are holophrastic units which ‘do not exhibit inflectional and derivational marking’, thus resisting processes of inflection and derivation otherwise operating in a language (see also Ameka 2006: 743; Ameka and Wilkins 2006: 5; Nübling 2004: 29; Velupillai 2012: 149). Interjections in the Paresi language of Matto Grosso in Brazil do not take affixes and are generally monomorphemic (Brandão 2014: 343), as are interjections (‘exclamations’) in the Hausa language of Northern Nigeria: They are typically ‘invariant and are used as full expressions in and of themselves’ (Newman 2000: 176).

Being morphologically unanalyzable is a property that is in fact pointed out in a number of studies of languages across the world. In Diyari, a Pama-Nyungan language of South Australia, there is a category of ‘interjections’ which in addition to interjections proper includes social formulae like the greeting *adu* ‘hello’, or the attention signal like *ayi* ‘hey’, thus representing three different types of interactives. Forming one of the six parts of speech, on the same level with nominals, pronouns, verbs, etc., they can comprise whole utterances by themselves, and are not inflected nor syntactically integrated with other linguistic material (Austin 1981: 36).

Note, however, that some interactives allow for variation, that is, instead of one invariable form there may be a set of functionally largely equivalent variants. For example, instead of saying *sorry!*, English speakers can draw on a range of other forms, like *I’m so sorry*, *I’m really sorry*, *Gee, I’m sorry*, or *I’m terribly sorry* (Tannen and Öztekin 1981: 38). And some interjections (or swearwords) can take an argument, as in *blast (it)!*, *(you) bastard!*, *fuck (you)!*, *screw you!*, or some other

modifier.<sup>17</sup> One may also wish to mention that within such sets, the invariable form conforming to the definition in (1) need not be the most salient or the most frequently used variant of the set.

Another kind of variation can be found in ideophones, which have been shown to have their own reduplicative morphology (see [Dingemanse 2017a](#)). However, we are not aware of any language where ideophones exhibit productive patterns of inflection and derivation; see Section 3.5.7.

Furthermore, in our data bank there are examples such as the following. Speakers of the Mayali language of Western Arnhem Land, Australia, use the expression *gebnguneng* to say ‘thank you’, which means literally ‘I ate your nose’. The meaning of the expression falls squarely within the type of what we propose to call social formulae (Section 3.9). However, as [Evans \(1992: 226\)](#) shows, the expression is transparently composed of a number of morphemes, having the structure *ø-geb-ngu-neng* (I/you-nose-eat-PAST.PFV). For this reason, [Evans \(1992: 226\)](#) concludes that the expression does not qualify as an interjection. To the extent that Mayali *gebnguneng* is conceived by its speakers as a uniform expression that is independent of its origin as a morphologically complex text piece, there would be reason to classify it as an interactive. The data provided by Evans, however, are not detailed enough to decide on this issue. One might wish to know, for example, whether Mayali speakers could extend this formula to also construct ‘I ate his nose’ or ‘I didn’t eat your nose’ for a similar functional purpose. In the absence of more detailed information, we prefer not to include cases such as *gebnguneng* within the category of interactives.

Among the different types of interactives there is one that is especially problematic, namely that of vocatives. Vocatives can be extended in various ways: They can be composed of a prefix and a last name (*Ms. Pandit*), a title and a first name (*Lady Diana*), the personal pronoun *you*, or a reduced relative construction (***You with the sweater on**, move about a foot to the left*), and there may be complex vocative forms like *Ladies and gentlemen*, *dear friends*, *dear children*, *dear lovers of classical music*, etc. Still, [Zwicky \(1974: 788, 791\)](#) argues that most vocative NPs are ‘idioms’. In accordance with our definition of interactives in Section 1.2, the concern of this book is exclusively with ‘idiomatic’, largely or entirely invariable vocatives, such as *darling*, *doctor*, *honey*, *sir*, *sweetie*, etc.

In the framework of Functional Discourse Grammar, the categories of ‘expressives’ (interjections) and ‘interactives’ (social formulae)—both being interactives in accordance with our definition in Section 1.2—have in common that they are invariable, in that an English formula like *Congratulations!* must be used

<sup>17</sup> We follow Neal Norrick (p.c. of July 7, 2021) in assuming that expressions like *fuck you*, *fuck yourself*, or *go fuck yourself* are perhaps better treated as invective formulas than as interjections. Furthermore, he draws attention to a similar variation of expressions in German: *leck*, *oh leck*, *leck mich*, *leck mich am Arsch* (‘lick, oh lick, lick me, lick my arse’).

in the plural, that is, there is no corresponding singular form *\*Congratulation!* (Hengeveld and Mackenzie 2008: 77).

Thus, there are a number of ways in which interactives may deviate from the format of invariable expressions. On the whole, however, most studies converge on describing interactives, such as interjections or ideophones, as holistic units having no internal morphology.

In concluding, mention should be made of the following important point. Kockelman (2003: 470) claims that interjections ‘cannot be further derived into another form class such as noun or verb’. That this claim is incorrect can easily be shown with English interjections like *ouch*, *wow*, *yuck*, or *yum*, which have acquired usages as verbs, adjectives, or nouns (cf. *to ouch*, *an ouchie*, *to wow*, *yucky*, *yummy*, etc.). The reason for this is that these interjections have undergone grammaticalization. In such usages they have turned into words of sentence grammar, and in the latter capacity they exhibit morphological features of the new word class they have joined; we will deal with this issue in Section 4.2. This fact, however, in no way contradicts the generalization proposed in this section: Interjections, such as *ouch* or *wow*, are invariable and once they are grammaticalized into verbs, adjectives, etc. they cease to be interjections, joining the relevant word classes of sentence grammar.

### 1.3.1.2 Meaning

The meaning of interactives may relate to that of the clause they are associated with but it does not normally modify the meaning of that clause, it has been portrayed as being located in a separate space of discourse processing. What Haselow (2016a: 82) writes with reference to discourse markers applies also to other types of interactives: They do not change the propositional content of a clause.

Dwyer and Moshi (2003: 177) suggest that interactives like ideophones belong to the expressive dimension of semiotics, having a different semantics than elements belonging to the analytic dimension:

[...] as expressives they seek to represent an event by evoking sensory or emotional image of it as opposed to breaking an event down into components of agent, action and the like. Expressives can evoke images from any of the five senses, seeing, hearing, touching, tasting, and smelling but in addition they also call on other sorts of feeling like fear, happiness, nausea, heat, cold, and so forth [...]. This characterization of expressives helps to explain why the meanings of expressives (including ideophones) are context dependent and why they are so elusive to define.

(Dwyer and Moshi 2003: 177)

In a similar way, Kita (1997: 379) distinguishes two kinds of semantic dimensions in his analysis of Japanese ideophones (referred to by him as ‘mimetics’). One dimension is called the analytic dimension, where meaning is represented as a hierarchical structure of decontextualized semantic primitives. The second dimension, called the affecto-imagistic dimension, represents meaning ‘in terms of affect and various kinds of imagery’, it relates to what in research on interactives is called the expressive function of language.

When Haiman (2018: 135) submits that ideophones and interjections are ‘dramatic performances of some kind’ then this applies to most interactives, with some possible exceptions, like discourse markers. Interactives, such as English *goodbye*, *hurray*, *oops*, and *ouch*, have been called ‘expressives’, that is, they are said to have use-conditional rather than truth-conditional content, as has been argued in some lines of philosophical and formal-semantic research (e.g., Kaplan 1999; Gutzmann 2019; Gutzmann and Henderson 2019; see also Dwyer and Moshi 2003: 174 on ideophones; see Section 2.1.2).

What according to Goffman (1978: 814) applies to interjections (his ‘semi-word response cries’) appears to also apply to a number of other kinds of interactives: They ‘are creatures of social situations, not states of talk’. And with reference to ideophones, Haiman (2018: 78) notes that ‘what conventional language *tells* us, an ideophone—exactly like a performative gesture—*shows* us’. The affinity of interactives to gestures is also pointed out in some other studies. Distinguishing between a descriptive and a mimetic mode of language use, Güldemann (2008: 291) suggests that, like non-linguistic gestures, ideophones operate in the latter mode.

Ideophones, as well as other interactives, are characterized as having depictive rather than descriptive meaning, or expressive rather than ‘plain’, ‘ordinary’, or ‘prosaic’ meaning (Dingemanse and Akita 2017: 505). While this characterization seems to be shared in some form or other by a number of researchers, the terms ‘expressiveness’ and ‘depiction’ have received a number of different interpretations, standing for notions such as ‘foregrounding’, ‘attracting attention’, ‘indicating noteworthiness’, ‘expressing emotional content’, or ‘framing something as a depictive performance’.<sup>18</sup>

Furthermore, interactives have been argued to have no denotative meaning, ‘their use is defined rather by their pragmatic function’ (Biber et al. 1999: 1082; but see also Section 2.1). They are not propositionally organized, their meaning can only be described in terms of the situation of discourse in which they are uttered rather than in terms of the structure of clauses (cf. Ljung 2011: 76); they usually have no bearing on the illocutionary force of their host (cf. Andrason, Fehn, and Phiri 2020: 21).

<sup>18</sup> See Dingemanse and Akita (2017: 505), who say: ‘We define the EXPRESSIVENESS of linguistic signs as the degree to which they are foregrounded as distinct from other items, for instance by special intonational or phonational features. This is in line with the established use of “expressive” as a term that contrasts with “plain,” “ordinary” or “prosaic.”’

Rather than being processed via sentence rules or norms, the use of interactives is contingent on the exigencies characterizing the situation of discourse, being determined by ‘pragmatic’ principles or rules (cf. Zwicky 1985: 304 on ‘discourse markers’) or by what Ferguson (1981: 26) calls ‘appropriateness conditions’.

The observation that interactives can be accessed immediately from the situation of discourse can be said to be captured in the framework of natural semantic metalanguage (Wierzbicka 1992, 2003a) by the difference between (2a) and (2b), where the meaning of the interjection *yuk* is represented simply by the sentence in (2a) whereas the representation of that sentence in (2b) is more complex, requiring an ‘I say’ component.

(2) English (Wierzbicka 1992: 162; boldening and italics in the original)

a ***Yuk!***

I feel disgusted.

b *I feel disgusted!*

I say: I feel disgusted

I say this because I want to say what I feel

The term ‘metatextual’, used here to describe a general function of interactives, is employed in work on discourse markers (Traugott 1995: 6; 2018: 27; Heine et al. 2021: 8) for a level of discourse processing that serves to monitor discourse processing and to provide instructions on how to interpret texts. The similar term ‘metadiscursive’ is also used in some research on text analysis, where the term stands broadly for discourse about discourse (e.g., Hyland 1998, 2005, 2017; Ädel 2006; Mauranen 2010; Zhang 2016).

Since the function of interactives is determined by the situation of discourse, they are highly dependent on the context in which they are used. Thus, in his treatment of ‘interjections’, which also include some other types of interactives, Wilkins (1992) observes that

[...] interjections are the most reduced form an utterance can take, and that the motivation for such reduction is to be found in the functional principle which determines that the more information that is recoverable directly from context the more reduced an utterance will be. Once again this relates to the context-bound nature of interjections; all their referential arguments are provided by extralinguistic context.

(Wilkins 1992: 153)

The range of context-dependent meanings expressed by interactives is in fact said to be richer than that of units of sentence grammar. Wierzbicka (1992: 180) points out that, depending on the context in which it is used, the English interjection *ah* expresses pain, surprise, pity, compassion, complaint, contempt, dislike, joy, and

exultation.<sup>19</sup> Furthermore, six different meanings are commonly distinguished for response elicitors like English question tags (Tottie and Hoffmann 2006: 300–1). Discourse markers have not uncommonly been found to distinguish more than a dozen of different meanings.

We will return to the meaning of interactives in Section 2.1.

### 1.3.1.3 Syntax

Interactives may exhibit various kinds of relations to the surrounding text, such as co-occurrence relations or forms of agreement, parallelism, or antithesis. But interactives are not constituents of syntactic hierarchies as they have been proposed in models of mainstream formal or functional linguistics. As Wharton (2003: 66) writes: ‘Interjections are, as it is often put, “thrown” (interjected) into utterances. They exist on the edges of utterances, always separated off from the main clause and rarely integrated into intonational units.’ Accordingly, Quirk et al. (1985: 853) note concisely with reference to English interjections that they ‘do not enter into syntactic relations,’ and much the same applies to other interactives.

Being syntactically independent, interactives are neither arguments nor adjuncts of a clause, they cannot normally be embedded as complements of a predicate, they cannot be the focus of a cleft sentence, nor can they normally be coordinated.<sup>20</sup> Thus, unlike the grammatical structures in (3), the corresponding structures in (4) do not seem to be well-formed.

- (3) English
  - a Jeff says that he likes cockroaches.
  - b It is cockroaches that Jeff likes.
  - c Jeff and the cockroaches.
- (4) English (# = non-felicitous utterance)
  - a #Jeff says that **wow/please/hi/boing/well**.
  - b #It is **wow/please/hi/boing/well** that Jeff likes.
  - c #**Wow** and **please**.

Evidence that interactives are syntactically unattached can also be seen in the fact that most of them can occur as stand-alones, forming utterances of their own, or occurring at the left periphery of an utterance, being prosodically separated from what follows, as in Table 1.2.<sup>21</sup>

<sup>19</sup> Franck Floricic (p.c. of July 22, 2021) rightly suggests that in addition to context there are also phonetic or intonational cues to distinguish meanings.

<sup>20</sup> The hedge ‘not normally’ draws attention to the fact that there are exceptions. As pointed out by Franck Floricic (p.c. of July 22, 2021), for example, response signals can well be coordinated, as in French *Oui ou non?* or English *Yes or no?*

<sup>21</sup> Rather than ‘periphery’, the term ‘edge’ is preferred in some of the formal linguistics work (e.g., Munaro 2010; Haegeman and Hill 2013).

**Table 1.2** Examples of English interactives in the left periphery slot of an utterance

| Type of interactive | Example                                             |
|---------------------|-----------------------------------------------------|
| Attention signal    | <i>Hey</i> , there is no alien in the fridge.       |
| Directive           | <i>Listen</i> , there is no alien in the fridge.    |
| Discourse marker    | <i>After all</i> , there is no alien in the fridge. |
| Evaluative          | <i>Great</i> , there is no alien in the fridge.     |
| Interjection        | <i>Oh</i> , there is no alien in the fridge.        |
| Response elicitor   | <i>Right?</i> There is no alien in the fridge?      |
| Response signal     | <i>No</i> , there is no alien in the fridge.        |
| Social formula      | <i>Sorry</i> , there is no alien in the fridge.     |
| Vocative            | <i>Darling</i> , there is no alien in the fridge.   |

There is in fact converging evidence that interactives of all types are normally syntactically unattached and that many can form utterances of their own. One of the features that [Biber et al. \(1999: 1082\)](#) use to define inserts, a category that largely corresponds to that of interactives (Section 5.1), is that ‘they may appear on their own, i.e. not as part of a larger grammatical structure’. In a similar way, [Zwicky \(1985: 303–4\)](#) concludes that ‘discourse markers and their kin are syntactically insulated from the rest of the sentences in which they occur’. Note that Zwicky’s category of discourse markers also corresponds largely to that of interactives (Section 5.1). And emphasizing that onomatopoeias (a group of ideophones; Section 3.5.2) need to be distinguished from interjections, [Meinard \(2015: 151\)](#) notes that the two have in common that they are syntactically isolated.

This has also been demonstrated by [Heine et al. \(2021\)](#) for discourse markers and by others for ideophones, which both are said to be ‘[s]et off from the rest of the sentence by a pause’ ([Childs 1995: 133](#)). Drawing attention to ‘the inability of ideophones to permit morphosyntactic integration’, [Haiman \(2018: 104\)](#) says that ideophones ‘are not a part of speech’, being set apart from the rest of the utterance. For problems with the view that interjections are a ‘part of speech’, see also [Meinard \(2015: 160–2\)](#). In a similar way, [Dingemanse \(2017b: 196\)](#) writes that ‘[i]deophones typically display a great degree of syntactic independence. They tend to occur at the edge of the utterance, unburdened by morphology and not deeply embedded in the syntactic structure of the clause.’ And [Evans \(1992: 242\)](#) takes ‘the formal test of syntactic independence as the paramount criterion for defining the class of interjections’.

Rather than in terms of syntax, [Mackenzie \(1998\)](#) discusses a set of English interactives with reference to their grammatical status. Describing social formulae like *Cheers!*, *How do you do?* and *Thank you* (‘fixed expressions’) and response signals like *yes*, *no*, *absolutely*, and *hardly* (‘indications of assent or dissent’) as



holophrases, he concludes that this kind of holophrases ‘can be safely left outside the grammar’ (Mackenzie 1998: 283).

Similar observations have been made in other languages across the world. In their grammar of the Salishan language Thompson (Nlaka’pamuctsin) of British Columbia, Thompson and Thompson (1992: 184–6) describe a category called ‘expletives’ which includes eight of the types of interactives distinguished in Table 1.1 (Section 5.1.1). ‘Expletives’ are defined by the authors thus:

[They occur] independently of [other words], constituting the whole of minor sentences. They also appear as parts of longer sentences, but they are not integrated in the structure, and they are regularly set off by intonational marking.  
(Thompson and Thompson 1992: 184)

What Andrason and Dlali (2020: 201–2) say about interjections in the South African Bantu language Xhosa typically appears to apply more generally to interactives cross-linguistically: ‘interjections cannot be governed by other constituents. They are not used as a predicate’s arguments, whether external (subjects) or internal (objects) [... and they] fail to function as adjuncts and/or nominal, adjectival, or verbal modifiers.’ And in Zulu, another South African language, interjections have been described as failing to exhibit grammatical relationship with or exert concordial effect on the adjacent sentence (Doke [1927] 1988: 285). Interjections in the Pama-Nyungan language Diyari of South Australia are neither inflected nor syntactically integrated with other linguistic material (Austin 1981: 36).

Describing the situation of ideophones in some languages of southwestern Cape York Peninsula of Australia, Alpher (2001: 9) concludes that ideophones in Yir-Yoront and neighboring languages ‘are not members of any part-of-speech class that participates in syntactic relations like head-attribute, verb-object, verb-adverb, or in derivational or inflectional morphological relations; they can be usefully viewed as belonging with intonational phenomena; they differ phonologically from words of other classes [...]’ (Alpher 2001: 8). And writing about Gooniyandi, another Australian language, McGregor (1990) identifies a category of ‘interjections’, which also includes interactive types such as attention signals, directives, discourse markers, response elicitors, response signals, and social formulae:

Interjections are words which stand outside of major clauses, and do not fulfil any role within them, or enter into syntagmatic relations with their constituents. Within an utterance they typically occur on their own intonation contour, and precede the remainder of the utterance.

(McGregor 1990: 227)

Similarly, many interjections in the Paamese language of Vanuatu ‘do not fit into the grammatical structure of an utterance’ (Crowley 1982: 76). And in the Dhaasanac language of northern Kenya and southwestern Ethiopia, ideophones and interjections form one of the eight closed ‘word’ categories, being extra-sentential elements or even forming sentences of their own (Tosco 2001: 249).

In Korean, interjections (‘sentential adverbs’) are frequently set off from the remainder of the utterance by a major phonological juncture or occur syntactically as independent utterances (Sohn 1994: 521–2), and in the more recent Japanese grammatical tradition, interjections are frequently set off from the rest of the utterance by some major type of juncture, and they frequently end in a ‘glottal catch’ (Hinds 1986: 443–4).

Most of the research mentioned in this section relates to interjections, but much the same applies to other types of interactives, like vocatives. For example, Zwicky (1974: 787) highlights the syntactic non-integration and non-argument status of vocatives, observing that an (English) vocative is ‘set off from the sentence it occurs in by special intonation [...] and it doesn’t serve as an argument of a verb in this sentence’, and for Levinson (1983: 71), vocatives are ‘noun phrases that refer to the addressee, but are not syntactically or semantically incorporated as the arguments of a predicate; they are rather set apart prosodically from the body of the sentence that may accompany them’.

Finally, drawing attention to ‘the inability of ideophones to permit morphosyntactic integration’, Haiman (2018: 105) notes:

While they inhabit ‘the same screen’ as conventional language, they are often both morphologically and syntactically ghettoized or quarantined, ‘aloof’ from the material that surrounds them.

(Haiman 2018: 105)

As we will see in Section 4.4, however, ideophones differ from all other types of interactives in their proclivity to grammaticalize on a large scale in the direction of sentence grammar—with the effect that in a number of languages they are morphosyntactically heterogeneous, in such cases no longer conforming to our definition of interactives in Section 1.2.

#### 1.3.1.4 Prosody

Interactives are likely to be set off prosodically from surrounding text material, being marked by special features of intonation and/or pause setting. Obviously, this applies most conspicuously to interactives forming utterances of their own.

That interactives are typically separated prosodically from their environment has in fact been observed in a number of studies. Being concerned with English

discourse markers in particular but also with other types of interactives, Zwicky remarks:

Unlike clitics, which are prosodically dependent, discourse markers and their parenthetical kin are prosodically independent. Typically, they are both accented and prosodically separated from their surrounding context, by pauses or intonation breaks or both.

(Zwicky 1985: 303)

Furthermore, Ameka (1992a: 108; 1994) notes that there is prosodic isolation of interjections by means of preceding and following pauses. The syntactic independence of interjections is reflected in their temporal isolation between preceding and following silent pauses.

Similarly, Andrason and Dlaki (2020: 194) observe from a comparative perspective: ‘Interjections are produced in isolation from the accompanying clause(s) and utterance(s), or their parts [...], thus constituting independent intonation units, separated from the remaining speech by pause’ (see also Andrason and Matutu 2019: 19). In the matter of the South African language Xhosa, these authors add: ‘What truly marks many interjections and distinguishes them from other lexical classes is greater energy or an increased volume with which they are uttered’ (Andrason and Dlaki 2020: 194). And with reference to ideophones, Andrason (2021a: 127) writes more generally that they are ‘accompanied by characteristic phonation and other prosodic phenomena such as breathy and/or creaky voice, whispering, intense airstream, and/or characteristic melody (pitch)’.

In the Tibetan language Denjongke of northeastern India, interjections usually occur at the beginning of a clause and are often followed by a pause (Yliniemi 2019: 133). In Korean, interjections (‘sentential adverbs’) are phonologically set off from the remainder of the utterance by a major phonological juncture (Sohn 1994: 521). They frequently carry a pitch level that is characteristically higher than that of the normal utterance. They are often pronounced with the first or last syllable prolonged for emotional effect, e.g., *kule::m* ‘indeed’ or *a::ni* ‘oh no!’. And they frequently end in a glottal stop. For example, the interjection *ak* ‘oh!’ is usually pronounced [aʔ].<sup>22</sup>

Much the same observations have been made on other types of interactives, such as discourse markers (Heine et al. 2021, Section 1.4.4), and Haiman writes on ideophones: They ‘are typically not influenced – either phonetically or morphosyntactically – by the structure of the utterances in which they are “embedded”’ (Haiman 2018: 111). In a similar way, Dingemanse (2011: 144) points out that an ideophone is not ‘part of the utterance that precedes it but forms its own intonation unit’. For Childs (1995: 133), ideophones in the West African Kisi language

<sup>22</sup> Concerning the role played by glottal stops in interactives, see Floricic (2015).

are simply different from surrounding text material—pronounced faster or slower than normal, overly long or short, and in a raised or a lowered register.

In the Gooniyandi language of Australia, a category of ‘interjections’ is distinguished which includes most interactive types in addition to interjections proper, and the members of this category are described as syntactically set off, and ‘they typically occur on their own intonation contour’ (McGregor 1990: 227). Similar observations have been made in languages from other parts of the world. In the Dhaasanac language of northern Kenya and southwestern Ethiopia, ideophones and interjections are treated as one ‘word’ category, being separated from the following sentence by a pause (Tosco 2001: 249). Crowley (1982: 76) notes that interjections of the Paamese language of Vanuatu ‘may have some abnormal intonation pattern’.

Similarly, the type of vocatives has been described as being ‘set off from the sentence it occurs in by special intonation’ (Zwicky 1974: 787) or ‘set apart prosodically from the body of the sentence that may accompany them’ (Levinson 1983: 71). In writing, prosodic separation tends to be signaled by punctuation marks, as in the following example, where the distinction between the sentence grammar unit *Peter* (a) and vocative *Peter* (b) is marked exclusively by commas:

(5) German (Sonnenhauser and Noel Aziz Hanna 2013: 8)

- a Was hast du Peter gesagt?  
    what have you Peter said  
    ‘What did you say to Peter?’
- b Was hast du, **Peter**, gesagt?  
    what have you Peter said  
    ‘What did you, Peter, say?’

Different intonations associated with one and the same interjection are fairly common—to the extent that Bolinger (1989: 265) claims that English interjections like *huh* and *hm* function as mere intonation carriers. Some German interjections can in fact be argued to be intonation carriers, exhibiting features to be observed in tone languages, where tonal distinctions signal meaning distinctions. Thus, the German interactives (‘interjections’) *ach* and *oje* and the response signal *hm* have been described as showing distinctions such as the ones presented in Table 1.3; for more fine-grained distinctions see Ehlich (1986) and Nübling (2001: 22). We agree with Ehlich (1986: 53–4) that such forms have features characteristic of tone languages, forming a ‘mini-tone system’ within an otherwise non-tonal language.

Similar observations on prosody distinguishing meaning have also been made in languages of the Americas. For example, in the Salishan language Thompson (Nlaka’pamuctsin) of British Columbia, a category called ‘expletives’ is proposed which includes eight of the types of interactives distinguished here (Section 5.1.1), and expletives ‘are regularly set off by intonational marking’ (Thompson and Thompson 1992: 184).

**Table 1.3** Tonal distinctions in the German interactives *ach*, *hm*, and *oje*

| Form        | Tone pattern   | Approximate meaning             | General meaning             |
|-------------|----------------|---------------------------------|-----------------------------|
| <i>ăch</i>  | Rising         | ‘really?’                       | Doubt                       |
| <i>âch</i>  | Falling        | ‘I wouldn’t have thought that!’ | Astonishment                |
| <i>āch</i>  | Level          | ‘stop that nonsense!’           | Annoyance, indignation      |
| <i>hîn</i>  | Falling        | ‘that’s strange’, ‘what now?’   | Perplexity                  |
| <i>hîî</i>  | Rising         | ‘why that?’, ‘really?’          | Doubt, divergence           |
| <i>hîî</i>  | Level          | ‘why that?’, ‘really?’          | Doubt, heralding divergence |
| <i>hîîî</i> | Falling-rising | ‘I agree’, ‘I’ve understood’    | Agreement, confirmation     |
| <i>ôjê</i>  | Falling        |                                 | Compassion                  |
| <i>ôjê</i>  | Level          |                                 | Disappointment              |
| <i>ôjê</i>  | Rising-falling |                                 | Dismay                      |

Sources: Based on Ehlich (1986: 50–1) and Nübling (2001: 22, 31).

Note: Tone marking is in accordance with that commonly employed for tone languages, differing slightly from that of the authors concerned.

Note, however, that acting as ‘information-carriers’ is not a general property of interactives; it appears to be primarily associated with interjections, while most other types of interactives have fairly fixed conventionalized prosodic contours.<sup>23</sup> Furthermore, interactives may adapt prosodically to their environment. When combining with other text units, especially in fast speech, prosodic boundaries may be weakened or disappear entirely, as has been observed especially in the study of discourse markers (Heine et al. 2021: 51–2).

O’Connell, Kowal, and Ageneau (2005) found that the highest percentage of interjections (45%) occurred without preceding or following silent pauses; there was no evidence for such temporal isolation of interjections in German- and English-language media interviews (see also O’Connell and Kowal 2005: 567). Rather than through articulatory isolation, O’Connell, Kowal, and Ageneau (2005: 166) found interjections to be singled out through the deliberate use of other prosodic changes, such as loudness, intonation, and/or articulation rate, frequently pronounced with greater energy and volume (cf. Andrason et al. 2020: 5).

Note also that Biber et al. (1999: 1082) observe that English inserts ‘have a tendency to attach themselves prosodically to a larger structure [...], and as such may be counted as part of that structure.’

### 1.3.1.5 Indexical status

Analyzing ‘interjections’ in the Mayan language Q’eqchi’ of Guatemala and Belize, Kockelman concludes:

<sup>23</sup> We are grateful to Gunther Kaltenböck (p.c. of September 26, 2021) for having drawn our attention to this fact.

Interjections are primarily *indexical*... in that they stand for their objects by a relationship of contiguity rather than by a relationship of convention (as in the case of symbols) or similarity (as in the case of icons).

(Kockelman 2003: 471; italics in the original)

Interactives are essentially deictic forms, or indexicals, that is, their interpretation depends on knowledge of the context in which the communication occurs—their usage and functions are anchored and typically restricted to the here-and-now of the situation of discourse. This was already pointed out in some way or other by earlier writers (e.g., Bally 1909; Bühler [1934] 1965)—accordingly, Ameka and Wilkins (2006: 3) state bluntly that ‘there is no dispute about the indexical nature of interjections’. For detailed analysis of interjections as deictics see Wilkins (1992: 133), where the term ‘interjection’ stands more broadly for a range of different types of interactives.

Deictic expressions are commonly understood to be expressions ‘determined in relation to features of the utterance-act: the time, the place, and the participants, i.e. those with the role of speaker or addressee’ (Huddleston and Pullum 2002: 1451). Interactives can be said to fall squarely in this category (Section 1.3.2.5). Accordingly, Ameka and Wilkins (2006: 3) write that interjections ‘are tied to specific situations and index elements in the extra-linguistic context’.

Clark and Fox Tree (2002: 77–8) note on the English interjection *ah*: ‘Each utterance of *ah* contains indices to the current speaker (I), the current addressees (you), the current moment (now), and other elements in the current common ground.’ The authors propose a temporal index *t* (utterance) for interjections which marks the precise moment at which an utterance is made. Thus, if in the exchange of (6) Sam had delayed the use of the interjection *ah* by one second, ‘that would have changed how soon he claimed to have been surprised and therefore, perhaps, what he was surprised about’.

(6) English (Clark and Fox Tree 2002: 77, (3))

William: I’m on the academic council,

Sam: **ah**, very nice position

Take also example (7) of the English discourse marker *well*. The first *well* of speaker B is a discourse marker, hence it is syntactically independent of the sentence. The second *well*, by contrast, is an adverb modifying the verb *feel*, being part of the meaning and the syntax of the sentence. The deictic time expressed by the second *well* is determined by the past tense form *did* of the predicate used by A—hence referring to the past. The deictic time of the first *well*, by contrast, is defined by the here-and-now of the situation of discourse, that is, by the time of speech.

- (7) English  
 A: Why did Liz not come?  
 B: **Well**, she did not feel **well**.

Consider also the following example, where the utterance in (8a) refers to the past. Yet, the social formula *sorry* refers to speech time. Hence, (8b) is appropriate while (8c) is not.

- (8) English  
 a **Sorry**, Kate missed the bus.  
 b Why are you sorry, it wasn't your fault.  
 c #Why were you sorry, it wasn't your fault.

It is throughout the various types of interactives that such a deictic behavior can be observed. In the following exchange, three types of interactives are involved, namely a social formula (*thanks*), a response signal (*yes*), and the discourse marker *well*, and all three refer to the time of the utterance whereas the remaining information of sentence grammar made by the two speakers refers to the past time.

- (9) English  
 Bill: Did you have a nice trip?  
 Sam: **Yes thanks**, but it was, **well**, a bit of a nightmare.

Example (10) concerns a question tag, that is, a response elicitor. Both the tag *didn't he?* and the preceding utterance are constructed in the past tense. But again, the tag refers to the here-and-now of the situation of discourse. Accordingly, (10b) is well-formed while (10c) is not.

- (10) English  
 a He tried to escape, **didn't he?**  
 b **Yes, you are right.**  
 c #Yes, you were right.

That they have an indexical status has in fact been pointed out for several types of interactives. Evans (1992: 229) concludes that '[i]nterjections are deictics in that they rely on context for their full interpretation', and Wilkins (1992: 135) concludes that discourse markers and determiners 'could also be argued to have essentially deictic semantics':

The solution to the problem of reference comes from recognizing that all interjections are indexical. Forms ranging from '**Ouch!**' to '**Yes!**', and from '**Huh?**' to '**Thankyou,**' which all meet the formal definition of an interjection [...] are all context-bound and they directly index entities in the extra-linguistic context as fillers of the argument positions in the proposition underlying the interjections.

(Wilkins 1992: 131–2; boldening in the original)

For Wilkins (1992: 132), interjections ‘share with words like *I, you, this, that, here*, and *now* the fact that they must be tied to the actual speech moment (i.e. the situation of utterance) before their complete interpretation (i.e. full referencing) can be made. When I say “**Yippee!**” I am indexing *myself* and something (i.e. “*this thing*”) *here* which just now made me aware of some proposition which has made me feel excited and more than happy (here and now), and so I say “[jɪpi:]!” in order to show how I’m feeling *right now*’ (boldening and italics in the original).

Wilkins (1992: 133) adds the following pieces of evidence. First, ‘interjections’ like *thankyou!*, *gimme!*, *welcome!*, and *dammit!*, ‘clearly incorporate the deictic forms *you, me, come*, and *it*, respectively.’<sup>24</sup> This observation can be supported by the kind of paraphrases used here (Section 2.1). Thus, *thankyou!* can be paraphrased as in (11a), which contains three deictic features, namely: *I*, standing for the speaker, *you*, standing for the hearer, and the time of reference is invariably that of speech time, that is, the form could not possibly be paraphrased as in (11b).

- (11) Paraphrases
  - a I am thanking you.
  - b #I was thanking you.

Second, the interactive *out!*, also classified by Wilkins as an ‘interjection’ (classified here as a directive), indicates that the hearer (‘addressee’) should move out of the place where the speaker is, that is, the here-and-now of the situation of discourse. And third, Wilkins draws attention to ‘the use of deictic gestures as part of, or as an accompaniment to, interjections’ (Wilkins 1992: 133), and he concludes:

In sum, then, the facts that (i) conventional deictic elements are commonly incorporated as part of interjections, (ii) deictic elements themselves may give rise to interjections, and (iii) deictic gestures may be built into, or may accompany, interjections, are here taken as strong evidence in support of the conjecture that all interjections contain basic deictic elements in their semantic decomposition. At the very least, these facts argue that interjections must be included in any serious investigation of deixis.

(Wilkins 1992: 134)

That they are deictic elements has also been pointed out for vocatives. Lambrecht (1996: 267) draws attention to the ‘the inherently deictic nature of vocatives’, and Noel Aziz Hanna and Sonnenhauser (2013: 9–10) conclude that ‘[v]ocative usage is indexical in nature’.

The deictic nature of interjections has also been pointed out in a number of studies of languages other than English (e.g., Evans 1992; Wilkins 1992). Evans (1992:

<sup>24</sup> Note that Wilkins’ notion of ‘interjections’ is a broader one: In the convention used here, the forms *thankyou!* and *welcome!* are not interjections but social formulae (see Section 1.5).



229) observes that in the Mayali language of Western Arnhem Land in Australia interjections are deictics in ‘that they rely on context for their full interpretation’. [Andrason and Dlali \(2020: 177\)](#) say about ‘interjections’ in Xhosa, which also include social formulae, vocatives, and some discourse markers, that ‘interjections refer to emotional and mental states occurring at the very moment of utterance. They are, in contrast, unable to refer to states that are located in the past or the future.’

But are ideophones also deictic elements? [Kilian-Hatz \(2001: 156\)](#) writes that ‘the ideophonic event happens simultaneously in the moment when it is uttered’, and [Haiman \(2018: 126\)](#) also concludes that ‘ideophones are always renderings of intensely felt impressions right here and now’. In contrast, according to [Dingemanse \(2011: 155\)](#) this question cannot clearly be answered in the affirmative. We will return to this issue in Section 3.5.7.

In addition to being deictics, interactives also do not allow for conceptual displacement or, in short, displacement (Section 1.3.2.6). Displacement is one of the 13 design features that were proposed by [Hockett \(1960\)](#) to define human languages. Displacement allows moving from the here-and-now to refer to persons, objects, places, times, and worlds outside of the situation of discourse. And in fact, it is a central feature distinguishing sentence grammar from interactive grammar. Whereas the former disposes of a wealth of lexical, grammatical, and syntactic means to refer to the world outside the situation of discourse, displacement is essentially absent in interactive grammar: Interactives, such as *bye-bye*, *no*, *please*, *shh*, *yuck*, etc., are as a rule restricted in their usage to the here-and-now of communication, that is, the situation of discourse.

With regard to displacement there is a further feature distinguishing interactive grammar from sentence grammar. The latter also contains sets of deictic forms, such as demonstratives (‘this’, ‘there’), adverbs (‘here’, ‘there’), and tense markers (past, future). Most, though not all of these forms can also be used for displacement: The use of demonstratives, like English *this* and *that*, or adverbs, like *here* and *there*, can be extended, for example, from space to time and discourse. This does not apply to interactives, which do not seem to allow for such displacement.

### 1.3.1.6 Negation

[Dwyer and Moshi \(2003: 176\)](#) argue that interactives like ideophones belong to the semiotic dimension of expressives, they are therefore ‘atactic’ rather than syntactic. This explains why they cannot combine with other signs, such as signs of negation. According to [Kilian-Hatz \(2001: 158\)](#), ‘ideophones are affirmative by definition’: ‘Ideophones in general are never negated and are rarely found in negated sentences.’

In fact, interactives cannot normally be negated. Thus, there are no well-formed corresponding negated counterparts to the list of interactives in Table 1.4 or, if

there are, they concern specific conceptual conditions or uses as units of sentence grammar.

**Table 1.4** English interactives and negation

|                 |                            |
|-----------------|----------------------------|
| <i>Goodbye!</i> | #Not/don't <i>goodbye!</i> |
| <i>Great!</i>   | #Not/don't <i>great!</i>   |
| <i>Hey!</i>     | #Not/don't <i>hey!</i>     |
| <i>Indeed.</i>  | #Not/don't <i>indeed.</i>  |
| <i>Oops!</i>    | #Not/don't <i>oops!</i>    |
| <i>Ouch!</i>    | #Not/don't <i>ouch!</i>    |
| <i>Shh!</i>     | #Not/don't <i>shh!</i>     |
| <i>Waiter!</i>  | #Not/don't <i>waiter!</i>  |

Accordingly, Pawley (2009: 7) notes with regard to social formulae like *Pleased to meet you* that they '[m]ust be an assertion in the present tense, cannot be questioned or negated'; Erman and Warren (2000: 32) point out that discourse markers ('epistemic pragmatic prefabs') like *I guess* cannot be negated (*\*I don't guess*), and so does Ljung (2011: 19) when he writes that (12a) is a good swear word but not (12b).<sup>25</sup>

- (12) English (Ljung 2011: 19)  
 a **Go to hell!**  
 b **\*Don't go to hell!**

Similarly, social formulae like English *How do you do?* cannot be negated (*#How don't you do?*). There are a few social formulae that are etymologically negative forms, and for them the opposite holds: They have no corresponding positive forms; thus *Don't mention it!* and *Never mind!* are common English response forms, while *Mention it!* and *Do mind!* are not social formulae.

That interactives cannot be negated has been documented in some form or other in a number of languages. Kita (1997: 389–90)<sup>26</sup> observes that Japanese ideophones ('mimetics') do not allow logical negation, and for Korean ideophones only 'metalinguistic negation' is possible (Diffloth 1972).

Part of the nature of interactives seems to be that they normally make a positive contribution, not a negative one. With regards to theticals ('parentheticals'), of which interactives are a part (see Section 5.2.3), Knowles (1980: 387, (36)) concludes more generally that a declarative parenthetical expression must imply a (positive) claim to the truth of the main proposition.

<sup>25</sup> The English negative comment clause *I don't think* is possibly an exception. But rather than expressing negation it is used pragmatically to express epistemic modality on a negative host, as in *John didn't go to Paris, I don't think* (Gunther Kaltenböck, p.c. of September 26, 2021).

<sup>26</sup> As described in Section 5.2.3, interactives form a subclass of 'theticals'.

A seeming counterexample to the present property of interactives can be seen in the structure of English question tags, which have both an affirmative and a negative variant, as in (13). In the sense of Knowles (1980) one may argue, however, that polarity is a fully grammaticalized device that serves essentially to express ‘a (positive) claim to truth of the main proposition.’ Accordingly, negation is possible as long as a positive appraisal of the main predication (*Ted is (not) an artist*) is not in doubt. Thus, rather than the grammatical category of negation, a semantic interpretation is at issue here: Interactives appear to imply a positive claim on some state of affairs.

(13) English

- a Ted is an artist, **isn’t he?**
- b Ted is not an artist, **is he?**

This constraint on negation has been emphasized especially in studies of ideophones. A number of authors suggest that ideophones display an ‘antipathy’ toward negation and questioning. Haiman (2018: 136) writes that the inability of ideophones to be negated recurs cross-linguistically, as has in fact been shown in a number of studies (e.g., Diffloth 1972; Kita 1997: 389; Nübling 2001: 25; Kilian-Hatz 2006: 510; Dingemanse 2017a: 363–4).

### 1.3.1.7 Interrogation

Interjections in a number of languages have been found to resist interrogation. Thus, as already noted, Andrason, Fehn, and Phiri (2020: 15) say that interjections of the Tjwao language of Zimbabwe are overall unaffected by syntactic operations of negation and interrogation.

In fact, what was observed about negation in the preceding Section 1.3.1.6 applies in a similar way to interactives used as interrogative forms: Part of the nature of interactives seems to be that they do not allow for positive–negative polarity. For example, Pawley (2009: 7) says that social formulae like *Pleased to meet you* can neither be questioned nor negated. Much the same has been observed for ideophones, interjections, and discourse markers.

There are, however, problems with this property of interactives. For one thing, some interactives are inherently interrogative speech acts due to their particular functions. This applies first and above to response elicitors, like *eh?*, *okay?*, *right?*, *see?*, including question tags (*isn’t it?*, *doesn’t he?*), but also to some social formulae like *how are you?* or *how do you do?*. Nevertheless, they do not allow for polarity in that they do not admit corresponding affirmative expressions (*#You are how.*).

For another, some interactives do allow for a contrastive affirmative vs. interrogative usage. This has been reported for some discourse markers that can be questioned, like *well* vs. *well?*, and also for the English response signal *huh*, which functions as an interjection in its affirmative form and as a kind of response elicitor in its interrogative form (Section 3.7.5). The latter, illustrated in the following

example, is discussed by [Dingemanse, Torreira, and Enfield \(2013: 7\)](#) under the term ‘other-initiated repair’, in which the speaker produces a turn at talk and the hearer then signals some trouble with this turn.

(14) English ([Dingemanse, Torreira, and Enfield 2013: 1](#))

G: It’s not too bad,

E: **Huh?**

G: ’S not too bad,

More research is needed on the nature of the constraints on negation and interrogation proposed here.

### 1.3.2 Further features

The definitional properties discussed in Section [1.3.1](#) do not exhaust all the features that have been mentioned to characterize interactives. Among the additional features that have been mentioned it is in particular the ones in (15) that have been adduced to distinguish interactives from forms of sentence grammar:

(15) Presumed special features of interactives

- a Their use is optional
- b They may be phonologically ‘unusual’
- c They show a pronounced affinity to gestures
- d They tend to be placed at the left periphery of an utterance
- e They show a distinction between primary and secondary forms
- f They do not allow for conceptual displacement
- g They include a set of conventionalized expressions for interaction with animals

These features are now looked at in turn.

#### 1.3.2.1 Their use is optional

This feature has most frequently been mentioned in connection with the meaning of the interactives discussed in Section [1.3.1.2](#): Since they do not change the meaning of their host sentence, it is claimed that interactives are ‘optional’ rather than obligatory elements. This claim was proposed most of all for discourse markers: Since they can frequently be omitted without loss of grammaticality or propositional content, their usage is said to be optional (cf. [Fraser 1988: 22](#); [1999: 944](#); [Schourup 1999: 242](#); [Jucker 2002: 212](#); [Barth-Weingarten and Couper-Kuhlen 2002: 352](#); but see also [Dér 2010](#)).

It would seem that this claim must be taken with care since it obfuscates the actual significance of interactives in linguistic communication. A distinction that needs further research is in particular that between grammatical (e.g.,

morphosyntactic) and discourse (e.g., social) ‘optionality’. For example, in his discussion of social formulae (‘politeness formulas’), [Ferguson \(1976: 24\)](#) observes with reference to the latter: ‘The importance of our trivial, muttered, more-or-less automatic polite phrases becomes clear when they are omitted or not acknowledged.’

1.3.2.2 They may be phonologically ‘unusual’

Phonologically unusual properties are certainly hard to find in sentence grammar, where words are commonly understood to conform to specific phonetic and phonotactic norms. In contrast, a number, though not all types of interactives are at variance with phonotactic constraints observed in sentence grammar. Thus, Haiman notes with reference to interjections (‘response cries’) and ideophones:

Not infrequently, response cries, like ideophones, violate canonical phonological constraints: either in terms of phoneme inventory (e.g., *brrr*) or of phonotactics (e.g., *mmm*), they are often marginal exemplars of words.

([Haiman 2018: 135](#))

It is the types listed in [Table 1.5](#) that are typically affected. Of these, especially ideophones stand out as being phonologically unusual, being ‘accompanied by characteristic phonation and other prosodic phenomena such as breathy and/or creaky voice, whispering, intense airstream, and/or characteristic melody (pitch)’ ([Andrason 2021a: 127](#)).

**Table 1.5** Unusual phonological features in English types of interactives

| Type              | Examples                                                     |
|-------------------|--------------------------------------------------------------|
| Attention signal  | <i>pst</i>                                                   |
| Directive         | <i>shh</i>                                                   |
| Discourse markers | <i>mh, mmm, uh, um</i>                                       |
| Interjections     | <i>brrr, hm, huh, mhm, mm, tt</i> (dental click), <i>ugh</i> |
| Response signals  | <i>uh huh, unh unh</i>                                       |

Cross-linguistically fairly widespread are interactives consisting of or containing the following features that may elsewhere be disallowed in the language concerned (see, e.g., [Aikhenvald 2008: 44](#) on the Manambu language of Papua New Guinea; see also [Joseph 1997](#) for examples):

- (i) Syllables without vowels
- (ii) Forms consisting only of non-syllabic consonants

- (iii) Sounds or sound combinations not occurring in sentence grammar
- (iv) Clicks

The English language provides examples of all features. With regard to (i), mention may be made of English words, which in general contain at least one vowel sound while response signals like *mhm* and *mm* are vowelless. (ii) can be observed for example in the English directives *pst!* and *shh!*, both consisting only of non-syllabic consonants. An English example of (iii) is *ugh* with its ‘un-English’ voiceless velar fricative [x] (Biber et al. 1999: 1083). An example of (iii) and (iv) is *tut-tut* with its dental click [ǀ], which is an interjection expressing disapproval, rendered in writing as *tt*, *tsk-tsk*, *tut-tut*, /tʌsktʌsk/, or /tʌttʌt/ (Wilkins 1992: 148, fn. 21). Note that this interjection is also found occasionally in written discourse:<sup>27</sup>

- (16) English (M. Gee, *Sole Survivor*, 1983, iv, 32—cited in *OED* 2nd edn. 1989; Wilkins 1992: 147, (9c))

The constable went **tsk, tsk** and looked at us sadly. There was nothing he could do so he sent us home.

Instances of (iii) can be found in some form or other across the globe, a few examples may illustrate this. Almost all Australian languages lack sibilant phonemes, and so does Dyirbal in Eastern Australia, but in Dyirbal ‘interjections’ sibilants do occur (Dixon 2010a: 283). Gooniyandi, a language of Western Australia, has general phonotactic constraints against open monosyllables with short vowels and against free morphemes which are closed monosyllables. But some interactives, like the directive *pa* ‘come on, let’s go’, violate these phonotactic constraints (Wilkins 1992: 122). In Mparntwe Arrernte of Central Australia, the interjection *eyyye!* [i:::] ‘yuck! that’s terrible, Oh no!’ is characterized in its pronunciation by a rising falsetto voice which is not used conventionally anywhere else in the language (Wilkins 1992: 122).

The Wampis language of Amazonian Peru has a two-mora requirement on phonological words, but the interjection *mah*, expressing surprise and admiration, does not conform to this requirement (Peña 2015: 412). Q’eqchi’, a Mayan language of Guatemala and Belize, has a commonly used alveolar glottalized stop, but its interjection *t’* is implosive (Kockelman 2003: 470), that is, it is possibly a dental click.

For some African languages, whispering or devocalization has been recorded in the production of interactives. In the Ndonga dialect of the Owambo language in northern Namibia, interjections such as the ones in (17) are formed by whispering. And similar observations have also been made in certain ideophones of

<sup>27</sup> Cf. *He tutted loudly* (Wharton 2003: 76). As Wilkins (1992) notes, this interjection has experienced a history from speech to writing and back to speech: ‘Thus “tsk-tsk” originally represented the conventional use of dental-alveolar clicks as a signal of disapproval, but, because of the written representation, a new interjectional form pronounced /trsktrsk/ has arisen’ (Wilkins 1992: 147).

some languages. For example, the ideophones *fa! fa! fa!* (depiction of drizzling) and *qhefe!* (depiction of crushing, e.g., a louse) of the Southern Sotho language in Lesotho are said to be whispered rather than expressed with voice (Guma 1971: 227).

- (17) Whispered interjections in Ndonga (Bantu, Niger-Congo;

Fivaz 1986: 169–70)

**aa!**        ‘good!’, on being impressed  
**ái?!**       an unacceptable situation  
**oosh!**     on making a mistake oneself  
**taataa!**   surprised satisfaction

In Korean sound patterns, nasalization is not a normal phenomenon, and final glottal stops are not phonemic units. But in interjections (‘sentential adverbs’) both can be observed: There frequently is nasalization of the vowel of interjections, and the latter tend to end in a glottal stop in discourse. Thus, *ak* ‘oh!’ and *mwe(s)* ‘what?!’ are usually pronounced [aʔ] and [mwəʔ], respectively (Sohn 1994: 521–2). And the Tibetic language Denjongke of northeastern India has the following phonologically distinctive characteristics in interjections: Emphatically long vowels, the use of the diphthongs /*ei*/ and /*ai*/, which with few exceptions do not occur elsewhere in word roots, and the otherwise non-occurring final consonant /*p*<sup>h</sup>/ [f] in the form *úf*, expressing discomfort (Yliniemi 2019: 133).

(iv), that is, use of clicks is most pronounced in one specific type of interactives, namely in response signals: In a substantial number of the world’s languages, forms for ‘yes’ and, even more commonly, for ‘no’ can be expressed by clicks (Gil 2013; Floricic 2015); see Section 3.8.4 for more details. But clicks are also widely observed in interjections, cf. the voiceless dental click ([l̥]) expressing disapproval in English (Wilkins 1992: 148, fn. 21), and also in other kinds of interactives such as animal calls. Speakers of the Muna language of Sulawesi in Indonesia, for example, may use an alveolar click to call a cat (van den Berg 1989: 189).

An additional ‘unusual’ feature is found in the way interactives such as ideophones are phonetically manipulated (Franck Floricic, p.c. of July 22, 2021). We will return to this subject in Section 3.5.3.3. As will be shown in Section 3.11, phonologically unusual features are particularly pronounced in animal directives.

### 1.3.2.3 They show a pronounced affinity to gestures

Gestures, that is, movements of the body, hands, arms, head, and/or eyes, are a feature of essentially every kind of speech: ‘Because gestures speak louder, and tell more, than words, we believe them’ (Janney 1999: 956)—oral language has been claimed to be rooted in gesture (see Armstrong, Stokoe, and Wilcox 1995). But, as has been pointed out time and again in many studies, its usage is excessive and frequently also of a different kind in interactives. For some authors, gestures are part and parcel of the existence of interactives like interjections and ideophones.

Wilkins (1992: 133) finds evidence for the fact that ‘interjections’ are complex predicational deictics in ‘the use of deictic gestures as part of, or as an accompaniment to, interjections’. Such observations have been made most of all with ideophones and interjections, both of which have been argued to be strongly related to gestures. Interjections, or a subset of them, have been called ‘response cries’ and ‘vocal gestures’ (Goffman 1978, 1981; Wierzbicka 2003a: 288), and Ameka (1992a: 106) characterizes interjections as being ‘relatively conventionalised vocal gestures (or more generally, linguistic gestures)’. Haiman (2018: 93) writes that ideophones ‘are often accompanied by visible gestures’, and that they ‘provide a plausible missing link between doing and saying.’

Wilkins (1992: 133–4) notes that ‘deictic gestures may be built into, or may accompany, interjections’, and interjections of ‘disgust and similar feelings’ have been analyzed by Wierzbicka (1992: 176) as being frequently linked with certain physical gestures.

Studying the East African language Swahili, Eastman (1992: 273) writes: ‘Interjections in Swahili, seen this way, range from verbal expressions alone, to verbal expressions accompanied by gesture, to gestures alone used with an interjectional function.’ This is in fact a situation that can commonly be observed in the usage of interactives across languages, it is captured by what we propose to call the voice-gesture model in (18). The model consists of a scale extending from vocal form (A) to gesture (C) with an intermediate stage (B) where the two are combined. It would seem that the model is part of the store of knowledge that interlocutors draw on in structuring their expression of a number of interactive concepts (cf. Haiman 2018: 298–9): Depending on their particular communicative intents, interlocutors may draw either on A, B, or C.

(18) A voice-gesture model for the expression of interactive concepts

| A    | B              | C       |
|------|----------------|---------|
| Form | Form + Gesture | Gesture |

Evidence in support of the model in (18) is not hard to come by. In order to ask for silence, a speaker of English may either say *shh!* (A) or place the index finger vertically on her lips (C) or, perhaps most commonly, use both (B). The Ewe language of southeastern Ghana and southern Togo has an attention signal, *kss...!*, used to call someone’s attention over a distance. It is employed, for example, by hawkers at lorry stations and along the streets to draw the attention of passers-by to their wares. In accordance with the model, *kss...!* may be used on its own (A), or it may be accompanied by a clap (B), but a clap alone without a vocalization (C) may also function as an attention signal (Ameka 1992b: 260).

Another, more common kind of example concerns an experience made by customers of a restaurant or bar in many different societies: To place one’s order the



customer may either draw on A, calling *Waiter!*, or on B, accompanying the call with a raised hand, or on C by simply using the raised hand. Instead of attention signals, various other types of interactives may be involved, such as social formulae. Meeting someone on the road may trigger either the greeting formula *Hi!* (A), or the formula combined with a gesture such as a wave of the hand (B), or by just using the latter (C).

There are some interactive functions that are primarily rendered by C rather than by A or B, that is, where gesture provides the main or even the only means of encoding. In the Tamil language of southern India and northern Sri Lanka, for example, begging forgiveness can be expressed with the social formula *mannikka Num* ('you should/must forgive me'), but by far the most common way to beg forgiveness for some minor affront is to use the kinesic gesture of prayer and supplication, a bow with hands together in prayer, or a reduced version, used to deities (Brown and Levinson 1987: 190).

Importantly, at least in the production of ideophones, 'speech and gesture are tightly coupled and alike in mode of representation: they perform the same role of depicting sensory imagery, albeit in different modalities'. In a Japanese corpus of cartoon retellings, Dingemanse and Akita (2017: 519) found as many as 94% of ideophones to co-occur with an iconic gesture, whereas this held for only 40% of verbs in a matched sample from the same corpus. We will return to this issue in Section 3.5.3.4.

To conclude, gesture constitutes an important component in the usage of various types of interactives, such as interjections, ideophones, attention signals, and some kinds of social formulae. Its contribution is clearly less pronounced in other types of interactives, especially in discourse markers. However, gesture also plays some role in sentence grammar, especially but not only in deictic categories such as demonstratives or personal pronouns—it can be said to be a general characteristic of spoken discourse. It therefore does not appear to qualify as a relevant definitional property of the category of interactives—the difference possibly being one in degree rather than in kind. The same applies to items like sighs, sniffs, and other gestures, described by Dingemanse (2020) as liminal signs. They are important elements in the structuring of linguistic discourse but do not seem to be consistent in their behavior to the extent that they should be taken into account in a definition of interactives.

Finally, there is also the following reason militating against proposing gesture as a definitional criterion. Interactives are mainly a feature of speech; nevertheless, they occur also in written discourse and, obviously, gesture is not normally part of the latter. Including gesture in a definition of interactives would therefore be tantamount to having to ignore written discourse in this study. Furthermore, the fact that interactives also occur in writing suggests that gesture is not an indispensable part of interactives.

### 1.3.2.4 They tend to be placed at the left periphery of an utterance

Being extra-clausal forms, interactives are highly flexible as to where they can be placed. But most commonly they appear at the left periphery of the utterance they are associated with.<sup>28</sup> The excerpt of an English conversation in (19) illustrates a common discourse pattern, where the following utterance-initial interactives are found: the discourse markers *well* and *now* and the response signals *uh-huh* and *yeah*.

(19) English (Norrick 2009: 871)

- [...]  
 Nancy: **well** I'm hoping.  
**yeah**, I would like to move back to Chicago or some place  
 in the fridge.  
 Margaret: **uh-huh, well, now**, in the area some way.  
 Nancy: **yeah**, I don't know.  
 [...]

Similar observations can be made with other types of interactives—in English or in other languages. From the perspective of formal syntax it is argued that elements like interjections are endowed with an ‘edge-feature’ that associates them with a clause (Munaro 2010). As Table 1.2 of Section 1.3.1.3 has shown, most interactives can occur in the same paradigm at the left periphery of an utterance, being prosodically separated from what follows.

There are exceptions, however: Response elicitors are typically placed to the right of their host utterance (Section 3.7), and the placement behavior of ideophones is complex (Section 3.5.3.2). Placement therefore does not qualify as a definitional property of interactives either.

### 1.3.2.5 They show a distinction between primary and secondary forms

Primary interactives differ from secondary ones in that they are not transparently related to any form of sentence grammar (see Section 3.6.4.1). For example, the English secondary interjection *shit* can be said to be related in some sense to the homophonous noun or verb, whereas the primary interjection *oops* does not appear to have a corresponding form in sentence grammar (but see also Section 4.2 for changes from interactive grammar to sentence grammar). In a similar way, secondary response elicitors like *right?* or *see?* have transparently related forms in sentence grammar, which is not the case with primary response elicitors like *eh?* or *huh?*.

Such a distinction is absent in sentence grammar, which does not dispose of forms that would be equivalent to primary interactives. But since the distinction

<sup>28</sup> Tom Givón (p.c. of January 27, 2022), however, questions the validity of this generalization.

is not observed in all types of interactives it does not qualify as a definitional property.

### 1.3.2.6 They do not allow for conceptual displacement

This feature was already mentioned in Section 1.3.1.5. Following Hockett (1960), displacement is argued to distinguish human languages from other kinds of communication systems, even if forms of displacement have also been observed in some non-human animals, like some bee and ant species. Its absence in interactive grammar is therefore conspicuous, being possibly suggestive of an archaic design feature of language evolution.

To be sure, interactives are commonly ‘displaced’ when being introduced by material of sentence grammar, for example when used in direct speech: *And then Br’er Rabbit whispered, ‘Yes, dad.’* In this text piece, the response signal *yes* and the vocative *dad* are part of the narrated event, referring to a time and apparently also a place and a world other than those of the speech event, that is, of the situation of discourse. But displacement in such cases is derivative of the time of the speech of the reporting clause *And then Br’er Rabbit whispered*, that is, of the context created by sentence grammar, it is not an inherent property of the two interactives themselves.

Nevertheless, in spite of the fact that none of the types of interactives can normally be displaced, this feature was not proposed as a definitional property in Section 1.2 but was discussed in conjunction with indexicality, suggesting that displacement can be treated as a feature defining a specific kind of indexical forms (Section 1.3.1.5).

### 1.3.2.7 They include a set of conventionalized expressions for interaction with animals

Sentence grammar is arguably the most powerful achievement in the evolution of the human species. It can be used for essentially any purpose, and it is also commonly employed in communicating with animals, most of all with domestic animals. But interactive grammar fundamentally differs from sentence grammar: It disposes of a set of function-specific conventionalized expressions—a kind of ‘mini-grammar’—dedicated to the interaction with animals (e.g., Bynon 1976; Amha 2013; Andrason and Karani 2021b).<sup>29</sup> This set is usually small in western industrial societies but can be fairly extensive in other societies, especially in pastoralist societies and among people dealing professionally with animals. We will return to this subject in Section 3.11.

<sup>29</sup> To be sure, the sentence grammar of many languages includes lexical and/or grammatical features derived from or referring to animals.

### 1.3.2.8 They may consist of clicks as distinct form–meaning units

Clicks serving as conventionalized form–meaning units are argued to belong to the earliest achievements in language evolution (Poulos 2021). In present-day languages, click consonants as parts of a phoneme system are only found in a number of Bantu languages of southern Africa and, most of all, in the Khoisan languages of southern and eastern Africa.<sup>30</sup> Beyond these African languages, no other languages in the world have been found to have phoneme systems that include click consonants. Nevertheless, clicks do occur quite commonly in interactives of many languages, most of all associated with emotional reactions of some kind. It is assumed, for example, that alveolar or dental clicks express irritation, sympathy, regret, disapproval, impatience, and exasperation; lateral clicks are said to encourage animals; while bilabial clicks are thought to be associated with kissing (Wright 2011: 208).

In interactives, usage of clicks is of a different kind than that in consonants of click languages, for a number of reasons. First, rather than being part of some meaningful form, the clicks themselves constitute meaningful forms without combining with any other phonological elements such as vowels—we therefore follow Brenzinger and Shah (2021; Matthias Brenzinger, p.c. of January 2, 2022) in referring to them as *click-only interactives*, consisting of a ‘naked click’. An English example of a click-only interactive is provided by the reduplicated dental click ([l-l]), which forms an interjection expressing disapproval, rendered in writing as *tt*, *tsk-tsk*, *tut-tut*, /tɪsktɪsk/, or /tʌttʌt/ (Wilkins 1992: 148, fn. 21). A different kind of an English click interactive serving as a discourse marker will be discussed in Section 3.3.5 (see Wright 2011).

Another example is the following: The Zulu language of South Africa has a number of click consonants in its phoneme system, being part of many words. But click-only forms do not seem to occur in this language—with one notable exception, namely that of interactives, like *x* ([ll]) ‘ach!’, an interjection consisting of a lateral click (Doke ([1927] 1988: 279–83).

Second, click consonants in the languages of southern and eastern Africa are complex phonological units. They consist, on the one hand, of a number of click types, namely dental (l), alveolar (!), palatal (ʡ), lateral (ll), and bilabial (⦿) clicks. On the other hand, each click type is subclassified in terms of its phonetic accompaniments. For example, the dental click type of the !Xóõ language of southwestern Botswana consists of altogether 17 different click phonemes depending on whether its accompaniment is voiceless, voiced, nasalized, aspirated, glottalized,

<sup>30</sup> The term ‘Khoisan’, formerly claimed to stand for one of the four language families of Africa (Greenberg 1963), is no longer taken to define a group of genetically related languages. It is used here instead as a convenient reference term for the languages distinguished by Greenberg, which include not only South African languages like !Xun, Khoekhoe, and !Xóõ, but also East African languages like Sandawe and Hadza.

ejected, uvular, etc. (Traill 1994: 11–13). Such diversity does not exist in interactives: Clicks are for the most part plain units not commonly involving any phonetic accompaniments.

Third, click-only interactives consist of any of the five types just mentioned, but these clicks may also contain specific phonetic features that set them off from the plain click types. For example, bilabial clicks (⊙) are attested in a number of Khoisan languages but, to our knowledge, a bilabial fricated click of the kind found in the Laal language of the Republic of Chad (Lionnet 2020) has so far not been observed in click languages (see below).

Fourth, click-only interactives are not always sensitive to distinctions between click types. Grenoble et al. (2015: 116) found in the Wolof language of Senegal that there is a response signal ‘yes’ which may be produced variously with either of three different click types.<sup>31</sup>

And fifth, click-only interactives are not only found in non-click languages but also in the click languages of southern and eastern Africa. With regard to the latter languages, this means that there are two kinds of forms using clicks, namely click-only interactives, on the one hand, and forms containing but not being restricted to a click consonant, on the other, where the latter include both interactives and forms of sentence grammar. For example, the click language Xhosa of South Africa contains two interactives containing a lateral click, namely *X!* ([ll]), expressing, e.g., disappointment or impatience, and *Xá!* ([llá]) ‘stop!’ (Brenzinger and Shah 2021). The former is a ‘naked click’, that is, a click-only interactive whereas the latter has a consonant-vowel structure as it can also be found in many words of sentence grammar in Xhosa. And in the Tswana language of Botswana, another click language, a series of four alveolar (‘prepalatal’) clicks ([!-!-!-!]) is used as an animal directive for calling fowls (Section 3.11), and a series of four lateral clicks ([ll-ll-ll-ll]) serves to express ‘excuse me please!’ or to coax or console a restless animal (Cole [1955] 1987: 396).

Click-only interactives can in fact be found in many languages. According to a world-wide survey carried out by Gil (2013), almost one third of the 143 languages analyzed by him include clicks for the response signals ‘yes’ and/or ‘no’ (see Section 3.8.4).

The following example may illustrate the nature of click use for interactive purposes. In Laal, a language isolate spoken by roughly 800 people in the Republic of Chad, four types of clicks are distinguished (Lionnet 2020): (a) A dental click, expressing negation or disagreement, (b) a lateral click, expressing approval, endorsement, or backchannel (‘yes, go on’ or ‘I get it’), (c) a back-released velar

<sup>31</sup> Wolof speakers refer to the click form of the response signal for ‘yes’ as *waaw*, the signal consisting either of a ‘palatal/velar click (mouth closed) single instance’, or a ‘lateral click, repeated’, or an ‘alveolar, click repeated’ (Grenoble et al. 2015: 116; see Brenzinger and Shah 2021 for more details on this general issue).

click, serving the same functions as (b), and (d) a bilabial fricated click, expressing dissent, disapproval, or contempt.

On the basis of the limited data available, the tentative conclusions in (20) can be drawn.

- (20) Hypotheses on click-only interactives
- a Click-only interactives are ‘naked clicks’ forming meaningful units of their own. Beyond interactives, ‘naked clicks’ have so far not been identified.
  - b They can be found in a number of types of interactives, but most of all in response signals (Section 3.8), interjections (Section 3.6), and animal directives (Section 3.11).
  - c Click-only interactives occurring in response signals usually express confirmation vs. disconfirmation or agreement vs. disagreement.
  - d The distinction confirmation vs. disconfirmation can be expressed either by using different click types or by click repetition, where—at least in a number of African languages—a single click expresses confirmation and a repeated click disconfirmation.
  - e Another function of repetition in click-only interactives is to express insistence or emphasis (see Section 3.8.4).

Ignoring the African click languages mentioned earlier, clicks are unlikely to occur in text pieces of sentence grammar, their usage being essentially a feature of the interactive domain of discourse processing. In ideophones, click-only interactives appear to be rare. This, however, does not apply to ideophones in click languages, where click consonants are as common as they are in the lexicon of sentence grammar. In the !Xóõ language just mentioned, for example, many ideophones (‘onomatopoeias’) contain a click consonant, the following are a few examples (Traill 1994; Brenzinger and Shah 2021):

- g|xòp** ‘sound made when someone walks through the bushes because the way is lost’
- llnùhm** ‘sound of a humming (e.g., of bees or a vehicle)’
- Øhòo** ‘sound of an object sliding into ash or sand with a flat trajectory’.

In spite of its wide cross-linguistic distribution, presence of clicks is not added to the list of definitional criteria in Section 1.2. First, because clicks are not found in all types of interactives and, second, because there is so far no information on whether they can be expected to occur in any language of the world. Nevertheless, one important generalization seems possible: Click-only forms, that is, forms consisting of a click only without an immediately preceding or following vowel and without any phonetic accompaniment, seem to be restricted to interactives—irrespective of whether the interactives occur in click languages, like Xhosa or

Xun, or in a non-click language like English or German (see also Section 3.3.5), or in animal directives (Section 3.11).

### 1.3.3 Conclusion

The catalog of features presented in this section does not exhaust all the phenomena that have been proposed to understand the behavior of interactives; suffice it to mention an additional feature that so far has found little attention in the relevant work and is in need of more research: Text pieces of sentence grammar can be varied in accordance with the parameters of modality and evidentiality, marking distinctions in the nature and the kind of information on which a given text piece is based. Such parameters do not clearly apply to interactives: Neither modality nor evidentiality are grammatical parameters of interactive grammar.

As conspicuous as the features discussed in Section 1.3.2 are, they are not generally found throughout the category as a whole—they are characteristic of most types of interactives but possibly not of all. Hence, they do not seem to qualify as definitional properties of the category.

To conclude, we are left with the list of properties discussed in Section 1.3.1 to define interactives, even if it may turn out at a more advanced knowledge that the list can be further extended.

## 1.4 Properties shared by the two domains of grammar

Interactives are widely held to be located at the periphery of language, they have been claimed to be intermediate between language and some other forms of human behavior, and they even have been denied a status within language or grammar (see Section 6.1 for detailed discussion). Accordingly, [Ferguson \(1981\)](#) writes with reference to politeness formulas, which are part of the domain of interactives:

They constitute another set of facts about human language which the linguists must somehow fit into their theories of grammar or turn over to others who are better able to deal with them ...

([Ferguson 1981](#): 34)

The preceding Section 1.3 was meant to demonstrate that there are reasons to define interactives as a domain of discourse processing of its own. At the same time, we argue that this domain belongs to grammar but needs to be distinguished from the domain of sentence grammar. For example, we saw that interactives, like

*hi, ouch, sorry, bang*, etc. can form utterances of their own without involving any material of sentence grammar. Conversely, sentences also can, and quite commonly do form utterances without any material from the grammar of interactives being involved. Accordingly, Wharton (2003: 84–5) concludes: ‘If one can have interjections, but not language, it is hard to see how the former can be viewed as part of the latter.’ The question then is: Do the two domains nevertheless have a common base?

The goal of the present section is to show that there is in fact a base shared by the two domains—a base that is referred to by Kaltenböck, Heine, and Kuteva (2011) and Heine et al. (2013) as discourse grammar, of which interactives are a part (see Section 5.2.3). This is suggested by the observation that the units of both sentence grammar and interactive grammar share most of the design features proposed by Hockett (1960) to characterize or define human languages, more specifically the features in (1).<sup>32</sup>

- (1) Features shared by interactives and forms of sentence grammar
  - a They are composed of form-meaning units (‘forms’) that must be learned.
  - b They are essentially arbitrary.
  - c They are discrete.
  - d They are transmitted from one generation to the next.
  - e Like sentences in sentence grammar, interactives can as a rule form utterances of their own.
  - f Like sentences, interactives are bounded, that is, they form distinct segments of discourse.
  - g They are both part of the linearized flow of discourse.
  - h They are sensitive to deixis.
  - i They can be linked by means of devices of cohesion.
  - j They are ‘universal’ in the sense that they can be expected to be found in languages across the world.
  - k They make use of gestures.
  - l They are subject to grammaticalization.

This catalog does not exhaust the range of similarities to be observed between the two domains. We are ignoring here problems that have been discussed controversially, such as the question of whether interactives such as interjections can form ‘speech acts’ (see Wilkins 1992: 152), or whether they are ‘words.’ The latter question, however, will be the subject of Section 6.2.

<sup>32</sup> This view is not the same as that expressed by Haiman (2018: 72), for whom ideophones ‘are performances and show moreover that they seem overall to lack precisely the “design features” that characterize human language.’



With reference to (1a), it has been demonstrated that interjections and other interactives are more or less conventionalized forms which are learned, stored, and retrieved like any other forms (Jespersen 1922: 415; Kryk 1992: 194). And they differ in their form from one language to another.

The form–meaning units of the two domains are widely assumed to be of the same kind, being referred to as ‘words’ or ‘lexical items’ (see, e.g., Wilkins 1992: 153, 155 on interjections). This, however, is not an unproblematic issue. Mention may be made, in particular of Goffman’s (1978) suggestion to call a group of English primary interjections ‘response cries’ or ‘semi-words’ rather than words, or of Wharton’s (2016: 21) characterization of interjections as ‘partly natural, partly conventional semi-words.’ Problems with treating interactives as ‘words’ or ‘lexical items’ are discussed in Section 6.2.

(1b) relates to another design feature mentioned by Hockett (1960). Rather than being motivated, most forms of both sentence grammar and interactive grammar are symbols, that is, arbitrary conventional units:<sup>33</sup> They exhibit no natural, e.g., iconic, connection between the forms (the signifiers) and the concepts (the signified) that they denote in a given language, or between forms in different languages denoting the same concept. The hedge ‘most forms’ indicates that there are exceptions, and the latter can be found in both kinds of grammar. For example, the signifier *miaow* is a word of English sentence grammar and can be said to exhibit a ‘natural connection’ with its signified, the sounds said to be normally emitted by cats. And this is also the case with signifiers of interactive grammar, like the English ideophones *bang*, *boing*, or *plop*.

There is, however, a difference: Nonarbitrary connections between the signifier and the signified can be observed clearly more frequently in interactive grammar, as has been pointed out in many studies of ideophones and interjections. For example, some interactives have been found to exhibit a regular connection between form and meaning across languages. This applies in particular to ideophones, some of which show sound symbolic features in many languages—note that ideophones are occasionally referred to as onomatopoeia (Section 3.5.2). But it also appears to apply cross-linguistically to response elicitors having the function of ‘other-initiated repair markers’ like English *huh?* (Dingemanse, Torreira, and Enfield 2013; see Section 3.7.5), or to attention signals like English *hey* (Section 3.1.5). Whether this difference is one of degree or of kind, however, is an issue that cannot be established at the present state of research.

(1c) is referred to by Hockett (1960: 6) with the term ‘discreteness.’ Both kinds of grammar rely on a set of distinct phonological units, even if that set tends to be larger in interactive grammar than in sentence grammar (Section 1.3.2.2).

<sup>33</sup> Arbitrariness means in the wording of Hockett (1960: 6) that ‘[t]he word “salty” is not salty nor granular; “dog” is not “canine”; “whale” is a small word for a large object; “microorganism” is the reverse.’

Discreteness is manifested in sentence grammar, for example, in the presence of minimally distinctive pairs, like English *pin* vs. *bin*. But the same kind of distinctions can be observed in interactives, for example, in the English interjections *ah* vs. *oh*, or *bang* vs. *boing* in ideophones, or *uh* vs. *um* in discourse markers (Section 1.5). And much the same applies to other languages, like ideophones in the African language Southern Sotho, where the following ideophones are distinguished only by their vowel quality: *fi* (image of suddenly becoming dark), *fo* (image of snuff coming out of a snuff box), and *fu* (image of cooking, e.g., meat for a short time) (Guma 1971: 229).

The effects of (1d) can readily be tested in our daily interaction: In the same way as young people acquire new words and constructions of sentence grammar, they acquire interactives, such as interjections, ideophones, or social formulae from previous generations—for example, at least since the period of Early Middle English, speakers of English have been learning the interjection *oh* and its uses from earlier generations of speakers, even if the uses are no longer exactly the same as those of eight or nine centuries ago (Jucker 2002: 217).

(1e) and (1f) are captured by Biber et al. (1999: 1070) succinctly in the following way, where ‘non-clausal units’ include interactives while ‘clausal units’ are restricted to sentence grammar: ‘Clausal and non-clausal units are maximal grammatical units in the sense that they cannot be *syntactically* integrated with elements which precede or follow them.’ According to (1f), both sentences and interactives form distinct segments of discourse. Both can be used as turns in conversational exchanges, and both are typically marked off from the rest of discourse by prosodic features like intonation and pause.

Even if interactives are not integrated in the syntax of a sentence, (1g) draws attention to the fact that both the forms of sentence grammar and interactives jointly contribute to the linearized arrangement of texts. The principles of linearization and placement though are different, resting on rules of sentence syntax in the case of sentence grammar but on functions of discourse organization in the case of interactive grammar.

According to (1h), deixis is a fundamental design feature of human languages. At least one part of linguistic forms can be interpreted meaningfully only with reference to the speech participants and the where and when of linguistic communication. However, this part is severely limited in the case of sentence grammar, where it is restricted to specific classes of forms such as person markers, demonstratives, adverbs, and tense markers. This is quite different in interactive grammar, which has deixis as one of its definitional properties (Sections 1.2, 1.3.1.5): Interactives are generally anchored immediately in the situation of discourse.

(1i) is a noteworthy feature, for the following reason. For sentence grammar, the structure of cohesion has been well described in the tradition of Halliday and Hasan (1976). This does not apply to interactive grammar; note further that interactives are syntactically independent forms (see the definition in Section 1.2).

Nevertheless, the observations that have been made so far suggest that interactives can be connected with text pieces of sentence grammar by means of cohesive devices such as agreement. This applies most clearly to English question tags: While being a more marginal class of interactives, more precisely of response elicitors, they provide an extreme example of the way in which cohesion can be established between a text piece of sentence grammar and an interactive, which in this example is provided by the type of response elicitors.<sup>34</sup> As the example below shows, cohesion manifests itself in multiple morphological agreement in person (third person), number (plural), tense (past), and polarity (negation) between the utterance of sentence grammar and the following response elicitor:

- (2) English (Biber et al. 1999: 433)

But they did say four to five weeks on the coat, **didn't they?** (CONV)

Considering that so far only a fraction of the world's languages is well documented, (1j) must be taken with care. Nevertheless, we are not aware of any well described language that would not dispose of some form of both sentence grammar and interactive grammar.

The significance of gesture (cf. (1k)) has been pointed out in particular in studies of interjections and ideophones, where it is argued, for example, that gesture is a characteristic of interactives (e.g., Wilkins 1992: 134; Haiman 2018: 93).<sup>35</sup> But gestures are also found in sentence grammar, even if distinctly less so and possibly of a different quality than in interactive grammar (Section 1.3.2.3). (1k) therefore is a robust design feature of both domains, and the difference between the two domains may be one of degree rather than of kind.

With regard to (1l) there is quite some evidence to show that grammaticalization is the main mechanism responsible for grammatical change and the development of new grammatical categories (e.g., Heine 1997; Hopper and Traugott 2003; Kuteva et al. 2019; Narrog and Heine 2021). No detailed research has so far been carried out on the grammaticalization of interactives, with the notable exception of discourse markers. That research has established that, even if the development of discourse markers cannot be reduced to it, grammaticalization is crucially involved in their development (see Heine et al. 2021 for more details). We will see in Sections 4.2 and 4.4 that this also applies to other types of interactives, but as will also be shown there, grammaticalization is not the only mechanism shaping the form and functions of interactives.

<sup>34</sup> Note that English question tags are not prototypical instances of response elicitors in particular and interactives in general because they are not invariable forms (see Section 1.2, see also Section 3.7).

<sup>35</sup> What is 'extra-linguistic' and what is not is to some extent controversial (cf. Floricic 2015). For the purposes of the present discussion, the term is restricted to non-vocal behavior such as body movements.

This catalog of design features does not exhaust all the phenomena shared by the two grammars. One may wish to draw attention, for instance, to conventionalized grammatical mechanisms at work in social interaction employed for achieving joint attention. As shown by [Evans, Bergqvist, and San Roque \(2018a; 2018b\)](#), such mechanisms serve what the authors call *engagement*, that is, the encoding of the relative accessibility of an entity or state of affairs to both the speaker and the addressee. As demonstrated by the authors, engagement surfaces in various formats in sentence grammar. But it is equally important in interactive grammar, which disposes of a wealth of forms and constructions that express functions relating to degrees of shared attention, knowledge or belief states in speaker–hearer interaction. Such forms include attention signals, response elicitors, response signals, directives, vocatives, and—perhaps most of all—discourse markers (see Chapter 3). For a rich data base on grounding and the synchronization of minds involving response elicitors and response signals, see [Wiltschko \(2021\)](#).

To conclude, the grammar of interactives and sentence grammar, on the one hand, differ significantly in their structure, as captured by the properties discussed in Section 1.3. On the other hand, the two share one and the same foundation, as the properties discussed in this section suggest: Both are part of language, and both are part of one encompassing grammar, referred to here as discourse grammar (Section 5.2.3). This conclusion is compatible with that reached by [Wiltschko \(2021: 3, 7\)](#), who suggests that ‘the same system that is responsible for computing propositional language is also responsible for computing interactional language’, and ‘that the language of interaction makes use of similar building blocks as the propositional grammar it embeds’.

In a number of studies, the question has been raised in some form or other whether interjections, ideophones, or some other types of interactives really belong to language, or to the grammar of a language (see Section 6.1). On the view proposed here, this question can be answered in the affirmative on account of the observations made in this section—the difference being that sentences and interactives constitute different domains of grammar and of discourse processing.

## 1.5 A classification of interactives

Interactives include a wide range of expressions covering quite different spaces of meaning. For example, you may want to teach your child to say *Thank you!* or *Sorry!*, but presumably not really *Bloody hell!* or *Shit!*. In Table 1.1 of Section 1.2, ten types of interactives were distinguished, and in Table 1.6, a list of these types with English examples is presented. The classification is based on the definitions of types to be provided in Chapter 3, where each of the types is looked at in its own right.

Interactives provide a set of conventional expressions dedicated to the coding of stancetaking in the sense of [Du Bois \(2007\)](#), covering all three key concepts of alignment, evaluation, and positioning. Paradigm English expressions are response signals like *yes* and *no* for alignment, evaluatives like *excellent* or *horrible* for evaluation, or interjections like *wow* and *yuck* for positioning, and positioning may be either affective (*phew*) or epistemic (*aha*).

No claim is made that the present taxonomy is less controversial than earlier ones—that is, it constitutes by no means the only way in which interactives can be classified. Being based on a prototype perspective, it assumes that a given interactive can belong to more than one type—for reasons to be discussed in [Chapter 4](#). Furthermore, relying on grammaticalization theory, it takes into account that double membership of a given interactive can in a number of cases be explained with reference to relative degrees of grammaticalization that the interactive concerned exhibits (see [Section 4.2](#) for detailed discussion).

Classifications of interactives into discrete categories, such as interjections or discourse markers, are notoriously controversial, for a number of reasons. For one thing, they are contingent on which forms a given author has in mind when establishing a type. A survey of classifications that exist suggests that there are hardly any two authors drawing on exactly the same set of forms to define a given type. And for another, the criteria used by different authors are frequently not the same.

The classification of types proposed in [Table 1.6](#) must be taken with care in particular for the following reasons. First, since they are not syntactically integrated, interactives are highly context-sensitive and, depending on the context in which they are used, they tend to be associated with a number of different functions. These functions may, and fairly commonly do relate to more than one type of interactive, and it is not always clear what the basic type is to which they belong, or whether there is a basic type at all. For example, depending on the context in which it appears, the interactive *boy* can be classified as a vocative, an interjection or a discourse marker, and *oh* as an interjection or a discourse marker ([Schiffrin 1987](#); [Fraser 1990](#): 385; [Section 4.2](#)). And much the same applies to interactives in other languages. The expression *ʔééka* of the Maale language in Ethiopia can function as a social formula for welcoming someone, as an interjection ('I am happy with what you said'), or as a discourse marker for backchanneling to signal that one is giving full attention to the speech of the previous speaker and encourage him or her to continue ([Amha 2001](#): 263, 265).

The reason for assigning a given interactive form to two different types A and B is that that form can be found to be used in at least some contexts where its function is in accordance with type A while there are also contexts where its function is in accordance with type B.

Second, while mirroring to some extent similar classifications by other authors or proposed in reference grammars of English, the criteria used here to define types are not always the same as those used by other authors. Take the example

**Table 1.6** The main types of interactives with English examples

| Type              | Sub-type                | English examples                                                                                                                                                                                                                                                                                                                                                         |
|-------------------|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Attention signals |                         | <i>ahem, come on, excuse me, hey, hoo, I say, listen, look, oh, oi, pardon me, pst, say, yo(o), you know</i>                                                                                                                                                                                                                                                             |
| Directives        | Imperative-derived      | <i>come, come in, careful, come on, fuck off, get off, halt, hang on, hold on, listen, look, piss off, shut up, stop, wait</i>                                                                                                                                                                                                                                           |
|                   | Other directives        | <i>chop-chop, out, pst, shh, shoo, shush</i>                                                                                                                                                                                                                                                                                                                             |
| Discourse markers | Plain                   | <i>actually, after all, anyway, as it were, besides, by the way, however, in fact, if you will, indeed, instead, I mean, now, okay, say, so, still, then, I think, well, what else, you know, you see</i>                                                                                                                                                                |
|                   | Fillers                 | <i>uh, uhm, um</i>                                                                                                                                                                                                                                                                                                                                                       |
| Evaluatives       |                         | <i>amazing, bloody brilliant, bullshit, excellent, fantastic, fine, glorious, good, good gracious, great, hilarious, horrible, how awful, incredible, irrelevant, lovely, lucky, marvelous, miraculous, nonsense, oh dear, ridiculous, right, rubbish, splendid, stupendous, super, terrible, tough, unbelievable, uncanny, (what a) pity, (what a) shame, wonderful</i> |
| Ideophones        |                         | <i>bang, boing, boom, gaga, helter-skelter, plip plop, pow, splish-splash, swish, ta-dah, thud, thump, tick-tock, vroom, zigzag, zoom</i>                                                                                                                                                                                                                                |
| Interjections     | Primary interjections   | <i>aargh, ah, aha, ahem, aw, cor, eek, ew, ha(h), ha-ha, heigh-ho, hey, hm, ho hum, huh, mhm, mm, oh, ooh, oops, ouch, phew, tsk, ugh, uh-oh, whew, whoa, whoopee, wow, wowee, yikes, yippee, yu(c)k</i>                                                                                                                                                                 |
|                   | Secondary interjections | <i>balls, blast, bloody hell, boy, bugger, Christ, crumbs, damn, (God) dang (it), fuck, (my) God, goddammit, heavens, hell, Jesus, Jesus Christ, man, shit</i>                                                                                                                                                                                                           |

Continued

Table 1.6 *Continued*

| Type               | Sub-type                    | English examples                                                                                                                                                     |
|--------------------|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                    | Moderated expletives        | <i>blimey, crikey, (gosh) darn (it), gee, geese, geeze, golly, good grief, goodness (me), (my) gosh, heck, jeez, my goodness</i>                                     |
| Response elicitors |                             | <i>eh?, huh?, isn't it?, okay?, pardon?, right?, see?, sorry?, what?</i>                                                                                             |
| Response signals   | Assent, agreement           | <i>absolutely, alright, aye, certainly, exactly, I see, indeed, mhm, mm, of course, okay, okie-dokie, right, sure, totally, uh huh, yeah, yep, yes, you bet, yup</i> |
|                    | Denial, disagreement        | <i>as if, certainly not, definitely not, hardly, huh uh, m'm, never, never ever, no, nope, not at all, not likely, not yet, unh unh</i>                              |
| Social formulae    | Greetings and leave-takings | <i>bye, bye bye, good morning, good-bye, hello, hey, hi, how are you?, how do you do?, yo</i>                                                                        |
|                    | Good wishes                 | <i>God bless you, good luck, good night, goodbye</i>                                                                                                                 |
|                    | Congratulations             | <i>congratulations, well done</i>                                                                                                                                    |
|                    | Thankings                   | <i>cheers, thank you, thanks [never mind, no problem, welcome]</i>                                                                                                   |
|                    | Polite speech formulae      | <i>please, sorry, welcome</i>                                                                                                                                        |
|                    | Apologies                   | <i>beg your pardon, excuse me, pardon, pardon me, sorry [no problem, that's okay]</i>                                                                                |
| Vocatives          | Conventional terms          | <i>Marjorie, Paul, Jennifer, Thomas; Marj, Paulie, Jackie, Tom</i>                                                                                                   |
|                    | Class terms                 | <i>boy, da, dad, daddy, doctor, driver, ma, madam, man, mom, mommie, mum, pa, pop, sir, waiter</i>                                                                   |
|                    | Epithets                    | <i>darling, honey, hon, jerkhead, (my) dear, love, sweetie (pie)</i>                                                                                                 |
|                    | Expletives                  | <i>asshole, bastard, bitch, bugger, dummy, faggot, idiot</i>                                                                                                         |

See Section 3.3.4 for a treatment of 'fillers'.

of the English forms *pst* ('come to me so I can talk to you confidentially') and *shh* ('be quiet'): Both are usually classified as interjections—more specifically as conative interjections. However, the term 'interjection' is reserved here strictly for items expressing emotive or cognitive states of the speaker, like *ouch*, *wow*, and *yuk* (Section 3.6) whereas the functions of *pst* and *shh* concern speaker–hearer interaction and seem to be classified more appropriately within the type of directives, *pst* also as an attention signal (Sections 2.1, 2.2, and 3.2).

And third, the list of types qualifying as interactives is far from exhaustive, for some reason or other. In particular, there is the class of exclamatives, such as the English *wh*-exclamatives illustrated in (1), which we decided to ignore here for the following reason. With reference to their semantic, syntactic, prosodic behavior, and their deictic anchoring, exclamatives are clearly part of the category of interactives (cf., e.g., Michaelis 2001; Rett 2011; Siemund 2015; Heine, Kaltenböck, and Kuteva 2020), and they can be argued to have a propositional structure (Franck Floricic p.c. of July 22, 2021). Furthermore, they resemble in particular interjections in expressing the speaker's affective response to a situation and in forming utterances of their own. In addition, they share with interactives that they belong to the domain of thetical grammar of Heine et al. (2013), as will be shown in Section 5.2.3.

- (1) English (ICE-GB, BNC)
  - a What a waste of money!
  - b How boring!

Nevertheless, exclamatives are not considered here for the following reason. The forms presented in Table 1.6 are restricted to those conforming to our definition of interactives, and exclamatives clearly contradict the definition of interactives (Section 1.2) in that they are not invariable forms. Rather, they are constructions consisting of a substantive part (*what a* and *how*, respectively in English) and a schematic part, and the latter part is inherently variable. For example, instead of *money*, the exclamative in (1a) could take a range of other nouns, like *time*, *resources*, or *energy*, and instead of *boring* in (1b), the substantive part could consist of a list of other adjectives such as *nice*, *awful*, or *wonderful*. To be sure, some degree of variability in form can also be observed in other kinds of interactives, most of all in vocatives (see Section 3.10), but variability in exclamatives is an inherent part of their structure. Exclamatives will therefore be reserved for a separate study, which will take their specific schematic structure into account. However, while ignoring exclamatives as a class, some of them exhibit a largely invariable structure and have a distinct discourse function and, hence, are included, such as *what a pity* and *how awful* (see Table 1.6, Evaluatives).

Furthermore, modal particles, like German *doch*, *etwa*, *ja*, *schon*, etc., are also not included in the category of interactives (see Foolen 2018 for a recent overview). While their functions are clearly interactional, they are syntactically integrated



in the structure of clauses and, hence, are at variance with our definition of interactives in Section 1.2.

A more detailed treatment of the ten types is found in Chapter 3.

## 1.6 This volume

The goal of the book is to describe the usage of interactive forms in grammar and discourse. To this end it is argued in Chapter 2 that one of their main characteristics is that they have an argument structure, that is, they differ from one another in which and how many arguments they take. The types of interactives distinguished in the preceding Section 1.5 also differ from one another in a number of other ways as will be shown in Chapter 3, where each of the types is looked at in turn. The chapter concludes by presenting a comparative overview in Section 3.12. An interesting phenomenon is discussed in Section 3.11: Interactives are not restricted to usage among humans; rather, they are cross-linguistically found to contain a set of conventionalized expressions to communicate with animals.

While Chapters 2 and 3 are descriptive in nature, dealing with the synchronic structure of interactives, Chapter 4 moves on to a diachronic perspective. Linguistic categories are notoriously ‘vulnerable’ in the sense that they are constantly affected and modified by other categories co-occurring with them in discourse, and this applies also to the category of interactives. As we will see in Chapter 4, much of that ‘vulnerability’ can be accounted for with reference to how interactives came to be what they are—in other words, with reference to their historical development.

The remaining chapters then are dedicated to more general issues. The question looked into in Chapter 5 is how the analysis proposed in this study relates to and is embedded in research presented by other authors. Sections 5.1 and 5.2 will be dealing with linguistic research, Section 5.3 with neurolinguistic research, and Section 5.4 with work carried out in social psychology.

Chapter 6 is then concerned with the ontological status of interactives, taking up some issues that had surfaced in earlier research, like where interactives are located, whether they can be called ‘words’ or form a system, how they behave in situations of language contact, etc.

The concern of the book is with interactive grammar, but Chapter 7 moves on to the coexistence of the two kinds of grammar distinguished, and more specifically to how the two perform both relative to one another and to discourse processing. In the final Chapter 8 then the findings made in the book are briefly summarized and some conclusions are drawn.

As this overview shows, there is a general directionality in the presentation of the chapters to follow, leading from the more specific to the more general. Starting out with description in Chapters 2 and 3, the book continues with diachronic

reconstruction in Chapter 4, and the remaining Chapters 5, 6, and 7 look at interactives in a wider perspective.

The methodology employed in the study rests on usage-based theory, and more specifically on the framework of discourse grammar (Kaltenböck, Heine, and Kuteva 2011; Heine et al. 2013, 2021; see Section 5.2.3 for a summary of the framework). Furthermore, the study draws on prototype theory to understand and describe both the internal and the external status of the category of interactives (Section 1.2). The category is looked at, on the one hand, as a ‘product’ or an ‘object’ of language use but, on the other hand, also as a tool for processing discourse and structuring social interaction.

The perspective adopted is a typological-comparative one, it is based on the analysis of grammatical descriptions of well over one hundred languages spoken in all major regions of the world. The generalizations proposed are therefore hypothesized to be cross-linguistically relevant. It goes without saying that this is a highly ambitious hypothesis, especially since for thousands of languages we have no or only inadequate descriptions. Furthermore, even the descriptions that we have are for the most part silent on the structure of interactives—the majority of grammars that we were able to consult contain little or no information on this subject. What this means is that these generalizations must be taken with care, subject to verification by future research.

In spite of this comparative perspective, however, exemplification relies more on data from English than from other languages: For one thing, in a volume written in English, examples from this language will presumably be most readily accessible to readers. And for another, research on interactives has so far focused more on English than on any other language, and many of the data discussed in that research are taken from English.

A few conventions adopted in the book need to be mentioned. First, an attempt is made to draw generally on attested data from text corpora for exemplification. But in a number of cases we were not able to find appropriate text data and had to rely on constructed examples. The latter examples differ from the former in being presented without reference sources.

Second, in order to make interactives more readily identifiable in the text material presented they are, unless indicated otherwise, generally printed in bold, even if this may be at variance with the convention used in the sources from which the material was taken.

Third, interactives have been referred to in many studies quite frequently as ‘interjections’, less commonly also as ‘discourse markers’, etc. In the present study, these terms are employed in a more restricted sense in accordance with the definitions proposed in Chapter 3. Thus, whenever referring to studies where these as well as a number of other terms do not match our definitions, inverted commas are used (e.g., ‘interjection’).

Previous treatments of interactives were for the most part restricted to one or a sample of the ten types of interactives distinguished here and in this chapter an effort was made to highlight the commonalities shared by the various types. Accordingly, there is no room for doing justice to the differences that also exist between these types. This will change in Chapter 2, and most of all in Chapter 3, where the wealth of potential that the grammar of interactives offers for structuring discourse is discussed.

## 2

# Argument structure

Chapter 1 was concerned with highlighting what the various types of interactives have in common. Having established that they can be defined as a category of their own, the present chapter moves on to showing that there are also differences between the various types. The discussion is restricted to one feature, namely to what is proposed to be called ‘the argument structure’ of interactives. However, the ‘arguments’ to be identified in Section 2.2 are of a different nature than those distinguished in sentence grammar. Rather than being entities of the syntax of the clause or sentence, or of the valence of a verb, their ‘valence’ is defined with reference to their functions in discourse, more specifically to the role they play in structuring interactive forms. The terms ‘argument’ and ‘valence’ as used in this study, thus, are shorthands for ‘discourse argument’ and ‘discourse valence’; they relate to the structure of meaning characterizing interactives.

To understand what arguments of interactives are about, more information on the meaning space covered by interactives is essential. This meaning space has been the subject of a number of studies, and its nature is discussed controversially. It will therefore be looked at in greater detail in Section 2.1, where our attention will be most of all with the notions ‘expressive meaning’ (Section 2.1.2) and ‘depiction’ (Section 2.1.3).

## 2.1 Meaning

One of the challenges that interactives present can be seen in their meaning. A view widely held by students of interactives is the following:

Although interjections are relatively easy to characterize from the standpoint of grammatical form, there is no framework in terms of which one may order and compare their meanings ...

(Kockelman 2003: 470)

It is fairly uncontroversial that interactives contrast starkly in their semantics with what can be observed in sentence grammar. But the question is: What exactly is the difference?

We will not be able to give a conclusive answer to this question, which should be the subject of a separate monograph. Our interest is restricted to identifying

certain features of meaning that are relevant for understanding the structural behavior of interactives in discourse. But what is argued in the present Section 2.1 is that the meaning of interactives is best analyzed in terms of three components, namely descriptive meaning, expressive meaning, and depiction. Discussion of these components will be restricted to the topic of the present chapter, that is, to what we propose to call the ‘argument structure’ of interactives.

### 2.1.1 Paraphrase

It is hard to compare the meaning space covered by the grammar of interactives with that of sentence grammar. As aptly observed by Haiman (2018: 138), ideophones and interjections ‘cannot be translated into propositional language.’ One reason for this problem of intertranslatability is volunteered by Goffman (1978: 126) with reference to exclamatory interjections (‘response cries’) when he says that they are ‘communications specifically designed for use outside states of talk.’

Conversely, one may add that it is equally hard, if not impossible, to translate concepts of sentence grammar, e.g., that of ‘tree’ or ‘believe,’ in terms of forms available in the grammar of interactives—in other words, the two grammars appear to be largely mutually incompatible in how they encode meaning (see also Section 7.3.1). The question therefore is: Is it at all possible to relate the two meaning spaces to one another in a principled way?

While there is so far no cogent answer to this question, there are some observations that may be of help to find at least a partial answer. Consider the following example of a paradigm instance of an interactive, namely a primary interjection: Are the two utterances in (1) equivalent, that is, does (1b) express the same meaning as (1a)—in other words, can (1b) be taken to be a suitable paraphrase of (1a), as has been argued by some students of interjections (e.g., Kaplan 1999: 11)?

- (1) English
  - a **Wow!**
  - b I'm delighted!

The paraphrase in (1b) is intended to interpret the communicative intents of the speaker by means of an alternative text piece having—at least in many contexts—a similar meaning as (1a).<sup>1</sup> One may wonder whether paraphrase in fact provides

<sup>1</sup> In presenting examples such as (1a) we are ignoring a number of things. First, we are ignoring that the meaning of (1a) can reasonably be interpreted only with reference to the particular situation in which it is used. Second, depending on the context in which it is used, *wow* could signal a variety of different meanings. It could signal, for example, any reaction extending ‘from surprise and being mildly impressed, through amazement and astonishment to jaw-dropping bewilderment,’ or ‘from satisfaction through enjoyment to absolute exhilaration’ (Wharton 2003: 46). And third, we are also ignoring here that the meaning of *wow* could also vary greatly depending on the prosody and the kind of facial expression and gestures associated with its usage.

an appropriate theoretical tool for capturing the meaning of *wow* and other interactives. A theoretically more sophisticated tool can be seen in what is portrayed by Wharton (2003: 40; 2016: 25) as the ‘conceptualist’ view of interjection analysis. Paradigm proponents of this view are Wierzbicka (e.g., 1992; 2003a; 2021), Ameka (1992a), Wilkins (1992), and Goddard (e.g., 2014) relying on the model of (natural) semantic metalanguage as developed by Wierzbicka and associates for describing or, more precisely, for explicating meaning. According to this model there is a set of 65 primitives or primes representing innate building blocks of meaning. Based on this model, Wierzbicka (1992) explicates the meaning of *wow* as in (2); for a slightly different explication of *wow* in the natural semantic metalanguage framework, see Goddard (2014: 57).

- (2) An explication of English *wow* (Wierzbicka 1992: 164)
- I now know something
  - I wouldn’t have thought I would know it
  - I think: it is very good
  - (I wouldn’t have thought it could be like that)
  - I feel something because of that

Natural semantic metalanguage provides a demanding tool for analyzing the structure of interactives, but the question is: How close do extant semantic models like that of natural semantic metalanguage come to describing and/or understanding the meaning of interactives like *wow* in (1a)? For some problems with explications such as (2), see Wilkins (1992: 150–1) and Wharton (2003). In particular, is the explication in (2) really specific enough to direct someone hearing or reading (1a) to the meaning of this interactive? One might wonder, for example, whether for an analyst not sufficiently familiar with the theoretical background of semantic metalanguage theory, (2) could not only be used for explicating *wow* but also a range of other items like, say, a new dress, some exotic fruit, or a painting that I come to admire on a museum visit.<sup>2</sup>

Another way of making the descriptive meaning of interactives like interjections explicit would be by translating it into that of interjections having ‘the same’ or nearly the same meaning, be that interjections of the same language or of a different language. For example, Haiman (2018: 93) suggests with reference to the English interjection *yuk*:<sup>3</sup>

*Yuk* may easily and felicitously be translated, but only into other ideophones that show the same feeling in both English and other languages: *Ew* or *Ick* are

<sup>2</sup> For reasons why he considers the ‘conceptualist’ view to be ‘highly problematic,’ see Wharton (2003: 44–9; 2021: 25).

<sup>3</sup> Haiman (2018: 93) treats English *yuk* as an ideophone rather than an interjection, but this is an issue that is of secondary import here. Note also that the boundary between ideophones and interjections is not always entirely clear (but see also Section 3.5.4).

acceptable translations or near synonyms within English, and *pfui* is as good as these for a translation into the ideophones of German.

(Haiman 2018: 93)

Such an approach, however, is not without problems. For one thing, many ideophones, or interjections, do not readily have clearly synonymous or antonymic counterparts, be that in the same language or in other languages. And for another, even if translations from one ideophone or interjection into another should be possible, it remains unclear how the resulting translations could be related in a principled way to concepts of conventional semantic analysis; note that Haiman (2018: 93) adds that we cannot ‘translate ideophones directly into conventional propositional utterances, either within a single language or across languages.’

In view of such problems we will draw on paraphrase in our analysis of interactives. The contribution that paraphrase can make in uncovering the descriptive content of interactive meaning has been acknowledged in some form or other by a number of authors. Evans (1992) in particular observes in his study of interactives (‘interjections’) in the Australian language Mayali:

Interjections are problematic for the usual methods of linguistics because their scant form tells us so little. Under this laconic exterior, however, we discover many of the semantic complexities familiar to us from the study of formally more complex expressions. It is only by bringing this semantic complexity to light through a system of explicit paraphrase that we shall accord interjections the attention they deserve.

(Evans 1992: 243)

For the purposes of the present study it would seem in fact that paraphrase such as that in (1b), while being a pre-theoretical technique, can be useful as a heuristic of elucidation. For one thing, it is of help in identifying some aspects of the meaning content of interactives. And for another, it also captures some of the relational properties that interactives have when combining with other text units, as will be argued in Section 2.2. Note, however, that there are limits to what paraphrase can achieve. First, it cannot claim and in fact does not achieve meaning identity with the item paraphrased. Second, considering that some interactives have fairly complex meanings, paraphrasing may require quite some descriptive expenditure. And third, and most importantly, paraphrase does not take care of the expressive meaning component of interactives, as we will see in the paragraphs to follow.

With these restrictions in mind, compare the interactives *wow* in (1) with *shh* in (3), where (3b) presents one of the possible paraphrases that may be proposed for *shh*: There are a number of features that distinguish the meanings of *wow* and *shh*,

but most of all the following: The paraphrase for *wow* in (1b) contains the subject pronoun *I*, standing for the speaker, and the same also applies to the paraphrase for *shh*. But there is also a difference: In the latter case, but not in the former, there is in addition an object pronoun *you*, standing for the hearer. What this difference in the paraphrases signals is that the two interactives differ in the number of speech participants they are associated with: The paraphrase ‘I’m delighted!’ of *wow* in contexts like (1) implies only one participant, the speaker, while *shh* implies two, namely a speaker and a hearer. We will return to this observation in Section 2.2.

- (3) English  
 a **Shh!**  
 b I ask you to be quiet!

The following example may give an impression of the relevance that paraphrases can have as a tool for reconstructing elements of discourse meaning (see also Section 2.4). In conventional frameworks of grammatical description, interactives tend to be portrayed as somewhat peripheral elements added to sentences, providing at best circumstantial information, as the response elicitor *right?* in (4a) or the interjection *wow* in (4b) are said to do.

- (4) English  
 a Chameleons are colorblind, **right?**  
 b You are such a great artist, **wow!**

From the perspective of discourse processing which aims at reconstructing the communicative intents of the speaker in structuring discourse, an alternative interpretation would suggest itself. Without any more detailed contextual information we are not able to exactly reconstruct what the intents of the speaker were when designing the utterances in (4). But one possible hypothesis would be that in her interaction with the hearer, the speaker was aiming at expressing what is captured by the paraphrases proposed in (5). On this hypothesis, one would interpret *right?* and *wow!* in (4) not as peripheral elements but rather as relevant means used by speakers to express their communicative intents—note that the meanings conveyed by the interactives *right?* and *wow!* in (4) surface in (5) as main clauses while the semantic content of the two clauses in (4) are rendered in the paraphrases of (5) by means of subordinate *that*-clauses.

- (5) Possible paraphrases of (4)  
 a Is it true that chameleons are colorblind?  
 b I am delighted that you are such a great artist.

To be sure, paraphrases such as the ones just proposed are neither definitions nor semantic equivalents of the material they paraphrase. Nor do we wish to claim



that the fact that the content of the interactives *right?* and *wow* in (4) is encoded in main clauses in (5) is suggestive of any privileged status of the interactives, e.g., as representing some underlying meaning, or a higher or dominating level of language processing or language competence.<sup>4</sup> What the paraphrases suggest however is, first, that on a syntactic analysis of (5), the text pieces corresponding to the interactives in (4) can be argued to present the main content whereas the actual main clauses in (4) present subordinate content in the form of subordinate complement clauses; we will return to this issue in Section 2.4. Second, it also suggests that interactives play a more important role in discourse processing than has been assumed in most of previous work and, third, that paraphrase provides a convenient heuristic for elucidating salient conceptual features of interactives.

Being used as a heuristic, paraphrases might in fact be viewed as revealing higher-level explicatures in the sense of Relevance Theory, that is, ‘representations in which the basic level explicature (the proposition expressed) is embedded under a speech-act or propositional-attitude description’ (Wharton 2016: 28). On the latter hypothesis it is assumed that an utterance like (6a) might lead a hearer to embed the proposition expressed under an attitudinal description such as that presented in (6b).

- (6) English (Wharton 2016: 28)
- a **Aha!** You’re here.
  - b The speaker is surprised that I am here.

But are paraphrases not essentially an arbitrary tool for elucidating conceptual information? Note, for example, that the English interjection *wow* is paraphrased by most authors as ‘I am delighted’ while Ameka (1992a: 113) paraphrases it as ‘I am surprised,’ and Ameka and Wilkins (2006) as ‘I feel surprised and I am impressed.’ While there is no entirely satisfactory answer to this question, it would seem that the question can be answered in the negative. A survey of the ways in which the meaning of interjections has been rendered suggests that there is a common core identified by different authors, irrespective of whether the meaning is framed in terms of a paraphrase, a definition, a description, or an explication. For example, for the meaning of the English interjection *ouch* (or *ow*) the predications in (7) have been proposed.<sup>5</sup>

<sup>4</sup> For example, Wiltchko (2021: 3) argues that interactional language is hierarchically structured in that ‘interactional structure dominates propositional structure.’

<sup>5</sup> The form of the interjection is *ouch* with some authors but *ow* with other authors.

- (7) Ways of rendering the meaning of English *ouch* or *ow*
- a I feel pain (Wierzbicka 1992: 162; Wharton 2003: 48)
  - b I am in pain (Kaplan 1999: 14; Wharton 2016: 25)
  - c I have just experienced a sudden and sharp pain
  - d I suddenly feel a pain (in this part of my body) right now that I wouldn't have expected to feel (Wilkins 1992: 149)
  - e I say '[auu!]' because I want to show that I am feeling pain right now [and because I know that this is how speakers of English can show (other speakers of English) that they are in pain (in a situation like the situation here)]. (Wilkins 1992: 149)<sup>6</sup>

The five predications in (7) differ in their relative degree of explicitness, where (7a) and (7b) are the least and (7d) and (7e) the most explicit predications. Nevertheless, all share a common core consisting of the speaker (*I*), a feeling (*feel* or *experience*), and the nature of the feeling (*pain*). This is especially noteworthy since the authors cited draw on contrasting models for analyzing interjections. Thus, we feel justified to take the most concise predications in (7a) and (7b) to present suitable paraphrases to elucidate the conceptual content or, as we will say here, (features of) the descriptive meaning of the two interjections.

Note that the term 'descriptive meaning' is not used here in the sense of Coulmas (1978: 242), for whom the term stands for the literal, or etymological meaning of an expression. For example, the 'descriptive meaning' of the English greeting *How do you do* consists for him of that denoted by the four words contained in the expression. In the present usage, by contrast, the descriptive meaning of *How do you do* concerns that of its actual function as a greeting, roughly paraphrasable as 'I am greeting you!'

The relevance of descriptive meaning can be demonstrated with the example in (8), which does not appear to present a felicitous utterance. While this observation can be explained in various ways, we argue that the interjection *ow* contains 'pain' as a descriptive meaning component, and that this component is incompatible with the concept 'pain' in the negated utterance *I'm not in pain* in (8). Accordingly, to the extent that the usage of interactives is sensitive to features of descriptive meaning in sentence grammar in the way *ow* is in (8) we assume that descriptive meaning is part of the usage of interjections and other interactives.

- (8) English  
#**Ow!** I'm not in pain!

<sup>6</sup> We are omitting here the subscripts used in Wilkins' (1992) data since their function is predictable from that of the participants concerned.

All the paraphrases proposed above are presented as sentences having a propositional format (see also [Munaro 2010](#), (27)). However, no claim is made here to the effect that interactives, like *wow*, *shh*, etc., have a propositional organization; rather, using the form of propositions is simply taken to provide a convenient tool of elucidation.

The observation that interactives have descriptive or conceptual meaning is not new. [Wilkins \(1992: 133\)](#), for example, describes ‘interjections’ as ‘complex predicational deictics which contain basic deictic elements in their semantic decomposition’:

Contrary to the view that interjections have no conceptual semantic structure, it has been demonstrated [...] that interjections tend to have quite a complex semantic structure which is amenable to semantic decomposition.

([Wilkins 1992: 153](#))

The perspective adopted here resembles that of [Wilkins \(1992\)](#) to the extent that it is argued that the meaning of interjections and other interactives can be ‘decomposed,’ even if decomposition as understood here is of a different nature than that proposed by Wilkins, or more generally in the framework of natural semantic metalanguage for that matter.

That interactives are sensitive to features of descriptive meaning is also suggested by the observation that in specific contexts they can be used referentially.<sup>7</sup> Take the example in (9), where the complex interjection *oh Lord* is presented by Bud in line 2 and taken up referentially by Carol in line 5. And much the same applies to the interjection *ouch* in the excerpt from a novel in (10): Introduced by B, it is referred to in A’s subsequent response.<sup>8</sup>

(9) *English*(*LSWEC*—AC145201; [Norrick 2009: 879](#))

- 1 Freda: was Chris married in ninety three.
- 2 Bud: **oh Lord.**
- 3 Carol: ((laugh))
- 4 ((?)): **yes.**
- 5 Carol: **oh Lord** is right.
- 6 that was the year of all the weddings.

<sup>7</sup> For example, ideophones (‘mimetics’) in Japanese have been described as being referential ([Akita 1997: 381](#)).

<sup>8</sup> More examples of this kind are not hard to come by, suffice it to mention the following text piece:

English (Nic34a: 167; [Stange 2016: 13, \(5\)](#))

INV: **Ow.**

CHI: What say **ow** for?

(10) English (Givón 2011: 407)

A: '[...] Four generals, three admirals, Colonel Sprague himself. That bad enough?'

B: '**Ouch.**'

A: Major Herskovitz nods. '**Ouch indeed.** Old Rodney's not mad at you though, if that's a consolation.'

But paraphrases, being confined to the descriptive content of interactives, cannot provide meaning identity with the interjection paraphrased. For example, the meaning of the English interjection *ouch!* in (11a) has been paraphrased by *I'm in pain!* (see (7) above). But the meaning of the interjection cannot reasonably be reduced to that conveyed by its paraphrase. For example, (11a) and (11b) are apparently not semantically exactly the same—what appears to distinguish them is that in addition to the descriptive meaning encoded in both, (11a) also has *expressive meaning*, immediately signaling the emotive state of the speaker.<sup>9</sup>

(11) English (Wharton 2016: 25, (20), (21))

a **Ouch!** I'm in pain!

b I'm in pain! I'm in pain!

According to Kaplan (1999: 11), in contrast, the information conveyed by *ouch* and *I am in pain* 'is the same, but they convey it through different modes of expression'—that is, he views the difference not as being semantic but, as he puts it, as 'more syntactical.'

The question then is: What is expressive meaning? This is the topic of the next section.

### 2.1.2 Expressive meaning

As observed in Section 2.1.1, we accept paraphrase as a useful tool of analysis. At the same time, we are aware that paraphrase fails to capture the interactive function of interjections, and this function is widely referred to as the expressive meaning of interactives. Referring to interjections and ideophones as 'expressives,' that is, as linguistic forms having expressive rather than descriptive, conceptual or referential meaning, has a long tradition (see Foolen 2016 for an overview).<sup>10</sup>

<sup>9</sup> We are ignoring here the possibility that the difference between (11a) and (11b) is possibly also due to the fact that, rather than *I'm in pain*, a more appropriate paraphrase of *ouch* would be *I have just experienced a sudden and sharp pain* (Franck Floricic, p.c. of July 22, 2021).

<sup>10</sup> Cf. the related distinction made in speech act theory between describing and indicating (e.g., Searle 1969). Discussing the structure of politeness formulas, that is, social formulae in the present terminology (Section 3.9), Davies (1987: 78) proposes a distinction between semantic or propositional

The term ‘expressive’ has acquired a wide range of different usages, appearing for example in combinations like ‘expressive word,’ ‘expressive intonation,’ ‘expressive lengthening,’ ‘expressive consonant,’ etc. (e.g., [Haiman 2018](#)). [Diffloth \(1972\)](#) treats ideophones including onomatopoeias (‘onomatopoeics’) as sub-classes of expressives, and in Southeast Asian linguistics, ideophones are widely referred to as ‘expressives’ ([Dingemanse and Akita 2017](#): 501).

But how can expressive meaning be made explicit and be defined? In some studies it is taken to refer to emotive content. For [Potts \(2007a: 180\)](#), for example, expressive content ‘is performative in this sense: quite generally, the act of uttering an expressive *is* the emotive performance.’ While emotion may in fact play an important role in the usage of interactives we converge with other authors on this issue in maintaining that expressive meaning and emotive content should not be equated. Interactives differ greatly from one another with regard to the size and the nature of the emotive content they express, being pronounced in interjections and ideophones but minimal, if at all present, in many usages of some other types of interactives.

According to a general characterization of it, proposed in some form or other in various accounts and adopted here, the expressive meaning of interactives *is anchored directly to the situation of discourse* rather than being processed by the structure of sentence grammar. [Jakobson \(1956: 82\)](#) characterizes the expressive function as ‘a direct expression of the speaker’s attitude toward what he is speaking about,’ it is said to be an aspect of meaning that ‘covaries with characteristics of the speaker’ (1977: 307). Rather than describing meaning, interactives *index* some state, as if the speaker were ‘inhabiting the frame’ of that state (cf. [Kockelman 2003: 471](#)).

Unfortunately, such characterizations are not specific enough to establish how exactly expressive meaning can be analyzed. And much the same applies to some other notions that have been proposed to understand expressive meaning, such as ‘foregrounding,’ ‘attracting attention,’ or ‘indicating noteworthiness.’ For example, [Dingemanse and Akita \(2017: 505\)](#) define the ‘expressiveness’ of linguistic signs as ‘the degree to which they are foregrounded as distinct from other items, for instance by special intonational or phonational features.’ But once again, such a definition does not allow either to determine to its core the semantic or functional substance of interjections, ideophones, and other interactives, especially since ‘foregrounding’ is a notion that has been applied to a number of different linguistic phenomena. A promising line of approach can be seen in the attempts made by [Potts \(2007a, 2007b\)](#) to propose indices of expressive meaning in terms

content and illocutionary force, which relates to that between descriptive vs. expressive meaning made in the present framework. The notion of illocutionary force is in the same way relevant to sentence grammar as it is to interactive grammar and, hence, is not taken here to be a discriminating feature.

of variables like ‘intensity,’ ‘negativity,’ and ‘positivity,’ even if Potts’ understanding of expressive meaning is a fairly general one and does not focus on interactives.<sup>11</sup>

In a number of studies, the situation of discourse is described with reference to the contexts in which interactives are used. Accordingly, Kaplan (1999) suggests approaching the study of expressive meaning of interjections like *ouch* in the following way:

The semantic information in the word *ouch* is represented by the set of those contexts at which the word *ouch* is expressively correct, namely, the set of those contexts at which the agent is in pain. That set of contexts represents the semantic information contained in the word *ouch*.

(Kaplan 1999: 10)

To be sure, on the basis of the observations made in Section 2.1.1, the semantic information contained in the word *ouch* cannot reasonably be reduced to context, but neither can it be reduced to descriptive or conceptual meaning.

The distinction descriptive vs. expressive meaning is defined by Kaplan (1999: 5) in the following way: Meaning is ‘descriptively correct’ if what it describes is the case. Descriptive meaning is sensitive to truth conditions, in Kaplan’s terms it belongs to the ‘semantics of meaning.’ Expressive meaning, by contrast, belongs to the ‘semantics of use,’ it is expressively correct since what it expresses or displays is the case. Rather than truth conditional it is said to be *use conditional*, where the latter is described ‘as a set of the contexts in which the sentence is uttered felicitously’ (Gutzmann and Henderson 2019: 119). Kaplan’s understanding of the distinction between descriptive and expressive meaning is summarized in Table 2.1.

For Kaplan (1999: 17), the two meanings ‘are informationally equivalent.’ This claim, however, is not without problems. For one thing, there is so far no empirical evidence to support the claim and for another, descriptive meaning is sensitive to truth conditions while expressive meaning normally is not. Thus, the utterances in (12) are not really felicitous (marked by ‘#’) since expressive meaning is said not to be sensitive to truth conditions (Gutzmann and Henderson 2019: 119; see also Gutzmann 2015).<sup>12</sup>

- (12) English  
       a #It is true that **wow**!  
       b #It is true that **shh**!

<sup>11</sup> For example, the German distinction between a familiar (*Du*) and a formal second person singular pronoun (*Sie*) is also described by him with reference to a difference in expressive setting (Potts 2007a: 190).

<sup>12</sup> Note, however, that the latter proposal does not seem to be entirely satisfactory since units of sentence grammar such as words can also be defined as a set of the contexts in which they are uttered felicitously.

**Table 2.1** Two kinds of meaning of interactives according to Kaplan (1999)

|                            | Descriptive meaning                                                                                                  | Expressive meaning                                                                                                |
|----------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| <i>Domain</i>              | ‘Semantics of meaning’                                                                                               | ‘Semantics of use’                                                                                                |
| <i>It is correct if:</i>   | What it describes is the case                                                                                        | What it expresses or displays is the case                                                                         |
| <i>It is sensitive to:</i> | Truth conditions                                                                                                     | Use conditions                                                                                                    |
| <i>Empirical basis</i>     | Tests of truth conditionality apply                                                                                  | The set of contexts in which the sentence is uttered felicitously<br>[Tests of truth conditionality do not apply] |
| <i>Examples</i>            | <i>I am in a state of joyful elation</i><br><i>I am in pain</i><br><i>you and I are now parting from one another</i> | <i>hurray</i><br><i>ouch</i><br><i>goodbye</i>                                                                    |

Whereas expressive meaning is said to be non-truth conditional, as (12) suggests, the descriptive component of the meaning of the interactives in (12) appears to be truth conditional, as is shown by the paraphrases of *wow* and *shh* in (13).

- (13) English  
 a It is true that I’m delighted.  
 b It is true that I ask you to be quiet.

Thus, what distinguishes interactives from most other linguistic forms is their expressive meaning: In addition to describing, they have been said to express, display, show, or depict meaning (Kaplan 1999: 5; Wharton 2016: 28–9; Gutzmann and Henderson 2019: 120; see also Dingemanse 2011: 17 and Haiman 2018: 57 on ideophones).

Mention should be made, however, that interactives, or at least part of them, can be used insincerely. For one thing, interjections can be used in rhetorical devices like irony or sarcasm. For example, saying *Wow!*, *that’s terrible* is odd in English but would be acceptable if intended sarcastically (Goddard 2014: 57). As observed by Stange (2019: 175–6), ironic uses constitute one of the four main discourse uses of interjections in spoken British English. And for another, speakers may employ an interjection not to express their own emotive state but rather the one conceived to be that of the hearer (cf. Goffman 1978: 805).

Part of the use conditions underlying expressive meaning in interactives consists of what following Evans (1992: 237) one may wish to call ‘social placedness conditions,’ that is, presuppositions about the social relationship between speaker, hearer, and setting. By means of social placedness conditions, constraints are imposed by the situation of social deixis on whether an interactive can be, must be,

or must not be used. The following example may illustrate the nature of such conditions. According to Wierzbicka (1992: 160), (14) is a conceivable conversation between two speakers of Yiddish in specific contexts:

(14) Yiddish (Wierzbicka 1992: 160; based on Rosten 1968: 26)

A: **Nu?**

B: **Nu-nu.**

The interactive *nu*, called an ‘interjection’ by Wierzbicka, is one of the most frequently used expressions of Yiddish, and Wierzbicka (1992: 160) observes that *nu* ‘has become the one word which can identify a Jew. In fact, it is sometimes used just that way.’ Thus, the exchange in (14) can in specific contexts be paraphrased as in (15).

(15) Possible paraphrases of (14)

A: Are you Jewish?

B: Yes, I am.

One way of interpreting (14) is to assume that with the use of the ‘interjection’ *nu*, both speakers A and B draw on descriptive meaning of the kind proposed in the paraphrases of (15). At the same time it would seem that without an appropriate setting of social deixis, (14) would not constitute an acceptable exchange. For example, Jim Matisoff (p.c. of December 3, 2020) observes that the usage illustrated in (14) is conceivably appropriate only in a context ‘where Gentiles were around and one wished to ascertain the stranger’s ethnic identity without being too obvious about it.’

The extent to which social placedness plays a role in shaping expressive meaning differs greatly from one context to another and from one type of interactives to another. It is presumably particularly strong in social formulae and weakest in discourse markers, but more research is needed on this issue.

In sum, interactives, such as interjections or ideophones, are assumed here to have two meaning components, namely a descriptive or conceptual and an expressive one. The former component is approached here by using paraphrase. But the exact nature of expressive meaning is still largely ill understood—note that expressive meaning has been declared to be ‘insusceptible to formalization’ (Kaplan 1999) or ‘ineffable’ (Potts 2007a; Blakemore 2011).

Furthermore, analyzing interactives as ‘expressives’ constitutes a necessary but not a sufficient condition, for the following reason: Expressive meaning is not restricted to the category of interactives. As the literature summarized in this section suggests, it is also found in English sentence grammar, as has been argued for with examples like the following, involving the adjective *damn* in (16a) and the noun *bastard* in (16b).



- (16) English (italics are ours)  
 a That *damn* Kaplan was promoted. (Kaplan 1999: 5)  
 b That *bastard* Kresge is famous. (Potts 2007a: 168)

To conclude, more research is needed on the status of expressive meaning in interactives considering, on the one hand, that ‘expressive’ meaning can also be observed in sentence grammar units and, on the other hand, that a number of interactives, especially discourse markers, exhibit little, if anything, in terms of such a meaning. But there is another concept that appears to be equally relevant for understanding the usage of interactives. This concept is that of depiction, which is the subject of Section 2.1.3.

### 2.1.3 Interactives as depictions

In 1765, the French writer Charles De Brosse wrote on interjections:

[...] les interjections sont courtes: elles partent du mouvement machinal & tiennent partout à la langue primitive. Ce ne sont pas des mots mais quelque chose de plus, puisqu’elles expriment le sentiment qu’on a d’une chose; & que par une simple voix prompte, par un seul coup d’organe elles **peignent** la manière dont on s’en trouve intérieurement affecté.

‘[...] interjections are short: they start from a mechanical movement and are everywhere in the primitive language. They are not words but something more, since they express the feeling one has of a thing; and that by a single prompt voice, by a single stroke of the organ, they **paint** the manner in which one is internally affected by it.’

(De Brosse 1765: 223; boldening ours)

The predicate *elles peignent* could in the sense of De Brosse (1765) be translated as both ‘they paint’ and ‘they depict.’<sup>13</sup> The view that interjections and other interactives ‘paint’ or ‘depict’ rather than describe meaning has in fact a long history, and it is also a view that is maintained in a number of contemporary studies: Ideophones, as well as other interactives, are widely characterized as depictions, they are said to display, show, demonstrate, or depict meaning (e.g., Dingemanse 2011: 17; Dingemanse and Akita 2017: 505; Haiman 2018: 57), or to be renditions of experiences rather than describing experiences (Dwyer and Moshi 2003)—in the wording of Kunene (1978: 3), to utter an ideophone is to ‘speak the act’ (see Haiman 2018: 79 for more detailed discussion).

<sup>13</sup> This interpretation was given to us by Franck Floricic (p.c. of August 7, 2021), who kindly provided us also with this quotation.

It would seem in fact that not only ideophones but also interactives in general are depictions. The question then is: What are depictions?

### 2.1.3.1 Depictions

The following observations are mainly based on [Clark \(2016, 2019\)](#); see also [Clark and Gerrig 1990](#); [Dingemanse 2013, 2015](#)), who distinguishes three basic methods of communication, each having different functions and displays, as shown in Table 2.2. Our interest is with (b) in Table 2.2, as well as with the way (b) can be distinguished from (c).

**Table 2.2** Methods, functions, and displays used in communication

|   | Method     | Function                                                            | Kind of sign |
|---|------------|---------------------------------------------------------------------|--------------|
| a | Indicating | To <i>locate</i> T for others in space and time (e.g., by pointing) | Index        |
| b | Depicting  | To <i>show</i> others what T looks, sounds, or feels like           | Icon         |
| c | Describing | To <i>tell</i> others about the properties of T                     | Symbol       |

Sources: [Clark 2016](#): 324, 342; [2019](#): 236

Note: T = the theme of a communication

On this view, a depiction is said to be the communicative display of an icon—that is, it is by definition an icon, and icons in the Peircean sense signify their objects via a perceptual resemblance to the object ([Peirce 1932](#)). Thus, an icon differs from an index, that is, a sign that signifies its object by means of a physical connection to the object ([Clark 2019](#): 236).

Depictions or, more narrowly, *performed* rather than exhibited depictions,<sup>14</sup> are physical scenes that people stage for others—by means of depictions people create one physical scene to represent another. Thus, they consist of something depicting (the ‘proximal scene’) and something depicted (the ‘distal scene’). They are described by Clark thus:<sup>15</sup>

Depictions, I argue, are *physical analogs* of what they represent. They are visible, audible, tactile, or proprioceptive models of things that one could actually see, hear, touch, or feel.

([Clark 2016](#): 327)

<sup>14</sup> Exhibited depictions include for instance paintings.

<sup>15</sup> Our understanding of depiction differs slightly from that of [Clark \(2016: 324\)](#), who treats nods and thumb-ups as descriptions while for us they are depictions.

Example (17), reported in the *New Yorker* of August 1, 2011, may illustrate what depiction in the sense of Clark (2016) is about.

- (17) English (Clark 2016: 325; the paraphrase of the depiction is printed in italics)  
 [Film director WK was speaking to reporter TF] ‘We were going to shoot on top of the MetLife Building ... but falcons had nested there, so they shut it down. In L.A., they would have—’ *He leveled a finger at some imaginary nestlings and made a gun-cocking sound.*

Although film director WK could have continued with the descriptive phrase *shot those falcons*, he continued with a depiction—a visible, audible model of shooting. The utterance printed in italics thus stands for a depiction whose descriptive content is paraphrased by Clark (2016: 324) as ‘shot those falcons.’

Example (17) shows what Clark (2016: 325–6) calls an embedded depiction, that is, it functions as part of the preceding utterance. Alternatively, depictions can be either adjuncts (adjoined as non-restrictive modifiers), indexed (introduced by ‘this long,’ ‘like this,’ or ‘here’), or independent (making an independent contribution to the discourse). Depictions often consist of manual or facial gestures, but they may also have various other manifestations. Thus, example (18) presents what Clark calls an ‘independent depiction,’ consisting of a musical performance by the student.

- (18) English (Clark 2016: 326; the paraphrase of the depiction is printed in italics)  
 [NT was asking a student about her interpretation of a musical phrase.]  
 Pianist NT: How does it manifest itself?  
 Student: *(plays a brief passage on the piano)*  
 Pianist NT: Oh! Okay.

The three kinds of signs, icon, index, and symbol, frequently overlap and are hard to separate from one another in depictions—they are generally mixed signs in the sense of Peirce (1932). For one thing, a given depiction may relate to more than one kind of sign. For example, an ‘indexed depiction’ (Clark 2019: 327) seems to manifest both an index and, since depictions are said to be icons (cf. Table 2.2), also an icon. Furthermore, Clark (2019: 325) defines a symbol as a sign that signifies its object by rule. A conventional gesture (an ‘emblem’ in his terminology), like crossing one’s fingers to signify good luck by rule, is therefore also a symbol, since it signifies its object by rule. But in its capacity as a depiction, a conventional gesture is also an icon. Furthermore, a conventional gesture is, as we argue, also defined with reference to its use in the here-and-now of the situation of discourse, hence it signifies its object by means of a physical connection to the object, that is, it is also an index.

In view of such problems, locating depictions with reference to the Peircean triad icon, index, and symbol must be taken with care.

### 2.1.3.2 Interactive depictions

We submit that interactives can be interpreted as a type of depictions in the sense of [Clark \(2016, 2019\)](#)—a type that is here called *interactive depictions*, that is, depictions using interactives as depicting material.<sup>16</sup> All other kinds of depictions will be called *non-interactive depictions*. This terminology differs from that proposed by [Clark \(2019\)](#), for the following reason. Clark distinguishes two kinds of depictions, namely codified and free depictions. The former include in particular ideophones and conventional gestures, while the latter include other kinds of depictions. Now, ideophones are a type of interactives (see [Table 1.6](#) of Section 1.5), hence they contrast with conventional gestures, which are non-interactive depictions.

Interactive depictions are staged scenes that use tokens of interactives as elements.<sup>17</sup> But are they physical analogs of what they represent, as stipulated for depictions (Section 2.1.3.1)? For one thing, we prefer to view them as linguistic rather than physical representations, thus distinguishing them from representations such as gestures.<sup>18</sup> And for another, it would seem that the notion ‘analog’ is not without problems when applied to interactives, including ideophones. For example, [Dingemanse \(2019: 15\)](#) observes with reference to ideophones that ‘they represent scenes by means of structural resemblances between aspects of form and meaning.’ The question may therefore be extended thus: Are interactives analogs that represent ‘structural resemblances’ between the depicting and the depicted item?

[Clark \(2019: 242\)](#) volunteers the following answer: Depictions are determined most of all by depictive and supportive aspects that apply to the depiction as a whole. He provides the following example concerning the English ideophone *thump*, as used in ‘codified depictions’ like *My heart went ‘thump’*. In the case of this ideophone, depictive aspects are said to be that it consists of (a) a closed syllable, (b) a low vowel, and (c) a nasal in the unit’s final cluster. And supportive aspects are that (d) it is an English form, and that it has both a (e) specific phonological form and (f) meaning (‘a dead, dull, somewhat hard sound’) ([Clark 2019: 242](#)).

One may wonder whether such aspect features are sufficient to establish that a codified depiction of this kind is in fact suggestive of an analog in the sense of

<sup>16</sup> However, interactives do not represent the only type of vocal depictions. A different type is represented by direct quotations, though not by indirect speech (cf. [Clark 2016: 324](#)). For a detailed analysis of direct quotations as demonstrations (or depictions) see [Clark and Gerrig \(1990\)](#). They are in need of a separate analysis within the framework of discourse grammar ([Heine et al. 2013](#); see Section 5.2.3).

<sup>17</sup> For [Clark \(2019: 243\)](#), codified depictions are staged scenes that use tokens of ideophones or emblems as elements, where emblems ‘belong to codes.’

<sup>18</sup> Obviously, in proposing this convention we are ignoring that linguistic activity is ultimately also physical activity.

signifying their objects via perceptual resemblance to the object (see Table 2.2 above). It would seem, for example, that there are a number of forms in the English language that appear to exhibit essentially the same set of aspect features as *thump*, such as *thumb*, *trump*, *pump*, etc., each with its own specific form and meaning. So what distinguishes the ideophone from such forms? The answer provided by Clark is the following:

The first three aspects are depictive, because they match the low muffled sound of a heartbeat, and the next three are supportive—necessary but not depictive. In German, Danish, and Turkish, the counterparts for *thump* are *bumm*, *dunk*, and *gum*. All four of these words have the same depictive aspects. Where they differ is in the supportive aspects ‘language’ and ‘phonological form.’

(Clark 2019: 241)

These observations are intuitively appealing, but are they sufficient to establish that there is perceptual resemblance between *thump* and its depicted element? Is there appropriate language-internal and/or cross-linguistic evidence beyond German, Danish, and Turkish in support of this assumption?<sup>19</sup> As for cross-linguistic evidence there is conceivably a number of equally promising, or even more promising interactives in addition to English *thump*, such as the response elicitor *huh?* (Dingemanse, Torreira, and Enfield 2013; Section 3.7.5) or the attention signal *hey* (Section 3.1.5). A central problem is how perceptual resemblance between a depiction and the depicted can be measured in terms of empirical techniques. The problem is nicely characterized by Clark (2019) himself in the following way:

*Oink-oink*, *cock-a-doodle-do*, and *bang* are ideophones in English; *guoin-guoin*, *cocorico*, and *pan* are the equivalents in French. People who don’t know English or French can only guess at what these words signify. And people’s guesses are, in general, terrible.

(Clark 2019: 241)

In short, whether or to what extent interactives ‘invite iconic associations’ (Dingemanse 2021a), that is, whether there is an analog between interactives and what they depict is a question that requires further information. Consider, for example, the following conclusion reached by Haiman (2018):

The vast majority of ideophones, however, are neither onomatopes or mouth charades, and relate to visual, tactile, olfactory, or motor impressions, which the

<sup>19</sup> The German equivalent of English *thump* appears to be *bums* rather than *bumm*, while the English equivalent of *bumm* is typically *bang*.

sounds and motions of the articulators simply are not able to iconically express ... But there is nearly nothing in the mere acoustic or motor phonetics of these dramatic 'depictive' words of showing that make them any more 'imitative' of their denotations than the neutral narrative 'descriptive' words of telling that correspond to them.

(Haiman 2018: 83)

To be sure, a number of authors point out that there is some amount of iconicity between the depicting and the depicted item in at least some ideophones but, as observed by [Güldemann \(2008: 281\)](#), 'the role of arbitrariness in this domain must not be underestimated.' Such iconicity is presumably more pronounced in gestures than in the interactives that they accompany. For example, the ideophone *kpiini-kpiini* 'stinky' of the West African language Kisi is accompanied by a crinkling of the nose and the same gesture is used for the ideophone *thuu* 'smelling horribly' of the Venda language of South Africa ([Childs 1994: 196](#))—in other words, nose-related gestures expressing stench appear to be fairly common while linguistic expressions for stench exhibit considerably more cross-cultural diversity.

Nevertheless, pending future research to provide such evidence, we follow [Dingemanse \(2013, 2015\)](#), [Haiman \(2018\)](#), [Clark \(2019\)](#) and others in assuming that ideophones and other interactives are in fact depictions of some kind. [Dingemanse \(2018: 19\)](#) notes that depiction is 'a mode of communication relying on perceptual analogies' and the question to be addressed in such research is: How can perceptual analogies be identified? There is so far no uniform answer, which is contingent on whether one relies on language-internal, cross-linguistic, and/or neurolinguistic evidence. It would seem that interlocutors tend to perceive expressions that are immediately anchored to the situation of discourse, that is interactives, as direct experiences and analogical representations.

Conventional gestures and interactives differ clearly from one another in their form. The former are physical in the sense that they involve body movements, while the latter are vocal depictions ([Dingemanse 2021a](#)), using linguistic substance like *thump*, *huh?*, or *hey*. But the two constitute functionally largely equivalent alternatives for expressing one and the same concept in many situations—to return to example (17): Instead of presenting a free depiction of shooting falcons, the film director might as well have drawn more concisely on the ideophone *bang* to achieve a similar effect. But what about other types of interactives, like the ones illustrated with the English examples in (19)? They involve a range of interactives other than ideophones: *brrr* and *yuck* are interjections, *shh* is a directive, *yes* is a response signal, while *bye* is a social formula (see [Table 1.6](#) of Section 1.5). It would seem that they also qualify as what Clark calls codified depictions.

- (19) English examples of gestural vs. interactive depictions
- a In order to express the concept of disgust, one may use, e.g., a facial gesture and/or a palm extended and turned forward (Haiman 2018: 71), or an interactive, using the interjection *yuck!*.
  - b To signal that one is feeling cold, one may draw on a gesture, using a shudder and/or wrapping one's arms about oneself as if to shelter oneself from the cold, or on an interactive, namely the interjection *brrr!*.
  - c To demand silence, one may draw on a gesture, placing one's index finger vertically on the lips, or on the directive *shh!*.
  - d To answer a question in the affirmative, one may draw on gesture, using a nod, or on the response signal *yes*.
  - e To take leave, one may draw on a waving gesture, or on a social formula like *bye!*.

Importantly, gestures and interactives do not commonly occur each on their own but rather tend to be produced concurrently, forming composites. Concepts like disgust, feeling cold, or demanding silence tend to be rendered each by means of a complex expression consisting jointly of a gestural and an interactive depiction. In their sample of Japanese ideophones ('mimetics'), Dingemanse and Akita (2017) found 85% of the codified depictions of ideophones to include gestures.

Thus, in accordance with the voice-gesture model proposed in Section 1.3.2.3, the two kinds of depiction, gestures and interactives, seem to be part and parcel of one and the same knowledge package that interlocutors dispose of for structuring their expression of interactive concepts.

In sum, while differing in their form from gestures and other kinds of physical depictions, interactives, being 'independent acts of communication' in the sense of Clark (2019: 246), also qualify as a type of depiction, especially for the reasons given in (20).

- (20) Features shared by gestural and interactive depictions
- a Both can be used as functionally largely equivalent alternatives in specific contexts
  - b Both are commonly used concurrently as composite expressions
  - c Both are indexicals
  - d Both are directly accessible from the situation of discourse
  - e Both have a descriptive meaning component which can be elucidated by means of paraphrase

(20a) and (20b) have just been discussed. Being indexicals (cf. (20c)) is a definitional property of interactives (Section 1.2) and like interactives, gestures are

normally anchored to the here-and-now of the situation of discourse.<sup>20</sup> (20d) needs more detailed discussion, for which see Section 2.1.3.3.

With reference to (20e), the significance of paraphrase for rendering the descriptive meaning component of interactives has been shown in Section 2.1.1. And much of what was said about interactives also applies to gestures and other physical depictions. For example, the descriptive meaning of the physical depiction in (17) was paraphrased by ‘shot those falcons,’ while that of the independent depiction in (18) is paraphrased by Clark (2016: 326) by ‘(student) plays a brief passage on the piano.’ In accordance with such observations, Dingemans and Akita (2017: 519) write that the two types of depiction ‘are tightly coupled and alike in mode of representation: they perform the same role of depicting sensory imagery, albeit in different modalities.’

### 2.1.3.3 Depiction vs. description

But how does depiction differ from description? The difference has been said to concern, on the one hand, knowledge and, on the other hand, processing. With regard to the former, Clark (2016: 324) observes that depiction and description are based on fundamentally different modes of thinking:

With descriptions, people rely mainly on their knowledge of the vocabulary, syntax, and semantics of, say, English or Japanese and their ability to categorize. With depictions, they rely mainly on their visual, auditory, tactile, and proprioceptive knowledge of physical scenes and on their ability to use one scene in imagining another. Describing depends on knowledge of a language or code. Depicting does not.

(Clark 2016: 324)

This characterization must be taken with care, however, since depicting of the kind proposed by Clark and adopted here also depends on knowledge of some ‘code’: Usage of interactive depictions, such as those of the English ideophone *thumb* or the interjection *brrr* depends on the ‘code’ of interactive grammar: Both are part of the inventory of English ideophones and interjections, respectively, and hence both have to be learned and are transmitted from one generation to the next. In this sense they can also be said to be ‘symbols’ since they signify the depicted objects by rule (Clark 2019: 235)—note that Güldemann (2008: 280) defines ideophones as ‘expressive and symbolic lexemes.’

And this also applies, at least to some extent, to gestural depictions. For example, Clark (2019: 235) notes that in certain cultures, ‘crossing one’s fingers signifies

<sup>20</sup> It goes without saying that the usage of interactives can be transferred from the here-and-now to the temporally displaced situation of narrated speech, but that also applies to the usage of gestures (see also Section 3.5.7).



good luck by rule, so it is also a symbol.' And much the same applies to gestural depictions of the kind mentioned in (19): Their usage is also conventional, being part of the knowledge acquired by speakers of English. To conclude, 'knowledge of a language or a code' remains a controversial criterion for distinguishing (performative) depictions from descriptions.

With regard to processing, Clark (2016: 342) argues that there are two fundamentally different processes involved. Descriptions are formed via linearization, that is, by arranging arbitrary symbols into linear strings. Depictions, in contrast, are said to involve spatialization, in that interlocutors spatialize the elements to depict 'by selecting a placement, orientation, scale, granularity, and dimensionality for them.'

On the perspective proffered here, manner of processing is an essential criterion for distinguishing depictions from descriptions. Take the interjection *brrr* just mentioned, commonly paraphrased by the sentence *I feel cold*. In order to process this sentence, interlocutors need (a) to be familiar with the grammar of English, (b) to take into account that this is a sentence (c) constructed in the present tense, (d) that *feel* is the predicate instantiating a specific class of verbs, (e) that *I* and *cold* are arguments of *feel*, (f) that *I* is a personal pronoun while *cold* is an adjective, etc. In sum, interlocutors need to push their way through the whole machinery of English sentence grammar to encode or decode a sentence like *I feel cold*, and this requires a knowledge of the structural templates of English word, phrase, and sentence formation.

This is different in the case of *brrr*: For processing, the task of the speaker is simply to search for an appropriate expression from the pool of interactives—in the present case the expression *brrr*—and an optimal placement slot for the expression to be presented in the discourse, rather than being constrained by the semantic and morphosyntactic templates mentioned.

And in the same way as semantic and syntactic templates, the processing of *I feel cold* also requires a knowledge of English phonological templates relating, for example, to how consonants and vowels are arranged into words. Again, this is different with the English interjection *brrr* just mentioned, or with the attention signal *pst*, the directive *shh*, or the response signal *mhm* for that matter: They are elusive to phonotactic templates of English word formation, apparently being processed by moving straight from the depicting item to the item depicted.

Interactive depictions thus are processed immediately without being constrained by the machinery of sentence grammar—as Kita (1997: 381) puts it with reference to Japanese ideophones ('mimetics'), 'speakers have the intuition that the sound–meaning relationship is direct, immediate, and nonarbitrary.' And much the same applies to gestural depictions used for interactive functions: They are also processed immediately from the situation of discourse (the context). Depiction in this sense, thus, refers to an immediate link between an expression and the item expressed, or between the depicting and the depicted item.

In sum, interactives and gestures conform to what we propose to call the *direct-access hypothesis*: Both are immediately accessible for discourse processing without being constrained by any intermediate mechanism such as sentence grammar.

What all this is meant to show is that there are three different means available to interlocutors for structuring their discourses, namely (a) appropriate physical resources used for gesturing, (b) interactive forms used for vocal depiction, and (c) the virtually unlimited potential provided by sentence grammar for description. General characteristics of these three modes are summarized in Table 2.3 (cf. Clark 2016: 342; 2019; Dingemanse 2015: 950–1).<sup>21</sup>

As was observed in Section 2.1.3.1 on the Peircean distinction of signs between icons, indexes, and symbols, the situation is complex and it is therefore not addressed in Table 2.3, except for the following: Being indexicals (cf. (20c)) is a definitional property of interactives (Section 1.2), and much the same seems to apply to gestures: Both are normally anchored to the here-and-now of the situation of discourse. This observation has been made in a number of studies; for example, Kita (1997: 392) notes that Japanese ideophones (‘mimetics’) reside in the same dimension of representation as the gestures associated with them and thus contrast with forms belonging to the descriptive dimension, that is, to sentence grammar. With their reliance on the descriptive mode, units of sentence grammar are for the most part symbols, even if they also have an indexical component, e.g., in personal pronouns, demonstratives, adverbs, tenses, etc.

**Table 2.3** Three basic modes of representing information for discourse processing

|                           | (a) Gestures                                                       | (b) Interactives    | (c) Descriptions                                      |
|---------------------------|--------------------------------------------------------------------|---------------------|-------------------------------------------------------|
| <i>Material used</i>      | Visual, auditory, tactile, and proprioceptive activity             | Interactive forms   | Lexical and grammatical resources of sentence grammar |
| <i>Knowledge required</i> | Conventional patterns of physical scenes                           | Interactive grammar | Sentence grammar                                      |
| <i>Mode of processing</i> | Depiction                                                          | Depiction           | Description                                           |
| <i>Nature of signs</i>    | Indexicals                                                         | Indexicals          | Mostly symbols                                        |
| <i>Example</i>            | [Facing the hearer, I place my index finger vertically on my lips] | <i>Shh!</i>         | <i>I ask you to be quiet!</i>                         |

The position just outlined can be reconciled with observations made in a range of other studies on interactives. On the one hand, it can be said to be compatible

<sup>21</sup> Dingemanse (2015: 951, Table 2) proposes the following criteria for distinguishing (linguistic) depiction from description: (a) form–meaning mappings, (b) building blocks, and (c) interpretation. We have adopted (c) but do not find (a) and (b) to be clearly significant parameters here.

with notions of a continuum extending from linguistic to non-linguistic phenomena that have been proposed by some authors. This applies in particular to Wharton (2003, 2016), who concludes that interjections are partly natural and partly coded, falling at various points along a continuum between language proper and display:

The continuum also allows us to capture the heterogeneity and marginal linguistic status of the class in general and makes it relatively easy to see how a given expression (e.g. an interjection) might move along the continuum from ‘non-linguistic’ to ‘partly linguistic’ to ‘linguistic’ without radically altering the type of information it conveys. Seeing interjections in this way, we should not be surprised either at their descriptive ineffability, or the fact that what they convey is sometimes too nebulous to be paraphrased in determinate conceptual terms: they are *partly* natural responses.

(Wharton 2016: 29; italics in the original)

In Wharton’s terms, (a) of Table 2.3 can be said to concern the non-linguistic and (c) the linguistic end of the continuum, while (b) relates to the intermediate stage of the partly linguistic and partly natural within the continuum.

On the other hand, the depictive nature of interactives has been characterized in terms of notions such as ‘showing,’ ‘displaying,’ ‘demonstrating,’ ‘experiencing,’ ‘performing,’ or ‘enacting’—all these notions contrasting with that of describing as the salient feature of sentence grammar. For example, dealing with expressives, which include social formulae (*hi, thanks*), interjections (*damn, wow*) and response signals (*yes*), Langacker (2008: 475–7) writes about the profiling of expressives:

In using one [an expressive; a.n.], the speaker is either performing a social action or vocally manifesting an experience – rather than *describing* a scenario, he *enacts a role* in it.

(Langacker 2008: 476; italics in the original)

Instead of ‘enacting,’ the notion of ‘experiencing’ is highlighted in other contributions to interactives. In their study of ideophones, Dwyer and Moshi (2003: 174–5, 179) propose a distinction between two ‘autonomous semiological systems,’ or between an analytic and an expressive dimension of semiotic categories. Whereas the former dimension is *about* a certain experience, the latter is *a rendition of the experience itself*. For example, ‘one may conceptualize unpleasantness in the analytic dimension without actually feeling any unpleasantness,’ while expressives represent ‘different facets of an experience’ (see also Bally 1909; Bühler [1934] 1965). And for Dingemanse (2011: 38), ideophones ‘enable others to experience what it is like to perceive the sensory imagery depicted.’

This distinction is informed by that of Kita (1997: 380) between an analytic and an ‘affecto-imagistic’ dimension of semantic representation. A paradigm feature distinguishing the two is that the former is ‘decontextualized’ where the latter is not. The distinction roughly echoes that between sentence grammar and interactive grammar and corresponds to that between descriptive and expressive meaning made here (see Section 2.1.2).

In sum, we follow Dingemanse and associates, Haiman (2018), and Clark (2019) in arguing that interactives are performative depictions of processes, states, objects, or qualities. But, on the one hand, we extend their analysis from ideophones to interactives in general. And on the other hand, we argue in accordance with the observations made in Section 2.1.1, that in addition to their expressive meaning, both kinds of depiction also have a *descriptive* meaning component—one that can be elucidated by means of paraphrase. For example, we saw in Table 2.3 that the process of asking someone to be quiet can be framed either as an interactive (saying *Shh!*) or as a gesture (placing the index finger vertically on the lips), and more likely by a composite of both, but an account of the process would be incomplete without knowing what the two kinds of depictions are about, that is, what their function is. This function can reasonably be determined only in terms of the descriptive meaning component, which is accessible via a paraphrase like ‘I ask you to be quiet!’.

Most important for the purpose of the present discussion, however, is the observation that like gestures, interactives differ from units of sentence grammar in being *immediately accessible*: What interlocutors need to do is simply draw on the pool of interactives, or on their ability to create new interactives by using the resources available in sentence grammar and deploy them in discourse unaffected by constraints imposed by the complex machinery of sentence grammar. It therefore comes as no surprise that interactives have the appearance of ‘displays,’ ‘vivid representations,’ ‘enactments,’ ‘staged scenes,’ or ‘dramatic performances.’

#### 2.1.4 Discussion

In Section 2.1.1 we drew attention to the hypothesis proposed by Kaplan (1999: 11) according to which English units like *ouch* and *I am in pain* convey the same information but do so ‘through different modes of expression.’ The observations made in the preceding Section 2.1.3 are in fact in accordance with this hypothesis: Unlike units of sentence grammar like *I am in pain*, interactives, like *ouch*, draw on the mode of depiction for their expression.

The approach sketched here is not the only one that has been proposed on the meaning of interactives. Whether interactives in fact have descriptive meaning has been discussed controversially (see, e.g., Padilla Cruz 2009; Wharton 2016). It has been argued, for example, that interactives such as interjections have procedural

rather than conceptual meaning and that they are non-referential, that is, they cannot be addressed in discourse (e.g., Wharton (2003, 2016: 26; Gehweiler 2008: 74). And according to Biber et al. (1999: 1082) interactives ('inserts') have no denotative meaning, in that 'their use is defined rather by their pragmatic function.' As we saw in Section 2.1.2, both assumptions are in need of qualification (see Meinard 2015: 160–2).

Wharton (2003: 61; 2016: 21, 25) argues that rather than describing meaning, interjections encode procedures, they involve displaying (or showing) instead of saying: Rather than encoding conceptual structure, Wharton (2003: 60) assumes, interjections 'encode procedural information which 'points' in the general direction in which relevance should be sought,' activating a range of attitudinal descriptions which involve delight, surprise, excitement, etc. (Wharton 2003: 59–60).

This view captures salient elements of the position maintained here but, as pointed out in Section 2.1.2, interjections or, more generally, interactives such as the ones discussed here are not restricted to encoding procedures—they also encode and depict concepts such as delight, surprise, disgust, or pain.<sup>22</sup> Accordingly, they cannot reasonably be reduced to one general function:<sup>23</sup> On the one hand, there is a descriptive component, also referred to as 'rational' or 'cognitive' meaning, relating to Bühler's ([1934] 1965) representational function (*Darstellungsfunktion*). It is captured here—in want of a more appropriate technique—by means of paraphrases like the ones proposed in Section 2.1.1: Paraphrase provides a way of getting in a propositional format as closely as possible at the descriptive content of interactives; see Munaro (2010) for a similar use of paraphrase to 'decompose' interjections. As was observed already in Section 2.1.1, however, this is by no means intended to imply that interactives have a propositional organization—paraphrase simply serves as a convenient heuristic for elucidating salient conceptual features of interactives.

On the other hand, there is an expressive meaning component, relating to Bühler's expressive function (*Ausdrucksfunktion*), which can be captured reasonably only with reference to the range of usages that interactives have. Interactives, like *oops*, *ouch*, *pst*, *shh*, *wow*, etc., have therefore been referred to in short as 'expressives'.<sup>24</sup> For example, distinguishing two semiotic dimensions, Dwyer and Moshi (2003: 174, (3a)) suggest that rather than to the analytic dimension of morphosyntax, ideophones belong to the expressive dimension (see also Kita 1997: 380).

<sup>22</sup> Whether disgust, pain, and surprise really qualify as 'concepts' must remain debatable though (Franck Floricic, p.c. of July 22, 2021).

<sup>23</sup> One is reminded here of the distinction made by speech-act theorists between describing and indicating (see Wharton 2016: 24).

<sup>24</sup> For example, Dwyer and Moshi (2003: 1) say: "Pure ideophones" stand outside syntactic grammar and belong instead to [the] distinct category of "expressives".

Defining expressives ‘as operators that actively change the context in specific ways,’ Potts (2007a: 166–167) identifies the following characteristics of expressive content: (a) independence (from descriptive content), (b) non-displaceability (expressives need to relate to the utterance situation), (c) perspective dependence (typically the speaker’s), (d) descriptive ineffability, (e) immediacy (the performative character of expressives), and (f) repeatability (repeating expressive items generally strengthens the emotive impact).

Note that Potts’ primary concern is with semantic content rather than with the nature of the items carrying that content. At the same time, his interest is quite different from the present one: Rather than dealing with interactives he is concerned with items that are anchored to the structure of a sentence. Thus, in the example *That bastard Kresge is famous* discussed by him, *that bastard* has expressive meaning but is encoded as a noun phrase having the syntactic function of the subject of the sentence.

Accordingly, Potts’ concept ‘expressive meaning’ is not exactly the same as the one to be observed in interactives. Nevertheless, there are similarities, most of all the following. For one thing, both approaches assume that expressive meaning is essentially different from descriptive meaning (cf. (a)). And for another, both have an indexical component, relating immediately to the situation of discourse (cf. (b), (e)) and typically resting on the speaker’s perspective (cf. (c)). Ignoring (f), that is, repeatability, which does not appear to be a characteristic of expressive meaning only, a difference to the approach preferred here can be seen in (d). While we admit that expressive meaning content is somehow ‘ineffable,’ interactives also contain a descriptive component, which can be elucidated via paraphrase (Section 2.1.1).<sup>25</sup> At the same time we agree with the following observation made by Potts: ‘Speakers are never fully satisfied when they paraphrase expressive content using descriptive, i.e., nonexpressive, terms’ (Potts 2007a: 166). There is an obvious reason for this: Even an optimal paraphrase remains restricted to the descriptive component of an interactive while the expressive component is beyond its scope.<sup>26</sup>

But what the research in the tradition of formal approaches like that of Potts (2007a, 2007b) and others (e.g., Gutzmann 2015, 2019; Gutzmann and Henderson

<sup>25</sup> We are ignoring the view that ‘a demand for a paraphrase of, say, *damn* is nonsensical – conceptually equivalent to asking for a paraphrase of *Chris Potts...*’ (Potts 2007a: 178) since this view is not appropriately substantiated.

<sup>26</sup> The following remark by Potts would seem to be in support of this interpretation: ‘I’ve only once been told that an uncontroversial expressive had an accurate paraphrase in descriptive terms: *bastard* was claimed to mean “vile contemptible person”. But this paraphrase misses its wide range of affectionate uses ...’ (Potts 2007a: 176–7). We suggest that such ‘affectionate uses’ are missing in the paraphrase because they concern the expressive component of the item *bastard*. Consider the following example. The English verb *look*, as in *Jane looked at me sadly*, has descriptive meaning, which can presumably be made explicit by means of appropriate paraphrases, e.g., in terms of alternative perception verbs. But what about the meaning of its interactive counterpart *look!*, as in the following example?

English (Biber et al. 1999: 216)

**Look**, do I have to tell you everything? (CONV)

2019) demonstrates is that ‘expressive meaning’ is not restricted to interactives, even if its manifestations in the latter appear to be of a somewhat different kind, being shaped by their function as depictions.

To conclude, it remains unclear at the present stage of research how exactly the expressive meaning space characterizing the use of interactives can be described in a principled way. Potts (2007a: 177) notes that ‘expressive content is not propositional, that it is distinct from the meanings we typically assign to sentences.’ In the tradition of Kaplan (1999) it can be explicated with reference to usage, especially to the contextual configuration and the specific social background underlying the usage of interactives. This task, however, is faced with two problems. First, it requires both a large corpus of data for analysis and a comprehensive analysis of these data—something that cannot be achieved here. The second problem concerns the fact that it is not only the meaning of interactives that is shaped by usage; rather, this applies in much the same way also to meaning in sentence grammar. Part of the task would therefore also be to determine how the usage of interactives can be distinguished and delineated from that in sentence grammar. It would seem that the distinction between depiction and description provides a means for tackling this task: Interactives are depictions in the sense proposed here whereas sentences are typically descriptions. But more research is needed on this issue.

More research is also needed on how exactly the notion ‘use conditional’ can be applied in the analysis of interactives.<sup>27</sup> In the tradition of Wittgenstein (1953), this notion has been employed in some form or other to understand meaning of all kinds, including the propositionally organized meaning of sentence grammar.

### 2.1.5 Conclusions

The concern of Section 2.1 was with finding an answer to the following question: What is the meaning of interactives? As the preceding sections have shown, we are not able to volunteer a satisfactory answer to this question. The magnitude of the problem can be illustrated by looking at one of the ten types of interactives, namely that of discourse markers. More than other types, they have been the subject of extensive research over the last decades, resulting in more than a hundred publications. But in spite of all the rich analyses that have been carried out in the course of that research, we are still far away from being able to describe the meaning of discourse markers in a coherent and concise format.

As is argued here, *look* in this sentence is an attention signal having an expressive component in addition to its descriptive one and being a depiction, which both account for its particular discourse function in this example.

<sup>27</sup> With reference to the meaning of a social formula like *goodbye*, Kaplan (1999: 3) distinguishes the semantics of meanings from the semantics of use, taking the social formula *goodbye* as an example and describing the latter thus: ‘I don’t ask, “What does *goodbye* mean?” Instead I ask, what are the conditions under which the expression would be correctly or accurately used?’

Thus, rather than proffering an answer to the question just raised we were satisfied with identifying some pieces of an answer, and these are in particular the ones discussed in the preceding sections of this chapter, summarized in (21).

- (21) Aspects of the meaning of interactives
- a Interactives have both an expressive and a descriptive meaning component and the latter can be elucidated by means of paraphrase.
  - b Interactives are depictions in the sense that they establish an immediate link between the depicting form and the item depicted without being edited and processed by the machinery of sentence grammar.
  - c In this respect they resemble conventional gestures used for the same or similar functions, which are also processed without involving any intermediate mechanism between the depicting and the depicted (see the direct-access hypothesis in Section 2.1.3.3).

Interactives are henceforth referred to as ‘expressive depictions.’ Note, however, that neither expressing nor depicting are restricted to interactives: Expressive meanings are also found in sentence grammar (Section 2.1.2), while depiction is a ubiquitous mechanism to be observed in many different domains of human activity (Section 2.1.3).

Being processed directly from the situation of discourse, interactives and conventional gestures are indexicals differing fundamentally from units of sentence grammar: Both are commonly characterized as ‘vivid representations,’ ‘enactments,’ ‘staged scenes,’ ‘displays,’ or ‘dramatic performances’—as if the speaker were ‘inhabiting the frame’ of the state concerned (cf. Kockelman 2003: 471). And such ‘performances’ can be reinforced by sound symbolic features that a number of interactives exhibit.

## 2.2 Identifying arguments

The main reason for discussing meaning in some detail in Section 2.1 was to prepare for the main subject of the present chapter, which is the status of discourse arguments or, in short, the ‘arguments’ that are part of the structure of interactives.

### 2.2.1 Introduction

Linguists commonly use notions such as constituent, thematic role, participant, argument, complement, etc. to describe salient features of sentence structure (or clause structure). For example, the sentence in (1) can be described in terms of grammatical or syntactic relations by saying that it consists of a subject (*David*), a verb (*kissed*), an object (*a mermaid*), and an adverbial phrase (*last Friday*).



## (1) English

David kissed a mermaid last Friday.

This is a garden-variety description—one of the possible morphosyntactic descriptions that there are. Alternatively, one could draw on semantic notions, as we will do here, by saying that (1) consists of a verb instantiating a predicate (*kissed*), an agent (*David*), a patient or undergoer (*a mermaid*), and a temporal participant (*last Friday*).

One can then go on and classify further in terms of semantic roles by adding that there is a predicate, *kissed*, which requires, licences or, as we will say here, takes two kinds of participants, namely two arguments (or core participants), that is, an agent and a patient, and one adjunct (or peripheral participant), namely a temporal participant. Arguments are obligatory participants, demanded by the predicate, which must be either stated or implied, that is, understood from the context. Adjuncts, by contrast, are optional participants, they can be omitted—thus, (1) would still be a well-formed sentence if *last Friday* were dropped.

For the purposes of the present discussion it is the general assumptions in (2) that are of particular interest. To be sure, (2) represents only one of the many ways in which the structure of a sentence can be and has been described.

## (2) Features characterizing sentence structure

- a A sentence consists minimally of a predicate and its arguments.
- b The predicate determines the type and the number of arguments demanded in the sentence.
- c Without the arguments stated or implied, the meaning of the predicate would be incomplete.
- d The number of arguments is limited, there are most commonly one, two, and normally no more than three arguments.
- e In addition to arguments there may be adjuncts, whose use is ‘optional’ in the sense that omitting them would not result in an infelicitous sentence.

Interactives are linguistic expressions that contrast sharply in their behavior with units of sentence grammar: They are syntactically and semantically independent from the sentence they may be associated with, they lack morphological devices to signal syntactic or semantic relations, and they are fairly free in their placement behavior (Section 1.2). As we argue in this section, however, there are significant parallels between the grammar of sentences and that of interactives, and we will draw on semantic notions used in the description of sentences to understand some salient features of the grammatical behavior of interactives. In particular we argue,

first, that interactives have an argument structure, second, that the distinction between arguments and adjuncts is also relevant to them and, third, that placement also plays an important role in describing the relation between interactives and surrounding text material.

Among the ‘interjections’ listed by Quirk et al. (1985: 853) there are the following forms: *ouch* (pain), *shh* (or *sh*, request of silence or moderation of noise), and *eh?* (impolite request for repetition). Ignoring the expressive meaning of these three forms, these interactives differ from one another in the participants they imply in order to be understood correctly. The participants implied are the speaker (S), the hearer (H), and some text piece, let us call it the theme (T), as can be shown by the paraphrases in (3) capturing at least one possible interpretation of the descriptive meaning of the three interactives (see Section 2.1.1).

(3) Possible paraphrases of three English ‘interjections’<sup>28</sup>

**Ouch** I (S) have just experienced a sudden and sharp pain!

**Shh** I (S) ask you (H) to be quiet!

**Eh?** Can you (H) repeat what you just told (T) me (S)?

It would seem that without information on who does what and to whom, that is, on what the participants are, use of the three interactives in (3) would be incomplete: The use of *ouch* presupposes someone having pain (S), *shh* presupposes not only someone making the request but also someone who is asked to be quiet (S,H), and *eh?* presupposes not only someone requesting and someone else requested to take action but also the theme of the action concerned (S,H,T). In other words, these participants can be said to be implied but obligatory parts of the forms. Therefore we propose to extend the term ‘arguments’ of sentence grammar to them, assuming that *ouch* has one argument (S), *shh* has two (S,H), while *eh?* has three (S,H,T).<sup>29</sup>

On this view, the generalizations proposed in (2) can be extended to the analysis of the interactives in (3), taking the following assumptions into account: First, instead of sentences, the structure of interactives is described in terms of utterances. And second, the units corresponding to predicates in sentence grammar are interactive forms in the grammar of interactives. Accordingly, (2) can be rephrased as in (4) to describe the semantic roles of the three forms in (3) and other interactives.

<sup>28</sup> The reason for presenting ‘interjections’ in inverted commas is that the classification by Quirk et al. (1985: 853) is not exactly the same as that proposed here (Section 3.6).

<sup>29</sup> Nevertheless, this analysis raises some problems, for which see Section 2.2.3.

- (4) Features characterizing the semantic roles of interactives
- a An utterance consists minimally of an interactive and its arguments.
  - b The interactive determines the type and the number of arguments demanded.
  - c Without the arguments stated or implied, the meaning of the interactive would be incomplete.
  - d The number of arguments is limited, there are most commonly one or two, and normally no more than three arguments.
  - e In addition to arguments there may be adjuncts, whose use is 'optional' in the sense that omitting them would not result in an infelicitous utterance.

We may illustrate the features listed in (4) with an example already mentioned in Section 2.1.1. The interjection *wow* in (5) forms an utterance of its own, expressing the emotive state of the speaker and having S as its argument (cf. (4a)). Part of the store of knowledge that interlocutors have when exposed to this utterance is, first, that interjections, such as *ouch* in (3) or *wow* in (5), take only one argument and, second, that that argument relates to the speaker, not to the hearer or someone else (cf. (4b)). Thus, a reasonable interpretation of the meaning of (5) would not seem possible without that knowledge (cf. (4c)). We observed that the interactives illustrated in (3) take either one, two, or three arguments. As we will see in the ensuing sections, three is actually the maximum number of arguments taken by interactives (cf. (4d)), but see also Section 2.2.3.

- (5) English  
**Wow!** That's absolutely amazing!

We are thus left with (4e). Notice that there is a second utterance in (5), namely *That's absolutely amazing*. This utterance elaborates further on the first utterance (*Wow!*) but its use is by no means required, it could be omitted without significantly affecting the meaning or the felicity of the first utterance. Thus, its use is optional and has the features corresponding to what in sentence grammar is called an adjunct (cf. (2e)).

To conclude, there are reasons for extending the notion of argument structure to the analysis of interactive grammar, arguing that the latter exhibits a 'valence' structure not unlike that of sentence grammar. In doing so, however, one has to be aware that there are limits to the parallels between the two domains of grammar, in particular the following.

First, rather than in the structure of a sentence, interactives are anchored to and shaped by the situation of discourse and their functions and placement are therefore derivative not of any syntactic constraints but rather of their role in discourse. For this reason, certain concepts used to describe sentence grammar are avoided here. This applies in particular to the term 'syntax,' for which there does not appear

to be an immediate equivalent in the grammar of interactives. Thus, instead of 'syntactic relation' we propose the less committal notion 'collocation pattern' to describe regularities of co-occurrence among interactives or between interactives and other text material. Accordingly, no claim is made that there is any parallel to notions of sentence grammar such as syntactic hierarchy; rather, we will be satisfied with pointing out constraints on co-occurrence relations which are derivative of the discourse functions of interactives.

Second, there is also a difference between the two domains of grammar in their fields of operation. In sentence grammar, it is the sentence and its propositional format that constitute the central units of operation. In the grammar of interactives, by contrast, the field of operation consists essentially of the discourse at large. Accordingly, argument structure in the latter does not relate to a clause or a sentence but may extend over larger pieces of discourse. Thus, in example (5) it involves two separate utterances but as we will see in the paragraphs to follow, it may as well extend, for example, over different turns of a conversational exchange.

And third, the conceptual basis of arguments is also different. In sentence grammar, arguments are shaped by a propositional understanding of events and states, involving semantic roles like agent, patient, experiencer, recipient, place, and time. In the grammar of interactives, the conceptual basis is different, being anchored to and shaped by the situation of discourse. Accordingly, the two main participant roles distinguished, namely S (speaker) and H (hearer), are defined deictically with reference to that situation. This means, on the one hand, that arguments are implicitly recoverable and, hence, need not and do not normally receive any formal encoding. On the other hand, participant roles relating to place and time would be redundant since the use of interactives is restricted to the here-and-now of the situation of discourse. It is only the third role T (theme) that relates interactives to the external world of discourse (Section 2.2.3).

For a more elaborate comparison of the argument structure in sentence grammar and interactive grammar, the reader is referred to Section 2.3.

### 2.2.2 The arguments

The examples just presented show that interactives can and quite commonly do occur as 'stand-alones', that is, as expressions forming utterances of their own. This does not clearly apply to all interactives but it is a conspicuous feature of most types of interactives. The examples in Table 2.4, reproduced for convenience from Section 1.2, provide a collection of them.

While being in fact independent units in the sense that they are normally prosodically and syntactically distinct from what precedes and follows them, the utterances in Table 2.4 are not 'self-contained' but rather require a specific situation of discourse for their interpretation. Nevertheless, each of the utterances

**Table 2.4** Types of English interactives

| Type              | English example |
|-------------------|-----------------|
| Attention signal  | <i>Hey!</i>     |
| Directive         | <i>Shh!</i>     |
| Discourse marker  | <i>Indeed.</i>  |
| Evaluative        | <i>Great!</i>   |
| Ideophone         | <i>Thud!</i>    |
| Interjection      | <i>Ouch!</i>    |
| Response elicitor | <i>Right?</i>   |
| Response signal   | <i>No.</i>      |
| Social formula    | <i>Bye!</i>     |
| Vocative          | <i>Mom!</i>     |

expresses information that speakers of English are normally aware of. Using paraphrase as a tool of analysis, one salient kind of descriptive content can, at least in specific contexts, be sketched as in Table 2.5.<sup>30</sup> As the paraphrases in Table 2.5 suggest, the interactives listed are associated with the three argument roles proposed in the preceding section, namely S (= speaker), H (= hearer), and T (= theme, i.e., information on the state of affairs implied by the usage of the interactive concerned).

**Table 2.5** Approximate paraphrases of the interactives in Table 2.4

| Example        | Possible interpretation                                                              |
|----------------|--------------------------------------------------------------------------------------|
| <i>Hey!</i>    | I (S) am calling your (H) attention (to the following (T)).                          |
| <i>Shh!</i>    | I (S) want you (H) to be quiet.                                                      |
| <i>Indeed.</i> | I (S) find that what you (H) say (T) is true.                                        |
| <i>Great!</i>  | What I (S) have just experienced (T) is great.                                       |
| <i>Thud!</i>   | Something (T) gives the impression of a heavy solid object falling to the ground.    |
| <i>Ouch!</i>   | I (S) have just experienced a sudden and sharp pain.                                 |
| <i>Right?</i>  | Am I (S) right in saying what I've just said (T) to you (H)?                         |
| <i>No.</i>     | I (S) disagree with what you (H) say (T).                                            |
| <i>Bye!</i>    | You (H) and I (S) are now parting from one another for a significant period of time. |
| <i>Mom!</i>    | I (S) am calling the attention of you ( <i>mom</i> , H).                             |

<sup>30</sup> See Kaplan (1999) for some details. It goes without saying that the paraphrases presented in Table 2.5 are far from taking appropriate account of the specific functions expressed by the interactives—a task for which there so far is no reasonable solution. However, Kaplan (1999: 17) argues that *goodbye* and *you and I are now parting from one another* are informationally equivalent.

These observations are suggestive of the following generalizations. First, without information of the kind presented in Table 2.5, a description of the forms in Table 2.4 would be incomplete. Second, the information sketched in Table 2.5 is not overtly expressed, it is implicitly contained in the interactives concerned. Third, the information conveyed by interactives is contingent on the nature of what we called in the preceding section the ‘arguments’ they contain; we are referring to these arguments in Table 2.6 with the labels ‘S,’ ‘H,’ and ‘T.’ Fourth, with one exception, all interactives have an obligatory S argument, the majority have two arguments and some even have three arguments. Their argument structure thus can be represented as in Table 2.6. And fifth, these arguments are not defined with reference to the structure of a sentence; rather, they are contextually determined, that is, they are anchored immediately to the situation of discourse.

**Table 2.6** Prototypical argument structure of the interactives in Table 2.5

| Form            | Type              | Arguments                            |
|-----------------|-------------------|--------------------------------------|
| <i>Hey!</i>     | Attention signal  | S,H,(T)                              |
| <i>Shh!</i>     | Directive         | S,H                                  |
| <i>Indeed.</i>  | Discourse marker  | T <sub>1</sub> ,S,(H,)T <sub>2</sub> |
| <i>Great!</i>   | Evaluative        | S,H,T                                |
| <i>Thud!</i>    | Ideophone         | T                                    |
| <i>Ouch!</i>    | Interjection      | S                                    |
| <i>Right?</i>   | Response elicitor | S,H,T                                |
| <i>No.</i>      | Response signal   | S,H,T                                |
| <i>Bye!</i>     | Social formula    | S,H                                  |
| <i>Darling!</i> | Vocative          | S,H                                  |

By calling the structures listed in Table 2.6 ‘prototypical argument structures’ we are adopting a simplified stance on the structures to be observed in interactives. First, we are reducing the information expressed by interactives to what may be called its ‘core,’ that is, to that of participant marking. This, however, is not the only information conveyed by them. Interjections like *ouch* in Table 2.6 are not restricted to signaling an S argument, in addition they may also index that there are specific sources being responsible for their usage. According to Kockelman (2003: 471), ‘interjections’ have a rich indexical relationship, indexing either objects, signs, internal states, or social relations. For example, the interjection *chix* of the Maya language Q’eqchi’ of Guatemala and Belize may index not only a loathsome object in the situational context but also a social relation, e.g., parent–child, or the internal state of disgust.

And second, these structures are not the only ones associated with a given type of interactives. As we will see in more detail in Section 2.2.4 (see also Chapter 3), interactives can be associated with more than one argument structure.

For example, we will see in Section 3.1 that the attention signal *hey* perhaps more commonly takes (S,H,T) as an alternative argument frame (as in *Hey, elephants can't kiss!*).

Nevertheless, it would seem that ignoring the argument roles in Table 2.6 would result in deficient understanding of the meanings of the interactives concerned. The arguments are inherent, that is, they are obligatory parts of interactives, as is suggested by the fact that without the arguments being implied, the meaning of an interactive would not be entirely well-formed. For example, use of the response signal *No!* or the response elicitor *Right?* of Table 2.6 would seem to be incomplete without knowing what the meaning content (T) is to which *No!* and *Right?* respond.

Since interactives are anchored directly to the situation of discourse, their arguments are recoverable from that situation and, hence, are as a rule covert, that is, they are implied rather than formally expressed. For example, if a speaker says '*Ouch!*' or '*Shh!*' then these utterances are only meaningful if the hearer assumes that the meaning expressed by these utterances relates typically to specific human argument roles. These roles are provided by the speaker (S) in the case of *ouch* but by both the speaker and the hearer (H) in the case of *shh*.<sup>31</sup>

But arguments of interactives can also be instantiated explicitly by means of additional text material. In this way, an H argument can be formally expressed, e.g., by adding a vocative interactive. Thus, both the directive *shh* in (6a) and the social formula *good morning* in (6b) have an implicit H argument, but that argument role can in addition be signaled by a textually instantiated vocative form whose use is optional, as *children* and *Nancy*, respectively, are in (6). Being optional, these forms have the status of adjuncts, that is, their presence is neither licenced nor do they necessarily contribute substantially to the content of the utterance;<sup>32</sup> we will return to this issue in Section 2.2.5.

- (6) English
  - a **Children, shh!**
  - b **Good morning, Nancy!**

To conclude, on the convention proposed here, valence is an intrinsic property of interactives—that is, without taking its argument structure into account, the use of a given interactive would be incomplete. On this account, interactives can be

<sup>31</sup> The form *shh*, or *shush*, is commonly classified as an 'interjection.' For example, Wierzbicka (2003b: 484) treats it as a 'volitive interjection.' On the account proposed here, it is a directive rather than an interjection (see Section 3.2.3).

<sup>32</sup> To be sure, the vocative forms *children* and *Nancy* may be of help in identifying the referents concerned, but this applies only to cases where alternative referents are at stake, and in such cases reference tracking could also be achieved by gestures.

classified with reference to whether they have one argument (S or T), two arguments (S,H) or three arguments (S,H,T). These arguments, or ‘participant roles,’ can be described in the following way:

S stands for the sender of a message, it may be a speaker or a writer and, accordingly, the message can be a spoken or a written text piece. Since interactives are the product of speaker activity, interactives, with the notable exception of ideophones, have an inherent S. This argument is covert, that is, it is never formally expressed.

Note that S stands for a general discourse argument, and this argument can further be subcategorized in terms of semantic roles (cf. [Andrason 2021c](#)). For example, whereas interjections (e.g., *Ouch!*) are suggestive of the role of S as an experiencer, directives (e.g., *Shh!*) appear to instantiate the role of S as an agent. A subcategorization of discourse arguments cannot be achieved in the present book, it is in need of a study of its own.

Interjections are typically restricted to S as their only argument and, hence, can be classified as expressing subjectivity, and in this respect contrast with types of interactives like directives or vocatives, which require an H argument and can be said to have an intersubjective function.<sup>33</sup>

H stands for the hearer, or the reader, or for hearers or readers. It is an inherent part of all interactives other than ideophones and interjections. In accordance with its function, use of H presupposes the presence of a hearer or hearers—in other words, its use would not normally be expected in soliloquy. Like S, H is covert, not formally expressed, but it can be repeated for example in the form of a vocative for specific discourse purposes like emphasis or participant identification for singling out the hearer, as we saw in the examples of (6).

H stands as a rule for a human participant, but it may and quite commonly does include animals treated in some way as humans, e.g., when a speaker of the Mayali language of Western Arnhem Land in Australia addresses a dog or a pig with *tja* ‘git!’ ([Evans 1992](#): 227); for detailed discussion of non-human H arguments, see Section 3.11.

T stands for the theme argument, that is, for information on the state of affairs expressed by a text segment that is implied by the usage of an interactive. In other words, T describes the circumstances that are responsible for the usage of the interactive concerned. The term ‘theme,’ thus, is used in a wider sense than in the relevant literature on semantic roles.

<sup>33</sup> Subjectivity, or subjectification, codes in some way or other the speaker’s beliefs and attitudes, whereas intersubjectivity, or intersubjectification, concerns an awareness of and attention to the addressee’s attitudes and beliefs, signaling cooperation or sharing (e.g., [Traugott 2003](#)). For [Nuyts \(2005: 14\)](#), ‘an evaluation is subjective if the issuer presents it as being strictly his/her own responsibility; it is intersubjective if (s)he indicates that (s)he shares it with a wider group of people, possibly including the hearer.’



T arguments differ from S and H in two important ways. First, both the meaning and the form that T takes can be complex: T may stand for an event, a state, an object, a quality, or a more complex content, and the text piece used for T can take on a range of different morphosyntactic forms, such as that of a phrase, a clause, or even a larger piece of discourse. And second, while also being an inherent part of a number of types of interactives, it differs from S and H in that most commonly it is overtly expressed. Thus, in the following example (7a), T is instantiated by the interrogative utterance *Have you ever seen a unicorn?* denoting an event, and it is an obligatory argument: Without that text piece, use of the response signal *no* would not be meaningful. In a similar way, the response elicitor *right?* in (7b) would be meaningless without a preceding T argument like *Jeff is British* in (7b).

- (7) English  
 a Have you ever seen a unicorn? – **No**.  
 b Jeff is British, **right?**

To be sure, arguments of sentence grammar can also take on a range of different morphosyntactic forms, such as that of a phrase or a clause. But as the examples above show, the linguistic form of T arguments is different from that of sentence grammar: Interactives are syntactically independent forms (see the definition in Section 1.2) and their arguments may, for example, belong to different turns of conversational exchange. Thus, the response signal *yep* in (8, B) has the preceding turn, that is, the question in (8, A) as its T argument, each uttered by a different speaker.

- (8) American English (Biber et al. 1999: 223)  
 A: **Hi Annie**, enjoyed your days off?  
 B: **Yep, lovely thank you**.

Which kinds of T argument a given interactive may or may not accept is contingent not only on the nature of the interactive used but also on the nature of the situation of discourse involved. The latter involves variables such as the intent and the state of knowledge of the speaker, the knowledge state that the speaker presumes the hearer to have, as well as information provided by the preceding text. A rich analysis of such variables and the constraints on collocation that they impose can be found in Wiltschko (2021) with reference to the grounding of a subset of response elicitors ('confirmational') and response signals ('response markers'). T arguments are not syntactic constituents of their interactive 'head,' but their selection is subject to constraints which can be analyzed as syntactic constraints, as Wiltschko (2021) in fact does (see also Section 5.2.4).

The analysis proposed here resembles that by Wilkins (1992) on 'interjections,' where the latter term is used by him in a wider sense for all kinds of interactive types like directives (*hold on*), response signals (*yes*), response elicitors (*huh?*),

social formulae (*thank you*), and interjections proper (*ouch, yippee*). This analysis is summarized by him thus:

I would claim that included in the semantic decomposition of all interjections are one or more of the following basic deictic referencing elements: *I, you, this, that, now*, and perhaps *here* and *there*. These elements in the decomposition provide the function of referencing in interjections and would account for the classification of interjections as (complex predication) deictics.

(Wilkins 1992: 132)

The deictic referencing elements that are at stake here are in particular *I* and *you*, which correspond to the argument roles S and H, respectively. There is no referencing element in Wilkins' (1992) approach corresponding to the T argument distinguished here. But Wilkins also assumes that, like deictic verbs, 'interjections' are two-argument forms:

In this sense I believe interjections are similar to deictic verbs like 'come,' or members of the grammatical category tense, since these elements also semantically predicate and refer indexically at the same time. That is to say, a past tense, for instance, can be analysed semantically as a two-argument predicate which takes an unspecified event as one argument and predicates that this event happened before the present speech moment, which is the indexed second argument of the predication.

(Wilkins 1992: 132)

However, what distinguishes interjections and other interactives from deictic verbs like 'come,' or past tense markers is that, unlike the latter, the former are expressive depictions, that is, they are immediately anchored to and shaped by the situation of discourse (see Section 2.1).

In concluding, two points need to be emphasized with reference to the valency analysis proposed here. First, the analysis takes care only of descriptive meaning, that is, the conceptual component of interactives, not of their status as expressive depictions. And second, the argument configurations of interactives proposed in Table 2.6 rest on prototypical uses of the forms concerned. Interactives are highly context-dependent grammatical forms, and this means that they are frequently associated with more than one argument structure, as will be shown in Section 2.2.4, and in more detail in Chapter 3. Furthermore, we will see in Section 2.2.5 that the structure associated with T arguments may even be more complex than has been sketched in this section.

### 2.2.3 More on T arguments

T arguments provide a paradigm tool for integrating interactives in the discourse. They are therefore distinctly more complex than the standard arguments S and H and, hence, require more attention. Three issues in particular need to be considered for a better understanding of this argument role. Therefore, they are the subject of the following sections.

#### 2.2.3.1 Implicit T arguments

Arguments of sentence grammar are commonly defined relative to their morphological, syntactic and semantic function within a sentence. This is strikingly different in the case of interactives: The latter are directly anchored to the situation of discourse and hence are indexicals. Accordingly, the main participants of discourse, which are the speaker and the hearer, need no participant marking: They are uniquely identifiable as the speaker (S) and the hearer (H) without being formally encoded. Obviously, this is different with the third argument, namely T. But many a time even T arguments are not formally encoded because the information they convey is also recoverable from the situation of discourse (the context).

What this means can be illustrated with the following example, which concerns social formulae expressing apologies, such as *excuse me*, *pardon*, *pardon me*, or *sorry*. The reason for apologizing is commonly encoded by means of a T argument, like *you are not getting your lip pierced* in (9)—a mock-apology which in this example prefaces a refusal. Thus, the interactive has the three-place argument structure presented in (10b), as is suggested by the paraphrase in (10a).

- (9) American English (Biber et al. 1999: 1093)

A: I'm getting my lip pierced.

B: **Excuse me**, you are not getting your lip pierced.

- (10) Argument structure

a I (S) apologize to you (H) for saying T.

b *excuse me* (S,H,T)

But quite commonly the T argument is not realized overtly, being recoverable from the context rather than being formally encoded. Thus in (11), T is justified by a social transgression by the speaker but is not overtly expressed. The argument structure of *pardon me* is the same as that of *excuse me* in (10), cf. (12).

- (11) American English (Biber 1999: 1093)

[Jeff is belching during lunch and says:]

**Pardon me!**

- (12) Argument structure  
 a I (S) apologize to you (H) for doing T.  
 b *pardon me* (S,H,T)

Social formulae for apologizing, thanking, or congratulating involve adjacency pairs (Schegloff and Sacks 1973; Schegloff 2007) which we propose to call *T-oriented*, they differ from most other social formulae, which take no T argument (Section 3.9.3).

### 2.2.3.2 Ideophones vs. interjections

In a number of studies, ideophones are not clearly distinguished from interjections (see Dingemanse 2021a for a discussion). But as we will see in Section 3.5.4, there is a catalog of features distinguishing the two. Perhaps the most conspicuous feature separating the two from one another can be illustrated with the following example. Suppose there is a gunshot upon which a person nearby produces the sequence of utterances in (13) as a reaction to the gunshot.

- (13) English  
**Bang!, Ouch!**

The utterance is composed of the ideophone *bang* followed by the interjection *ouch*, both of which depict an expressive response to the gunshot. The responses differ most of all in the following: Ideophones typically refer to some state, event, object, or quality external to the speaker (Section 3.5.2), that is, to what is described here as a T argument standing for a theme, which in this case is the gunshot. This is quite different in the case of interjections: They refer to the emotive or cognitive state of the speaker, that is, to the pain experienced by the speaker as a result of the gunshot in this example, which is an S (speaker) argument, and S is typically the only argument of interjections (Section 3.6). This situation is captured succinctly by Dingemanse (2021a) when he notes that '[i]deophones are typically depictions *of* events, while interjections are typically responses *to* them' (italics in the original).

Accordingly, paradigm utterances introducing the two in the form of direct speech, like the ones in (14), take as their subject the 'event' (= *the car* in (14a)) in the case of ideophones but the speaker (= *the reader* in (14b)) in the case of interjections.

- (14) English (Dingemanse 2021a)  
 a The car went **vroom**.  
 b The reader went **ow!**

To be sure, usage of the interjection *ouch* in (13) also presupposes something that caused the emotional reaction expressed by it, and usage of *bang* also presupposes a speaker. But what appears to be central to distinguishing the descriptive meanings of the two is that the interjection focuses on the state

of the speaker whereas the ideophone focus on some specific theme, which in this case is a gunshot (see Meinard 2015: 158 on onomatopoeia). Note, however, that in some cases the T argument of ideophones refers to a human being, and in such cases the distinction between ideophones and interjections can be blurred.

We will return to these examples in Section 3.5.7 when dealing with the indexical status of ideophones, and in Section 3.5.3.2, where the encoding and placement of T arguments is discussed.

### 2.2.3.3 Discourse markers

We saw in Section 2.2.1 that two of the three arguments distinguished, namely S and H, are fairly straightforward, being uniquely identifiable as referring to the speaker and the hearer, respectively, and being implied rather than formally expressed. The structure of the T argument is more complex and, as the analysis of discourse markers suggests, it is even more complex than described in Section 2.2.1. According to the classic definition by Schiffrin (1987: 31), discourse markers are sequentially dependent expressions which bracket units of talk. A more elaborate definition of this kind is the following:

[Discourse markers] impose a relationship between some aspects of the discourse segment they are a part of, call it S2, and some aspect of a prior discourse segment, call it S1. In other words they function like a two-place relation, one argument lying in the segment they introduce, the other lying in the prior discourse.

(Fraser 1999: 938)

In accordance with this definition (for alternative definitions, see Section 3.3.2), paradigm usages of discourse markers are said to connect two segments of discourse, namely the arguments S1 and S2 in the terminology of Fraser (1999). In the convention proposed in this chapter, the two are instantiations of the T argument and S1 and S2 are referred to here as T<sub>1</sub> and T<sub>2</sub>, respectively. Thus, in the following example, the DM *well* takes two T arguments, namely a preceding T<sub>1</sub>, instantiated in (15) by *You are always hungry*, and a following T<sub>2</sub>, namely *I'm not now*. The function of the discourse marker *well* in this example is to establish a relation of coherence between T<sub>1</sub> and T<sub>2</sub>, it can therefore be sketched as in (16a) (underlined), and its argument structure is described as in (16b).

(15) British English (Biber et al. 1999: 1087)

A: You are always hungry.

B: **Well**, I'm not now.

## (16) Argument structure

- a Paraphrase: I (S) need to correct you (H) because your claim (T<sub>1</sub>) is not in accordance with my present situation (T<sub>2</sub>).
- b Argument structure: DM (T<sub>1</sub>,S,H,T<sub>2</sub>)

With the phrasing ‘paradigm usages of discourse markers,’ we wish to draw attention to more recent findings according to which definitions such as those by [Schiffrin \(1987: 31\)](#) and [Fraser \(1999: 938\)](#) take care of canonical cases but not of all usages of discourse markers. In addition, Heine et al. (2021: 253–6) propose to distinguish what they call *constituent anchored discourse markers*. The latter need not have semantic-pragmatic scope over two segments of discourse (that is, T<sub>1</sub> and T<sub>2</sub>) but only over some specific constituent of an utterance. In such usages of discourse markers, the T argument is restricted to that constituent. The following is an example of a constituent anchored usage of the discourse marker *well*: Unlike the usage of this discourse marker in examples like (15) it has semantic-pragmatic scope only over the noun phrase *a sex drive*, and this text piece forms its T argument. In such usages, thus, there is no distinction between T<sub>1</sub> and T<sub>2</sub>; rather, there is only an ordinary T argument like in other types of interactives.

(17) English (*Larry King Live*, March 14, 2004; [Furkó 2014: 294](#))

You’re not going to have quality if you can’t sleep and you itch and you bitch and you weep and you cry and you bloat and you can’t remember anything and you don’t have a, **well**, sex drive.

We will return to the grammar of discourse markers in Section 3.3.3.

## 2.2.4 More than one argument structure

It is well-known that there are verbs which are associated with more than one argument structure—in English more so than in many other languages. Thus, the verb form *wake up* can be used intransitively or transitively, taking an object in (18b) but not in (18a), acting like a one-place verb in the former case but as a two-place verb in the latter.

## (18) English

- a Liz woke up.
- b Roger woke Liz up.

Much the same applies to interactives: Depending on the context in which they are used, they may be associated with more than one argument structure. One case concerns vocative interactives (Section 3.10), as illustrated in (19).

(19) American English (Biber et al. 1999: 1108, 1111)

a **Mom!**

b **Honey**, can I use that ashtray **please**?

In (19a), *Mom!* forms an independent utterance—a self-contained message which can be glossed somehow as in (20a). As this paraphrase suggests, there are two arguments, namely an inherent speaker (S) and a hearer argument (H), both formally unexpressed. This also applies to the vocative *honey* in (19b), but there is also a difference, as the paraphrase in (20b) suggests: The vocative in (19b) also has both S and H, but in addition takes a third argument, T, without which the use of the vocative form would somehow be incomplete in this context. Thus, vocatives are associated with two different argument structures, namely the ones in (21), where (21a) can be said to relate to its call function and (21b) to its address function (Zwicky 1974: 787; see Section 3.10.3), behaving like an attention signal (Section 3.1).

(20) Paraphrases

a I (S) am calling you (*mom*, H).

b I (S) am asking you (*honey*, H) if I can use that ashtray please (T)?

(21) Argument structure

a Vocative (S,H)

b Vocative (S,H,T)

Another case concerns directive forms (Section 3.2), which like vocatives have as a norm a two-argument structure (S,H), as *Listen!* does in (22a).<sup>34</sup> But quite commonly, the directive is followed by another text piece which assumes the role of a T argument, as *keep this quiet* does in (22b). In the latter case, *listen* assumes the function of an attention signal (Section 3.1). As the approximate paraphrases of (22) in (23) suggest, directives are also characterized by the presence of two argument structures, as depicted in (24).

(22) English (Biber et al. 1999: 210)

a **Listen, please!**

b **Listen**, keep this quiet, **won't you?** (FICT)

(23) Paraphrases

a I (S) am asking you (H) to listen please!

b I (S) am asking you (H) to listen that you should keep this quiet (T),  
won't you?

<sup>34</sup> The directive *Listen!* is followed by the social formula *please*, which is an optional participant, that is, an adjunct.

## (24) Argument structure

a **listen**<sub>1</sub> (S,H)b **listen**<sub>2</sub> (S,H,T)

It was observed at the beginning of this section that whether or not a given interactive is associated with more than one argument structure is contingent on the context in which it is used. Since interactives are anchored to the situation of discourse, context is in fact an important determining factor in shaping their valence. Frequently, however, context turns out to be insufficient for reconstructing the communicative intents of interlocutors.

The problem of reconstruction can be illustrated with the following example discussed by Ljung (2011: 91). In the conversational exchange of (25), speaker A uses the complex interjection *Oh damn it*. The interjection has the appearance of the one-argument structure in (26b), where speaker A signals her emotive state, and one way of paraphrasing its conceptual content is sketched in (26a). The reaction by speaker B is of interest in two ways. For one thing, it suggests that the interjection has a referential function, being the subject of the response elicitor *what?*. For another, it also suggests that B assumes that there is a reason for A to use the interjection and that she is entitled to an answer, and the answer is in fact subsequently volunteered by A in the third turn of (25).

## (25) English (KBA 46; Ljung 2011: 91, (25))

A: **Oh damn it.**B: **What?**

A: This one doesn't seem to want to come out.

## (26) Argument structure

a Paraphrase: I (S) am angry

b Argument structure: *damn it* (S)

The question that this example raises is the following: Did speaker A use the interjection in (25) as a complete contribution to the discourse without any intention to reveal the reason for its use—in other words, did s/he have the argument structure in (26b) in mind, perhaps in a kind of soliloquy? Alternatively, when using the interjection, did speaker A have an implicit argument T in mind, of the kind presented in the third turn in (25), possibly planning to disclose that argument on request? In other words, would (27) be more appropriate to capture the communicative intent of speaker A?<sup>35</sup> On the basis of the contextual evidence available there is no definite answer available.

<sup>35</sup> The problem is actually even more complex. Franck Floricic (p.c. of July 22, 2021) draws our attention to the fact, for example, that the usage of any interjection can be taken to have some cause.



## (27) Argument structure

- a Paraphrase: I (S) am angry for a specific reason (T)
- b Argument structure: *damn it* (S,T)

Interjections may also be used to draw attention to a related problem, also concerning contextual variation in the usage of interactives. It was observed in Table 2.6 of Section 2.2.2 that interjections are prototypical one-place interactives, taking only an S argument. But there may also be a context where—for whatever reasons—the speaker adopts the perspective of the hearer. In doing so, the speaker might be inclined to react to a situation by saying ‘*yuck!*’ but actually says ‘*wow!*’ since she believes this to be the reaction of the hearer to the situation (see Meinard 2015: 158). Should one conclude that in such contexts *wow* is a two-place interjection (S,H)?

There is no appropriate information in the data available on how to answer these questions, and in fact, this applies to a number of the text data that we were able to survey in the course of our research. What this means is that in quite a few cases discussed in this book, there is no conclusive answer to whether the use of an interactive in a given context involves only one or more than one argument structure.<sup>36</sup>

## 2.2.5 Adjuncts

There is also another problem in addition. The analysis proffered in Section 2.2.4 raises the following question: Is the T argument postulated in (27) of the preceding section really an argument, or could it not be interpreted also as an adjunct? In other words, is it possibly an optional participant whose use is neither licenced by the valence of the interactive concerned and/or which does not contribute substantially to the content of the utterance—being omissible without affecting the well-formedness of the utterance?

Without more information allowing to reconstruct the communicative intents of speaker A in (27) there is no conclusive answer to these questions. This problem though does not exist in many other text data that we were able to consult. For example, interjections express an emotive or cognitive reaction by the speaker, and S is therefore typically their only argument (Section 2.2.1). However, there are also contexts where interjections take a T argument, as in (28), where the utterance made by A is a prerequisite for B’s use of *wow*.

<sup>36</sup> Gunther Kaltenböck (p.c. of September 26, 2021) suggests that a way of resolving this issue is ‘by assuming an implicit T argument which may or may not be instantiated/realised.’

(28) English

A: Jane has been appointed to the board of directors.

B: **Wow**, that's great news!

B': **Wow!**

The text piece *that's great news* following the interjection, by contrast, is argued to be an adjunct: It could be omitted without affecting the status of the interjection—that is, its presence is not licenced by the use of the interjection and it does not contribute substantially to the content of the utterance. Adjuncts quite commonly present additional information, such as an elaboration on or a justification for the use of a preceding interjection, but they are not required. Accordingly, in (28) there could also have been B' instead of B.

It is most of all response signals like *yes* and *no* and their variants that are prone to take adjuncts, contributing little or no new information and being omisable. Typically, adjuncts present information that is largely recoverable from the situation of discourse. However, this need not be the case. For example, by adding the adjunct *she's going by train* to her response in (29), speaker B provides information that must be entirely new to speaker A. That *she's going by train* nevertheless appears to be an adjunct is suggested by the fact that it provides additional information that was not solicited by speaker A.

(29) English (Biber et al. 1999: 460)

A: Is she using her car? (Constructed)

B: **No**, she's going by train. (CONV)

Adjuncts are commonly provided by text pieces of sentence grammar, but this need not be so: They may as well consist of interactives. Thus, in the following example, the social formula *sorry* takes the vocative form *Liz* as an adjunct. This text piece is not required since in the context in which the relevant exchange takes place the referent for which *Liz* stands is already uniquely identifiable: The form *sorry* has an implicit H argument. Accordingly, omitting the vocative form *Liz* would not intrinsically affect the use of *sorry*.

(30) English

Liz: **Paul**, you missed our farewell party last night.

Paul: **Sorry, Liz!**

To conclude, an argument is an inherent part of the function of an interactive, *irrespective of whether it is formally coded or only implied*. An adjunct, by contrast, can be omitted without substantially affecting the function of the interactive concerned.

### 2.2.6 How to identify arguments

In a similar way as in sentence grammar, there are cases where there may be problems with the reconstruction of argument roles. For example, a given linguistic form may take an obligatory participant, that is an argument, in some contexts but not in others. So the question is, how do we establish whether some participant is obligatory or optional? According to the procedure adopted here, an argument entails some specific referent, or class of referents, and without such a referent being implied or overtly expressed, usage of the interactive concerned would not be felicitous. We may illustrate the relevance of this procedure with examples of the interjection *ouch* and the response signal *yes*, cf. (31).

(31) Paraphrases of *yes* and *ouch*

a **Yes.** I (S) agree with what you (H) say (T).

b **Ouch!** I (S) have just experienced a sudden and sharp pain.

Both interactives in (31) entail S, a speaker, even if S is never formally expressed. And the utterance *Yes* in (31a) would presumably not be felicitous without implying both a hearer (H) to whom the utterance is directed and some text piece (T) providing the meaning content to which the response signal *yes* refers. Accordingly, a hearer exposed to (31a) might ask: What does ‘Yes’ refer to?

One may wonder whether there is not also a T argument entailed by *Ouch!* in (31b) in that on being exposed to (31b) a hearer might ask: What is causing you to say ‘*Ouch!*’? Our justification for assuming that there is a T argument in (31a) but not in (31b) is that, unlike (31a), (31b) is likely to be taken by the hearer as a felicitous utterance, accepted as a self-contained contribution to the discourse. Furthermore, since there is no H argument, (31b) is also felicitous if there is no hearer present, e.g., if produced in soliloquy, which is less likely to be the case in (31a).

Another issue relating to the reconstruction of arguments can be illustrated with reference to the H argument in (31a). This utterance typically entails a situation where the speaker responds to a previous discourse contribution by the hearer. But there may also be contexts where there is no hearer involved, for example, when the speaker reacts to a TV show she is watching by saying *Yes*. We will nevertheless assume that response forms like *yes* take a T argument since ‘prototypical’ uses of such forms presuppose a situation where the hearer responds to a discourse contribution T just made by the speaker.<sup>37</sup>

<sup>37</sup> Franck Floricic (p.c. of July 22, 2021) rightly observes that a response signal like *yes* may also be used cataphorically, referring to a following rather than a preceding T.

## 2.3 Discussion

In addition to the problems mentioned there is one of a more fundamental nature, namely the question of whether there is sufficient empirical support for relating two quite different phenomena to one another, namely sentence grammar and interactive grammar: Is there in fact sufficient evidence for extending the notion of syntactic relations within a clause or sentence between a verb and its arguments to relations on the level of discourse processing, as we observe them in the structure of interactives? While this issue is in need of further research, it would seem that this question can be answered in the affirmative for the following reasons.

First, in both cases there is a relation of dependency involved. An English verb like *buy* instantiating a predicate typically implies two arguments, namely an agent A ('someone who buys') and a patient or undergoer P ('something that is bought'), and without any knowledge of these arguments, the use of the verb would be incomplete in many contexts. Similarly, we saw that an interactive like *shh* also requires knowledge of the two arguments of the interactive, namely the speaker (S) and the hearer (H).

Second, both verbs and interactives differ in the number of arguments each of them take: Whereas the verb *buy* takes two arguments (*Jeff bought a chicken*), a verb like *die* takes only one (*The patient has died*). This is exactly the situation that we observed with interactives: An interactive like *shh* takes two arguments while *ouch* takes only one.

Third, verbs normally take at least one argument, and this also applies to interactives: All interactives identified so far take at least one argument being implied or formally expressed.

Fourth, we saw that the arguments of interactives are quite commonly implied rather than formally expressed. Thus, an utterance like *shh* involves the two arguments S and H, but both remain as a rule unexpressed. A similar situation can be observed in sentence grammar: While the arguments A and P of the verb *buy* are commonly expressed, this need not be so, especially when the verb is used in its imperative form, as in (32), where neither A nor P are formally encoded.

- (32) English  
Buy now!

Fifth, differences in the argument structure of one and the same form may account for differences in the meanings expressed by the form concerned. For example, the English ambitransitive verb *smell* occurs both in a transitive form (33a) and an intransitive form (33b), that is, the former has two arguments (agent and patient), whereas the latter has only one (patient). This difference in argument structure is reflected in the meanings of the verb, which, while similar, are not exactly the same in (33): While *smell* relates to the *jasmin* in (33a), in (33b) it relates to *Tom*, and it was presumably not the same kind of smell in the two cases.

- (33) English
- a Tom was beginning to smell the jasmin.
  - b Tom was beginning to smell.

In a similar way, some interactives can be used in a ‘transitive’ (S,H) and an ‘intransitive’ format (S), referring to the hearer in the former and to the speaker in the latter case, and the social meaning conveyed would also not be exactly the same. For example, the interactive *bugger* would presumably be quite acceptable in its ‘transitive’ form (S,H), referring to the hearer, but also in its ‘intransitive’ form (S), referring to the speaker.

Sixth, while the nature of the arguments distinguished is not the same, differences to be observed between arguments of interactives and of verbs can be related to the different worlds in which the two grammars operate. Since the grammar of interactives is immediately anchored to the here-and-now of the situation of discourse, and hence, is inherently indexical in nature (Section 1.2), the predominant argument roles are also deictic ones, namely S (speaker) and H (hearer). In sentence grammar, on the other hand, deixis plays only a limited role, and the most common argument roles relate to non-deictic event structure and concepts like agent (A), patient (P), and recipient (R), or syntactic functions like subject and object.

And finally, the difference just mentioned also accounts for the fact that implicit rather than explicit argument coding is distinctly more important in interactives than in sentence grammar: Due to their anchoring to the situation of discourse, S and H arguments of interactives can be recovered from the context, hence they can be and, in fact, are obligatorily suppressed. This is quite different in sentence grammar with its distinctly less pronounced context dependence, where the nature of arguments cannot easily be inferred from the context and, hence, omission of arguments is clearly less common.

To conclude, there are reasons to make use of the notion ‘argument structure’ or, more precisely, the structure of discourse arguments for a better understanding of the behavior of interactives, and to account for differences between them in a principled way.

## 2.4 Contrasting ways of discourse coding

There remains another question, namely whether the kind of similarity between the two grammars that are available to interlocutors extends beyond the structure of argument presentation. One of the concerns of the present study is to determine how information is coded in discourse and, more specifically, how interactives contribute to coding in discourse. We saw in Section 2.2 that the two grammars dispose of a similar mode for coding dependencies as they obtain between a head,

which in the case of sentence grammar is a verbal predicate but an interactive form in the case of interactive grammar, and its dependent elements, that is, its arguments. But beyond this observation, similarities between the two grammars are few. Differences concern in particular the way information is presented in discourse, as can be illustrated with the following case, which was alluded to already in Section 2.1.1.

Take the examples in (34). Each of the five utterances contains a combination of a clause of sentence grammar and an interactive, namely a discourse marker (*I think*), a response elicitor (*okay?*), an interjection (*aha*), a response signal (*no*), and an attention signal (*hey*). The interactives are presented at either the right or the left periphery of the clause to which they relate and in fact, they somehow have the appearance of peripheral appendages of the relevant clause. Compared to the conceptual information provided by that clause they might well be interpreted as coding more marginal information, and omitting them would not result in deficient utterances.

(34) English

- a Paul keeps earthworms as pets **I think**.
- b I'll leave her a message, **okay?** (Biber et al. 1999: 1089)
- c **Aha!** You're here. (Wharton 2016: 28)
- d **No**, I don't keep crocodiles in my study.
- e **Hey**, can you pass me a bulldozer?

But from a perspective of discourse processing which aims at reconstructing the communicative intents of interlocutors in structuring discourse there is also an alternative view. While we are not able to reconstruct exactly what the intents of speakers may be when designing the utterances like those in (34), an alternative interpretation arises when drawing on paraphrase as a tool for elucidating the descriptive meaning content of interactives, as described in Section 2.1.1. Thus, on the paraphrase approach one may say that speakers uttering (34) are aiming at conveying conceptual information of the kind paraphrased roughly in (35).

(35) Possible paraphrases of (34)

- a I think that Paul keeps earthworms as pets.
- b Is it okay if I leave her a message?
- c I am surprised that you are here.
- d It is not true that I keep crocodiles in my study.
- e I am calling your attention to asking you if you can pass me a bulldozer?

In the paraphrases of (35), the information expressed by the interactives in (34) constitutes the semantic core of the main clause while the conceptual information contained in the clauses of (34) appears in the form of subordinate clauses in (35), somehow suggesting that they present less central information. The two analyses just sketched are hard to reconcile with one another:

What appears to be the main information in (34) is coded in (35) as subordinate and, conversely, what appears in (34) to be peripheral or more marginal surfaces in (35) as the central information of the utterances concerned (see also Section 6.4). Thus, a comparison of (34) and (35) suggests the hypothesis in (36).

- (36) In combinations of interactives with pieces of sentence grammar, the former tend to present foregrounded information while the latter present backgrounded information.

With the term ‘foregrounding’ we are alluding to the function of interactives in discourse processing and speaker–hearer interaction, but there may also be phonological correlates, such as ‘unusual’ phonetic and prosodic features (see, e.g., Section 1.3.2.2).

The hypothesis in (36) can be interpreted as lending support to the view that interactives represent higher-level explicatures, being suggestive of ‘representations in which the basic level explicature (the proposition expressed) is embedded under a speech-act or propositional-attitude description’ (Wharton 2016: 28). And for Wiltschko (2021: 3), ‘interactional language,’ which in her framework consists of response elicitors and response signals, is hierarchically structured in that ‘interactional structure dominates propositional structure.’ Such views can be taken to imply that paraphrases such as the ones in (35) in fact capture features that might be of help for describing the relationship between interactives and sentence grammar. But more research is needed, especially on whether syntactic relations as they surface in paraphrases are in fact relevant for reconstructing communicative intents of interlocutors.

## 2.5 Conclusions

The main goal of Chapter 2 was to propose a theoretical frame for describing interactives. To this end, a distinction between two meaning components was introduced in Section 2.1, namely a descriptive or conceptual and an expressive component. It was argued that the former component is amenable to description in terms of conventional notions of linguistic analysis, and it will be in the focus of ensuing discussions.

The expressive component, by contrast, turned out to be extremely complex—for example, how could the difference in expressive content between the English interjections *ouch* and *wow* be described exhaustively in an appropriate format—if there is one in the first place? To be sure, there are a number of attractive theoretical frameworks that provide relevant pieces of insights, as we saw in Section 2.1. But neither work on natural semantic metalanguage, nor on the relevance-theoretic notion of procedural meaning, nor an analysis with reference to use conditional

rather than truth conditional logic have been able to provide a comprehensive framework to this end.

Equally important for understanding the nature of interactives was the distinction between description and depiction made in Section 2.1.3. In this respect they behave like gestural depictions, which rely on body movements and other kinds of physical activity rather than on linguistic material. Interactive and gestural depictions—unlike descriptions—have in common that they are accessed immediately from the situation of discourse without being affected by any intermediate mechanism. As was argued in Section 2.1.3.3, depiction in this sense stands for an *immediate link between an expression and the item expressed*, and it therefore tends to be conceived of as a ‘vivid representation,’ an ‘enactment,’ a ‘staged scene,’ or a ‘dramatic performance.’

Section 2.2 formed the main part of the chapter, where interactive grammar was shown to exhibit a discourse valence structure akin to that of sentence grammar. Differences between the two structures were argued to be due to the latter being anchored to the structure of sentences whereas the former is immediately anchored to and shaped by the situation of discourse, not being processed by the machinery of sentence grammar. In accordance with the direct-access hypothesis (Section 2.1.3.3), the two main arguments of interactive grammar have a deictic foundation, highlighting the central role played by the speaker (S) and the hearer (H) as the main protagonists of discourse processing. In addition to these two, a third argument was distinguished, called the theme (T), which relates interactive expressions to the surrounding text.

To conclude, interactives can be shown to dispose of a valence structure that determines their behavior in discourse. But this is conceivably not the only feature that allows classifying them conceptually. As will be shown in Section 3.5.3.1, ideophones seem to be sensitive to aspectual distinctions and this might also apply to some other types of interactives.



# 3

## Types of interactives

In his reference grammar of Japanese, [Martin \(1975: 1041\)](#) discusses a group of ‘words’ which are used to call attention, to respond to a question, a command, or a social interaction, to hold the floor when fluency fails, or to express the speaker’s emotive reaction. He calls this group ‘interjections’ or interjectional phrases. This convention, on what interjections are, is in fact in accordance with a widespread understanding of the term in linguistics and language typology.

This is not the convention proposed here: To aim at getting attention is the function of *attention signals*, to respond to a question is that of *response signals*, commands are the domain of *directives*, social interaction is most of all the domain of *social formulae*, managing discourse is mainly that of *discourse markers*, while the term *interjection* is reserved here for linguistic forms expressing emotive and cognitive reactions by the speaker. Nevertheless, Martin’s understanding of ‘interjections’ overall reflects what is here referred to as interactives as defined in Section 1.2, and much the same applies to a number of other authors, as we will see in Chapter 5.

The goal of the preceding chapters was to understand what interactives have in common. In the present chapter, our interest is in what distinguishes the various types of interactives beyond what they share. The world of interactives can be classified in various ways. As is argued in this chapter, the space of discourse functions covered by interactive grammar can be cut up into ten salient types, namely the ones listed in [Table 3.1](#). This table also summarizes what appear to be the main functions expressed by these types. Note, however, that the functions listed here constitute only one component of the definition of the interactives concerned, the second component was contributed by the general definition of interactives proffered in Section 1.2.

In the remainder of this chapter, the various types of interactives are discussed in turn. Like the general category of interactives (see Section 1.2), the types distinguished are conceived as prototypical entities (see, e.g., [Andrason 2020; 2021a; 2021b; Andrason and Dlali 2020; Andrason, Fehn, and Phiri 2020](#); see also [Childs 1994: 181](#) and [Haiman 2018: 104](#) on ideophones). This means, first, that not all the properties of interactives listed in Section 1.2 may be present in a given instance of an interactive form. Second, this also means that each of the types distinguished has prototypical, less prototypical, and peripheral members. Third, interactives are highly context sensitive and depending on the context in which they are used they can fulfill quite different functions each associated with different types of

**Table 3.1** Main functions of the ten types of interactives

| Type               | Main function                                                                                                                           |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Attention signals  | To call the hearer's attention (to something)                                                                                           |
| Directives         | To express a request or command                                                                                                         |
| Discourse markers  | To express metatextual functions relating to the organization of texts, the attitudes of the speaker, and/or speaker–hearer interaction |
| Evaluatives        | To express a concise appraisal of what the previous speaker has just said                                                               |
| Ideophones         | A vivid depiction of sensory imagery of a state, event, object, or quality                                                              |
| Interjections      | To index a sudden change in the emotive or cognitive state of the speaker                                                               |
| Response elicitors | To seek the hearer's agreement to or confirmation of a remark just made                                                                 |
| Response signals   | A concise response to a previous remark by a different speaker, typically expressing agreement or disagreement                          |
| Social formulae    | To foster social relations with the hearer                                                                                              |
| Vocatives          | To call or address the hearer, asking for their attention                                                                               |

interactives. For example, the Swahili expression *ehē* 'I understand,' classified by Eastman (1992: 277, (2)) as an 'interjection,' may in fact act as an interjection, signaling the speaker's emotional reaction to what has just been said, but it has also the function of a discourse marker, marking the end of the previous turn and inviting a new turn, and it also may have features of a social formula, encouraging the previous speaker to continue and to know that attention is being paid.

And fourth, and most importantly, boundaries between types are gradual rather than discrete. Membership within a given type can be, and has been a matter of debate—some ideophones have been classified as interjections, interjections as discourse markers, discourse markers as interjections, etc. Rather than discrete taxa, which can clearly be distinguished from other entities by means of a set of necessary and sufficient criteria, the types of interactives form an interlocking network across the boundaries of types (cf. Section 4.2).

One reason for this structural behavior can be seen in the role played by context in the usage of interactives. Context is a ubiquitous phenomenon in linguistic discourse, both in sentence grammar and interactive grammar. But in the latter it is distinctly more important: Being directly anchored to the situation of discourse, interactives are produced by circumventing the machinery of sentence grammar, that is, their usage is not affected by syntactic constraints, hence, it is highly variable and context-sensitive (cf. the direct-access hypothesis, Section 2.1.3.3). Accordingly, quite a number of interactives exhibit context-related usages cutting across boundaries of interactive types.

Both the terminology and the classification of types proposed here owe a great deal to the seminal work on inserts by Biber et al. (1999: 1069, 1082). Differences

concern mainly some types which are not distinguished by Biber et al. (1999). These differences are far from substantial, as Table 3.2 suggests; see Section 5.1.2 for more details.

**Table 3.2** Correspondences between inserts (Biber et al. 1999) and interactives

| Type of inserts (Biber et al. 1999) | Roughly corresponding type of interactives |
|-------------------------------------|--------------------------------------------|
| Attention signals                   | Attention signals                          |
| -                                   | Directives                                 |
| Discourse markers                   | Discourse markers                          |
| -                                   | Evaluatives                                |
| Expletives                          | (Secondary) interjections                  |
| Greetings and farewells             | Social formulae                            |
| Hesitators                          | Discourse markers                          |
| -                                   | Ideophones                                 |
| Interjections                       | (Primary) interjections                    |
| Polite speech-act formulae          | Social formulae                            |
| Response elicitors                  | Response elicitors                         |
| Response forms                      | Response signals                           |
| -                                   | Vocatives                                  |

In the remainder of this chapter, each of the ten types of interactives is looked at on its own. For each of the types, a working definition is proposed, designed in accordance with the framework proposed here for analyzing interactives. Thus, the definitions are not intended to replace earlier definitions to be found in the relevant literature. Discussion is restricted to highlighting specific features characterizing the types. This means that features shared by the various types are ignored since such features were already the subject of Chapters 1 and 2, and information on the diachronic behavior of the types is reserved for Chapter 4.

3.1 Attention signals

The term is taken from Biber et al. (1999: 1088–9, 1097), who write that attention signals, also called attention getters or seekers, or ‘attention grabbers,’ ‘have the main function of attracting the attention of addressees – for example, when speakers want to make clear that they are addressing a particular person.’ Elsewhere in the literature on interactives, they are not commonly distinguished as a distinct type of discourse processing or discourse grammar.

3.1.1 Introduction

Attention signals are classified in most works as ‘interjections’ but here they are classified as a type of their own, differing in their argument structure from

interjections (see Sections 3.1.2, 3.6.2), even if they may have exclamatory functions. In English, usage of attention signals is not seldom subject to social constraints, being felt to be inappropriate in specific contexts. Thus, Biber et al. (1999: 1089) point out that attention signals ‘are familiar and often impolite in their effect – especially when followed by a directive (e.g. *Hey stop bitching* (BrE))...’ (Biber et al. 1999: 1088).

Similarly, usage of the attention signal *kss...!* of the West African language Ewe ‘may be perceived as rude if a young person uses it to get the attention of an older person’, and more generally, people lower in social status should not use this form to their superiors (Ameka 1992b: 260). Such social constraints seem to be legion in languages across the world.

### 3.1.2 Working definition

Attention signals are frequently referred to as attention getters, occasionally also as ‘vocatives’ (Section 3.10). They are defined here as in (1).

- (1) An attention signal is an interactive used to call the hearer’s attention (to something).

The following English items tend to be included in this type: *come on, excuse me, hey, hoo, I say, listen, look, oh, oi, pardon me, say, pst, and yo(o)* (cf. Zwicky 1974: 787; Biber et al. 1999: 1088–9). In addition, some primary interjections and vocatives can assume the function of attention signals (Section 4.2). Furthermore, a range of other interactives can in specific contexts be ‘exapted’ for uses as attention signals.<sup>1</sup>

Like interjections, attention signals exhibit a distinction between primary and secondary forms in English (Section 3.6.4.1).

### 3.1.3 Grammar

Attention signals are associated with two argument frames (see Section 2.2): (a) They are either used as stand-alones, even if they may be followed by a vocative form, cf. (2a) or, more commonly, (b) they take a T argument immediately following them, cf. (2b). The examples in (3) from the West African language Ewe relate to the attention signal *he!* ([fɛ]) ‘hey!’, where it also occurs with both argument structures. When functioning as a stand-alone, or followed by a vocative, as in (2a) and (3a), the attention signal usually has a calling (‘summoning’) function, whereas it functions as an address when taking a T argument, encoded as a text

<sup>1</sup> Franck Floricic (p.c. of July 22, 2021) notes that the response signals *no* of Spanish and Italian can also function as attention signals, and so can the French response elicitor *n'est-ce pas?* ‘isn’t it?’.

piece, as in (2b) and (3b) (see Zwicky 1974: 787; see also Section 3.10.3, where vocatives are discussed).

- (2) English (Biber et al. 1999: 463; Hengeveld and Mackenzie 2008: 78–9)  
 a **Hey Bert!**  
 b **Hey**, did you read through this yet?
- (3) Ewe (Kwa, Niger-Congo; Ameka 1992b: 262)  
 a **He Kofi!**  
   'Hey Kofi!'  
 b **hé** **ɖe**      ne-ga-tró              gbɔ-na              **láo?**  
   *he*    Q        2.SG-REP-turn    come-HAB    Q  
   'Hey! Are you coming back or what?'

The conceptual content of the two usages of Ewe *he* can be paraphrased as in (4a) and (5a), respectively, and the corresponding argument structures as in (4b) and (5b), respectively,

- (4) a Paraphrase:                    I (S) want to call your (H) attention  
     b Argument structure: Ewe *he* (S,H)
- (5) a Paraphrase:                    I (S) want to draw your (H) attention to what I  
                                               am going to say now (T)  
     b Argument structure: Ewe *he* (S,H,T)

Attention signals differ greatly from one another in which argument structures they select. As we saw in (3), the Ewe form *he* can take either the 'intransitive' structure (S,H) in (3a), or the 'transitive' structure (S,H,T) in (3b). Other forms again may take only one structure. For example, in addition to *he*, Ewe has another attention signal form, *kss...!*, which appears to be restricted to the 'intransitive' structure: It can combine neither with a vocative nor with a T argument (Ameka 1992b: 262)—that is, it is associated only with the structure in (4). Furthermore, there are also attention signals that cannot occur as stand-alones, being restricted to the 'transitive' structure in (5). T is always overtly marked, and it may take the form of either a declarative sentence (6a), an interrogative sentence (6b–c), or a hortative sentence (6d):

- (6) English (Biber et al. 1999: 1088–9)  
 a **Hey you**, I buttered that bread for you and you didn't eat it. (BrE)  
 b **Say, Mom**, have you got any paint rollers. (AmE)  
 c **Hey Rob** do you want to lend me fifty p? (BrE)  
 d **Hey** let's read a story. (BrE)

Attention signals quite commonly give rise to discourse markers (Section 3.3). This happens when the signal is extended to contexts where it is interpreted as introducing a response to the preceding discourse, typically acting as a topic shift

marker (Section 4.2.2). The result is that the attention signal now takes two T arguments, its argument structure being  $T_1, S, H, T_2$  (cf. 2.2.3.3). Thus, in the following example, the attention signal *hey* marks discontinuity between the preceding and the following text, the latter introducing a new topic.

- (7) English (Levinson 1983: 313, (32); boldening ours, upper case and italics in the original)  
 R: It's o – it's okay we'll pop down tomorrow **Gertrude**  
 C: You sure you don't, it is an awful lot of it, you want to quickly nip down now for it.  
 R: **Okay** I will. **Er HEY** you **hmm** that is have you been lighting a fire down there?

The hearer argument (H) of attention signals is in principle covert, but quite frequently it is also expressed overtly by a vocative interactive, which is pronominal (*you*) in (6a) and nominal (*Mom*, *Rob*) in (6b) and (6c). That an overt H is optional can be concluded from the fact that the vocatives *you*, *Mom* and *Rob* in (6) could be omitted as long as the referent can be identified from the context. Vocatives frequently follow the attention signal. Thus, a fairly common collocation pattern, as it also surfaces in the examples of (6), is the following:

- (8) Attention signal—vocative—T

In the Tjwao language of western Zimbabwe, the only regular collocation of interactives is that of attention signals ('interjections') and vocatives. Thus, Andrason, Fehn, and Phiri (2020: 15) say that 'interjections form coherent units with vocative nouns,' the two often resisting phonological separation, as in (9).

- (9) Tjwao (Khoe ('Central Khoisan'); Andrason, Fehn, and Phiri 2020: 16, (20))  
**Yeyi Balisi**, tcaa ti.a maa.  
 ATT Balisi water 1:SG.O give  
 'Yeyi Balisi, give me water!'

Note that what the authors call an 'interjection' is classified here as an attention signal in usages such as *yeyi* in (9): The Tjwao form *yeyi* is used to catch the hearer's attention, or to express admiration, the latter also being a function not infrequently expressed by attention signals.

Attention signals are largely restricted to the position at the left periphery of their host utterance, which happens to be their T argument. However, they are occasionally found also in other positions. Thus, in the following text piece from the Tjwao language, the attention signal *yeyi* follows its T argument *kx'ui ti* 'don't talk,' but is followed by the vocative *Balisi*:

- (10) Tjwao (Khoe, “Central Khoisan”; [Andrason, Fehn and Phiri 2020](#): 16)  
 Kx'ui ti **yeyi Balisi.**  
 speak NEG ATT Balisi  
 ‘Don’t talk, *yeyi* Balisi!’

In his discussion of vocatives, [Zwicky \(1974: 787\)](#) distinguishes vocatives used as calls and as addresses. The former, illustrated in (11), ‘are designed to catch the addressee’s attention,’ and he adds that they are similar in function to the following: *hey, yoo, hoo, oh, I say, excuse me, and pardon me*. In such usages we will include them among the forms serving as attention signals.

- (11) English ([Zwicky 1974: 787](#))  
**Hey lady**, you dropped your piano.

The boundary between attention signals and directives (Section 3.2) is fluid. It is especially directives derived from verbs of perception, such as *listen* and *look*, that are frequently followed by another text piece, and in specific contexts that text piece can be interpreted as an obligatory participant of the directive, that is, as a T argument; see Section 3.1.3 for discussion. Thus, in (12a), *listen* is a directive, paraphrasable as in (13a), whereas in (12b) the text piece *keep this quiet* can be taken to be a T argument of the directive *listen*, and (12b) can be paraphrased roughly as in (13b).

- (12) English ([Biber et al. 1999: 210](#))  
 a **Listen, please!**  
 c **Listen**, keep this quiet, **won’t you?** (FICT)

- (13) Paraphrases of (12)  
 a I (S) am asking you (H) to listen please!  
 b I (S) am asking you (H) [to listen] to keep this quiet (T), won’t you?

In usages such as (12b), where the verbal semantics of perception may be backgrounded, the directive can in much the same way be interpreted as an attention signal, presented to draw attention to what follows, and there clearly is some overlap between the two types of interactives (cf. the overlap model of [Heine 1993: 48–53](#)).

In fact, there appears to be a cross-linguistically regular pattern of usage where a directive acquires features of an attention signal once it takes a T argument. An interesting example is the following. The form *ayya* of Tunisian Arabic appears to be an attention signal implying a following action, which can be expressed by many different types of verbs, as in (14). When used to address a second person referent, *ayya* denotes ‘come on’ or ‘hurry,’ and when used with inclusive ‘we’ it expresses hortative ‘let’s (do)’.

(14) Tunisian Arabic (Veselinova 2006:144)

- a **ayya** (i:ja) ‘come on, (come here)’
- b **ayya** (kul) ‘come on, (eat)’
- c **ayya** (?qra) ‘come on, (read)’

In neighboring Algerian Arabic, by contrast, the same form *ayya* has the function of a directive expressing ‘come,’ that is, it is used for deictic motion (see Veselinova 2006: 145).

Another example is provided by the directive *hima* ‘come on, let’s go’ of the Arawá language Jarawara of Brazil. This directive can be followed by a clause instantiating a T argument. And in this case, *hima* has the features of an attention signal drawing attention to the invitation to do the action described in that clause, e.g., *hima! ee mii ka-na-hi!* ‘Come on! let us both shit!’ (Dixon 2004: 262, 389–90).

### 3.1.4 Discussion

Being a device for attention getting is a function that is also part of the usages of interactives beyond the ones mentioned above. One of these interactives is the English discourse marker and response signal *right*, which also serves to signal an interactive relationship between speaker, hearer, and discourse. Its functions include that of an attention catcher, as in (15) (Heine et al. 2021: 126–7).

(15) English (FLOB L13: 96; Méndez-Naya 2006: 145, (11))

- “**Right**,” he said briskly, “I’m going now. I’m sorry we’ve had this unfortunate scene but –”

Attention signals are quite commonly primary forms, like English *hey*, that is, elements that are not transparently related to other forms (Section 3.6.4.1). A somewhat unusual situation is found in Korean, where an attention signal (‘discourse initiator’) like *iss-ci?* ‘Look!’ developed from a rhetorical question, used by interlocutors to initiate a discourse:

(16) Korean (Rhee 2004: 414)

- iss-ci?**  
exist-Q  
‘Look!’ (Lit. ‘Does (it) exist?’)

In Ewe, a language spoken in eastern Ghana and southern Togo, a distinction is made between visible and invisible hearers (Ameka 1992b: 257–60). On the one hand, there are the forms *ú:ru!* and *u:wúí!* which are used to call people across a distance in the bush or on the farm, and they are invisible to the speaker. On the other hand, there are the attention signals *kss...!* and *he!* ([hɛ]), the latter presumably functionally equivalent to English *hey*: It is used to call people’s attention, and these people are typically visible. All four are called ‘conative interjections’ by



Ameka, English *hey* but it seems that *kss...!* and *he!* are straightforward attention signals (see 3.1.5) whereas *ú:ru!* and *u:wúi!* appear to be more strongly related to vocatives (Section 3.10).

The boundary between attention signals and some other interactive types such as directives, vocatives, and discourse markers is fluid. For example, English directives like *look*, *look here*, or *listen* have in some contexts more in common with attention signals than with directives.

Hengeveld and Mackenzie (2008: 78–9) classify the English form *hey* as a paradigm instance of a vocative, as in *Hey!*. This is not the convention proposed here. For one thing, vocatives typically name some specific referent as addressee, most commonly by means of a proper noun, while this does not clearly apply to *hey*, where the identity of the addressee can only be recovered from the situation of discourse. And for another, the two appear to occupy different placement slots; for example, the combination *hey* + vocative forms a highly frequent collocation pattern (e.g., *Hey, Mike!*); see the examples in Section 3.1.3. Vocatives in the sense of the term as used here, by contrast, do not normally form a collocation pattern with other vocative forms (e.g., *#Mike, Mr Smith*). We therefore prefer to classify *hey* as an instance of the type of attention signals, even if the boundary between the two types is at times hard to locate.

### 3.1.5 The attention signal *\*hey*

There is conceivably a sound symbolic feature associated with attention signals, namely a general form–meaning unit found in many languages that we tentatively label *\*hey*.<sup>2</sup> The instantiations of this unit can be characterized phonetically thus: They tend to consist of (a) a voiceless aspirated fricative ([h]), (b) followed by a mid front vowel ([e, ε]) or a low vowel, and (c) followed by a high back vowel ([i]) or a corresponding semivowel ([j]). Furthermore, this form (d) is typically monosyllabic and (e) does not end in a consonant, and the only obligatory element of the unit is (b). Note that (d) and (e) are shared by other-initiated repair markers (*\*huh?*), as identified by Dingemanse, Torreira, and Enfield (2013: 7), see Section 3.7.5.

Table 3.3 presents examples of this form from a range of languages. The compilation of these examples rests on a more or less casual search, a thorough worldwide survey is urgently needed.

The following examples may show that *\*hey* forms are, on the one hand, cross-linguistically similar in form and meaning and, on the other hand, exhibit ranges of related functions. For example, the attention signal *he* of the Ewe language,

<sup>2</sup> The asterisk indicates that *\*hey* is a reference form standing for a set of expressions that are similar in form and meaning across languages. No claim to the effect is made that the similarities are due to genetic or areal relationship.

**Table 3.3** Corresponding forms of the attention signal \**hey* in different languages

| Language        | Genetic affiliation             | Form                         | Published sources                        |
|-----------------|---------------------------------|------------------------------|------------------------------------------|
| Akie            | Southern Nilotic, Nilo-Saharan  | <i>hēi</i>                   | König, Heine, and Legère (2015a: 117–18) |
| Albanian        | Indo-European                   | <i>hej</i>                   |                                          |
| Ambrym          | Oceanic, Austronesian           | <i>hɛ</i>                    | (Paton 1971: 94)                         |
| Bulgarian       | Slavic, Indo-European           | <i>hey</i>                   |                                          |
| Danish          | Germanic, Indo-European         | <i>hey</i>                   |                                          |
| Dciriku         | Bantu, Niger-Congo              | <i>hé, hēwa</i>              | (Möhlig 1967: 237)                       |
| Dhaasanac       | Cushitic, Afro-Asiatic          | <i>ʔey</i>                   | (Tosco 2001: 249)                        |
| Diyari          | Pama-Nyungan                    | <i>ayi</i>                   | (Austin 1981: 36)                        |
| Dutch           | Germanic, Indo-European         | <i>hé</i>                    |                                          |
| English         | Germanic, Indo-European         | <i>hey</i>                   |                                          |
| Ewe             | Kwa, Niger-Congo                | <i>he</i><br>([ <i>hɛ</i> ]) | (Ameke 1992b: 262)                       |
| French          | Romance, Indo-European          | <i>eh</i>                    |                                          |
| German          | Germanic, Indo-European         | <i>he</i>                    |                                          |
| Gilaki          | Iranian, Indo-European          | <i>hey</i>                   | (Zahra Mirkhaef p.c.)                    |
| Gooniyandi      | Bunuban (Australia)             | <i>yay</i>                   | (McGregor 1990: 228)                     |
| Italian         | Romance, Indo-European          | <i>eh, ehi</i>               |                                          |
| Jarawara        | Arawá                           | <i>ai</i>                    | (Dixon 2004: 388–94)<br>[1]              |
| Kabiye          | Gur, Niger-Congo                | <i>hét</i>                   | (Lévikaza 1999: 526)                     |
| Kalenjin        | Southern Nilotic, Nilo-Saharan  | <i>ɛ:</i>                    | (Creider and Creider 2001: 83)           |
| Korean          | Unaffiliated                    | <i>ei</i>                    | (Sohn 1994: 524)                         |
| Lakota          | Siouan                          | <i>heī</i>                   | (Buechel 1939: 127) [2]                  |
| Logba           | Kwa, Niger-Congo                | <i>hey</i>                   | (Dorvlo 2008: 248) [3]                   |
| Ma Manda        | Trans-New Guinea                | <i>aai</i>                   | (Pennington 2016: 198)                   |
| Muna            | Malayo-Polynesian, Austronesian | <i>ee</i>                    | (van den Berg 1989: 188) [4]             |
| Nepali          | Indo-Aryan, Indo-European       | <i>ai, e, he</i>             | (Acharya 1991: 141–2)<br>[5]             |
| Portuguese      | Romance, Indo-European          | <i>ei</i>                    |                                          |
| Romanian        | Romance, Indo-European          | <i>hai</i>                   | (Haegeman and Hill 2013) [6]             |
| Russian         | Slavic, Indo-European           | <i>ej</i>                    |                                          |
| Slovak          | Slavic, Indo-European           | <i>hej</i>                   |                                          |
| Sotho, Southern | Bantu, Niger-Congo              | <i>he, hei</i>               | Guma (1971: 246)                         |
| Thompson        | Salishan                        | <i>hey'e</i>                 | Thompson and Thompson (1992: 184)        |
| Tigemaxo        | Mande, Niger-Congo              | <i>ɛ</i>                     | (Blecke 1996: 190)                       |
| Tjwao           | Khoe ('Central Khoisan')        | <i>yeyi</i>                  | (Andrason et al. 2020: 16)               |
| Tswana          | Bantu, Niger-Congo              | <i>hêê</i>                   | Cole ([1955] 1987: 395)                  |

Continued

Table 3.3 *Continued*

| Language | Genetic affiliation       | Form          | Published sources              |
|----------|---------------------------|---------------|--------------------------------|
| Ulithian | Micronesian, Austronesian | <i>aay</i>    | Sohn and Bender (1973: 87)     |
| Wintu    | Wintuan                   | <i>he.he.</i> | Pitkin (1984: 261) [7]         |
| Xhosa    | Bantu, Niger-Congo        | <i>heyi</i>   | (Andrason and Dlali 2020: 205) |
| Zulu     | Bantu, Niger-Congo        | <i>he:</i>    | (Doke [1927] 1988: 279–83)     |

Notes: [1] The Jarawara form *ai*, translated by Dixon (2004: 388–94) with ‘hey,’ expresses surprise when one is impressed. [2] The Lakota attention signal *hēi* is used only by men (Buechel 1939: 127). [3] This attention signal of Logba is used to call the attention of a person but is considered to be impolite (Dorvlo 2008: 248). [4] The Muna interjection *ee* is an attention signal expressing disagreement or amazement, e.g.: *ee, miina nao-ghosa* (hey, not 3.SG.irrealis-strong) ‘hey, he is not strong’ (van den Berg 1989: 188). [5] Acharya (1991: 141–2) presents the following ‘vocatives,’ which he also calls attention drawers: *ai, e, he, o, oi*. Their main function seems to be to introduce a following vocative expression. [6] Seeking attention is only one of the functions of Romanian *hai*, cf. *Hai șa citim* (*hai* SUBJ read.1.PL) ‘Come on, let’s read’ (Haegeman and Hill 2013: 273, (2)). [7] The ‘exclamative’ *he.he.* of the Californian language Wintu is translated by Pitkin (1984: 261) as ‘see here, well, well!’

classified by Ameka (1992b: 257–61) as a ‘conative interjection,’ is functionally and semantically similar to English *hey*, being used to call people’s attention and it may also be used to alert an addressee who is in the same place as the speaker and with whom the speaker may have had prior interaction. But a number of the forms listed in Table 3.3 have a highly specific function. The ‘interjection’ *?ey* of the Dhaasanac language of southern Ethiopia, for example, serves to attract the attention of persons (‘hey, you!’) rather than to a piece of text, and it is used by women also to reply to a call (Tosco 2001: 249). The ‘interjection’ *ε:* of the Nandi dialect of Kalenjin is used to get a person’s attention when handing something to him or her, e.g., *ε: kim-yé:t* ‘Here is the food (come and take it)!’ (Creider and Creider 2001: 83).

Some of the attention signals listed are designed to give the hearer a warning on what is expounded in the following predication. This applies, for example, to the form *héi* of the Kabiye language of Togo (Lébibkaza 1999: 526), and the attention signal *hēi* of the Akie language of north-central Tanzania is mainly used to induce to give the hearer a warning or to take some action, as in the following example:

- (17) Akie (Southern Nilotic, Nilo-Saharan; König, Heine, and Legère 2015a: 121)  
      **hēi**, pákaai.  
      ‘Hey, leave it/stop it!’

Other forms, in contrast, have a wider range of functions. For example, the form *hey* of the Gilaki language of Iran serves a number of different purposes. First, it is used to grab the attention of the hearer, or to greet friends. Second, it is a groan to

show sadness when losing someone or something. Third, it is an animal directive (Section 3.11), used by shepherds to call their animals for eating grass or to lead them to the barn. Fourth, its prolonged form *heeeeeeeey!* is used to express one's anger and take the turn to speak, also to express 'Watch your mouth! You are so rude to me!'. And finally, there is the form *ey*, used to ask the hearer to repeat a word or a sentence (Zahra Mirkhaef, p.c.).

A comparative study of the distribution and functions of the attention signal *\*hey* should also take account of other kinds of interactives showing cross-linguistic regularities in their phonetic form. One example is discussed in Section 3.7.5, namely what Dingemanse, Torreira, and Enfield (2013) call 'other-initiated repair' markers, that is, response elicitors, which also exhibit significant similarities in form and meaning. Another example is provided by response signals. For one thing, expressions for 'yes' and 'no' include forms in quite a number of languages that consist of a click sound (Section 3.8.4). And for another, instances of forms for 'yes' show a marked tendency to contain the laryngeal phonemes /ʔ/ or /h/ in the languages of the world (Parker 2021).

### 3.2 Directives

Searle (1983: 166) writes that by means of directives 'we try to get our hearers to do things.' His is a fairly general understanding of the term, and the same also applies to the seminal study by Mauri and Sansò (2011), where the following definition is presented:

By directive situations we mean all those situations in which the speaker wishes a state of affairs (henceforth SoA) to become true and conveys an appeal to the addressee(s) to help make this SoA true.

(Mauri and Sansò 2011: 3491)

This definition as well is too general for the purposes of the present study. As used here, 'directive' applies strictly only to linguistic expressions that conform, on the one hand, to the definition of interactives in Section 1.2 and, on the other hand, to the definition in Section 3.2.2 below.

Directives are commonly treated as 'interjectives' in grammars of South African Bantu languages, called by Doke and Mofokeng ([1957] 1985: 368) 'imperative interjectives' in their grammar of Southern Sotho (see also Cole [1955] 1987: 396 on Tswana). Many, but by no means all instances of directives are in fact provided by grammaticalized imperative forms and our concern in this section is mainly with directive forms derived from imperatives. But rather than imperative forms or constructions at large, the present section is restricted to conventionalized, formulaic second person singular imperatives like English *Listen!* or *Come!*.

Henceforth, when using the term ‘imperative’ we are referring exclusively to this kind of formulaic imperative forms.

Note that directives are by no means restricted to communication among humans, they can also be used in exchanges with animals. This, however, is a special kind of usage that will be treated separately in Section 3.11.

### 3.2.1 Introduction

Like vocatives, directives have been classified within a higher category of ‘appellatives’ (Sonnenhauser and Noel Aziz Hanna 2013: 12). Both vocatives and directives in fact bear some of the hallmarks of the sentence grammar categories from which they are derived. Whereas vocatives exhibit nominal properties, directives frequently show some verbal properties. At the same time, vocatives and directives contrast sharply with the categories from which they have been coopted, as has been pointed out by a number of authors (Jakobson 1971: 10; Sonnenhauser and Noel Aziz Hanna 2013: 11): Both are said to be located outside the ‘grammatical sentence,’ ‘share the same intonation,’ and both can form unanalyzable utterances.

Imperative constructions exhibit a wide variety of structures in the languages of the world (see Aikhenvald 2010 for a convenient survey). Our interest here is severely restricted to what Andrason and Dlaki (2020: 212) call ‘significantly grammaticalized imperatives,’ that is, invariable forms like *hold on* in (1) which conform to our definition of interactives in Section 1.2. Following Heine (2016a) we are dealing with what Aikhenvald (2010: 4) calls canonical imperatives, and more narrowly with formulaic positive imperatives typically addressing a singular referent as a hearer (or reader) and expressing commands or requests directed to the hearer.

- (1) English (Heine 2016: 248, (6d))

**Hold on**, are we late?

Directives, like English *shh*, have most commonly been classified as interjections (but see Section 3.2.3). This does not apply to most imprecations cast as imperatives, like English *Fuck you*. Unless such expressions are clearly intended to express a request or command (see the definition in Section 3.2.2), they are ignored here—as aptly pointed out by Aikhenvald (2020: 54–5), imprecations express emotional states rather than directive speech acts.

### 3.2.2 Working definition

The term ‘directive’ is defined here as in (2).

- (2) A directive is an interactive used to express a request or command.

Common English directives are *believe me, bugger off, careful, come, come in, come on, fuck off, halt, hang on, hold on, listen, look, out, piss off, pst, shh, shush, stop, and wait*.

In the spirit of [Mauri and Sansò \(2011: 3492\)](#), the term ‘directive’ refers ‘to the set of forms that encode positive directive situations in a language.’ However, in accordance with the definition in (2), the present usage is restricted to a subset of the directives discussed by those authors, namely to those that qualify as interactives.

The majority of directives are conventionalized imperative forms of verbs, but there are also other kinds of directives. For example, English directives include expressions that have traditionally been treated as interjections, such as *shh*, which according to their function belong to the type of directives as defined here.

The form *out*, as in (3), represents another kind of directive, apparently coopted from the adverb or preposition, typically used to indicate ‘that the addressee should move out of the place where the speaker is’ ([Wilkins 1992: 133](#)). Thus, its conceptual content can be paraphrased as in (4a), its argument structure being that in (4b).

- (3) English  
**Out!**

- (4) a Paraphrase: I (S) ask you (H) to move out of where you are presently.  
b Argument structure: *out* (S,H)

Directives are especially common in the closing section of conversations (e.g., *Take care!, Enjoy!, Have fun!*), and their effect there is for the most part intended to be in the interest of the hearer.

Not uncommonly, directives develop discourse functions that are fairly remote from the literal, lexical meaning. One may assume, for example, that someone exposed to the expression in (5) is unlikely to interpret (5) as a directive in its literal sense, or generally as a directive in the first place (see Section 3.2.1).

- (5) English ([Ljung 2011: 19](#))  
**Go to hell!**

Directives are typically grammaticalized imperative forms, hence, they differ from the latter in having undergone decategorialization (Section 4.2), that is, they are invariable expressions that have lost the ability to take complements, to be negated, and they have been generalized to the extent that they are no longer restricted to second person singular referents as hearers (or readers), but they may as well refer to more than one hearer (cf. [Mauri and Sansò 2011](#)). But distinguishing directives from plain imperatives is in many cases hard, if not impossible, since the

boundary between the two is fluid and the transition from imperative forms used productively and imperatives used as conventionalized interactives is gradual.

The target of a directive, that is, the hearer (H), is instantiated typically by one or more human addressees. But there may also be other kinds of addressees, such as supernatural powers. For the !Xóõ people of the Kalahari Desert in Botswana, for example, an important addressee is provided by thunder, and their language contains a range of directives used to plead the thunder for rain (Traill 1994; Brenzinger and Shah 2021).

### 3.2.3 Grammar

Directives, like *shh*, can be paraphrased as in (6), taking two arguments, both covertly expressed.

- (6) Paraphrase: I (S) ask you (H) to be quiet  
Argument structure: *shh* (S,H)

We noted in Section 2.2.2 that S arguments can be subcategorized into a number of what have the appearance of ‘semantic roles’ (cf. Andrason 2021c). For example, interjections are suggestive of the role of an experiencer (Section 3.6.3), that is, a participant typically expressing the psychological or physical state of a speaker who is not in control of that state. This is different in the case of directives, which can be argued to stand for a participant aiming to initiate or instigate an action, that is, agent. More research is needed on this issue.

In the tradition of Ameka (1992a), *shh* has been called a ‘conative interjection’ (see Section 3.2.1). Stange (2016) adopts this terminology and classification but observes:

Conative interjections may have an emotive component in that their utterance may express annoyance at the noise level (*Shh!*), or the speaker’s desire to share delicate news (*Psst!*), but the predominant pragmatic element is that the speaker expects the listener to react in a certain way.

(Stange 2016: 13)

Directives like *shh* or *shush* have also been classified as ‘volitive interjections’ (e.g., Wierzbicka 2003b: 484). In the framework proposed here, however, volition is interpreted as being derivative of their directive function, the latter distinguishing *shh* or *shush* from emotive (e.g., *yuck*) and cognitive interjections (e.g., *aha*). In other words, *yuck* and *aha* are interjections (Section 3.6) while *shh* or *shush* are classified here as directives.

Directives quite commonly form utterances of their own as stand-alones. However, they may as well be followed by another utterance, cf. (7), typically providing

an elaboration or justification for the request or command expressed by the interactive.

- (7) English (Biber et al. 1999: 407)
- a **Come on**, tell me about Nick. (CONV)
  - b **Hold on!** What are you doing there? (FICT)
  - c **Oh shut up!** You're so cruel. (CONV)

Cases such as (7a) mark the gradual transition from directive to attention signal (Section 3.1), where the semantic content of the form is backgrounded in favor of a text structuring function, namely that of drawing attention to the following text piece. Especially directives derived from verbs of perception, such as *listen* and *look*, are frequently followed by another text piece, and in specific contexts that piece can be interpreted as an obligatory participant of the directive, that is, as a T argument; see Section 3.1 for discussion. In fact, there appears to be a cross-linguistic regular pattern of usage whereby directives acquire features of attention signals taking a T argument. This issue will be discussed in Section 4.2.

### 3.2.4 Discussion

On the analysis proposed here, interjections are one-place expressions, being restricted to S as their only argument, whereas directives are two-place expressions, taking an S and an H argument (see Section 3.2.2). The difference in valence between interjections and directives can be illustrated with the following example volunteered by Ameka and Wilkins (2006): The English expressions *Fuck!* and *Bugger!* are not directed at anybody, they are secondary interjections taking S as their only argument (Section 3.6). The expressions *Fuck off!* and *Bugger off!*, by contrast, are directed at an interlocutor, they are directives, hence they take two arguments, namely S and H.

This also appears to apply to the interactives *Me!* and *Ngke!* of the Mparntwe Arrernte language of Australia, classified by Wilkins (1992: 130) as 'interjections' but clearly having the functions and the two-place valence (S,H) of directives:

The step from imperatives to interjections may be taken by noting that certain types of '**Me!**' 'Take this!' and '**Ngke!**' 'Give that here!' have no cognate forms in the language, but are cognate with the Warlpiri imperative form of the verbs for 'take,' '**Ma-da!**' 'take it!', and 'give,' '**Yu-ngka!**' 'give it here,' respectively. These Mparntwe Arrernte interjections only occur with the intonation pattern given to imperative sentences in the language.

(Wilkins 1992: 130; bold print in the original)

The kind of activities figuring in directives can be illustrated with the Australian examples in Table 3.4. The directives listed in the table relate to general activities,



but in many languages there are also directives specifically dedicated to structuring the speech event concerned. In the Akie language of north-central Tanzania, for example, the latter include expressions such as *ńtân* ('wait!'), used to interrupt the hearer's flow of speech, and *ng'óloon* ('say (it)!'), asking the hearer for (more) information (König, Heine, and Legère 2015a: 129).

Table 3.4 Functions expressed by directives in two Australian languages

| Language   | Genetic affiliation | Form          | Meaning             | Source              |
|------------|---------------------|---------------|---------------------|---------------------|
| Gooniyandi | Bunuban             | <i>pa</i>     | 'come on, let's go' | (Wilkins 1992: 122) |
|            |                     | <i>kaj</i>    | 'O.K., you can go'  |                     |
| Mayali     | Gunwinyguan         | <i>ba</i>     | 'hurry on!'         | (Evans 1992: 277)   |
|            |                     | <i>boi</i>    | 'hey, come here!'   |                     |
|            |                     | <i>me:d</i>   | 'wait on!'          |                     |
|            |                     | <i>ndjudj</i> | 'blow your nose!'   |                     |

As observed in various parts of this book, there are no discrete boundaries separating the various types of interactives. For one thing, there are some interactives located at the boundary between two types, sharing properties with both types. For another, and presumably more commonly, a given interactive form may exhibit properties of type A in one context but of type B in another context.

The following example may illustrate the latter case when a given interactive form may exhibit properties of type A in one context but of type B in another context. In Thai there are a number of abusive expressions with imperative forms, accompanied by terms for diseases to be inflicted upon the person being sworn at. These expressions are suggestive of being essentially directives, as in the following example. The expression *ta:y hàa* (die cholera.spirit), lit. 'die of cholera!,' is frequently used in Thai as a rude curse, 'directed at a person, but more often used merely as an untargeted exclamation when things suddenly go wrong, for instance, while driving on a highway, one rounds a curve and sees large rocks on the road, or one is pounding nails and accidentally pounds one's own finger' (Aikhenvald 2020: 54).

What this account suggests is that *ta:y hàa* is basically a directive taking an S (speaker) and an H (hearer) argument (Section 2.2). But in contexts such as the ones described by Aikhenvald (2020: 54) the expression can be used in soliloquy without implying an H argument, no longer serving speaker–hearer interaction. And in such contexts it turns into an 'untargeted exclamation,' expressing the emotive state of the speaker and being used as an interjection (see Section 4.2).

The communicative significance that directives have can be seen, for example, in the role they play in situations of intense language contact. For example, in the English discourse of Hebrew–English bilinguals in Jerusalem there are also directives taken from Hebrew such as *tagidi li* 'tell me', *ta'amini li* 'believe me',

*tir* ‘look,’ *tishme* ‘listen,’ and *xaki* ‘wait a second,’ facilitating ‘the negotiation of interpersonal closeness vs distance between participants’ (Maschler 1994: 341).

### 3.2.5 Canonical imperatives

Directives are most commonly derived from verbs, but this need not be so. For instance, Wilkins (1992: 133) notes that there is *There, there* in English, which is used to calm someone down and/or give him or her consolation, or *Now, now!*, which may be used to placate or reprove someone. And the imperative form *ayya* ‘come!’ of Algerian Arabic appears to owe its origin not to a verb but to an interjection-like form *ayya* ‘come on!’ (Veselinova 2006: 144–5).

But by far the largest number of directives are derived from what Aikhenvald (2010: 18) calls ‘canonical imperatives,’ that is, forms that have an (implicit) second person singular subject referent as a hearer (or reader or signee) and express commands or requests directed to the hearer.

In many languages across the world, canonical imperatives constitute the simplest verb form. Perhaps their most conspicuous typological characteristics are: They tend to be (a) non-finite, (b) short, and (c) suppletive.<sup>3</sup> Most directives are in fact derived from canonical imperatives and usually conform to these properties (Heine 2016a: 253).

In about one-third of the languages of the world, imperative directives coincide with the verb root or stem, and there are even languages where imperative directives are shorter than the verb stem. For example, in (Logudorese) Sardinian, monosyllabic words are avoided; nevertheless, there are some monosyllabic and monomoraic imperative forms, as Table 3.5 shows; for more examples, see Heine (2016a).

Table 3.5 (Logudorese) Sardinian directives

| Imperative form | Corresponding imperative directive | Meaning |
|-----------------|------------------------------------|---------|
| <i>Mira</i>     | <i>mi</i> ’                        | ‘look!’ |
| <i>Tene</i>     | <i>te</i> ’                        | ‘hold!’ |
| <i>Tocca</i>    | <i>to</i> ’                        | ‘take!’ |

Source: Floricic and Molinu (2012: 6–7)

Furthermore, imperative directives may have special, suppletive forms differing from those of corresponding verbs, and their meanings need not be exactly the same. As the examples from the Khoisan language !Xun of northern Namibia show

<sup>3</sup> Alexander Andrason (p.c. of September 6, 2021) though argues that imperatives are not necessarily non-finite.

(Table 3.6), the imperative forms *khò* ‘wait for me!’ and *nà* ‘give me!’ differ from corresponding verbs not only in their form but also in the fact that they are deictics restricted to first person object referents.

**Table 3.6** Suppletive imperative stems in !Xun (E1 dialect, Kx’a family)

| Verb meaning | Declarative verb form | Imperative directive      |
|--------------|-----------------------|---------------------------|
| ‘come’       | <i>hòě</i>            | <i>tsí</i>                |
| ‘go’         | <i>ú</i>              | <i>tò’á</i>               |
| ‘wait’       | <i>!háít</i>          | <i>khò</i> ‘wait for me!’ |
| ‘give’       | <i>/’àn</i>           | <i>nà</i> ‘give me!’      |

Source: Heine (2016a: 258)

Similarly, Table 3.7 shows that a number of imperative directives in the East African Swahili language differ in their form from their lexical counterparts, even if the two may be etymologically related. For example, the suppletive *nipe* ‘give me!’ can be traced back to the subjunctive form *u-ni-p-e* (2.SG-1.SG.O-give-SUBJ) ‘you should give me.’

**Table 3.7** Partial suppletism in Swahili imperative forms

| Verb meaning | Declarative verb stem | Imperative directive   |
|--------------|-----------------------|------------------------|
| ‘come’       | <i>-ja</i>            | <i>njoo</i>            |
| ‘go’         | <i>-enda</i>          | <i>nenda</i>           |
| ‘give’       | <i>-pa</i>            | <i>nipe</i> ‘give me!’ |

Source: Heine (2016a: 259–60)

Note that neither in !Xun nor in Swahili is there an imperative form for the verb ‘give’; rather, there is only the imperative directive having an inherent first person singular object referent referring to the speaker.

### 3.2.6 Special domains of usage

Typically, directives are used in social interaction among adult persons. But there are also domains of usage where the subjects addressed are not adult persons. One prominent domain concerns interaction with children, such as ‘baby-talk.’ In a number of languages there exists a special set of directives used to encourage young children to perform certain actions. Schrock (2014: 462) found such a set to be distinguished by the speakers of Ik, a language spoken in northeastern Uganda.

A collection of expressions, referred to by the author as ‘infantile imperatives,’ is presented in (8).

Interaction with young children constitutes in fact a domain of language use where interactives play a prominent role; see for example Section 3.11.7, Table 3.29.

- (8) Directives used with young children in Ik (Nilo-Saharan; [Schrock 2014](#): 462)
- |              |                                                   |
|--------------|---------------------------------------------------|
| <b>ḃá</b>    | Used to call a child to come eat                  |
| <b>bubú</b>  | Used to encourage a child to go to sleep          |
| <b>dí</b>    | Used to encourage a child to defecate (‘poo-poo’) |
| <b>kókó</b>  | Used to warn a child not to touch something       |
| <b>kukú</b>  | Used to encourage a child to climb on one’s back  |
| <b>kwaa</b>  | Used to encourage a child to urinate (‘pee-pee’)  |
| <b>mama</b>  | Used to encourage a child to eat                  |
| <b>noónó</b> | Used to encourage a child to nurse                |

A second, even more prominent domain of usage concerns interaction with animals. Since this involves a fairly complex field of usage, we reserve a separate section for it (Section 3.11).

### 3.3 Discourse markers

One out of every 60 words we say is ‘um’ or ‘uh.’

([Enfield 2017](#): 2)

Fillers, like English *um* and *uh*, are useful forms (Section 3.3.4) but, as we will see in this section, they are not the most prototypical instances of discourse markers.

Work on discourse markers has developed into a flourishing field over the last decades, having been the subject of a wide range of studies, although most studies focus on European and some East Asian languages (see, e.g., [Schourup 1999](#); [Aijmer 2002](#); [Brinton 2017](#): 2–37; [Dér 2010](#) for overviews). What has been established in previous research, however, is that discourse markers form a ‘universal category’ in the sense that any given language can be expected to distinguish them.

#### 3.3.1 Introduction

Serving to monitor the production and comprehension of texts, discourse markers are known under a variety of different terms, such as discourse particles, pragmatic markers, formulaic theticals, discourse organizers, discourse signals, etc. (see [Heine et al. 2021](#): 6). Not all terms have the same meaning, but the terms refer to at least a set of the forms that are classified here as discourse markers.

Whether discourse markers in fact form a type of interactives is questioned by John Haiman (p.c. of July 5, 2021). To be sure, discourse markers do not constitute prototypical instances of interactives. For one thing, they rely in their usage—more than most other types of interactives—on the surrounding text. For another, they rarely occur as stand-alones forming utterances of their own. And finally, they lack the expressive power characterizing some other types such as interjections or vocatives.

It would seem, however, that there are reasons to insist that discourse markers are interactives. First, and most important, they conform in every respect to the definition of interactives in Section 1.2; for a definition of discourse markers, see Section 3.3.2. Second, they belong to the same network of grammaticalization as other interactives (Section 4.2). And third, and relatedly, there are a number of interactives that in some contexts behave like discourse markers but in other contexts like interjections, vocatives, or attention signals—that is, there is no clear boundary separating discourse markers from other types of interactives.

There is a group of interactive forms, represented in English by *uh* and *um*, which are classified by some authors as discourse markers but by others as interjections. These forms, most commonly called ‘fillers,’ are treated here as discourse markers and will be discussed in Section 3.3.4.

### 3.3.2 Working definition

There is a plethora of definitions of discourse markers, we mentioned two of them in Section 2.2.3.3 (Schiffrin 1987: 31; Fraser 1999: 938). Most of these definitions highlight the fact that in many of their usages, discourse markers are used to relate two pieces of discourse to one another, as the following more recent definition does:

By a DM I mean a metatextual marker that signals some kind of relationship between clauses/utterances.

(Traugott 2018: 27)

As pointed out in Heine et al. (2021: 1.1.2), however, not all discourse markers, or uses of discourse markers, do signal such a relationship and in the present study we will rely on the following definition instead:<sup>4</sup>

<sup>4</sup> This definition seemingly differs from that stipulated by Heine et al. (2021: 6, (7)): ‘Discourse markers are (a) invariable expressions which are (b) semantically and syntactically independent from their environment, (c) set off prosodically from the rest of the utterance in some way, and (d) their function is metatextual, being anchored to the situation of discourse and serving the organization of texts, the attitudes of the speaker, and/or speaker–hearer interaction.’ As a matter of fact, however, it is essentially the same, the difference being due to the fact that the definition is phrased here with reference to the framework of interactive grammar.

- (1) A discourse marker is an interactive used to express metatextual functions relating to the organization of texts, the attitudes of the speaker, and/or speaker–hearer interaction.

The definition has two components: On the one hand, it locates discourse markers within the category of interactives, that is, as forms conforming to the properties of this category (Section 1.2); on the other hand, it highlights the three functional domains that discourse markers serve, where text organization is clearly the most central domain.

The following are common examples of English discourse markers: *actually, after all, anyway, as it were, besides, however, indeed, if you like, if you will, in fact, instead, I mean, now, okay, say, so, still, then, I think, well, what else, you know, you see*. Note, however, that these forms are not discourse markers in all of their usages. In the following example, it is only the usage of *still* in (2b) that instantiates a discourse marker, while the form *still* in (2a) is an adverb of sentence grammar, not a discourse marker since it does not conform to the definition in (1).

- (2) English (Heine et al. 2021: 3)  
 a Jeff is still sick.  
 b **Still**, Jeff is sick.

English discourse markers have been classified in various ways vis-à-vis other discourse structuring elements. Schiffrin (1987) treated them as a fairly inclusive category, distinguishing three kinds of discourse markers, namely social markers (e.g., *please*), epistemic markers (e.g., *y'know*), and textual markers (e.g., *and, but*). Other authors, in contrast, consider discourse markers to be part of some more inclusive category. For Fraser (1990, 1999), they form a sub-type of pragmatic markers (see Section 5.1.1), and Traugott (2022: 5) includes them in the class of discourse structuring markers, which again form a sub-class of pragmatic markers. Other authors again use the term ‘pragmatic marker’ as an equivalent of ‘discourse marker’ (e.g., Claridge 2013; Beeching 2016; Brinton 2017). Such classifications show that the term ‘discourse marker’ is not employed by all authors in the same way—what Schiffrin (1987) calls discourse markers are not discourse markers but rather pragmatic markers in the terminology of Traugott (2022: 85), etc. In the present study, it is only elements that conform to the definition in (1) that are classified as discourse markers.

Like interjections, discourse markers also exhibit a distinction between primary and secondary forms in English (see Section 3.6.4.1).

### 3.3.3 Grammar

The conceptual content of discourse markers is frequently hard to describe, especially since it is susceptible to contextual manipulation on a large scale, and in

many cases there is no discernable content that would allow for an appropriate paraphrase. Whether discourse markers can be said to have a ‘core meaning’ has been discussed controversially. What is uncontroversial, however, is that discourse markers are notoriously polyfunctional; for example, Beeching (2016: 53–6) identifies the nine different functions in (3) for the English discourse marker *well*, and more functions in addition have been mentioned by other authors (e.g., Jucker 1993, 1997; Aijmer and Simon-Vandenberghe 2003; Furkó 2005; Cuenca 2008; Beeching 2016: 51–75). The example in (4) illustrates the other-correction function of *well* in (3g).

(3) Functions subserved by the DM *well* (Beeching 2016: 53–6)

- a Hesitation
- b Transitional *well*
- c Changing the topic
- d Raising an objection
- e Prefacing a dispreferred response
- f Taking a turn/polite interruptions
- g Other-correction
- h Self-correction
- i Quotative *well*

(4) English (Beeching 2016: 55, 3.7)

- A: but if it involves earning nothing then I’m gonna have to rule that out  
 B: **well** you’ll get expenses paid

A rough paraphrase of the function of *well* in (4) is proposed in (5a) (underlined) and its argument structure in (5b). As pointed out in Section 2.2.3.3, many discourse markers have a complex T argument: Being connective forms that typically relate two pieces of discourse to one another, they have two thematic roles, where one (T<sub>1</sub>) precedes and the other (T<sub>2</sub>) follows the discourse marker. Thus in (4), T<sub>1</sub> consists of the utterance made by A, and T<sub>2</sub> is instantiated by B’s response *you’ll get expenses paid*.

- (5) a Paraphrase: I (S) want to slightly correct what you (H) have just said (T<sub>1</sub>) by drawing your attention to the fact that there is also T<sub>2</sub>.  
 b Argument structure: *well* (T<sub>1</sub>, S, H, T<sub>2</sub>)

Similarly, the English interactive *oh*, commonly classified as an interjection, has a range of usages as a discourse marker involving a question (T<sub>1</sub>) followed by a response (T<sub>2</sub>) prefaced by *oh*, where ‘T<sub>2</sub> indicates that T<sub>1</sub> has occasioned a shift in the respondent’s attention’ (Heritage 1998: 293, 326).

Note that  $T_1$  can be implied rather than being formally expressed. For example, the discourse marker *after all* typically establishes a link between  $T_1$  and  $T_2$ , creating a justifying, concessive, or other relation between the two. But in example (6),  $T_1$  is not formally expressed, it is only recoverable from the situation of discourse rather than being overtly marked, as provided by the contextual information summarized below in square brackets. On this interpretation, (6) can be paraphrased as in (7a) (where the paraphrase of the discourse marker is underlined). This paraphrase suggests that, unlike in (4), there need be no H argument involved—that is, (6) could have been uttered in soliloquy.

- (6) English (Lewis 2000: 133, (11))  
 [The speaker takes an extremely large slice of cake]  
**After all**, it is my birthday.
- (7) a Paraphrase: I (S) am entitled to an extremely large slice of cake ( $T_1$ ) because it is my birthday ( $T_2$ ).  
 b Argument structure: *after all* ( $T_1, S, T_2$ )

As was observed in Section 2.2.3.3, constituent anchored discourse markers need not have semantic-pragmatic scope over two segments of discourse (that is,  $T_1$  and  $T_2$ ) but only over some specific constituent of an utterance. In such usages of discourse markers, the T argument is restricted to that constituent. Constituent anchored discourse markers most commonly precede their T argument, but they may also follow it. For example, the discourse marker *if you like* precedes its T argument *of tribes* in (8), while in (9) the discourse marker follows its T argument *middle-aged*.

- (8) English (Claridge 2013: 174, (22b))  
 As I have said, I am not explaining the origin of races, but of nations, or, **if you like**, of tribes.
- (9) English (Claridge 2013: 174, (22d))  
 ‘What!’ I exclaimed. ‘Lady Y –: that funny old woman!’ ‘No – middle-aged,’ he corrected. [...] ‘Very well, middle-aged **if you like**.’

As we will see now, however, the structure of discourse markers is more diversified than that. Even constituent anchored discourse markers occasionally have a double T argument. For example in (10), the speaker corrects the segment *college* by *uni* ‘using *well* to flag a hesitation about the use of the word ‘college’ and the transition to the correction ‘uni’ (Beeching 2016: 56). Rather than other-correction, the function of *well* in (10) concerns self-correction (cf. (3h)), and *well* could be replaced by some other discourse marker used for this function, in particular by *I mean*. In this kind of context, the segment *college* can be said to constitute  $T_1$ , while  $T_2$  consists of the segment *uni*, and the DM has a connective function, namely that



of relating  $T_1$  and  $T_2$  to one another. This kind of constituent-anchored usage of the discourse marker *well* can be paraphrased roughly as in (11).

(10) English (Beeching 2016: 56, 3.8)

**hi um well you know** we finish college **well** uni, in two weeks

(11) I (S) need to correct myself in replacing *college* ( $T_1$ ) by *uni* ( $T_2$ )

To conclude, prototypical instances of discourse markers differ from other types of interactives in distinguishing two T arguments ( $T_1, T_2$ ) each instantiating a separate text piece. Such a text piece may extend over a larger discourse unit, but it may also be extremely short, consisting of a single word, as (10) shows, or it may be implied rather than being encoded, as (6) shows. But there also is a second structure where there is only a simple T argument (cf. (12c)), as we saw in Section 2.2.3.3. In short, discourse markers typically take either of the argument structures in (12). There may also be an H (hearer) argument, cf. (12a), but frequently there is no identifiable H argument, cf. (12b).

(12) Argument structures of discourse markers

a ( $T_1, S, H, T_2$ )

b ( $T_1, S, T_2$ )

c ( $S, T$ )

All observations that have been made on discourse markers suggest that the left periphery of a sentence or utterance is cross-linguistically their favored position (Onodera 2011; see Heine et al. 2021: 155 for discussion). Accordingly, Aijmer (2002: 29) concludes that the initial position ‘functions as a clue to discourse particle status.’ However, there are many exceptions to this placement preference across languages, paradigm examples being provided by utterance-final *then*, *though*, *anyway*, or *but* in English (e.g., Haselow 2011, 2013, 2016a) or Korean *kuntey*, marking disagreement, rejection, or refusal (Kim and Sohn 2015).

Discourse markers are prototypical instances of text-organizing devices and as such they belong to those interactives that do not typically qualify as stand-alones, that is, as occurring as utterances of their own. Nevertheless, a number of them can occur as stand-alones, forming, e.g., distinct turns in conversational exchanges. English examples are *anyway*, *okay*, *right*, *sure*, and *well*, and similar cases are found in other languages. But there are also languages where the left periphery does not seem to enjoy a privileged role in the placement of discourse markers. Dér and Markó (2010: 144–5) found that of the 14 most frequently used Hungarian discourse markers, only 39.6% occurred in a turn-initial position while 40.2 % occurred turn-medially, 8.3% turn-finally, and 11.9% were stand-alones.

### 3.3.4 Fillers

Why fillers in spontaneous spoken discourse have had such an erratic research history while at the same time being on the receiving end of both the opprobrium and the neglect of the linguistic community remains a mystery.

(O'Connell and Kowal 2008: 126)

Also called hesitation fillers, hesitation markers, hesitators, interjective hesitators, filled pauses, or pause fillers, fillers form a cross-linguistically fairly common discourse-structuring device (see Heine et al. 2021: 260–6 for more details). The term 'filler', if taken literally, may be somewhat misleading since it does not very well reflect the functions that such expressions have. Thus, Enfield (2017) notes about the English fillers *uh* and *um*:

They show the curious property, unique to human language, of being part of a communication system that can communicate about itself. They don't add information about the current topic of conversation; they add information about the state of mind of the speaker and about the flow of the conversation itself.

(Enfield 2017: 106)

Such functions have been associated in particular with monitoring the flow of the production of utterances or, more generally, with progressivity in linguistic interaction. It is most of all the following functions, or tasks, that are mentioned in the relevant literature (e.g., Clark and Fox Tree 2002; Hayashi and Yoon 2006; O'Connell and Kowal 2008: 127–9; Watanabe 2010; Tottie 2014, 2015):

- (13) Functions frequently expressed by fillers (Heine et al. 2021: 262)
  - a Pause filling
  - b Marking of hesitation and/or delay
  - c Searching for what to say next, or prefacing a new conversational move
  - d Correcting (e.g., self-repair)
  - e Signposting speaker turns (turn taking, turn holding, turn yielding)
  - f Inviting the hearer to reply
  - g Commenting on the preceding discourse
  - h Breaking the continuity with preceding discourse (e.g., topic change)
  - i Preparing the hearer for what is going to happen next
  - j Attracting attention or highlighting
  - k Mitigating undesirable effects of the message
  - l Being polite

The following example relates to (13c), where the filler *um* seems to signal that Judy needs processing time since she cannot find a simple answer to the foregoing question:

- (14) English (SBCSAE 048; Norrick 2009: 875)
- Lea: how'd you open it.
- Judy: **um** it's,  
let's see it,  
**well**,  
it's supposed to ... open on the side.

English has a range of elements that have been classified as fillers (see O'Connell and Kowal 2008: 126–7), but two forms are generally considered paradigm exemplars of fillers, namely *uh* ([ʔə(:)]) and *um* ([ʔə(:)m]), usually transcribed *er* and *erm*, respectively, in British English (e.g., Biber et al. 1999: 1092; Clark and Fox Tree 2002; O'Connell and Kowal 2004, 2005; Tottie 2011, 2014, 2015). The example in (15) illustrates the English filler *uh*, occurring three times in one utterance.

- (15) English (ICE-GB: s1b-075-68; Heine et al. 2021: 271, (17))
- We did feel **uh** union council was two or three weeks ago when this was put to us that **uh** the increase from seven to sixteen **uh** was actually a very good idea.

It is uncontroversial that fillers are interactives, showing the definitional properties presented in Section 1.2: They are (a) invariable forms primarily serving the organization of texts, (b) their meaning is not part of the meaning of the sentence they may be associated with, (c) they are syntactically unattached, (d) they tend to be prosodically set off, and they are (e) indexicals which cannot be (f) negated or (g) questioned. There is, however, a slight difference to other types of interactives in that, more than other interactives, they may be found being prosodically integrated in the surrounding text.

Most commonly, fillers have been classified either as discourse markers (e.g., Hayashi and Yoon 2010: 43; Tottie 2011, 2014, 2015) or as interjections (e.g., Biber et al. 1999: 1054; Clark and Fox Tree 2002: 76–80; Norrick 2009). We follow Tottie (2014: 21) in treating them as discourse markers since they conform to the definitional criteria of the latter (Heine et al. 2021: 260).

Like 'canonical' discourse markers, fillers are used to monitor the flow of the delivery of an utterance (Hayashi and Yoon 2010: 58) and, unlike interjections, both fillers and other discourse markers are usually anchored to some piece of text, that is, they are not normally found as 'stand-alones.' The boundary between 'canonical' discourse markers and fillers is frequently fuzzy, and one and the same expression not uncommonly serves both functions. For example, the discourse markers *t'a* and *tɛ* of the Tibetic language Denjongke of northeastern India can

as well do service as fillers, e.g., when the speaker is not yet sure of what to say (Yliniemi 2019: 134).

There also is frequently no clear boundary between interjections and fillers either (O'Connell and Kowal 2004: 469–70; 2005: 571–2), but O'Connell, Kowal, and Ageneau (2005) point out that fillers need to be distinguished from interjections not only in their inability to occur as free-standing turns but also in their placement, in prosodic features, and in their lack of affective and emotional significance.

Unlike some other discourse markers, fillers are almost entirely a feature of spoken language: 'We think it is quite safe to say that the filler is not at home on the printed page' (O'Connell and Kowal 2008: 126). But, as the following example shows, they can also be observed occasionally in writing, even in a headline of a newspaper article, as in (16).

- (16) English (*The Los Angeles Times*, December 17, 2009; Tottie 2015, (13))  
Ashley Dupre: From prostitute-to-politicians to, **uh**, paid journalist

Like a number of other types of interactives, fillers show a distinction between primary and secondary ones (Section 3.6.4.1). Primary fillers, such as English *uh* and *um*, are not transparently related to any units of sentence grammar, and they are etymologically opaque. Secondary fillers by contrast have identical or similar forms in sentence grammar, and all evidence that there is suggests that they are historically derived from the latter. There are languages that are largely or entirely restricted to primary fillers, such as English (*uh*, *um*), Dutch (*uh*, *um*), German (*äh*, *ähm*), and French (*eu*, *euh*, *em*, *eh*, *oe*, *n*, *hein*).

Other languages again draw overwhelmingly or entirely on secondary fillers. The main sources of the latter are provided by demonstrative pronouns, question words, and adverbs. In the languages of East Asia it is demonstrative pronouns that appear to be the main source of fillers. Japanese uses the adnominal form of the distal demonstrative *ano* 'that', Korean the adnominal forms of the medial (*ku*), the distal (*ce*), and the locative demonstratives (*ceki* 'that (far away)'), and Mandarin Chinese speakers the proximal demonstrative (*zhe-ge*) and the distal demonstrative (*na-ge*) as common fillers. Thus, in the following example, Hayashi and Yoon point out that the Japanese demonstrative-derived filler *ano* has a function similar to English *um*:

- (17) Japanese (Hayashi and Yoon 2010: 43; excerpt only)  
iya konkai **ano** hashittemo [...].  
well this.time DEM run:even.if  
'Well, this time, **ano** [= *um*], even if ((you)) run ((in a race)), [...].'

Finally there are also languages combining both primary and secondary fillers. Spanish appears to be such a language: As the data in Clark and Fox Tree

(2002, Table 8.1) suggest there are, on the one hand, the primary fillers *eh* and *em* and, on the other hand, secondary ones, namely *este* (< *este* ‘this’) and *pues* (< *pues* ‘then’).

### 3.3.5 A note on clicks as English discourse markers

Based on the analysis of telephone interactions in British and American English, Wright (2011) finds that clicks are regularly produced in new sequence boundaries of discourse. Alongside the phonetic parameters of pitch, articulatory segmental features, and voice quality, the ‘new sequence indexing clicks’ function to demarcate the onset of new and disjunctive sequences, frequently followed by the discourse markers *anyway*, *right* as well as fillers (Wright 2011: 224).

Wright identifies four different click types that instantiate an English discourse marker (called by the author a ‘particle’), namely a bilabial one ([⦿]), an alveolar/dental one ([!]), a labiodental one ([⦿]), and an alveolar lateral one ([!l]), where the first two account for 97% of all 86 occurrences of the marker that the author found in her text sample.<sup>5</sup> The following example illustrates the use of the discourse marker with two instances of the bilabial click ([⦿]), produced by the same speaker.

(18) English (Holt.SO.88.1.2/bath/; Wright 2011: 221, Fragment 7)

1. Bill: hello:
2. Gor: [⦿] hi Bill
3. Bill: hi Gordy
4. Gor: [⦿] uh:m (0.4) are you going tonight
5. (.)
6. Bill: mm
7. Gor: .hhh (0.2) would you mind giving me a lif[t
8. [no that’s alright

The function of the marker is to demarcate the onset of a new and disjunctive sequence in discourse. But boundary marking is not the only function of the marker, another paradigm function is word searching: The marker is used systematically in environments where the speaker is actively searching for a word (Wright 2011: 224).

On account of these two functions we suggest treating the ‘new sequence indexing’ clicks as an English discourse marker, bearing the following problems in mind: First, it does not become entirely clear whether the marker is conceived by its users as a conventionalized grammatical form. And second, more information on the role played by the four different click types is needed: Is each of the

<sup>5</sup> Wright (2011: 214–15) points out, however, that she had problems establishing whether a click was produced with alveolarity or dentality.

four variants associated with its own functional and/or contextual features or are the four free variants?

Importantly, the clicks are ‘consistently produced as singletons, without voicing or nasality and with a high amplitude’ (Wright 2011: 209)—that is, they differ from the click consonants of the click languages of Africa (Section 1.3.2.8). This is suggested by the following: First, rather than being part of some meaningful form, the clicks themselves can be meaningful forms without combining with any other phonological elements such as vowels. And second, the clicks do not seem to take phonetic accompaniments relating to voice, nasalization, aspiration, glottalization, etc. as they are generally found in click languages. Thus, the English discourse marker lends further support to the hypothesis that click-only forms are a salient feature of interactive grammar.

### 3.3.6 Discourse markers in language contact

One way of finding out whether a person is or is not a native speaker of English is by listening to how they use discourse markers: Where does the person place them, and for what functions are they used? Learning to use discourse markers appropriately is in fact a task that requires a fairly advanced command of the language concerned, and formal language teaching is of little help in this task since discourse markers tend to be discouraged in formal education. We noted in the introduction to Section 3.3 that discourse markers can be understood to form a ‘universal category.’ But exactly the way they are used is highly language specific and presupposes a pronounced familiarity with the relevant language.

It may therefore come as a surprise that discourse markers move fairly easily from one language to another—that is, they are borrowed more readily than most other forms of a language. It would seem in fact that discourse markers are amongst the first grammatical items that speakers borrow or code-switch in situations of intense language contact—accordingly, Matras (2009: 193–7) draws attention to the ‘contact-vulnerable’ and ‘almost epidemic status of discourse markers’ in bilingual situations. A paradigm example is provided by the English DM *okay*, which has found its way into many languages across the globe.

The present section is restricted to a few general observations; the reader is referred to Heine et al. (2021, chapter 7) for more detailed treatment. Two examples may illustrate the magnitude of areal distribution that discourse markers may acquire.

The first example concerns the Spanish discourse marker *entonces*, derived from the adverb *entonces* ‘then, therefore, thus’ (Stolz 2007). Following the rise of the Hispanic Monarchy from the end of the fifteenth century onwards, Spain controlled vast portions of the Americas, the Philippines, various islands in the Pacific as well as in Africa, and the Spanish influence in these regions is reflected

in the worldwide distribution of the discourse marker *entonces*. As the research carried out by [Stolz \(2007: 77–8\)](#) demonstrates, the discourse marker is now found, for example, in languages such as Guaraní in Paraguay, Totonac of Mexico, Hiligaynon of the Philippines, or Rapanui of Easter Island.

One of the main functions subserved by the borrowed discourse markers is to introduce consecutive connections between pieces of discourse. Thus, in the following example from the Philippinian language Hiligaynon, the marker *intonsis* (<*entonces*) establishes causal relations between two clauses:

- (19) Hiligaynon (Austronesian, Philippinian; [Stolz 2007: 78](#))

Madamo ang bulak dira' **intonsis**  
 many DET flower at.there therefore  
 manguha kita.  
 FUT.get 1.PL  
 'There are lots of flowers there, so let's get some.'

The second example is about the Arabic discourse marker *yaa'ni* ([jæʕni] or [jaʕni]), which is derived from the third person masculine singular imperfective of the defective verb *'anā* 'he means,' whose meanings include that of 'that is.' The marker is found across northern Africa and the Middle East where Arabic is spoken as a first language, its reflex being [jæʕne] in Libanese, Syrian, and Palestinian Arabic, serving as a discourse marker to signal functions such as reformulation, correction, and background repair.

But in addition, *yaa'ni* has also been borrowed in genetically unrelated languages of Africa, Europe, and Asia where Arabic spread as the language of Islamic religion and culture. Such languages include Turkish, where it takes the form is *yani* 'that is, in other words,' Swahili in East Africa (*yaani* 'that is, that is to say, I mean'), and Persian in Iran (*yæ'ni* 'that is'). In all these languages, the discourse marker exhibits a wide range of functions; in Persian, for example, *yæ'ni* includes the following main functions ([Noora and Amouzadeh 2015](#)): (a) filler or hesitation marker indicating ongoing planning, (b) repair, serving as a 'mistake editor,' or marker of (self)-initiated (self)-repair of a preceding utterance, (c) elaboration, clarification, expansion, explanation, or reformulation of the preceding utterance, (d) a voluntary marker of 'imprecision,' (e) non-equivalence, (f) presenting a 'further instance' where the general is made more specific, and (g) summing up ('the point is').

Presumably the most persistent function across these languages is that of (c), illustrated in the following example by the reformulation use of *yæ'ni* in Persian:

- (20) Persian (Indo-Iranian, Indo-European; [Noora and Amouzadeh 2015: 99](#))

Ali! **yæ'ni**, âqâye Ahmadi, ketabe- toon.  
 Ali, meaning Mr. Ahmadi book.EZ- your  
 'Ali! I mean, Mr. Ahmadi, here is your book!'

Why are discourse markers constantly transferred from one language to another? It would seem that the main reason can be seen in their independent holophrastic nature. Elements of sentence grammar, such as nouns, verbs, etc., are syntactically, prosodically, and semantically integral parts of sentence structure, functioning as arguments or predicates, and they may have a complex internal morphological structure. This is fundamentally different with discourse markers and other interactives: They are invariable expressions, syntactically unattached, typically separated prosodically from surrounding text material, and not being a semantic part of sentence structure. Accordingly, they are technically much easier to handle and to extract and integrate in another language. For the learner of another language, therefore, interactives are most readily identifiable to be deployed for discourse processing.

At the same time, discourse markers acquired via language contact tend to remain psychologically associated with their donor language, evoking cultural and/or emotional features believed to characterize the donor language community. These markers have therefore been characterized as signaling ‘emblematic switches’ (Poplack 1980: 614). But rather than constituting an obstacle to borrowing, this may in the right circumstances also be viewed as a factor favoring rather than discouraging transfers from one linguistic community to another.

### 3.3.7 Grammaticalization

Of all types of interactives, discourse markers are the most grammaticalized ones. As will be shown in Section 4.2.2, they constitute the endpoint in the development of interactives. This means that there is a one-way traffic: Discourse markers are commonly derived from attention signals, directives, interjections, and vocatives while they themselves do not give rise to any other interactives.

The result is that discourse markers are special in a number of ways, in particular for the following interrelated reasons. First, they are most strongly integrated in the structure of texts. Second, they differ from all other interactives in commonly taking two text pieces as their T argument, as we saw in Section 3.3.3 (but see especially Section 2.2.3.3). And third, being strongly dependent on the surrounding discourse, they also differ from other interactives in not normally occurring as stand-alones, that is, forming utterances of their own.

However, interactives constitute only one of the ways in which new discourse markers can arise. A clearly more common source is provided by text pieces of sentence grammar such as reduced clauses (e.g., English *you know*) and adverbs (e.g., English *well*). In Heine et al. (2021: 245, Table 8.1, 23) general sources are



listed for discourse markers in the languages of the world, of which only three are interactives, namely directives ('imperatives'), ideophones, and vocatives. This is an issue for which there is abundance of studies (see especially [Brinton 2008, 2017](#); [Heine et al. 2021](#)), but since this issue is beyond the scope of the present study, the reader is referred to the studies mentioned.

### 3.3.8 Discussion

In some publications, discourse markers are called interjections, and interjections are called discourse markers. A number of the expressions of the Mayali language in Western Arnhem Land in Australia are referred to by [Evans \(1992\)](#) as 'interjections'; yet on the basis of the definition proposed in Section 3.3.2 they might more profitably be re-classified as discourse markers. This applies, for example, to the 'organizing interjections' *bonj* and *wanjh*, which Evans portrays thus:

The most important conversational organizers in Mayali are *bonj*, which signals the completion of one topic or organizational unit, and *wanjh*, which shows a new direction for the conversation and can also begin a new turn or a whole conversation. Each can work at a number of levels, signalling episodes within a story or turn, turns in a conversation, or boundaries to a whole activity, conversation or narrative.

([Evans 1992](#): 230)

All the functions distinguished by Evans are paradigm instances of discourse marker usages and we therefore suggest re-classifying *bonj* and *wanjh* as discourse markers, and much the same applies to many other instances of 'interjections' appearing in the relevant literature. In doing so we rely, on the one hand, on the definition in Section 3.3.2 and, on the other hand, on the argument structure of the expressions concerned: Canonical interjections have only one argument (Section 3.6.3), namely S (speaker), whereas discourse markers have a more complex argument structure, as we saw in Section 3.3.3.

The problem that this issue raises, however, is more complicated. For one thing, the boundary between discourse markers and interjections is not seldom fluid. For another, a number of expressions can be found cross-linguistically which have the structure of interjections in some contexts but that of discourse markers in other contexts. The main, if not the only reason for this fact is that there is a common process of grammaticalization leading from interjection to discourse marker, as we saw in Section 3.3.7. And since this process is gradual it may cut across the boundary between the two types of interactives, where less grammaticalized uses of an expression are suggestive of interjection status while more strongly

grammaticalized uses have the appearance of discourse markers. This situation will be looked at in more detail in Section 4.2.

### 3.4 Evaluatives

Of all types of interactives, evaluatives have found the least attention in studies of interactives. Yet it would seem that they play an important role in linguistic discourse. Speakers not only confirm or disconfirm what the previous speaker has said or written (Section 3.8), they also comment on the quality of a preceding contribution or some state of affairs, most commonly using evaluative interactives for this purpose.

Evaluatives are, as we argue, stance markers. In the framework of stancetaking proposed by Du Bois (2007: 144), three main concepts of stancetaking are proposed, namely evaluation (*that's horrible*), positioning (*I'm glad*), and alignment (*I agree*). Evaluatives are instantiations of the concept of evaluation, which Du Bois considers to be the most salient and widely recognized form of stancetaking.

The term 'evaluative' has found many different applications in the linguistic literature, relating for the most part to specific discontinuities observed in sentence grammar (see, e.g., Orozco 2016). In the present study it is used strictly only for a set of interactives to be defined in Section 3.4.2.

#### 3.4.1 Introduction

Evaluatives are routine formulae presenting a concise assessment of what the speaker just experienced. The assessment expressed by an evaluative is for the most part positive, less commonly negative, and it may also signal responses like excitement, surprise, disappointment, or fright.

The transition from evaluatives to exclamatives (e.g., English *how nice*, *what a shame*) and interjections (*yuck*, *shit*) is gradual. Like other types of interactives, evaluatives exhibit significant overlaps with other types, most of all with response signals but also with discourse markers, exclamatives, interjections. Hence in many of their usages they cannot easily be delineated from those types. Nevertheless, prototypical instances, like *wonderful* or *great*, express epistemic functions that are sufficiently different from those of other types of interactives to justify treating them as a type of their own.

Response signals, like *yes* or *no*, are usually triggered or invited by a previous remark or a question by a different speaker. As a rule, this is also the case with evaluatives, but it need not be so. Thus, in the exchange of (1), the evaluative *brilliant* used by Alina appears to refer not primarily to what Lenore just said but rather to the remark she herself has made earlier, namely that some specific referent is going to CSUN (California State University).

- (1) English (SBCSAE 0006; Norrick 2009: 876)

ALINA: and he's going to CSUN,  
because that's the only place you can go at night.

LENORE: **wow.**

ALINA: **brilliant.**  
he's been in there since nineteen eighty,

Biber et al. (1999: 520) treat evaluatives as 'adjectives used as exclamations,' especially in conversation and fictional dialogue, using examples like the following:

- (2) English (Biber et al. 1999: 520)

**Good!** I like that. (CONV)

While evaluatives are in fact commonly derived from adjectives, they need to be distinguished from the latter because they unambiguously conform to the definition of interactives (Section 1.2), differing from adjectives, e.g., in the fact that they are not part of syntactic constituents like noun phrases. Biber et al. (1999: 520) add, however, that '[a] number of these could alternatively be regarded as inserts,' giving the following examples from fiction: *Ridiculous!*, *Irrelevant!*, *Stupendous!*, *Good gracious!*, *Miraculous!*, *Glorious!*, and *Marvelous!*. Note that 'inserts' belong in fact to our category of interactives (Section 5.1.2). Furthermore, evaluatives like *Excellent!* have been classified as truncated sentences (Meinard 2015: 164) but again, to the extent that their usage is in accordance with our definition of interactives they are included here.

### 3.4.2 Working definition

The term 'evaluative' has found many different usages in linguistics. It is used here strictly in the sense defined in (3).

- (3) An evaluative is an interactive used to express a concise appraisal of what the previous speaker has just said.

Common English examples of evaluatives are *amazing*, *bloody brilliant*, *bullshit*, *excellent*, *fine*, *glorious*, *good*, *good gracious*, *great*, *hilarious*, *horrible*, *how awful*, *incredible*, *irrelevant*, *lovely*, *lucky*, *marvelous*, *miraculous*, *nonsense*, *oh dear*, *ridiculous*, *right*, *tough*, *splendid*, *stupendous*, *super*, *terrible*, *unbelievable*, *uncanny*, *what a pity*, *tough*, and *wonderful*. Furthermore, Mackenzie (1998: 282) found in English fictional work evaluatives such as the following: *bloody shame*, *how interesting*, *how nice*, *very funny*.

The definition in (3) takes care of prototypical usages of evaluatives. As we will see in Section 3.4.3, there may also be less prototypical situations where instead

of a comment by the previous speaker there is a non-verbal stimulus, and that stimulus need not be caused by another interlocutor.<sup>6</sup>

### 3.4.3 Grammar

Evaluatives commonly form utterances of their own. But when combined with other text pieces the clearly preferred placement slot of an evaluative is at the left periphery of an utterance, cf. (4).

- (4) English (Biber et al. 1999: 520)

**Great!** I need some of those. (CONV)

The conceptual content of evaluatives, like *great* in (4), can be paraphrased as in (5a), suggesting that there are three arguments, cf. (5b).

- (5) a Paraphrase: I (S) find great what you (H) have just said (T).  
b Argument structure: *great* (S,H,T)

Evaluatives can be uttered in soliloquy without a hearer being present or implied. In such cases there is no H argument, that is, the structure in (5) is reduced to (S,T).

There is no information on the T argument in the text example of (4) provided by Biber et al. (1999: 520)—such an argument is implied in the usage of the evaluative *great*. But (4) contains an adjunct, namely *I need some of those*. The adjunct is an optional participant which is not required by the valence of *great*: If it were omitted then the utterance in (4) would nevertheless remain well-formed. Evaluatives are quite often followed by such adjuncts, whose function it normally is to provide further information relating to the usage of the evaluative.

Evaluatives may collocate with interjections, and in this case, they tend to follow the latter. Thus, in example (6), the evaluative *good* follows the interjection *oh*, where the T argument is provided by Kevin's response *hell yeah*.

- (6) English (SCoSE (Football); Norrick 2009: 880)

Lana: did you burn anything?

Kevin: **hell yeah**.

Lana: **oh good**.

((laughs))

Quite commonly, evaluatives are used as reactions to what the previous speaker just said (T), less commonly also to what the speaker herself has just said. However, there are various other kinds of situations that may trigger the use of evaluatives,

<sup>6</sup> This observation was made to us by Gunther Kaltenböck (p.c. of September 26, 2021).

and such situations need not involve a previous speaker, as the following example may show.

(7) English

[Paul and Carla are standing at the train station. Noticing that their train is arriving, Paul says:]

**Excellent.**

In such situations the T argument does not refer to any previous text piece but stands for some experience that the speaker has just had. In (7) it is the arrival of the expected train that induces Paul to produce the evaluative *excellent*. If Carla were to inquire about T by asking, ‘What is excellent?’, Paul may simply point at the arriving train. Thus, the T argument need not be overtly expressed if it is implied by the situation of discourse.

As just noted, evaluatives are frequently followed by some comment rephrasing or elaborating on the meaning of the evaluative. However, such comments are not part of the argument structure of evaluatives. While they may present useful additional information, they are not entailed in the meaning of evaluatives and omitting them does not normally result in a non-felicitous utterance.

### 3.4.4 Discussion

An interesting example of an evaluative in Japanese is provided by the expression *yokumo*, which when used as a declarative transforms an ordinary assertion into a kind of negative exclamative expressing surprise (Gutzmann and Henderson 2019: 120).

(8) Japanese (Gutzmann and Henderson 2019: 120, (62); McCready 2010: 40)

**Yokumo** Dallas to kekkon shita na!

*yokumo* Dallas with marry did PTCP

‘He did an amazingly stupid and shocking thing by marrying Dallas!’

While showing a structural format that differs from the examples looked at above, Japanese *yokumo* appears to be an evaluative presenting a comment and taking two arguments, S and T; whether there is an H (hearer) argument implied does not become clear from the data provided. T follows the evaluative and is described as belonging to the truth conditional layer while *yokumo* is said to be located at the ‘use conditional’ layer or dimension:<sup>7</sup>

That is, *yokumo* does not leave anything meaningful behind in the truth-conditional dimension.

<sup>7</sup> Concerning the notion ‘use conditional’ meaning, see Section 2.1.

Accordingly, a *yokumo* sentence cannot be used to make an assertion, since it lacks the necessary propositional content; rather, it can only be used to perform an expressive speech act.

(Gutzmann and Henderson 2019: 120)

On the interpretation proffered here, the conceptual content of *yokumo* can, in accordance with the description provided by Gutzmann and Henderson (2019: 120), be paraphrased roughly as ‘I (S) think that T is bad and I am surprised and shocked by T,’ where T stands for an argument instantiated by the propositional content of the sentence following *yokumo*.

While evaluatives are found most commonly in speech, they also appear in writing. A special usage is pointed out by Biber et al. (1999: 520): Evaluatives, referred to by them as ‘adjectives as exclamations,’ also occur in headlines and captions in news writing, such as *Blinding!* or *Hair-raising!*. What is unusual in this usage is that, like in the Japanese example of (8), it is not retrospective but prospective, that is, the T argument provided by the following text follows rather than precedes the evaluative form.

### 3.5 Ideophones

[N]o grammar of an African language with ideophones can be considered adequate without a treatment of ideophones.

(Childs 1995: 132)

One is inclined to add to this claim that so far, no African language—or any other language for that matter—has been found to lack ideophones (Section 3.5.1). But of all interactives, ideophones are the most marginal and this raises the question of whether they in fact qualify as a type of interactives anyway. The answer proffered here is in the affirmative for reasons given in Section 3.5.7.

With regard to their diachronic status, ideophones are somehow double-faced. On the one hand, they are for the most part etymologically opaque—more so than other types of interactives. On the other hand, they have the appearance of a transit category: As we will see in Chapter 4, words of sentence grammar can become ideophones and ideophones can become words of sentence grammar. While this also applies to other types of interactives, such a bi-directional traffic of grammatical change is clearly most pronounced in ideophones. What this means with reference to the grammatical structure of ideophones is that in some languages they may have as much in common with words of sentence grammar as they have with interactives (Section 4.4).

Ideophones have been covered in great detail and in all its various aspects of behavior most of all in the work of Mark Dingemanse, and the reader is referred

to this work for more information (see References). This section is therefore restricted to features of ideophones that immediately relate them to the grammar of interactives.

### 3.5.1 Introduction

Ideophones have been called ‘linguistic rebels’ (see [Andrason 2021a](#): 122), and they have been portrayed as playful enactments or ‘theatrical performances’ of events or states (Voeltz and [Kilian-Hatz 2001](#) and the contributions therein; [Haiman 2018](#)). [Güldemann \(2008: 280\)](#) defines them as ‘expressive and symbolic lexemes’ belonging to what he calls the mimetic mode of language use. They are in fact also called ‘mimetics,’ e.g., in Japanese linguistics, or ‘expressives’ ([Dwyer and Moshi 2003](#): 174), especially in Southeast Asian linguistics ([Dingemanse and Akita 2017](#): 501). For example, [Kita \(1997: 381\)](#) writes: ‘Japanese mimetics are a class of words that not only are referential but also evoke a vivid at-the-scene feeling.’ For a list of terms that are employed for ideophones, see [Dingemanse \(2018: 2, Table 1\)](#).

While ideophones ‘are conspicuously undeveloped and poorly structured in the languages of Europe’ ([Diffloth 1972](#): 440), they belong to the most numerous form classes in some languages of Africa and eastern Asia, their number being close to that of nouns and verbs: There are thousands of ideophones in the spoken usage of a number of languages. According to a compilation presented by [Dingemanse \(2018: 15, Table 4\)](#), Basque has more than 4500 of them, Japanese 4500, Gbeya of central Africa over 3000, and Zulu 3000. Comparable numbers of ideophones appear to exist in the Asian languages of the Austroasiatic family, spoken on the Malay Peninsula. For example, the Semai language of western Malaysia has an ideophone class of the same magnitude as nouns and verbs ([Dingemanse 2019: 24](#)).

But arguably the largest inventory of ideophones is found in Korean, which has been shown to have 5249 of them (Seongha Rhee, p.c. of June 10, 2021), and a lexical collection of Korean ideophones lists 29,600 entries ([Koo and Rhee 2018](#)). Ideophones contrast in this respect conspicuously with other types of interactives. This fact is in need of explanation, and in the final part of Section 3.8 an explanatory account is proposed.

Are ideophones a universal category? [Ameka \(2001: 26\)](#) notes that they ‘are first and foremost a type of words – a lexical class of words – which need not belong to the same grammatical word class in a particular language nor across languages.’ Nevertheless, important aspects of their form and function appear to be predictable across languages ([Dingemanse 2017b](#): 196) and it seems in fact that any given language can be expected to dispose of a set of units conforming to the definition to be proposed in Section 3.5.2.

Having been one of the pioneers in research on ideophones in Shona, a Bantu language of Zimbabwe, Fortune (1962: 3) concludes:

[...] the special features of ideophones in Shona lead us to recognize in this language a subsystem of syntactical, morphological, and phonological units and structures, peculiar to ideophones, and characteristic of a style of speech which can loosely be called ‘free expression.’

(Fortune 1962: 3)

The grammatical status of ideophones, more than that of most other types of interactives, has been a challenge to researchers. With reference to its relative degree of grammaticalization (see Section 4.4), that status differs considerably from one language to another. A survey by Dwyer and Moshi (2003) shows that ideophones have been classified as adverbs, adjectives, pseudo adjectives, onomatopoeic adjectives, verbs, derived verbs, predicatives, etc.

Usage of ideophones is presumably more common in spoken than in written discourse. Nevertheless, this is not always the case, as the analysis of ideophones in Korean by Koo and Rhee (2018) suggests. The authors found only a relatively moderate frequency of occurrence in conversations, where only 1.48% of the sentences contained an ideophone. The largest number of ideophones was observed in media essays (19.79%) followed by novels (17.45%), newspapers (7.62%), and movie or drama scenarios (6.91%).

For a distinction between primary and secondary ideophones in English, see Section 3.6.4.1.

### 3.5.2 Working definition

An important line of traditional research can be traced back to Africanists working on Bantu languages in southern Africa (see, e.g., Voeltz and Kilian-Hatz 2001). It was Doke (1935: 118) who volunteered the by now classic definition of ideophones:

A vivid representation of an idea in sound. A word, often onomatopoeic, which describes a predicate, qualitative or adverb in respect to manner, colour, sound, smell, action, state or intensity.

(Doke 1935: 118)

Another aspect of the function of ideophones is highlighted by Kunene (1978: 3) when he says that ‘[a]n ideophone is a dramatization of actions or states.’ For Dingemanse (2019: 18), ‘ideophones are conventionalised words that combine iconic and arbitrary form-meaning mappings.’ Having earlier used a working



definition of ideophones as ‘marked words that depict sensory imagery’ (Dingemanse 2011: 25; 2012: 655), he later proposed the following definition: ‘[...] a canonical ideophone is a member of an open lexical class of marked words that depict sensory imagery’ (Dingemanse 2019: 13).

We use the definition of ideophones below, which rests on the one hand on that proposed by Dingemanse but, on the other hand, also emphasizes that ideophones are interactives, that is, they have the properties listed in Section 1.2 in addition:

- (1) An ideophone is an interactive used for a vivid depiction of sensory imagery of a state, event, object, or quality.

English examples of ideophones include *bang, boing, boom, gaga, helter-skelter, plip plop, pow, splash-splash, swish, ta-dah, thud, tick-tock, vroom, zigzag, and zoom*. Concerning the notion ‘depiction’ in (1) see Section 2.1.3. The phrase ‘state, event, object, or quality’ refers to the fact that ideophones can depict a range of quite different concepts. With regard to the aspectual (*Aktionsart*) categories distinguished by Vendler (1967), ideophones can depict the aspectual contours of accomplishments, achievements, activities, or states (Section 3.5.3.1).

The term ‘ideophone’ might—in accordance with its Greek etymology—suggest that ideophones concern sound (cf. ‘phone’). As a matter of fact, however, ‘sensory imagery’ covers both perceptions of the external world and inner sensations and feelings, it includes in much the same way the sensory domains of sight, touch, hearing, kinaesthesia, taste, color, texture, and inner physiological conditions (Dingemanse 2018: 3; 2019: 15). In Korean, for example, which disposes arguably of the largest number of ideophones in the world (Sohn 1994: 496), the following general types are distinguished with regard to sensory imagery: (a) phonomimes (imitating human, animal, or natural sounds), (b) phenomimes (describing the manner or look of a situation in the external world), and (c) psychomimes (expressing internal feelings, conditions, or sensations). Psychomimes are suggestive of human referents acting as T arguments and in this capacity they may have the appearance of interjections (Section 3.6).

Dwyer and Moshi (2003: 174, (2)) describe ideophones in terms of six properties. A more extended list of properties is proposed by Andrason (2021a: 124–31). Describing them as a prototype, he presents six phonological, five morphological, and five syntactic properties characterizing this type.

Ideophones are frequently treated in tandem with *onomatopoeias* (or ‘onomatopoeics,’ or ‘onomatopes’), that is, forms characterized by a sound symbolic relationship between the depicting and the depicted item, or syntactically isolated forms that ‘phonetically imitate noises’ (Meinard 2015: 151). In many treatments it remains largely unclear what exactly the status of the two is relative to one another. For example, according to Bussmann (1996: 532) an ideophone is ‘an onomatopoeic representation of a concept.’ In other studies again, onomatopoeia is treated not as a feature of ideophones but rather as a separate, even if related class of

forms. Fortune (1962: 5) suggests that ideophones need to be distinguished from onomatopoeias because, unlike the latter, some ideophones do not depict sound but rather the opposite, namely absolute silence. The ideophone *kananaa* of the Siwu language of Ghana for instance means ‘quiet, silent’ (Dingemanse 2011: 43). In accordance with the convention proposed by other authors (e.g., Dingemanse 2018: 3; Haiman 2018; Andrason, Phiri, and Fehn 2021) and adopted here, thus, *kananaa* qualifies as an ideophone but not as an onomatopoeia.

For the purposes of the present work, we follow a number of students of ideophones in assuming that onomatopoeias are located within the semantic space covered by ideophones, and that they form a ‘subset of ideophones imitative of sound’ (Dingemanse 2018: 19): ‘All onomatopes are ideophones, but not all ideophones are onomatopes’ (Haiman 2018: 82). And more specifically, Andrason, Phiri, and Fehn (2021) maintain:

Onomatopoeias occupy the first (lowest) position on the implicational hierarchy of ideophones that leads from ideophones denoting auditory sensations (onomatopoeias) to ideophones denoting visual sensations, and, further, to ideophones denoting psychological and cognitive states.

(Andrason, Phiri, and Fehn 2021: 4; see also Dingemanse 2012: 663)

Observations made in African languages suggest that onomatopoeias form a fairly small sub-set of ideophones (Samarin 1965: 119; Childs 1995: 134). For example, Lébikaza (1999: 448) found only around 7% of ideophones in the Kabiye language of Togo to have an onomatopoeic base.

### 3.5.3 Grammar

#### 3.5.3.1 Argument structure

As was argued in Section 2.2, ideophones focus on a theme (T) rather than on the speaker or the hearer (cf. Meinard 2015: 158 on onomatopoeia). T is described in the relevant literature in a number of different ways. For example, whereas Doke (1935: 118) characterizes it in grammatical terms (‘a predicate, qualificative or adverb’), Kunene (1978: 3) uses general semantic terms (‘actions or states’), and in the definition in Section 3.5.2 the T argument is referred to as a ‘state, event, object, or quality.’

An example of the English ideophone *thud* is presented in (2), and (3a) proposes a paraphrase and (3b) the argument structure of it. There is only one argument, namely T, taking on the form of the state, event, object, or quality describing the circumstances for the usage of the ideophone. Note that ideophones, as is claimed here, have no S argument and in this respect differ from all other types of interactives. However, in some cases the T argument of an ideophone may also refer to

an object consisting of a human being, and in such cases the boundary between ideophones and interjections can be blurred.

(2) English

And the vase went **thud!**

- (3) a Paraphrase: There is something (T) giving the impression of a heavy solid object falling to the ground  
 b Argument structure: *thud* (T)

While ideophones can form independent utterances, they differ, e.g., from interjections in that they need to be ‘supported by other talk’ (Dingemanse 2011: 145), that is, they require what is called here a T argument providing the ‘other talk,’ that is, the conceptual anchor for their usage. What Meinard (2015: 167) observes on onomatopoeias can equally be applied to other kinds of ideophones and the way they differ from interjections: Whereas interjections focus on the subject of conceptualization, ideophones focus on an object of conceptualization, even if that object may also be a human referent. The latter applies, for example, to the Japanese ideophone *ki* ‘being determined and defensive’ (Kita 1997: 404–5) or the Swahili ideophone *bwe bwe* ‘talk nonsense, talk noisily’ (Baldi and Wamitila 2014: 41–57).

Without reference to some kind of T stated or implied, the use of ideophones would normally not be meaningful. Since T need not be expressed, ideophones are not uncommonly used as stand-alones, forming utterances of their own, with T being implied rather than encoded, being recoverable from the preceding context, as in the English example in (4).

(4) English

**Ta-dah!**

The T argument can take on a number of different forms. In (5) from the Bantu language Shona of Zimbabwe, it is the whole sentence preceding the ideophone *páyi* ‘hang up’ that instantiates the T argument, describing the circumstances for the use of the ideophone.

- (5) Shona (Bantu, Niger-Congo; Fortune 1962: 7; italics in the original)

Nd-a-tí            bhátýe rángu **páyi**  
 1.SG-PAST-do jacket my hang.up.IDE  
 ‘I hung up my jacket.’ Lit. ‘I said [to] my jacket: hang up.’

Another widespread form that T can take is that of a word denoting a state, event, object, or quality followed by the ideophone. Thus, in example (6) from the East African language Swahili, the ideophone *tititi* ‘pitch-black’ follows the adjective *mweusi* ‘black,’ which forms its T argument. The way T arguments of ideophones can be instantiated is examined in Section 3.5.3.2.

(6) Swahili (Bantu-Niger; Baldi and Wamitila 2014: 59)<sup>8</sup>

Watu humcheka Oduor kwa kusema kuwa haonekani  
 People keep.laughing Oduor for to.say that he.is.invisible  
 ‘People laugh at Oduor by saying that he cannot be seen  
 gizani kwa kuwa ni mweusi **tititi**.  
 in.darkness because is black pitch.black  
 in darkness because he is very dark/black.’

Beyond their argument structure, ideophones can also be likened to verbal predicates in their internal descriptive content in the way they are structured in relation to time. Based on Vendler (1967), Kita (1997: 404–5) shows that Japanese ideophones (‘mimetics’) can be described with reference to the aspectual features they exhibit. Table 3.8 shows how the Japanese adverbial ideophones examined by Kita can be classified with reference to the aspect (or more precisely, *Aktionsart*) categories distinguished by Vendler (1967), namely accomplishment, achievement, activity, and state (see also Section 4.1.2 on Swahili).

**Table 3.8** Aspectual features of Japanese adverbial ideophones (‘mimetics’)

| Ideophone       | Meaning                                                         | Aspectual feature expressed   |
|-----------------|-----------------------------------------------------------------|-------------------------------|
| <i>bitya</i>    | ‘a liquidy object hitting a flat surface, splashing the liquid’ | Achievement                   |
| <i>Goro</i>     | ‘heavy object rolling once’                                     | Accomplishment or achievement |
| <i>Gorogoro</i> | ‘heavy object rolling twice’                                    | Activity or accomplishment    |
| <i>Gosigosi</i> | ‘rubbing a hard surface, involving friction’                    | Activity                      |
| <i>ki</i>       | ‘being determined and defensive’                                | State                         |

Source: Kita (1997: 404–5)

It would seem in fact that ideophones, conceivably also some other types of interactives, dispose of aspectual contours as maintained by Kita (1997: 404). For example, English ideophones such as *tick-tock* and *vroom* are suggestive of an aspectual feature of activity, or ideophones like *bang* and *thud* of an achievement feature. Similarly, ideophones in Swahili of East Africa express the whole gamut of aspectual distinctions proposed in the tradition of Vendler (1967) (cf. Section 4.1.2, Table 4.1):

|                  |                                 |                |
|------------------|---------------------------------|----------------|
| <b>paa</b>       | ‘burst of something very tight’ | Achievement    |
| <b>nyaku</b>     | ‘take by force’                 | Accomplishment |
| <b>bwe bwe</b>   | ‘talk nonsense, talk noisily’   | Activity       |
| <b>lewa lewa</b> | ‘be totally drunk’              | State          |

<sup>8</sup> Glosses are ours, no glosses are provided by the authors.

Interestingly, the Swahili ideophones *paa* and *nyaku* express temporally bounded events whereas the reduplicated ideophones *bwe bwe* and *lewa lewa* express unbounded events.

Ideophones are likely to reduplicate—to the extent that some authors take the presence of a reduplicated form as evidence to identify it as an ideophone. However, in some languages there is a marked tendency to avoid reduplication when an ideophone expresses an achievement or accomplishment. For example, in the Sepik language Iatmul of Papua New Guinea, ideophones are almost invariably reduplicated, but they are not reduplicated when the ideophone expresses a punctual event (Jendraschek 2012: 315).

### 3.5.3.2 Three main processing constructions and placement

Depending on their exact function, ideophones may appear in various slots of a sentence, or they may form utterances of their own, but most commonly they have been said to be placed utterance-finally or at the right periphery of an utterance. In this respect they differ from most other types of interactives, whose preferred position is at the left periphery of an utterance. For example, in the Kivunjo language of northeastern Tanzania ideophones are found sentence-finally, and in Cambodian, it seems that ideophones also occur most frequently as final elements of the sentence (Dwyer and Moshi 2003). Similarly, in the Kisi language of Guinea, Sierra Leone, and Liberia (Childs 1995: 137) and in the Siwu language of southeastern Ghana they typically occur in sentence-final position (see Haiman 2018: 111–12 for more details).

This, however, is not the situation to be found in all languages. Studying ideophones in the Baka language of southern Cameroon, Kilian-Hatz (2001: 162) found that they may occupy any position in the sentence, and in Japanese, ideophones usually precede the main verb (Kita 1997).<sup>9</sup>

It would seem that placement of ideophones is most of all a derivative of the discourse constructions employed for their presentation. Our observations suggest that it is especially the constructions in Table 3.9 that determine the way ideophones are presented in discourse. These observations are based on findings made in a number of genetically unrelated languages spoken in various regions of the world. But the extent to which they apply generally across languages is an issue that is in need of further research. It would seem, for example, that the situation in Japanese is more complex (see Dingemanse and Akita 2017 for details).

The three constructions in Table 3.9 can be illustrated with the English examples in (7), where (7a) illustrates the free construction, (7b) is an example of the quotative construction, and (7c) of the modifier construction.

<sup>9</sup> The situation in Japanese, however, is not really telling since ideophones are fairly strongly grammaticalized (see Section 4.4) and, hence, are not characteristic of paradigm uses of ideophones across languages.

**Table 3.9** The main discourse constructions for presenting ideophones

| Construction           | Discourse status of ideophone | Marking                                                     | Placement                              |
|------------------------|-------------------------------|-------------------------------------------------------------|----------------------------------------|
| Free construction      | Independent utterance         | No marking                                                  | No constraints                         |
| Quotative construction | Discourse complement          | Verb for ‘say,’ ‘do,’ ‘go,’ etc., and/or a quotative marker | In the slot reserved for direct speech |
| Modifier construction  | Appositional parenthetical    | No marking                                                  | After the text piece they refer to     |

## (7) English

a **Thud!**b And the vase went **thud!**c The vase fell to the ground – **thud!**

The free construction is of low frequency in the texts examined by us. In this construction, the ideophone is used as an independent utterance of its own, that is, its T argument is not realized overtly, even if the ideophone may be preceded and/or followed by other discourse segments. Thus, in example (8) from the Khwe language of northern Namibia, the ideophone *bwww* makes an independent contribution to the discourse, even if its descriptive content relates to both the preceding and the following utterances. But even when forming an independent utterance, ideophones in the free construction tend to relate to what follows and, hence, they may have the appearance of being placed initially relative to what follows.

## (8) Khwe (Khoe, ‘Central Khoisan’; Kilian-Hatz 2001, (14))

Áta kx’á í nà-m àni-mà? **Bwww**, Ci nò  
 thus vulture which-3.M.SG side-3.M.SG *bwww*.IDE Arrive and  
 cì-llà-à-xú.

NEM-land-JII-CPL

‘This vulture – in which direction [can he fly]? *Bwww* [he tries to take off by moving his wings], however [he fails again and again], he lands immediately.’

In the quotative construction, the ideophone is presented as a discourse complement of its T argument, which forms the main clause containing a quotative index (Güldemann 2008).<sup>10</sup> The latter frequently takes on the form of a verb for ‘say,’ ‘do,’ ‘go’ or of other verbs whose only or main function it is to introduce direct

<sup>10</sup> It is important to note that a discourse complement is not a syntactic complement, it serves the organization of texts rather than that of sentences and may consist of a word, an utterance, or more than one utterance, or even a contribution made by an interlocutor other than the speaker. A ‘quotative index’ is a linguistic expression which signals reported speech (Güldemann 2008: 1).

speech. English speakers frequently use *be like*, or *go* as a quotative index instead, cf. (7b), and the ideophone is presented like a direct speech utterance. In (7b), the T argument of the ideophone *thud* is instantiated by the preceding predication (*And the vase went*) introducing the ideophone via direct speech. The verb *went* can be interpreted simultaneously as the main verb and as a quotative marker introducing the ideophone.

Alternatively, the quotative index consists of a function-specific quotative marker. In Japanese, for example, the latter is instantiated by the quotative particle *-to* or its colloquial counterpart *-te*, e.g., *bakibaki-to* ‘with a cracking sound’ (Dingemanse and Akita 2017: 514).

Being presented as a discourse complement, the ideophone cannot be omitted, its use is obligatory in the quotative construction. Güldemann (2008: 212) found eighteen out of the twenty African languages surveyed by him to use the quotative construction for presenting ideophones.

The modifier construction is of a different kind, as (7c) shows: The ideophone follows its T argument, being presented as a kind of afterthought or appositional parenthetical, that is, it is an omissible text segment having the appearance of a discourse modifier of its T argument. The modifier construction can be illustrated also with the following example from the Siwu language of southeastern Ghana, where the T argument of the ideophone *fututu-tutututu* ‘pure white’ appears to consist of the semantically related verb *-fudza-* ‘be white,’ and the two form a common collocation pattern (Dingemanse 2017a: 366–7). Over time, the ideophone may gradually grammaticalize into a syntactic modifier of its preceding T argument, assuming features of an adjectival or adverbial, as has happened frequently in Siwu—in fact, (9) illustrates an instance of early grammaticalization, as we will see in Section 4.4.1.

- (9) Siwu (Kwa, Niger-Congo; Dingemanse 2017a: 366, (1))<sup>11</sup>  
 i-tì            si i-fudza-ɔ                    ↑*fututu~tutututu*↑  
 NC1-head   if S.NC1-be.white-SG.O   IDE.pure.white~EMPH  
 ‘That your head may become white *fututu-tutututu* [pure white].’

Typically, interactives are semantically independent from the utterance they are associated with. This, however, does not apply in the case of the modifier construction, where T must be semantically compatible with the meaning of the ideophone—in their grammar of Southern Sotho, Doke and Mofokeng ([1957] 1985: 337) characterize the function of the ideophone in this construction ‘as an intensifier or cognate descriptive.’ Accordingly, Dwyer and Moshi (2003: 176) note that the English utterance #*The pillow slipped to the floor, pow!*, instantiating the modifier construction, is not well-formed because the T argument *The pillow*

<sup>11</sup> Arrows (↑) mark the beginning and the end of prosodic foregrounding, such as markedly high pitch.

*slipped to the floor* and the ideophone *pow* are not semantically compatible with one another.

The modifier construction thus differs from the other two constructions in imposing remarkable co-occurrence constraints on the usage of ideophones. Japanese has a plethora of selection restrictions on which kind of T argument ('semantic role') an ideophone ('mimetic') is allowed to take in this construction. For example, the adverbial ideophone *gorogoro* describes an event involving continuous rolling of a heavy object. It imposes the following selection restriction on the kind of T arguments it can take: The motion (a) must involve rotation, (b) the event described must be iterative, and (c) the rolling object must be heavy, that is, it could be an iron ball but not a pearl (Kita 1997: 402–3).

Not uncommonly, usage of ideophones based on the modifier construction is restricted to very few, or even to only one T argument token, that is, the two form fixed collocations. The following example from the West African language Hausa seems to be characteristic of similar situations in other languages: The ideophone *fat* 'pure white' co-occurs only with the adjective *farī* (*farā*) 'white,' hence *farā fat* 'snow white' is a fixed collocation (Newman 2000: 253).

The difference between the three main constructions is also reflected in their grammaticalization behavior. On the basis of the evidence that has become available so far it can be assumed that the modifier construction differs from the other constructions in being most likely to develop in the direction of sentence grammar (Section 4.4).

The distinction between the quotative and the modifier construction roughly corresponds to that made by Dingemanse and Akita (2017) between the quotative and the collocational construction in Japanese. And it also corresponds to the 'lexical' distinction predicative vs. intensifying ideophones made by Amha (2001: 253–5) in the Omotic language Maale of Ethiopia. Predicative ideophones, instantiating the quotative construction, take mainly the verb *geʔ-* 'say' as a quotative index and *geʔ-* follows the ideophone—in accordance with the verb-final syntax of Maale, cf. (10). Unlike predicative ideophones, intensifying ideophones, being instances of the modifier construction, are few in number, they co-occur with a preceding adjective and express intensity of color, size, shape, etc.

- (10) Amharic (Semitic, Afroasiatic; Amha 2001: 255)

**hírc'** geʔ-í hant-é  
 IDE say-CNV walk-2.SG:IMP  
 'Be alert on the way!'

Such observations suggest, first, that the way ideophones are presented in discourse is contingent on which of the constructions in Table 3.9 is involved. And second, they suggest that the preference for placing ideophones utterance finally is not due to some general discourse constraint but is construction-specific—and, hence, this preference can be, and in fact frequently is contradicted.



With regard to the quotative construction in (7a), for example, we would expect that in languages having a rigid verb-final syntax (SOV), ideophones do not appear finally since direct-speech utterances function as discourse complements: In such languages, the complement precedes its head and, being discourse complements, ideophones are likely to precede the clause introducing them. Accordingly, in the following example from the verb-final language Turkish, the complement expressing direct speech precedes the ‘say’-clause introducing it, cf. (11a), and so does the ideophone *zmk* in (11b): It also precedes the verb of saying *diye* introducing it, thus occurring utterance initially rather than finally.

- (11) Turkish (Jendraschek 2001: 92–5; Haiman 2018: 81)
- a ‘Kue harfi laazım’ diye tuttur-du  
 Q letter necessary saying mutter-PAST.3.SG  
 ‘“We need the letter ‘q,” he muttered’
- b ..**zmk** diye dur-du  
 stop.IDE saying stand-PAST.3.SG  
 ‘(the car) stopped abruptly’

With regard to the modifier construction, ideophones occur after the text piece they refer to—hence, their position is usually utterance final. Thus, in the following example from Kivunjo of northern Tanzania, the ideophone *katsu* must occur after the T clause it refers to.

- (12) Kivunjo (Chaga dialect, Bantu, Niger-Congo; Dwyer and Moshi 2003)
- N-á-lé-fúnjá ki-zri **katsu**  
 FOC-3-PAST-chase NC7-stick crack.IDE  
 ‘S/he broke the stick with a crack.’

In accordance with these constraints associated with the two constructions, placement of ideophones in the Maale language of Ethiopia is also predictable (Amha 2001: 253–5): As was just observed, predicative ideophones, based on the quotative construction, usually take the verb *ge?*- ‘say’ as a quotative index and since Maale is a verb-final language, the ideophone *precedes* the verb *ge?*-, cf. (10). Intensifying ideophones, by contrast, involve the modifier construction and, in accordance with this construction, *follow* their T argument, that is, the adjective which they moderate. Note that, like in Turkish, the syntactic order in Maale is normally modifier-modified, but since the modifier construction is based on principles of discourse organization, this syntactic constraint is ignored.

The quotative construction is the most common one in some languages while in other languages it is the modifier construction that is predominant. The former applies to Maale, while the latter is the case, for example, in the Siwu language of Ghana. According to Dingemans (2011: 145–8; 2017a: 366–70; 2017b), five morphosyntactic constructions involving ideophones need to be distinguished in

Siwu, namely the ones in [Table 3.10](#). On the perspective proposed here, the criteria used by the author for classification consist, on the one hand, of the discourse construction employed, and on the other hand, of the relative degree of grammaticalization. As we will see in Section 4.4.1, Siwu has grammaticalized the modifier construction to various degrees, where each degree is manifested in the form of a distinct construction (the Roman figures in parentheses refer to relative degrees of grammaticalization described in Section 4.4.1, where II = weakly grammaticalized and IV = most strongly grammaticalized).

**Table 3.10** Ideophone constructions in Siwu

| Construction<br>( <a href="#">Dingemanse 2011</a> ) | Number of<br>tokens | Percentage  | Discourse construction                                                  |
|-----------------------------------------------------|---------------------|-------------|-------------------------------------------------------------------------|
| Adverbial                                           | 101                 | 46%         | Modifier construction, slightly grammaticalized (Stage II)              |
| Complement                                          | 56                  | 26%         | Quotative construction                                                  |
| Holophrase                                          | 27                  | 12%         | Free construction, or ungrammaticalized modifier construction (Stage I) |
| Adjectival                                          | 13                  | 6%          | Modifier construction, moderately grammaticalized (Stage III)           |
| Predicative                                         | 11                  | 5%          | Modifier construction, strongly grammaticalized (Stage IV)              |
| Others                                              | 11                  | 5%          |                                                                         |
| <i>Total</i>                                        | <i>219</i>          | <i>100%</i> |                                                                         |

Source: Based on [Dingemanse \(2011: 145–8; 2017a: 366–70; 2017b\)](#)

Ignoring the category ‘Others’,<sup>12</sup> the findings presented in [Table 3.10](#) can be summarized as follows. According to the examples given by [Dingemanse \(2017a\)](#), the free construction is ambiguous, it can as well be interpreted as a non-grammaticalized modifier construction. Accordingly, the modifier construction is clearly predominant, accounting for more than two-thirds (69%) of all tokens of use, whereas only every fourth use of ideophones in Siwu appears to involve the quotative construction (26%). Thus, as we will see in our diachronic account in Section 4.4.1, the Siwu language represents what may be called a ‘mixed’ situation, where ‘pure’ manifestations of the two constructions are represented by what [Dingemanse \(2011\)](#) calls the holophrase (= free construction) and the complement construction (= quotative construction).

<sup>12</sup> The category ‘Others’ ‘lumps together a small number of cases of ideophones occurring in syntactic environments that do not slot neatly into the five types described above’ ([Dingemanse 2017a: 370](#)).

### 3.5.3.3 Phonological features

In more general terms, [Dingemanse \(2017a\)](#) writes about phonological features of ideophones:

Cross-linguistically, some recurrent ways in which ideophones are structurally marked include skewed phonotactic distributions, feature harmony, more possible syllable structures, special word forms, susceptibility to expressive morphology, relative syntactic independence, and in actual use, foregrounded prosody.

([Dingemanse 2017a](#): 365)

Observations on the phonological behavior of ideophones suggest, on the one hand, that they are somehow unusual, as noted by [Childs \(1995\)](#):

Ideophones have no morphology and little syntax. Their phonology is unusual, characterized both by sounds found nowhere else in the language and by violations of phonological constraints.

([Childs 1995](#): 132)

On the other hand, there are also findings suggesting that ideophones are not generally phonologically ‘unusual.’ For example, [Diffloth \(1994: 109–10\)](#) writes the following on ideophones (‘expressives’):

It is important to notice that, in most languages, expressives use only phonic material which is already given by the phonological system of the language. This is true not only in terms of inventory, but also, for the most part, in terms of distribution. Expressives are phonologically well-formed words of the language, and when there are unusual sound combinations, the phonological violations appear limited, and clearly iconic.

([Diffloth 1994](#): 109–10)

Thus, it seems that at least in some languages, ideophones do not differ dramatically in their phonotactic or phonetic form from other linguistic forms. For example, [Newman \(2000: 242\)](#) notes that ideophones in the West African language Hausa ‘generally employ the same inventory of consonants and vowels as non-ideophonic words, i.e., while they are phonologically aberrant in some respects, they are not wildly so.’ And [Kita \(1997: 382\)](#) asserts that no ‘exotic’ phonetic material is found in Mon-Khmer ideophones (‘expressives’): ‘They seem only to elaborate on the existing sounds,’ and with reference to Japanese ideophones (‘mimetics’) he notes:<sup>13</sup>

<sup>13</sup> An anonymous reviewer of Oxford University Press (February 3, 2022), however, observes: ‘While it is true that all the phonemes of mimetics can be found in the phonetic inventory of Japanese, mimetics violate important phonotactic rules of Japanese. One is that the phoneme “p” cannot appear word

What is important is the fact that mimetics are not phonologically eccentric. All the phonemes and their combinations found in mimetics are found elsewhere in the Japanese lexicon.

(Kita 1997: 382)

In spite of such contrasting observations, there can be little doubt that, at least in some languages, ideophones exhibit phonological features that are absent elsewhere in the language. For example, words in the Maale language of Ethiopia end in a vowel, whereas ideophones may end not only in a single consonant but even in a geminated consonant (Amha 2001: 254), and exactly the same applies to ideophones in the Northern Sotho language of South Africa (Poulos and Louwrens 1994: 352–3). Examples of this kind of contrast in fact abound in the languages of the world.

More than other types of interactives, ideophones exhibit an extraordinary amount of phonotactic manipulation. One of the characteristics of ideophones can be seen in the phonetic prolongation of forms to express higher degrees of intensity of meaning. For example, the ideophone *mwí:* ‘quiet’ of the Shona language of Zimbabwe can form an intensive variant by ‘being higher in pitch and much longer in final syllabic length’ (Fortune 1962: 33). Or ideophones in the Supyire language of Mali, Ivory Coast, and Burkina Faso ‘feature ad lib iconically motivated prolongation of their vowel’ (Haiman 2018: 90). This characteristic is by no means restricted to ideophones, neither in Shona nor in Supyire, but it is distinctly more common and more pronounced in ideophones than in other linguistic expressions (see especially Dingemanse 2017a).

Another morphophonological feature that is mentioned in some studies concerns the length of ideophone forms—a feature that does not appear to be typologically entirely consistent. In the Siwu language of eastern Ghana, ideophones are on average longer than nouns and verbs, and this constitutes one of the most reliable cues to ideophone status in Siwu (Dingemanse 2017a: 365). In fact, excessively long interjections seem to be a characteristic of some West African languages. Take the example of the ideophone *sinisinisinisini* (it holds itself together) ‘closely woven’ of the Siwu language, or that of *jededededededede* of the Koromfe language of northern Burkina Faso, denoting the animals’ footsteps during a tug-of-war: It consists of nine syllables, spoken in fast succession and accelerating towards the end (Rennison 1996: 472):

initially in ordinary native Japanese words while it is a very common word initial phoneme in mimetics (e.g. *pika-pika* ‘glittering/shiny’). Also, the glottal stop is non-functional in ordinary words, but appears frequently at the end of a mimetic to register a sudden stop of an action or event (e.g. *pika?* ‘a flicker of light’). However, it is true that Japanese mimetics participate fully in sentence grammar. Most conventional mimetics can be used as adverbs, nouns and adjective, and some have become a full-fledged verb (e.g. *pika?* ‘a flicker of light’ > *hikar-u* ‘to shine’ – note the use of “h” in this word, not “p” which violates a phonotactic rule.)

- (13) Koromfe (Gur, Niger-Congo; Rennison 1996: 472)

a jemdi tik(ɪ) a tife  
 ART hippopotamus.SG do.DUR ART elephant.SG

**jededededededede.**

IDE

‘The elephant also did the same to the hippopotamus: *jededededededede*.’

And extremely long ideophones are not restricted to West African languages, they can also be found in some other languages. The mechanism responsible for the presence of such long forms is almost invariably repetition, be that full or partial repetition. Speakers of the Tjwao language of western Zimbabwe employ the ideophones (‘onomatopoeias’) in (14) to mimic experiences they commonly make in their environment, where each of the three forms consists of eight syllables.

- (14) Some ideophones (‘onomatopoeias’) of Tjwao (Khoe; Andrason, Phiri, and Fehn 2021; [ɿ] = alveolar click, [ɿ̤] = dental click)

*Form*

*Phenomenon depicted*

**!ha!ha!ha!ha!ha!ha!ha!ha** ‘lightning’

**!hu!hu!hu!hu!hu!hu!hu!hu** ‘joint thunder and lightning’

**|ha|ha|ha|ha|ha|ha|ha|ha** ‘sparking fire’

In other languages, the situation is quite different. With regard to the Turkana language of northwestern Kenya, Dimmendaal (1983: 381) writes that ‘it should be noted that few words in Turkana are mono-syllabic, whereas almost all ideophones in the language are. Elsewhere there is a tendency to avoid mono-syllabic units.’ In the Bantu language Shona of Zimbabwe, genetically not related to Turkana, ideophones are also for the most part monosyllabic while words of sentence grammar are generally at least disyllabic (Fortune 1962: 29). Likewise, in the Northern Sotho language of South Africa, monosyllabic words are rarely found, but a large number of ideophones consist of one syllable only (Poulos and Louwrens 1994: 352). And Swahili of East Africa has a constraint on syllable structure in that words are normally at least disyllabic. Many ideophones, by contrast, are monosyllabic (Baldi and Wamitila 2014: 41).

Another area where more cross-linguistic research is needed concerns the tonological contours of ideophones. Observations made by Haiman (2018: 301) suggest that in tone languages, ideophones exhibit a more reduced format compared to the tonological structure to be observed in sentence grammar. For example, in the Diding’a language of Southern Sudan, ideophones have almost always low tone, while in Mundang, spoken in the border area of Cameroon and Chad, they tolerate only low tones or only high tones throughout.

### 3.5.3.4 On the role of gesture

Gesture, that is, movement of the body and/or any part of it, is an important feature of many types of interactives, and it seems that ideophones are particularly sensitive to it. Thus, [Dingemanse and Akita \(2017: 519\)](#) observe that in the production of ideophones, speech, and gesture, while operating in different modalities, perform the same role of depicting sensory imagery. For example, they found in a Japanese corpus of cartoon retellings as many as 94% of ideophones co-occurring with an iconic gesture. In the Bulu language of Cameroon, ideophones are not only set off by a pause and preceded by the particle *ne*, they can be preceded by a snap of the fingers or a click ([Childs 1994: 196](#)).

In Section 1.3.2.3, a voice-gesture model was proposed, consisting of a scale extending from vocal form (A) to gesture (C)—that is, an interactive concept can be expressed either by a linguistic form (A), a combination of form and gesture (B), or only by a gesture (C). This may mean that stage B of the model represents the default use of an ideophone, and it has some disambiguating potential. [Dingemanse \(2011: 220\)](#) provides the following example from the Siwu language of Ghana to illustrate this point. The ideophones *giligili* and *minimini* are similar in meaning, both expressing some kind of roundness. To distinguish the meanings of the two, speakers may rely on gestures, drawing a circle with their index finger for *giligili*, while in the case of *minimini* they produce a two-handed gesture depicting a sphere. The gestures in fact provide an optimal means of distinguishing the two, in that *giligili* depicts two-dimensional, circular roundness and *minimini* three-dimensional, spherical roundness.

But do gestures constitute an ancillary element in the usage of ideophones—or vice versa? Obviously, an answer to this question is contingent on the perspective one may wish to adopt, and it is also in need of more empirical input. [Dingemanse and Akita \(2017: 519\)](#) assert that gestures occurring with ideophones coincide temporally and the authors add:

[...] speech and gesture are not loosely aligned and complementary, but tightly coupled and alike in mode of representation: they perform the same role of depicting sensory imagery, albeit in different modalities and therefore also with different affordances for iconicity.

([Dingemanse and Akita 2017: 519](#))

However, there are also reasons to assume that there is some degree of conceptual independence between gesture and ideophone. While the two are in fact tightly coupled, they appear to be conceptually separate from one another to some degree. For one thing, ideophones can occur without gestures and vice versa—in accordance with the voice-gesture model in Section 1.3.2.3. And for another, the two do not seem to necessarily coincide in time.

Take the following example. Observations on ideophones in the languages of the world suggest that stench belongs to the concepts especially widely distinguished in some form or other. One of the ideophones for this concept available to

speakers of the Northern Sotho language in South Africa is *hôtôtô*, translated as 'smelling horribly'.<sup>14</sup> The gesture expressing this concept is described by Poulos and Louwrens (1994: 346) thus: 'The speaker lifts both hands to face level and then moves them up and down continuously with palms directed downwards or towards the face to supply fresh air. At the same time the lips are lifted upwards to narrow the stricture of air getting into the nostrils.'

That the vocal and the gestural representation of stench in Northern Sotho are conceptually independent from one another is suggested by the following: First, as just observed, each of the two can be produced in the absence of the other—in accordance with stages A and C of the voice-gesture model (Section 1.3.2.3). Actually, the Northern Sotho gesture is distinctly more likely to be presented on its own without the ideophone, the reason given being the following: Unlike the ideophone, the gesture can be produced without the polluted air entering the mouth of the speaker. Second, production of the two kinds of expression need not coincide in time: An ideophone can and frequently is reduplicated, nevertheless the time span employed for its production is limited. This is somewhat different with gestures, which can easily be extended for a larger period of time. With the expression of stench signaled by *hôtôtô* in Northern Sotho this means that the speaker may extend the use of the gesture until the nasty smell subsides, way after the ideophone had been presented (Poulos and Louwrens 1994: 346).

Irrespective of how such observations are to be interpreted, they suggest—pending further analysis—that both the two modalities of expression and their interrelationship be each treated best on its own.

### 3.5.3.5 Stringing ideophones

According to Childs (1995: 132), '[i]deophones perform expressive rather than purely informational functions.' Like other interactives, ideophones have in fact expressive meaning, but as is argued here, they also have descriptive meaning content. That they have descriptive meaning can be shown, on the one hand, by the fact that their meaning can be paraphrased (Section 2.1). On the other hand, this is also suggested by the fact that they can be stringed into sequences to describe short narrative events in some languages. We saw already an example from Baka in Section 1.1; the following are examples from other African languages. (15) is taken from the Shona language of Zimbabwe (Fortune 1962: 17), where there is another 'mini-narrative' consisting of what appears to be a discourse marker (*iko*) plus a string of three ideophones (*símu*, *shápu*, and *tóro*).

- (15) Shona (Bantu, Niger-Congo; Fortune 1962: 17)  
**íko símu, shápu, tóro**  
 then get.up.IDE go.off.saucily.IDE disappear.IDE  
 'Then he got up, went off saucily and disappeared.'

<sup>14</sup> There are two other ideophones in Northern Sotho having the same meaning as *hôtôtô*, namely *phuu* and *fong* (Poulos and Louwrens 1994: 354).

The ‘mini-narrative’ in (16) comes from Hamar of southern Ethiopia, a language unrelated to Shona, exhibiting essentially the same structure: Following what appears to be a discourse marker (*länsi*) there is a sequence of three ideophones (*zäp*, *torsh*, and *lïmm*).

- (16) Hamar (Omotic, Afroasiatic; [Lydall 2000](#))

**länsi zäp torsh lïmm**  
 then take.hold.IDE spear.IDE life.extinguished.IDE  
 ‘Then (he) grabbed it, speared it, and (it) died.’

The examples just presented illustrate another noteworthy feature of ideophones, namely that they appear to dispose of what one might call an elementary asyndetic ‘coordinating syntax,’ allowing interlocutors to combine ideophones to express sequences of events. Another example is the following from the Siwu language of eastern Ghana, where the combination of two ideophones expresses the causal sequence of hitting someone with high intensity (*teteree*) so that it produces the ‘sound of dry hit’ (*kpɔ*) ([Dingemanse 2011](#): 241–2).

- (17) Siwu (Kwa, Niger-Congo; [Dingemanse 2011](#): 242, (128))

**lowĩ-ũ teteree, kpɔ!**  
 1.SG.hit-3.SG.O IDE.intense IDE.impact  
 ‘I hit him *teteree* (and it sounded) *kpɔ*!’

A similar example of ‘coordinating syntax’ can be seen in the following example from Arusa Maasai, a Nilotic language of northern Tanzania: There is a sequence of the two coordinated ideophones *firtit* ‘slipping’ and *buss* ‘thudding sound made when falling’ being in a causal relation, introduced by the quotative form *ne-jo*, and the two ideophones each constitutes an utterance of its own. Note that this example seems to illustrate the fairly common aspectual contour pattern activity-accomplishment, in that *firtit* signals an activity and *buss* a subsequent accomplishment (Section 3.5.3.1).

- (18) Arusa Maasai (Eastern Maasai, Nilo-Saharan; [Karani and Andrasen 2021](#), (5))

**ne-jo firtit buss.**  
 SUB-say IDE.slipping IDE.falling  
 ‘He slipped and fell down.’

Our final example (19) is taken from a song in the Gola language of Liberia, where a sequence of three units used as ideophones depicts the situation after the birth of a child: *bunyā* describes a person who has passed the time in a secret society on the way to becoming an adult, *ja* stands for something small, typically a child, and *kwe kweye kwe* is an expression signaling extreme delight ([Koroma 1994](#): 68, (76)). The text piece can therefore be translated approximately as in (19).



- (19) Gola (Atlantic, Niger-Congo; [Koroma 1994](#): 68, (76))

**Bunyã ja kwε kwεyε kwε**

‘We are excited about the arrival of the newborn child!’

Note, however, that such instances of ‘coordinating syntax’ are not restricted to ideophones; rather, they can also involve other types of interactives. Thus, in the following example from the Tswana language of southern Africa, there is a sequence of three interactives, namely the social formula *tšwââtšwââ* ‘please,’ the vocative *rra* ‘o father!’, and the directive *nthusa!* ‘help me!’

- (20) Tswana (Bantu, Niger-Congo; [Cole \[1955\] 1987](#): 400)

**Tšwêêtšwêê, rra, nthusa!**

‘Please, father, help me!’

### 3.5.4 Ideophones vs. interjections

Ideophones have been classified as interjections and interjections as ideophones, and in fact, the boundary between the two is far from clear. Different views have been voiced on this issue (see, e.g., [Dingemanse 2011](#): 155–6; [Haiman 2018](#): 134ff.). It would seem that the features in (21) are among the less controversial ones separating the two.

- (21) Features distinguishing ideophones from interjections

- a Interjections commonly involve emotive states, ideophones by contrast do not seem to necessarily involve emotions.
- b Whereas interjections typically have the speaker (S) as their only argument, ideophones have a T argument (T), referring to some state, event, object, or quality (Section 2.2.3.2).
- c When combined with a sentence, interjections most likely will precede the sentence. Ideophones, by contrast, most likely follow the sentence (cf. [Dingemanse 2011](#): 35; [Haiman 2018](#): 105; but see also Section 3.5.3.2).
- d Interjections do not commonly reduplicate whereas ideophones quite commonly do reduplicate (Section 3.12).
- e Interjections have cross-linguistically a limited membership. In contrast, ideophones may have a large membership, in some languages running in the hundreds or even thousands (Section 3.5.1).
- f Whereas interjections participate in grammaticalization processes within the category of interactives, ideophones commonly grammaticalize in the direction of sentence grammar (Sections 4.2, 4.4).
- g Interjections tend to include forms consisting of a click only. Such an observation has so far not been made in ideophones.<sup>15</sup>

<sup>15</sup> However, in the Bulu language of Cameroon, the production of ideophones can be preceded by a snap of the fingers or a click ([Childs 1994](#): 196).

Note, however, that ideophones are said to depict sensory imagery, and sensory imagery covers not only perceptions of the external world but also inner sensations and feelings (Dingemanse 2018: 3; 2019: 15; see Section 3.5.2). In the latter capacity ideophones refer to states of human beings and this also applies to interjections (Section 3.6). There is thus an overlap area between the two types of interactives and the boundary between the two is not entirely clear in such cases, as pointed out by Haiman (2018). We will return to this issue in Section 3.5.7, (h).

In addition to the features in (21), a number of other features have been mentioned, but most of them are of a more controversial nature (Section 3.5.5). This applies in particular to the question of whether ideophones are deictics (see also Section 1.3.1.5). This question will be taken up in Section 3.5.7.

For the present study, which is mainly concerned with the commonalities shared by the various types of interactives, distinguishing ideophones from interjections is not a big issue since both are part of the interlocking network of interactives. Nevertheless, distinguishing the two types is also essential for understanding the nature of interactives.

### 3.5.5 Grammaticalization

Ideophones differ from other types of interactives most of all in the diversity of grammatical structures that they exhibit. The reason for this is mainly due to their propensity to grammaticalize in the direction of sentence grammar. Accordingly, Dwyer and Moshi (2003) distinguish between primary and grammaticalized ideophones,<sup>16</sup> and they summarize their findings thus:

[...] there has been considerable variability in the grammatical characterization [of] ideophones: some claiming that ideophones do not fit any known syntactic category and others claiming that ideophones are adverbs, adjectives, verbs and so forth. While allowing for some difference of opinion [...] we would in general attribute the first category to observations made about primary ideophones and the latter category to observations made about grammaticalized ideophones. The differing categories to which ideophones are assigned are due to the fact that languages can grammaticalize ideophones, as they do any linguistic material, differently, according [to] the resources available within the language.

(Dwyer and Moshi 2003: 181)

This issue will be the subject of Section 4.4.

<sup>16</sup> The adjective ‘primary’ must not be confused with its usage elsewhere in this study as described in Section 3.6.4.1, where it figures in the distinction ‘primary interactive’ (e.g., *wow*) vs. ‘secondary interactive’ (e.g., *hell*).

3.5.6 A note on sound symbolism

Of all types of interactives, ideophones stand out with reference to the magnitude and quality of sound symbolic behavior they exhibit. This is a complex issue—one that would need a separate treatment. One salient kind of sound symbolic feature is pointed out by [Haiman \(2018: 118\)](#) when he argues that ideophones are characterized by ‘sound iconism’ according to which they tend to be characterized by the contrasts depicted in [Table 3.11](#).

**Table 3.11** Correlations in sound symbolism  
observed in ideophones

| Meaning:     | Small size          | Large size       |
|--------------|---------------------|------------------|
| <i>Form:</i> | High tone           | Low tone         |
|              | High front vowel    | Back vowel       |
|              | Voiceless consonant | Voiced consonant |

Source: [Haiman \(2018: 118\)](#)

The correlations summarized in [Table 3.11](#) build on the seminal work by [Westermann \(1927, 1937\)](#) on West African languages, who had proposed features such as the ones in [Table 3.12](#) to be observed in ideophones and some other forms (see [Dingemanse 2018: 8](#), Table 2; see also [Sohn 1994: 496–502](#); [1999: 96–102](#) on Korean). In the Maale language of Ethiopia, ideophones with high vowels (*i* or *u*) are often associated with lightness, smallness, or fast speed whereas those with non-high vowels (*e* or *o*) are associated with heaviness, bigness, or slowness ([Amha 2001: 255](#)).

**Table 3.12** Iconic sound–meaning distinctions in West  
African languages

|   |                        |                                    |
|---|------------------------|------------------------------------|
| a | Low tone:              | big, bulky, dull, wide             |
|   | High tone:             | small, tender, tight, tense        |
| b | Back (‘dark’) vowel:   | heavy, thick, plump, bloated, dark |
|   | Front (‘light’) vowel: | light, thin, tall, fine, bright    |
| c | Voiced consonant:      | heavy, soft                        |
|   | Voiceless consonant:   | light, hard                        |

Source: [Westermann \(1927, 1937\)](#)

Sound symbolism in fact seems to be a common phenomenon in ideophones; for a detailed analysis of Korean, see [Rhee \(2019\)](#). At the same time, correlations between ideophones and tone pitch do not seem to be cross-linguistically entirely

consistent. One feature observed in some West African languages is that ideophones tend to be strongly associated with high tone syllables. In the East African language Diding'a of the Southern Sudan, by contrast, the tone melody of an ideophone is almost always low (see [Haiman 2018](#): 115). In Logba of southeastern Ghana, ideophones show a tonal distinction expressing the following contrast: Low tones are associated with bad, unpleasant or amorphous features and high tones with nice, pleasant, small and cute features ([Dorvlo 2008](#): 241).

Haiman concludes that 'ideophones tend to exhibit greater freedom in their segmental inventory and phonotactics, lesser freedom in their suprasegmental melody, than is exhibited in prosaic vocabulary,' that is, in the vocabulary of sentence grammar. An interesting example of segmental sound symbolism in ideophones ('expressives') is provided by [Diffloth \(1994\)](#) from the Mon-Khmer language Bahnar of Vietnam, where the distinction high vowel vs. low vowel correlates productively with the conceptual distinction big size vs. small size and the author draws attention to the claim made by some earlier authors according to which the expected correlation in English or other languages would be the opposite, namely high vowel = small size and low vowel (*a*) = big size ([Diffloth 1994](#): 112), see also [Table 3.11](#).

Another phonetic characteristic of ideophones pointed out by a number of authors consists in the prolongation of forms to express higher degrees of intensity of meaning (see [Haiman 2018](#): 117 for discussion; Section 3.5.3.3). This feature is by no means restricted to ideophones, even if it is distinctly more common and more pronounced in ideophones than in other linguistic expressions. In a similar way, sound symbolism as it surfaces in the vowel polarity structure of Korean in (22) is not restricted to ideophones.

(22) Vowel polarity in Korean ([Rhee 2019](#): 112, (3))

'Positive' (or light) vowels: /a/, /o/, /æ/, and diphthongs involving them.

They describe small, delicate, and bright objects, and movements arousing such sensation

'Negative' (or dark) vowels: /ʌ/, /u/, /e/, and diphthongs involving them.

They describe big, crude, and dark objects, and movements arousing such sensation

Commenting on the correlation high pitch and high vowels = smallness vs. low pitch and back vowels = largeness, [Childs \(1995](#): 135) writes: 'The evidence for such correlations is not extensive, and thus we must be chary of claiming too much structure to the ideophonic subsystem.' What this suggests is that more research is needed on the role played by sound symbolism in ideophones.

In concluding, mention may be made that sound symbolism is not unheard-of in other interactives either. For example, [Westermann \(1912: 33\)](#) reports that in the Shilluk language of Southern Sudan, interjections exhibit a meaningful tonal distinction: Low tone tends to be associated with the expression of vexation, disappointment, and contempt and high tone with joy, astonishment, and admiration.

### 3.5.7 Are ideophones interactives?

In the introduction to Section 3.5 we raised the question of whether ideophones qualify as a type of interactives. In the present section, an affirmative answer is proposed.

Crucial for our decision on the status of ideophones is the following question: Do ideophones conform to the prototypical definition volunteered in Section 1.2? In order to answer this question, each of the definitional properties is now looked at in turn. Like in the case of other interactives, we are restricted to expressions that have not undergone grammaticalization, that is, to what [Dwyer and Moshi \(2003\)](#) call primary ideophones (see Section 4.4). The reason for this is nicely described from a different perspective by [Dingemanse \(2017b: 198\)](#) thus: ‘The tell-tale signs of depiction that occur when ideophones are morphosyntactically independent all disappear when ideophones lose their freedom and are assimilated to become more like normal words.’

This note is important because some generalizations that have been proposed for ideophones do not clearly distinguish between primary (non-grammaticalized) and secondary (grammaticalized) ideophones.

#### (a) Interactives are invariable

English ideophones, like *bang*, *plop* or *thud*, illustrate this property: They are typically holophrases—unanalyzable expressions which do not allow for any modification by means of inflectional or derivational morphology.<sup>17</sup> Accordingly, [Haiman \(2018: 104, 265\)](#) says that ideophones are holistic (unparsable) signs and for [Childs \(1995: 132\)](#), ideophones ‘have no morphology.’ Accordingly, [Dingemanse \(2017b: 196\)](#) concludes that ideophones are ‘unburdened by morphology,’ and [Haiman \(2018: 126\)](#) draws attention to ‘the inability of ideophones to bear inflectional affixes.’

An overview of the literature on ideophones suggests, however, that this issue requires more detailed discussion. Ideophones in fact make massive use of reduplication and lengthening for expressive purposes (Section 4.3.1), showing a range of

<sup>17</sup> As was just observed, this does not apply to secondary ideophones, that is, to expressions that have been grammaticalized, like *bang* (>to bang), *plop* (>to plop), or *thud* (>to thud, a thud); see Section 4.2.

patterns of relative productivity, but also of frozen reduplication, as demonstrated in the seminal study of [Dingemanse \(2015\)](#). Salient functions of ‘repeated talk’ expressed by ideophones are plurality, duration, and intensity ([Dingemanse 2015: 955, 959](#)). Reduplication usually has semantic motivation, but it may also be influenced by phonotactic factors. In the Northern Sotho language of South Africa, for example, monosyllabic ideophones can be triplicated whereas ideophones having more than one syllable are duplicated ([Poulos and Louwrens 1994: 357](#)).

One may hesitate, however, to accept reduplication, or repetition of any kind as a challenge for (a). First, in spite of its outstanding role in the production of ideophones it does not appear to be an obligatory property of ideophones. Second, it is also not an intrinsic feature of ideophones—accordingly, [Dingemanse \(2015: 947\)](#) writes that ideophones ‘often show reduplication, though this is neither a necessary nor a sufficient feature.’ For this reason we prefer to view reduplication, or repetition, as a mechanism that tends to but need not apply to interactives (see Section 4.3.1).<sup>18</sup>

Third, reduplication is not only a characteristic of ideophones; rather, it seems to be equally common in animal directives (Section 3.11.5). The latter are not only reduplicated but also quite commonly triplicated and quadruplicated sequences across languages, as shown by [Andrason and Karani \(2021b: 36\)](#). And finally, functions expressed by reduplication, such as plurality, duration, and intensity, can also be observed in reduplication as it is found in sentence grammar.

[Dingemanse \(2015: 958\)](#) shows that variation in ideophones of the Siwu language analyzed by him can be classified into nine distinct morphophonological templates of ‘expressive morphology.’ The templates are distinguished mainly on the basis of differences in the kinds of reduplication (e.g., full vs. partial reduplication) and lengthening they show. The ideophone *sinisini* ‘woven tightly,’ instantiating Template 1, for example, allows for the expressive variant *sinisini~sinisinisini*, signaling intensification, but not for *\*sinisini~si*.

What this suggests is that Siwu exhibits noteworthy regularities and constraints in the expressive usage of its ideophones. Whether this qualifies as morphological variation might be a matter of debate, but if it is it seems to be of a different nature than that to be observed, for example, in the morphology of inflection or derivation. With reference to the definitional property in (a), however, the following observations need to be considered. First, modification for expressive purposes is far from being a feature restricted to ideophones, it is also found in sentence grammar, even if of a different and a less pronounced kind. And second, it is also a feature commonly observed in interactives other than ideophones—a paradigm example was provided in Table 1.3 of Section 1.3.1.4, involving the German

<sup>18</sup> According to [Kita \(1997: 400\)](#), the main distinction between reduplication and repetition is the following: A reduplicated unit forms one phonological word, whereas a repeated unit and the unit it repeats each forms a phonological word of its own.

interjections *ach*, *hm*, and *oje*, each of which has several variants distinguished by intonation, and each expressing a different interjective function.

Accordingly, even in languages like Siwu, cases such as that of ideophones would not seem to pose a serious problem for the property in (a): Siwu ideophones like *sinisini* ‘woven tightly’ do not take affixes—quite unlike Siwu words of sentence grammar such as nouns or verbs, which do. But perhaps even of more interest is the following observation: When Siwu ideophones are grammaticalized, they in fact lose their invariable format, being integrated in categories of sentence grammar and acquiring the affixal morphology of their new category, as will be demonstrated in Section 4.4.

Another kind of stem variation can be observed in Korean. For example, ideophones distinguish shades of meaning by using tensed consonants for stronger hue vs. non-tensed consonants for weaker hue, and ‘positive’ vowels for bright hue vs. ‘negative’ vowels for darker hue (see Section 3.5.6), as illustrated in (23).

(23) Korean (Koo and Rhee 2018)

|   |                 |                            |                          |                |
|---|-----------------|----------------------------|--------------------------|----------------|
| a | <b>ppalkah-</b> | ‘be red’                   | Tensed initial consonant | Positive vowel |
| b | <b>palkah-</b>  | ‘be reddish’               | Non-tensed consonant     | Positive vowel |
| c | <b>ppelkeh-</b> | ‘be dark red (not bright)’ | Tensed initial consonant | Negative vowel |

To be sure, such kinds of sound symbolic stem alternation are at variance with the notion of invariability as stipulated by our definition in Section 1.2. Still, variability is for the most part of a kind that is different from canonical morphological variation as it is manifested, e.g., in the affixal morphology of both Siwu and Korean. However, some languages also exhibit a limited amount of affixation in ideophones—for example, in South African Bantu languages, ideophones can take a few affixes (see, e.g., Doke and Mofokeng [1957] 1985: 338–40 on Southern Sotho). Interestingly, ideophones (‘expressives’) in the Mon-Khmer language Bahnar of Vietnam—the language distinguishes around 2000 ideophones—take plural and dual suffixes while neither nouns nor verbs take plural or dual suffixes, nor suffixes in general (Diffloth 1994: 109).

In want of more detailed information we assume that such cases are due to processes of grammaticalization and, thus, contrast with prototypical instances of ideophones, being suggestive of ‘secondary ideophones’ rather than of ‘primary ideophones’ in the sense of Dwyer and Moshi (2003: 180–1). Such processes are highly common in ideophones, as will be shown in Section 4.4. In sum, we concur with Haiman’s (2018: 104, 265) conclusion that ideophones are essentially holistic (unparsable) signs that exhibit no internal morphosyntax.

(b) The meaning of interactives may affect but does not change the meaning of a clause

Ideophones may contribute to but are not part of the meaning of the clause they are associated with. First, clauses are restricted to descriptive meaning. Ideophones, by contrast, are depictions, not descriptions (e.g., [Dingemanse 2011](#) and elsewhere; see also Section 2.1.3)—accordingly, they cannot reasonably be part of the descriptive meaning of a clause. And second, and relatedly, the meaning of ideophones is commonly classified as expressive rather than descriptive (Section 2.1.2) and, hence, contrasts with that of sentences or their parts.

There is one notable exception though. As was observed in Section 3.5.3.2, there are three main discourse constructions commonly associated with the usage of ideophones in the languages of the world. The exception concerns the modifier construction, which imposes a semantic constraint of the following kind on the T argument of the ideophone: The meaning of the T argument, which typically is a text piece of sentence grammar, must be compatible with that of the ideophone.

(c) Interactives are syntactically detached

Words of sentence grammar, like nouns, verbs, adjectives, adverbs, etc., are part of syntactic constituents of noun phrases, verb phrases or clauses. Ideophones and other interactives, in contrast, are not syntactic constituents, they cannot be syntactic modifiers, nor can they be subordinated. This has been emphasized in work on ideophones, for example, when [Dingemanse \(2017a: 373\)](#) concludes that ideophones are ‘devoid of syntactic integration’ in their free constructions, that is, as primary ideophones—they ‘are set apart from the rest of the utterance’ ([Dingemanse 2011: 35](#)). And taking grammaticalized structures of ideophones into account, he writes:

In many languages, ideophones do in fact participate in sentential syntax to varying degrees. A common enough response is to ignore this: we know that ideophones are supposed to have no syntax, most data appear to confirm this, so we discount the few remaining exceptions.

([Dingemanse 2017b: 197](#))

In a similar way, [Nuckolls \(1996\)](#) states that ideophones are ‘syntactically isolated’ ([Nuckolls 1996](#)), or as [Haiman \(2018: 115\)](#) notes, ‘while they inhabit “the same screen” as conventional language, they are often both morphologically and syntactically ghettoized or quarantined, “aloof” from the material which surrounds them.’

Such observations have also been made in various individual languages. For example, in the Dhaasanac language of northern Kenya and southwestern Ethiopia, ideophones and interjections are extra-sentential elements or form ‘sentences’ of their own and are separated from the following sentence by a pause ([Tosco 2001: 249](#)).



Interestingly, ideophones change dramatically when they are grammaticalized, as will be shown in Section 4.4: In this case they gradually develop into words like verbs, adjectives, or adverbs and, hence, are no longer syntactically detached, but this also applies to other types of interactives (Section 4.2).

(d) Interactives are set off prosodically from surrounding text material

This property does not seem to need any further discussion: It has been mentioned and described time and again in studies of ideophones—as pointed out for example by [Dingemanse \(2011: 144; 2017a: 373\)](#) when he notes that ideophones feature ‘prosodic foregrounding’ in their free constructions, forming their own intonational units (see also [Haiman 2018: 89–91](#) for cross-linguistic data).

That ideophones are set off from their environment has also been observed in a number of other studies. In the Bantu language Shona of Zimbabwe, ideophones ‘are marked off from the embedding text by longer pauses than other words’ ([Fortune 1962: 8](#)), and ideophones in the Turkana language of northwestern Kenya are often formed with ‘expressive intonation’ ([Dimmendaal 1983: 381](#)).

(e) Interactives are indexicals

The status of ideophones as indexicals has been discussed controversially (cf. Section 1.3.1.5), as the following account suggests. [Kilian-Hatz \(2001: 156\)](#) writes that ‘the ideophonic event happens simultaneously in the moment when it is uttered,’ and [Haiman \(2018: 126\)](#) concludes (where his ‘propositional language’ corresponds to what is called here sentence grammar):

Where propositional language is displaced and is designed to talk not only about the here and now but all places and times, and the imaginary as well as the factual, ideophones are always renderings of intensely felt impressions right, here and now, where the play is the only thing.

([Haiman 2018: 126](#))

According to [Dingemanse \(2011: 155\)](#), in contrast, this question cannot clearly be answered in the affirmative:

The key difference is that ideophones are not immediate responses to events; they are the depictions of events... [and as such] they do not have the necessary indexical connection to the here and now; they are used freely for past events, future events, and imagined events, as for events in the immediate context.

([Dingemanse 2011: 155](#))

To this, [Haiman \(2018\)](#) responds thus:

[...] it is also not as clear as Dingemanse wishes to emphasize that ideophones are not anchored to the here and now. They can just as easily dramatize the ‘there and

then,' but while the speaker is dramatizing them, s/he is on stage. It is precisely because they are dramatizations, and because their producer steps on stage, that they so typically resist the metalinguistic embeddings of assessment, negation, and interrogation.

(Haiman 2018: 136)

Similar observations can be made on interactives other than ideophones. First, interactives have generally been described as depictions in Section 2.1.3. And second, interjections and other interactives can also enact a narrated event in the here-and-now of the speech event. Andrason and Dlali (2020: 178) provide such an example from the South African Xhosa language, where some interjections may indicate a state of affairs experienced at a specific moment in the past or the future. Thus, in (24), a woman refers to an event that occurred a month earlier. When telling the story, she uses *yho* to indicate her surprise experienced in the past, where *yho* is an interjection expressing the speaker's surprise or displeasure. In such a situation the interjection refers not to the speech event but rather to the narrated event.

(24) Xhosa (Bantu, Niger-Congo; Andrason and Dlali 2020: 178, (10b))

|                 |                  |                    |
|-----------------|------------------|--------------------|
| Ndandi-si-ya    | e-venkile-ni,    | – <b>yho!</b> –    |
| 1.SG.TMA-SIT-go | LOC-NC9.shop-LOC | <i>yho!</i> .INT   |
| nda-chola       | imali            | e-sitratw-eni      |
| 1.SG.PAST-pick  | NC9.money        | LOC-NC7.street-LOC |

'I was going to the shop, and – *yho!* – I found money in the street.'

The question then is: Are ideophones indexicals like other interactives? This question can be answered by looking at example (25), which was discussed already in Section 2.2.3.2 when comparing the argument structure of ideophones with that of interjections. The example concerns a constructed situation where a gunshot provokes a person to produce a sequence consisting of an ideophone (*bang*) and an interjection (*ouch*). As was observed there, the two interactives differ in their argument structure, in that the ideophone refers to the situation depicted, that is, a T argument, whereas the interjection depicts the emotive state of the speaker, that is, an S argument.

(25) English (repeated from Section 2.2.3.2)

**Bang!, Ouch!**

In spite of the difference in argument structure, however, the two seem to be the same with reference to the kind of indexicality expressed by them: Both present vocal depictions anchored to the here-and-now of the speech event, that is, the situation of discourse, expressing experiences located right within the moment of speaking. To be sure, by saying *Bang!* the speaker is referring to an event, such as a gunshot, that took place prior to speech time. But much the same can be said to

apply to *Ouch!*, in that the speaker is referring to an experience made immediately prior to the time when the utterance was made.

Furthermore, we just saw that interjections, like *yho!* of Xhosa in example (24), can be used for displaced reference, depicting a situation located outside of the speech event. This applies in the same way to ideophones, as the example in (26) shows: While the two situations in (26) describe events located in the past, both the ideophone *vroom* and the interjection *ow* are depictions enacted as ‘dramatic performances’ in the sense of Haiman (2018: 135), enabling the hearer to experience them in the here-and-now of the situation of discourse.

(26) English (Dingemanse 2021a; repeated from Section 2.2.3.2, (14))

- a The car went **vroom**.
- b The reader went **ow!**

What this suggests is that ideophones do not seem to differ fundamentally from interjections in their value of indexicality: Both are vocal depictions anchored to the here-and-now of the speech event even if the time of the event depicted need not coincide with that of the speech event. If in fact there should be a difference then it would seem to be derivative of a difference in argument structure: Unlike the other nine types of interactives, ideophones lack an S (speaker) argument (Section 2.2.3.2), which means that they do not focus on the situation of the speaker in the way that other interactives do.

Finally, the two contrasting positions can conceivably be reconciled if one relates them to one specific characteristic that sets ideophones off from other types of interactives, namely that they tend to undergo grammaticalization on a massive scale. As we will see in Section 4.4, this process leads almost invariably from interactive grammar to sentence grammar, and the latter is with few exceptions non-indexical in nature.

#### (f) Interactives cannot be negated

This property is mentioned in a number of treatments of ideophones. Kilian-Hatz (2001: 158) writes that ‘ideophones are affirmative by definition,’ and she adds: ‘Ideophones in general are never negated and are rarely found in negated sentences.’ Haiman (2018: 126) asserts that ‘the inability to be negated is a very specific absence, which recurs cross-linguistically,’ and he cites a wide range of references in support of this assertion. Ideophones, he adds, ‘so typically resist the metalinguistic embeddings of assessment, negation, and interrogation.’ This observation echoes that made by a number of other authors, such as Diffloth (1994: 108), who writes on ideophones (‘expressives’): ‘Expressives, on the other hand, though they may on occasion be used in sentences containing negatives, cannot be themselves negated.’

Interestingly, when ideophones are grammaticalized into verbs, adjectives, or adverbs they lose their inability to be negated, now allowing negation like other

words of sentence grammar. For example, Siwu ideophones are not negated, but when being grammaticalized they are no longer immune to negation (Dingemanse 2017a: 371).

(g) Interactives cannot be turned into questions

In much the same way as for negation, it seems fairly uncontroversial that ideophones are not normally allowed to be used as questions—in the wording of Haiman (2018: 136), they resist metalinguistic embeddings of interrogation. With regard to the Turkana language of northwestern Kenya, Dimmendaal (1983) states:

A common syntactic feature of ideophones in Turkana as well as in many other African languages is that they are only used in affirmative sentences, often with expressive intonation.

(Dimmendaal 1983: 381)

There are some data suggesting that this generalization can be contradicted (see Dingemanse 2017a: 363–4); but it would seem that all these data concern ideophones that have undergone grammaticalization and, hence, can be predicted to have lost the constraint on disallowing interrogation (see the introductory note to this section).

What was observed on negation in Siwu in (f) also applies in much the same way to interrogation: Siwu ideophones are not used in question formation. But when they are grammaticalized they lose their immunity to be questioned, that is, they can now turn into questions (Dingemanse 2017a: 371).

In addition to the seven definitional features mentioned there are even more reasons for relating ideophones to other types of interactives, suffice it to mention the following reasons.

(h) There is no clear boundary separating ideophones from interjections

As observed by Dingemanse (e.g., 2011: 155), ideophones can clearly be distinguished from interjections in that the former ‘are not immediate responses to events; they are the depictions of events.’ Similarly, it is argued here that the two need to be kept apart for the reasons given in Section 3.5.4. Nevertheless, the boundary between the two is far from clear. Thus, Haiman (2018: 135) observes with reference to ‘response cries’, that is, a group of interjections:

Response cries and ideophones are clearly alike in many ways. First, note that all hands agree that both categories are dramatic performances of some kind, thus clearly satisfying the first item on the a priori list of properties of missing links. All hands also agree that both categories are perfectly capable of acting as

complete non-propositional utterances (response cries always, ideophones subject to restrictions that vary from language to language). Not only this, they resist incorporation into larger units, whether lexical, phrasal, or clausal. Not infrequently, response cries, like ideophones, violate canonical phonological constraints: either in terms of phoneme inventory (e.g. *brrr*) or of phonotactics (e.g. *mmm*), they are often marginal exemplars of words. Again, there is probably general agreement on the not universally shared feature of gestural accompaniment in both categories. Finally, although both ideophones and response cries are words, they can be called ‘semi-words’ insofar as they share to some extent the properties of lacking some of the design features of language. They are not entirely conventionalized, they exhibit analog patterning, and they are less standardized than conventional words.

(Haiman 2018: 135)

According to the definitions proposed here, interjections express a sudden change in the *emotive or cognitive state* of the speaker (Section 3.6.2) while ideophones are used for vivid depictions of sensory imagery (Section 3.5.2). And ‘sensory imagery’ in the sense of Dingemanse (2018: 3; 2019: 15) includes not only perceptions of the external world but also *inner sensations and feelings*, as is the case, for example, with the Japanese ideophone *ki* ‘being determined and defensive’ (Kita 1997: 404–5) and the Swahili ideophone *lewa lewa* ‘be totally drunk’ (Baldi and Wamitila 2014: 41–57; Section 4.1.2, Table 4.1). Thus, there is significant overlap in the semantic space covered since both the S argument of interjections and the T argument of ideophones may refer to human participants.

As we will see in Section 4.2, semantic overlaps are not only found between interjections and ideophones, they are also found between interjections and other types of interactives. In this sense therefore, ideophones can be argued to be part of the overall network characterizing interactive grammar.

In concluding, mention may also be made of features that are not part of the definition of interactives but generally apply to interactives, namely features relating to historical development.

(i) Interactives are commonly coopted from words of sentence grammar  
Discourse markers like *in fact*, directives like *listen*, evaluatives like *bullshit*, interjections like *hell*, response elicitors like *right?*, response signals like *sure*, social formulae like *please*, etc. all have their origin in corresponding forms of sentence grammar. And there are examples suggesting that this also applies to some ideophones. In a number of Bantu languages, ideophones have been coopted on a large scale from verbs of sentence grammar. This mechanism will be discussed in Section 4.1.2; suffice it here to give a couple of examples from the Swahili language of East Africa: The verb *ku-twang-a* ‘to pound’ has given rise to the ideophone *twaa* ‘hit strongly’ and the verb *ku-wash-a* ‘to ignite, kindle, light’ was coopted

as the ideophone *waa* ‘light at once’ (Baldi and Wamitila 2014: 41–57). And the mechanism of shift from verb to ideophone is also not unheard of in English. Thus, English verbs like *sweep* or *sniff* are also found to be used as ideophones (*sweep*, *sweep!*, *sniff*, *sniff*, *sniff!*, respectively; Clark 2019: 243), that is, the latter have their origin in the former.

(j) Interactives can develop in the direction of sentence grammar

We will see in Section 4.4 that ideophones do not uncommonly undergo grammaticalization in the direction of sentence grammar, acquiring properties of verbal, adverbial, adjectival, or other words. This is in fact a noteworthy observation, but again, it does not appear to be significant to the extent that it would justify separating ideophones from other types. As will shown in Section 4.2, grammaticalization processes leading from interactives to forms of sentence grammar are by no means restricted to ideophones. For example, English interjections like *wow*, *yuck*, and *yum* can also be found to have acquired uses as verbs (*to wow*) or adjectives (*yucky*, *yummy*) (Section 4.2.2.5).

An overview of the relevant literature suggests that there are even more features in addition that ideophones have in common with other kinds of interactives—suffice it to mention two further features that ideophones share with interjections: One concerns an observation made by Haiman (2018: 135), mentioned in the quote presented earlier, according to which ideophones and interjections are ‘dramatic performances of some kind’ and similar observations can be made in fact in a number of other interactives. According to the second feature, certain ideophones are produced by whispering in some languages and much the same applies to interjections (Section 1.3.2.2).

Nevertheless, there is a conspicuous feature distinguishing ideophones from interjections, response signals, and some other types of interactives, namely that forms consisting of clicks appear to be blatantly absent in ideophones. However, the same also applies to some types of interactives such as discourse markers and vocatives.

In sum, there are differences between ideophones and other interactives, but for one thing, these appear to be differences in degree, not in kind. And for another, much the same applies to the other types of interactives: As can be seen in the various sections of this chapter, each of the ten types has its own distinctive features.

Dingemanse and Akita (2017: 508) note that ideophones ‘have to be detached to some degree from the prosaic descriptive linguistic material that may surround them,’ and as we saw in the preceding paragraphs, they also conform in other respects to the structure of interactive grammar. And we concur with Mark Dingemanse when he emphasizes that ideophones constitute an extraordinary class of linguistic expressions. Nevertheless, it is hoped that the observations made in this section show that ideophones, while being special in some ways, exhibit

sufficient grammatical properties justifying their inclusion in the category of interactives. This is also suggested by the following characterization of ideophones by [Güldemann \(2008\)](#), which applies in much the same way to other types of interactives, as we saw in Section 1.3.1:

[I]t can be stated for ideophones, too, that they can be inserted in discourse without any indexing construction of a predicative or other nature,<sup>19</sup> that they are often set apart from the cotext by a pause or other suprasegmental features and thus constitute independent and focal intonation units, and that the clause indexing their presence does not bear a syntactic relation to them.

([Güldemann 2008](#): 294)

To be sure, ideophones are conceivably not indexicals in the way other interactives are considering that they have no indexical participant, S or H, in their argument structure. But there are also other types not fully conforming to the set of properties listed in Section 1.2—note that the definition proposed there rests on a prototypical understanding of the category rather than on principles of discrete categorization.

### 3.5.8 Ideophones as a potentially open-ended class

If there is one feature that dramatically distinguishes ideophones from other interactives, then this is their potentially unlimited class membership. Most types of interactives have a limited membership, typically ranging from a handful to a few dozen and maximally a few hundred.<sup>20</sup> This is different with ideophones. As was observed in the introduction to Section 3.5, the number of ideophones in languages like Korean, Japanese, Bahnar, Basque, Gbeya, or Zulu ranges in the thousands. The question then is: Why should this be so? Ideophones resemble interjections in a number of ways, but why is the number of the latter relatively small whereas that of the former is at times close to ‘excessive’?

The answer we submit is the following. We saw in Section 2.2.3.2 that interjections differ from ideophones crucially in their valence: The former have S (speaker) as their only argument while the latter have T (theme) as their only argument. For the structure of these two types of interactives this has noteworthy implications. With their focus on S, interjections have the inner world of the speaker in their scope and the way this world is reflected in vocal depictions of the emotive and cognitive states of the speaker. Accordingly, the domain of expressions covered by interjections is restricted to this inner world.

<sup>19</sup> As we saw in Section 3.5.3.2, however, this does not apply to all ideophones: Those introduced by means of the quotative construction do in fact require an indexing construction, typically having a verb meaning ‘say’ or ‘do’ as a quotative index.

<sup>20</sup> We are ignoring here the special case of vocatives (Section 3.10).

This is quite different in the case of ideophones. Since T stands for themes of any kind, it can be instantiated by expressions for any phenomena in the outside world. Accordingly, unlike S, T provides a pool for sensory imagery that is virtually unlimited, extending from states, events, and objects to qualities of all kinds. This means, on the one hand, that ideophones not uncommonly express functions similar to lexical distinctions made in sentence grammar. In many societies, ideophones accompany the people from morning to evening, assisting them in voicing in a concise format the sensory experiences they make in their daily activities. The Tjwao people of western Zimbabwe, for example, dispose of the set of ideophones presented in (27) that they produce or are exposed to day in, day out—expressions many of which have counterparts in lexical words of sentence grammar.

- (27) Some ideophones for human activities in Tjwao (Khoen ('Central Khoisan'); [Andrason, Fehn, and Phiri 2021](#). [l] = dental click, [!] = alveolar click)
- |                                              |                                  |
|----------------------------------------------|----------------------------------|
| <b>ooooaa, oooooe</b>                        | yawning                          |
| <b>atsu</b>                                  | sneezing                         |
| <b>phu</b>                                   | spitting                         |
| <b>xuaxuaxua</b>                             | scratching                       |
| <b>ahahahaha</b>                             | laughing                         |
| <b>ehehehehe</b>                             | laughing lightly                 |
| <b>ohohohoho</b>                             | laughing heavily                 |
| <b>buu, piii, or tccc</b>                    | passing wind                     |
| <b>tcwaaa</b>                                | passing water                    |
| <b>tcuaxx</b>                                | having diarrhea                  |
| <b>[ha ha ha</b>                             | walking                          |
| <b>tcatcatca</b>                             | walking on leaves                |
| <b>[hap hap hap hap, khapakhapakhapa</b>     | walking on shallow water         |
| <b>gubugubugubu</b>                          | walking on deep water            |
| <b>g ug ug u</b>                             | jumping                          |
| <b>!ha!ha!ha</b>                             | clapping                         |
| <b>ng eng eng eng eng eng e</b>              | clapping concurrent with dancing |
| <b>mmm, tũthuthutũthuthutũthuthutũthuthu</b> | singing                          |
| <b>dididididididi</b>                        | dancing                          |
| <b>tcatcatcatcatcatcatcatcatcara</b>         | dancing by modern music          |
| <b>txitxixi</b>                              | heartbeat                        |

On the other hand, this also means that ideophones provide semantically rich depictions of phenomena in the world (see [Dingemanse 2011](#) for a seminal semantic analysis of ideophones). The following are a few examples from Korean illustrating the nature of meanings expressed by ideophones. Note that in accordance with the analysis proposed in Section 2.1.1, ideophones, like other interactives,



have both a descriptive and an expressive meaning component. We are restricted here to the former component and assume that the translations provided by [Koo and Rhee \(2018\)](#) in (28) include the main features of the descriptive meanings involved.

(28) Examples of Korean ideophones ([Koo and Rhee 2018](#))

- a **huntul** Movement of an object loosely fixed upon the ground or floor rocking or swinging on a horizontal axis (e.g., a tree in the wind, an unstable chair on the floor)
- b **wumcik** Movement of an object normally restricted to a limited scope (e.g., a person moving their body slightly while sitting on a chair or floor)
- c **chikchik** Sound of a wet stubby object brushing on a surface repeatedly (e.g., of a paint brush on canvas, or of an aerosol paint sprayed onto canvas)
- d **swuk** Sound of an object moving into an object with minimal friction for certain depth (e.g., a well-sharpened knife deeply driven into a pumpkin)

As the translations in (28) suggest, each of the ideophones is composed of a number of conceptual elements: They capture not only everyday situations commonly experienced by the speakers of the language but also objects, processes, and qualities and the relations holding between them. It may therefore come as no surprise that the number of ideophones can be extraordinarily large in some given language compared to that of interjections, or of other interactives for that matter, in some languages being comparable to that of nouns or verbs.

These observations raise questions such as the following:

- (a) Why have some languages of eastern Asia such as Korean and Japanese, Asian languages like Semai, Mon-Khmer languages like Bahnar, or African languages such as Gbeya and Zulu exploited their potential for conventionalizing ideophones on a large scale whilst most European languages, including English, have made only minimal use of this potential?
- (b) Ideophones are not the only type of interactives having T in their argument structure, so why does what was just observed on ideophones not also apply to other interactives taking a T argument?

We have no answer to question (a), which is in need of further investigation. Note, however, that (a) is not restricted to the European languages mentioned, much the same also applies to some African languages. Our research on the Akie language in north-central Tanzania, for example, suggests that in this language as well, ideophones are equally poorly represented. Furthermore, (a) does not apply

to all European languages—Basque is claimed to have more than 4500 ideophones (Dingemanse (2018: 15, Table 4).

A conclusive answer to question (b) is also awaiting further research but at least a partial answer seems possible along the following lines. Ideophones take a T argument, but this also applies to a number of other types of interactives, namely attention signals (e.g., *hey*), discourse markers (*after all*), evaluatives (*great*), response elicitors (*right?*), and response signals (*yes*).<sup>21</sup> What these five types have in common is that their argument configurations are more complex, consisting in addition to T of an S (speaker) and an H (hearer) argument (see Table 2.6 of Section 2.2.2). And the functions of these interactives are procedural, used by the speaker to guide the hearer on how to interpret or frame some specific part of the discourse. This is different with ideophones. Rather than having procedural functions, they immediately express at times complex sensory imagery of states, events, objects, and qualities, having rich conceptual contents, and the number of conceivable phenomena they may take to instantiate T is virtually unlimited.

### 3.6 Interjections

Of all types of interactives, interjections have found the largest amount of scholarly attention and they will be treated in more detail. The ontological status of interjections has been discussed controversially, especially the question of whether they have illocutionary force, qualify as illocutionary acts or speech acts (see, e.g., Meinard 2015: 155). In the model of Functional Discourse Grammar, interjections are treated as a lexeme class that can occur in ‘expressive discourse acts,’ giving direct expression to the speaker’s feelings (Hengeveld and Mackenzie 2008: 63). One may in fact wonder if interjections qualify as ‘interactives’ since they can well be uttered in isolation without any other interlocutor being present. The main reason for extending the term to interjections is that in most of their occurrences in texts they seem to involve or imply interactive situations.

Interjections, in the sense of the term as used here, largely correspond to what Doke and Mofokeng ([1957] 1985: 365) call ‘radical interjectives’ in their grammar of Southern Sotho. Together with vocatives and verb imperatives, radical interjectives form the category of interjectives proposed by these authors.<sup>22</sup> Rather than as a homogeneous set, ‘interjections’ tend to be treated as a complex and diversified type (see, e.g., Andrason and Matutu 2019). As will be shown in Section 5.1.1.1, ‘interjection’ has quite commonly been used as a cover term for a range of different

<sup>21</sup> As was shown in Section 3.1, T arguments are not an obligatory feature of attention signals.

<sup>22</sup> The same kind of classification is proposed by Cole ([1955] 1987: 394) in his grammar of Tswana, a language very closely related to Southern Sotho. But rather than ‘radical interjectives,’ Cole calls them ‘primitive interjectives,’ and ‘vocatives’ are referred to by him as ‘vocative interjectives’ and ‘verb imperatives’ as ‘imperative interjectives.’

phenomena. In fact, we are not aware of any two treatments that would include exactly the same range of expressions within their coverage of ‘interjections.’

In order to avoid confusion on what does and what does not qualify as an ‘interjection,’ we propose to restrict the term to what is widely held to be its ‘core’ or ‘prototypical’ function, namely that of expressing emotive and cognitive states, as it will be defined in Section 3.6.2. This proposal is in accordance with conventions observed in many classical studies, such as that by Jakobson (1960: 354), for whom interjections represent the ‘purely emotive stratum in language’ (see also Nübling 2004: 17). And it is also to quite some extent in accordance with the tradition of grammar writing in the languages of the world. A look at grammatical descriptions for well over one hundred languages across the world suggests, on the one hand, that the term ‘interjection’ tends to be applied as a cover for a range of different types of interactives. On the other hand, however, more than two-thirds of all definitions volunteered in those descriptions are in accordance with the convention proffered here, restricting the function of interjections to that of expressing emotive and cognitive states (Section 3.6.2)—for example, interjections in the Paamese language of Vanuatu ‘are nothing more than automatic, emotional responses to some kind of situation’ (Crowley 1982: 243).

This means that forms that in some earlier studies were subsumed under the term ‘interjection’ such as discourse markers, attention signals, response elicitors, response signals, or social formulae, are each allocated here to separate types of interactives treated in the relevant sections of this chapter. What seems to be fairly uncontroversial across those studies, however, is that most or all of these types belong to one and the same general category, described here as the category of interactives.

In proposing this classification, we are aware that it is not in accordance with some earlier classifications. Bloomfield ([1933] 1962: 176), for example, included vocatives such as *ma’am* or *sir* or social formulae like *goodbye* and *thank you* but no discourse markers in his class of ‘interjections.’ In contrast, the category of ‘interjections’ proposed by Norrick (2009: 868) consists mostly of discourse markers, like *anyway*, *I mean*, *well*, and *y’know*, but includes neither vocatives nor social formulae. It is hard to see what vocatives, social formulae, and discourse markers have in common other than the fact that they all operate in the space of social interaction generally characterizing interactives as a category.

### 3.6.1 Introduction

Linguists of both the past and mostly also of the present have with few exceptions not been very kind to interjections. It is therefore almost a commonplace in contemporary work on interjections to deplore that this category has not received the kind of attention it deserves. In this vein, O’Connell and Kowal (2008: 133) characterize interjections as ‘a phenomenon that has been historically a marginal,

thoroughly neglected linguistic category,' and similar views have been voiced by a number of other scholars (e.g., [Ameka 1992a](#); [Andrason and Dlali 2020](#): 159).

Interjections can in many of their uses be classified as stance markers. In the framework of stancetaking of [Du Bois \(2007\)](#) three main concepts of stancetaking are distinguished, namely alignment, evaluation, and positioning. Interjections, like 'wow' and 'yuck,' express functions of positioning, and the functions are typically affective (e.g., 'phew') but less commonly also epistemic (e.g., 'aha').

Interjections are first and foremost a feature of speech. In written texts they are rare, except for specific genres like novels, drama plays, emails, WhatsApp messages, and comic strips ([Andrason and Matutu 2019](#): 2).

### 3.6.2 Working definition

In many grammars of languages across the world, interjections are treated as a highly inclusive category, or 'part of speech' (see [Borchmann 2019](#)). Problems with the nature of interjections start with their definition. For good reasons therefore, [Cuenca \(2000\)](#) gave her paper on interjections the title 'Defining the indefinable? Interjections.' In fact, coming up with a definition that would be subscribed to by all scholars that have worked on this phenomenon would seem to be near to impossible: What counts as an interjection is a question that is answered somewhat differently by each researcher. All would agree, however, that there are a number of fairly uncontroversial examples such as English *ouch*, *wow*, and *yuck*, which are generally agreed upon to be paradigm instances of interjections. But there are different views on the boundary areas, which remain somehow fuzzy.

Pioneers of modern research on interjections decided to ignore meaning and restrict their definitions essentially to formal features (cf. [Fraser 1990](#): 391; [Ameka 1992a](#); [Wilkins 1992](#): 124; [Evans 1992](#): 226), as the following examples may illustrate:

- (1) [Interjections are words or short phrases] which can constitute an utterance by themselves and do not normally enter into constructions with other word classes. ([Ameka 1992a](#): 105)
- (2) [An interjection is] a conventional lexical form which (commonly and) conventionally constitutes an utterance on its own, (typically) does not enter into construction with other word classes, is (usually) monomorphemic, and (generally) does not host inflectional or derivational morphemes. ([Wilkins 1992](#): 124)
- (3) I define the prototypical interjection as a word that is (a) monomorphemic, (b) capable of making up an independent, non-elliptical utterance, and (c) not used to represent a non-speech sound.<sup>23</sup> ([Evans 1992](#): 226)

<sup>23</sup> Evans' criterion (c) may sound strange. It is meant to exclude ideophones and onomatopoeic words from the list of interjections.

Such definitions have been adopted by a number of later authors, like [Dingemanse \(2017b: 198\)](#), who follows [Wilkins \(1992\)](#) in asserting that interjections are ‘defined as conventional lexical forms which are monomorphemic and typically constitute an utterance of their own.’

But there are also alternative definitions that are not restricted to formal criteria (cf. [Borchmann 2019: 8](#)). A number of authors in fact converge on proposing a combination of both formal and functional parameters. Already a century ago, [Jespersen \(1922\)](#) proposed the following definition:

[Interjections are] abrupt expressions for sudden sensations and emotions [...] isolated in relation to the speech material used in the rest of language.  
([Jespersen 1922: 415](#))

For [Ameka \(1992a: 106\)](#), interjections are also characterized as ‘lexical utterances,’ being ‘relatively conventionalised vocal gestures (or more generally, linguistic gestures) which express a speaker’s mental state, action or attitude or reaction to a situation.’<sup>24</sup>

Authors combining formal and semantic criteria in defining interjections can also be found in handbooks of English grammar. [Quirk et al. \(1985: 853\)](#), for example, say that they ‘are purely emotive words which do not enter into syntactic relations,’ and for [Biber et al. \(1999: 1083\)](#) they are ‘inserts which have an exclamatory function, expressive of the speaker’s emotion,’ where ‘inserts’ are defined for the most part in terms of formal features. And in their grammar of English, [Carter and McCarthy \(2006: 224\)](#) describe interjections as expressing ‘positive or negative emotional reactions to what is being or has just been said or to something in the situation,’ giving examples like *bother*, *crikey*, *damn*, *heavens*, *hooray*, *jeez*, *ouch*, *ugh*, *wow*, and *yuk*.

[Huddleston and Pullum \(2002: 1361\)](#) define interjections as ‘a category of words that do not combine with other words in integrated syntactic constructions, and have expressive rather than propositional meaning.’ They propose interjections as a kind of prototypical category, which has both central members like *ah*, *hey*, *oops*, *ouch*, *shh*, *wow* ‘which in their sole or primary meaning are used as expressive exclamations, on their own, or as supplements with clausal anchors’ and less central members such as *blast*, *bugger*, *damn*, *fuck* which are supplements having been reanalyzed as interjections.

Ignoring the fact that interjections are commonly defined as ‘words’ (see Section 6.2), the definition by [Huddleston and Pullum \(2002: 1361\)](#) captures paradigm

<sup>24</sup> Doubts have been raised, however, whether this semantic characterization is specific enough to characterize interjections. [Wilkins \(1992: 125\)](#), for example, draws attention to the fact that in ‘most theories of speech acts, all (true) utterances convey something about the “speaker’s mental state”.’

features of interjections in particular and of interactives in general, namely their syntactic detachment, the prototypical organization of interjections as a category, as well as the relationship of interjections with swear words, in that some of the former can be said to be derived from the latter via ‘reanalysis’ or, as we say here, via grammaticalization (see Section 4.2).

The definition proposed here, while highly restrictive, also combines formal with semantic features (see also Wharton 2003: 42–3):

- (4) An interjection is an interactive that expresses a sudden change in the emotive or cognitive state of the speaker.

The formal features contained in this definition are those classifying interjections as a type of interactives, that is, as having the definitional properties of this category (Section 1.2). For English examples of interjections, see Table 1.6 of Section 1.5.

On this definition, interjections may express either emotive or cognitive states. This distinction was already highlighted by Ameka (1992a), according to whom the former express the speaker’s state with respect to emotions and sensations, while the latter ‘pertain to the state of knowledge and thoughts at the time of utterance’ (Ameka 1992a: 113). The former, including English forms such as *ouch* and *yuck*, appear to be cross-linguistically more common than those expressing cognitive states, like English *ah* and *aha*.

In accordance with the general theme of this study, interjections are treated here as belonging to *emotive* rather than emotional communication. According to Caffi and Janney (1994: 328–9), emotive communication is ‘inherently strategic, persuasive, interactional and other-directed by its very nature,’ while emotional communication is a ‘type of spontaneous, unintentional leakage or bursting out of emotion’ (see Stange 2016: 29, Borchmann 2019: 12–15 for discussion).<sup>25</sup> This, however, is an issue that is in need of more research.

In addition to emotive and cognitive interjections, Wierzbicka (2003b: 484) proposes ‘volitive interjections’ as a third type. Expressions of this last type are treated here as directives (Section 3.2).

Not everybody might be satisfied with the definition in (4), and in fact, none of the components of the definition is entirely uncontroversial. The first component stipulates that interjections express ‘sudden changes.’ However, ‘sudden change’ does not necessarily entail that that change is non-intentional or spontaneous since, as (Evans 1992: 237) points out, ‘some assessment of the social setting must be made before the appropriate interjection can be uttered.’ And Haiman (2018: 135–46) emphasizes that interjections (‘response cries’) ‘are in fact public performances that are anything but spontaneous. Note the telling distinction he

<sup>25</sup> Stange (2016: 29) observes on this issue: ‘As a result, emotive interjections like *Ow!*, *Whoops!*, or *Ugh!* are on the borderline between emotional and emotive communication, with their affiliation to one or the other type of communication depending on their context of use.’

draws between saying *Oopsy-daisy!*, a response cry uttered in a nursery, and *Shit!*, the appropriate version of this cry in a foundry: ‘far from being ‘exuded,’ these ‘spontaneous responses’ are tailored for their audiences’—that is, the speaker ‘is in total control of the situation’ (Haiman 2018: 136).

There is, however, also an alternative view according to which interlocutors are not, or not entirely in control when producing interjections. Stange (2016: 18) calls English interjections ‘semi-automatic utterances in that they are (totally) spontaneous exclamations, apparently produced without the speaker’s full intention,’ and Borchmann (2019: 36–7) assumes that an (emotive) interjection is a function and that ‘the function is a spontaneous and non-communicative relief and gratification of emotions and sensations.’

One problem with our definition in (4) concerns the term ‘emotion,’ which has turned out to be close to indefinable considering the plethora of studies that have been devoted to this topic without reaching any noteworthy agreement. In this study we subsume the following phenomena distinguished by Ekman, Friesen, and Ellsworth (1972: 60–1, Table 2) within the term ‘emotion’: anger, disgust/contempt, fear, happiness, interest, sadness, and surprise. Note, however, that the main concern of these authors is not with linguistic expressions but rather with facial behavior. And note further that analyzing interjections as involving emotion is also not entirely uncontroversial. Wharton (2009: 88), for example, suggests that what is communicated by the interjection *yuk* in (5) can be said to be a feeling or sensation rather than an emotion.

- (5) English (Wharton 2009: 87, (93))  
 Child: (taking foul-tasting medicine) *Yuk!*

In fact, not all authors would consider emotion to be an essential part of the meaning of interjections, and some even propose to omit reference to emotions altogether. Kockelman (2003) disputes the claim that interjections exclusively indicate emotions (see Schourup 2003: 483). And Norrick (2009: 868) argues that ‘many primary interjections do not express emotions ... but rather information states.’ The reason for this discrepancy is as a rule that the range of expressions that a given author may have in mind when defining interjections differs greatly from one author to another. Thus, the items classified by Norrick (2009) as ‘primary interjections’ also include English items like *uh* and *um* which are considered here to be discourse markers rather than as interjections (see Section 3.3), or only as fairly marginal instances of interjections (cf. O’Connell and Kowal 2005).

Rather than ‘emotion,’ some authors use the phrase ‘attitude’ or ‘mental state’ in their definition. For Ameka and Wilkins (2006: 2), for example, interjections ‘express a speaker’s current mental state or reaction or attitude towards an element in the linguistic or extra-linguistic context.’ Nevertheless, most authors converge on including ‘emotion’ in their definition of interjections. They are ‘purely emotive words’ according to Quirk et al. (1985: 853), they ‘have an exclamatory function,

expressive of the speaker's emotion' according to Biber et al. (1999: 1083), for Nübling (2001: 21) they serve primarily the spontaneous expression of subjective emotionality, for Dixon (2010b: 27), an interjection is 'a conventionalized cry, typically indicating the speaker's emotional response,' and for Andrason and Dlaki (2020:171), they express the current 'emotional mental state of a speaker.'

A more refined characterization is proposed by Kockelman (2003: 471), who argues that 'interjections' may index objects, signs, internal states, and social relations and he refers to these four distinct types, respectively, as situational, discursive, expressive, and social indexical objects.

Equally controversial is the third component of the definition in (5), namely that interjections are restricted to attitudes of the speaker. The definition thus excludes what is usually called 'conative interjections' (Ameka 1992a), or 'volitive interjections' (Wierzbicka 1992: 165), like English *hey*, *shh*, German *psst* 'be quiet,' Polish *sza* 'be quiet.' And social formulae like *thank you* and (*good*) *bye* are sometimes also classified as interjections (Wilkins 1992: 142), all of which are directed to the hearer and are not considered interjections here (see Sections 3.1, 3.2, 3.9).

Usage of interjections in accordance with (5) is perfect in soliloquy and in contexts regardless of who is there, even if no one is expected to respond (cf. Goffman (1978: 794). This does not apply clearly, or less so, to interactives involving a hearer, that is, having an H argument (Section 2.2). Accordingly, forms like English *shh*, German *psst* 'be quiet,' or Polish *sza* 'be quiet' are classified here as directives (Section 3.2), while *thank you* and (*good*) *bye* are social formulae (Section 3.9) and, as just observed, *uh* and *um* are discourse markers (Section 3.3).

Cross-linguistically, the meaning space of 'emotive or cognitive states' covered by interjections relates most of all to the following concepts: admiration, amazement, anger, annoyance, appraisal, comfort, compassion, consent, contempt, disappointment, disapproval, dissatisfaction, excitement, fear, frustration, joy, pain, regret, refusal, relief, resignation, satisfaction, shock, sorrow, surprise, sympathy, and wonder (Andrason and Dlaki 2020: 161–2).

Rather than emotive states, some interjections express primarily cognitive states. For example, following Ameka (1992a: 113) we include the 'cognitive interjections' *aha* 'I now know this' of English, or *ehé* 'I now remember' of the West African language Ewe within the type of interjections.<sup>26</sup>

The meaning space expressed by a given interjection is in many cases fairly restricted. But it may as well include a variety of the concepts just mentioned. For example, the interjection *kuma mayo* (lit. 'your mother's cunt') used in the East African language Swahili mainly by women, young men, or children, is an all-purpose exclamation for surprise, annoyance, pain, and a number of other emotive states.

<sup>26</sup> 'Cognitive interjections are those that pertain to the state of knowledge and thoughts at the time of utterance' (Ameka 1992a: 113).



Some discussions of interjections are restricted to primary interjections (Section 3.6.4.1). This applies, for example, to Quirk et al. (1985: 853) and Wierzbicka (1992: 164), where primary interjections, like *wow!*, *oops!* or *aha!*, are included while secondary interjections, like *Good Lord!*, *Good heavens!*, *Christ!* or *Hell!*, are excluded.<sup>27</sup> This is not the convention proposed here: To the extent that a given expression conforms to the definition in (5), it will be classified as an interjection, irrespective of whether it is a primary or a secondary interjection.

Finally, some authors follow Goffman (1981) in calling interjections, or a substantial part of them, ‘vocal gestures’ (e.g., Ameka 1992a: 106; Wierzbicka 1992: 163). While intuitively appealing, this term is avoided here because to our knowledge there is no appropriate description of the features that would allow classifying interjections as gestures. However, as was pointed out in Section 2.1.3, interjections as well as other interactives share with gestures that both are depictions that are immediately anchored to the situation of discourse.

### 3.6.3 Grammar

Paradigm instances of interjections, like *ouch* in (6), are immediately anchored to the situation of discourse and prototypical instances of them take only one argument, namely S (the speaker). Thus, the utterance in (6) is paraphrasable as in (7a), hence its argument structure can be described as in (7b).

(6) English

**Ouch!**

(7) a Paraphrase: I (S) have just experienced a sudden and sharp pain.

b Argument structure: *ouch* (S)

Being typically restricted to S as their only argument, interjections can be said to express subjective meaning, thus contrasting with a number of other types of interactives, which require an H argument and hence can be said to have intersubjective meaning (e.g., Traugott 2003; Nuyts 2005).

As was observed in Section 2.2.2, S arguments can be subcategorized into a number of what have the appearance of ‘semantic roles,’ and interjections are suggestive of the role of an experiencer, that is, a participant typically expressing the psychological or physical state of a speaker who is not in control of that state.

<sup>27</sup> Wierzbicka (1992: 164) defines interjections thus: ‘An interjection can be defined as a linguistic sign expressing the speaker’s current mental state (1) which can be used on its own, (2) which expresses a specifiable meaning, (3) which does not include other signs (with a specifiable meaning), (4) which is not homophonous with another lexical item that would be perceived as semantically related to it, and (5) which refers to the speaker’s current mental state or mental act ...’

Usages of interjections as in (6) may have a smack of self-talk: Since they have no H (hearer) argument they can be uttered in soliloquy, without any person other than the speaker being present. A general quantitative estimate suggests, however, that such ‘prototypical instances’ do not constitute the most frequent usage of English interjections. First, interjections appear to be rarely produced as ‘response cries’ in isolation to express the speaker’s attitude; more commonly, they tend to be used to impress other people around in the situation of discourse. This has been demonstrated nicely in [Taavitsainen’s \(1995\)](#) account of primary interjections in texts from Early Modern English, where she concludes:

Thus, in contrast to the definition based on Present-day conversation, interjections are not only reactions to situations but are used to appeal to addressees. In addition, they are deliberate devices in manipulating reader involvement.

([Taavitsainen 1995](#): 463)

And second, and more importantly, interjections are quite commonly accompanied by some piece of text acting as their T argument—in the wording of [Kockelman \(2003: 473\)](#), they ‘index signs in the *discursive* context in that they co-occur with (or serve as) a *response* to an addressee’s previous utterance.’ Such a text piece is provided by *I’ve got 30 in tens* in (8), forming a separate utterance and conversational turn. Note that *Ah Jesus!* is a combination of a primary (*ah*) and a secondary interjection (*Jesus*).

(8) English (KBD 378; [Ljung 2011](#): 90 (22))

A: I’ve got 30 in tens.

B: **Ah Jesus!**

That T is an obligatory argument in (8) is suggested by the following observations. The interjection complex *Ah Jesus!* serves as a reaction to the information given in the previous turn. In the conversation concerned, B had asked A to lend her/him 40 pounds and when it turned out that A had only 30 pounds, B uttered *Ah Jesus!*, probably as an expression of disappointment (see [Ljung 2011](#): 90). One possible reading of the interjection can therefore be paraphrased as in (9a), being suggestive of the argument structure in (9b).

(9) a Paraphrase: I (S) am somehow disappointed by what you just told me (T).

b Argument structure: *Ah Jesus!* (S, T)

To conclude, in addition to (7) there is a second argument structure commonly associated with interjections. We will return to the distinction between the two in Section 3.6.6. As was observed in Section 2.2.4, the reconstruction of arguments is not always easy to achieve, being dependent on the availability of appropriate contextual information.

Interjections are positionally mobile, but they are less so than some other types of interactives. Their clearly preferred position, not only in English, is at the left periphery of an utterance, preceding other types of interactives (see, e.g., [Sohn 1994](#): 521–2 on Korean or [Andrason, Fehn, and Phiri 2020](#): 16 on Tjwao). However, being essentially mobile, interjections may also appear in other positions vis-à-vis some host utterance. For example, in the South African language Xhosa, interjections are placed most of the time utterance-initially, less commonly also utterance-finally. But they may also circumfix an utterance, appearing both at the beginning and the end of an utterance and some, especially expletives, may occur internally ([Andrason and Dlali 2020](#): 206).

Of all English interjections, *oh* is the most frequently used, it ‘frequently occurs with – and always precedes – other interjections’ ([Norrick 2009](#): 877; see also [Heritage 1998](#); [2002](#) for detailed discussion). It is described by [Biber et al. \(1999\)](#) thus:

*Oh* is by far the most common interjection. Its routine use to introduce utterances or to respond to the utterances of others often gives it the character of a discourse marker. *Oh* also frequently combines with other inserts; some of the most common combinations are: *Oh yeah*, *Oh yes*, *Oh no*, *Oh aye* (regional BrE), *Oh well*, *Oh God*, *Oh I see*, *Oh right*. Although the use of *oh* is highly conventionalized, its core function appears to be to convey some degree of surprise, unexpectedness. ([Biber et al. 1999](#): 1083; boldening and italics in the original)

*Oh* is also the most likely to appear in the left-most position of an utterance. Example (10) shows its combination with both the secondary interjections *jeez* and *god* (lines 2, 7) and the response signal *no* (line 5). These combinations have level stress across both units and signal empathy with the person(s) in the situation described in the foregoing turn:

(10) English (SCoSE, Addie and Brianne; [Norrick 2009](#): 886)

1. Brianne: it took me an hour to get out to O’Hare.
2. Addie: **oh [jeez]**,
3. Brianne: [and then] I was standing on the aisle for an hour.
4. ‘cause it’s packed.
5. Addie: **oh no**.
6. Brianne: and-
7. Addie: **oh [god.]**
8. Brianne: [then-] okay, we left on time.

Like other types of interactives, interjections are most of all a matter of spoken discourse. Accordingly, examples like the following, where the English primary interjection *yuk* has made its way into print are fairly rarely encountered: In a letter to the editor of the newspaper *The Australian*, a reader reacts to new guidelines for

gender-inclusive language with the remark in (11), where *yuk* presents an emotive comment on the preceding topic, consisting of the noun *feminists*.

- (11) English (*The Australian*, 25-10-1988: 14; [Wierzbicka 1992](#): 168)  
FEMINISTS - **YUK!** As a woman I feel ashamed.

As was observed in Section 2.1.2, interjections can be used insincerely, for example, in ironic uses or when the speaker employs an interjection not to express their own emotive state but rather the one conceived to be that of the hearer. Note, that ironic uses constitute one of the four main discourse uses of interjections in spoken British English ([Stange 2019](#): 175–6).

### 3.6.4 Sub-types

The ensuing discussion of taxonomic frameworks that have been proposed for interjections are especially of interest here in that they are useful to establish what is treated here as ‘interjections’ and what is not.

#### 3.6.4.1 Primary vs. secondary interjections

Classifying interjections into primary and secondary is presumably the most widely accepted division in work on this type of interactives. As we will see later on in this section, however, this distinction is by no means restricted to interjections.

Primary interjections cannot be used in any sense other than as interjections (‘they are not used otherwise’; [Ameka 1992a](#): 105). They are said to have no homonyms in other word classes ([Gehweiler 2008](#): 73; [Ljung 2011](#): 78) or are not homophonous with another lexical item that would be perceived as semantically related ([Wierzbicka 1992](#): 164).

Secondary interjections, by contrast, ‘are forms that belong to other word classes based on their semantics and are interjections only because they can occur by themselves non-elliptically as one-word utterances and in this usage refer to mental acts’ ([Ameka 1992a](#): 105). This definition is largely adopted by [Evans \(1992](#): 226), even if the latter applies his phrase ‘independent, non-elliptical utterance’ not to secondary but rather for ‘prototypical interjections’ (see his definition in Section 3.6.2 above).

But using ‘non-ellipsis’ as a criterion raises a problem. Consider the following constructed example.

- (12) English  
A: Yesterday I saw a five-legged rabbit.  
B: **And?**

On [Ameka’s \(1992a](#): 105) view, B’s response *And?* would be an ‘elliptical utterance’ usually carrying a rising intonation. Questions that one might wish to be answered are the following: Do speakers conceive B’s usage of *And?* as being elliptical, or

incomplete? Assuming that ellipsis were in fact involved, what is the mechanism and/or function that would account for it? Until these questions are answered one might hesitate to adopt an account of secondary interjections in terms of (non-) ellipsis.<sup>28</sup>

For the purposes of the present study we will say that primary interactives differ from a secondary one in that they are not transparently related to any form of sentence grammar. The reason for proposing this convention is that labels such as ‘word class’ or ‘lexical form’ are not part of the technical terminology used here (Section 6.2) whereas the distinction between sentence grammar and interactive grammar is. Nevertheless, the effect of both conventions is essentially the same. For example, the English secondary interjection *shit* can be said to be related in some sense to a ‘homophonous’ noun or verb of sentence grammar, whereas for the primary interjection *oops* there does not appear to be a corresponding form in sentence grammar—thus, *shit* is a secondary and *oops* a primary interjection in both conventions.

Neither of these conventions, however, is entirely unproblematic, especially for the following reasons: First, one may slightly hesitate to call sentence grammar *shit* and the secondary interjection *shit* ‘homophones’ or ‘homonyms’ considering that the two are usually distinguished prosodically and that the function of the interjection has as a rule not much in common with the lexical content of the noun or the verb *shit*. And second, some forms that are commonly classified as primary interjections have acquired uses as words of sentence grammar and, hence, can be said to have ‘homophones’ or ‘homonyms,’ or to also be ‘used otherwise.’ For example, the primary interjection *wow* can also occur as a verb (*to wow*) and the interjections *yuck* and *yum* have acquired uses as adjectives, cf. *yummy* and *yucky* (see Section 4.2.2.5).

A distinction between the two kinds of interjections was already made by earlier authors. It appears to have been coined by Wundt (1911: 319–20). For him, primary interjections are those that have completely retained ‘the character of sounds of nature’ (*den Charakter von Naturlauten*) whereas secondary interjections are ‘those expressions of feeling that have been couched in other linguistic forms’ (*diejenigen Gefühlsäußerungen, die in andere sprachliche Formen eingekleidet wurden*) (see also Schwentner 1924; Nübling 2001: 27; Stange 2016: 8–9). Jespersen (1924: 90) proposed to separate interjections which are never used otherwise, and those containing words from the ordinary language, like *well* or *why*. For Bloomfield ([1933] 1962: 176), secondary interjections are phrases, ‘often of peculiar construction,’ such as *dear me*, *goodness gracious*, *by golly*, *you angel*, *please*, *thank you*, and *goodbye*, while ‘canonical’ interjections are called ‘special words’: *ouch*, *oh*, *sh*, *gosh*, *hello*, *sir*, *ma’am*, and *yes*.

<sup>28</sup> In Section 4.1, an alternative account is proposed in terms of cooptation which suggests that ellipsis might not be a necessary mechanism to be invoked in such cases.

Table 3.13 provides lists of English forms illustrating the distinction; Norrick (2009: 869) argues that the two constitute open classes of interjections in English. Note that secondary interjections are further classified into taboo expletives, such as *damn*, *fuck*, and moderated expletives, e.g., *darn* (<*damn*), *flip* (<*fuck*) (see Section 4.3.2; Biber et al. 1999: 1094).

**Table 3.13** English forms that have been classified as primary and secondary interjections

|                         |                                                                                                |
|-------------------------|------------------------------------------------------------------------------------------------|
| Primary interjections   | <i>ah, hey, hm, huh, mhm, mm, oh, oho, ooh, oops, ouch, uh, um, wow</i>                        |
| Secondary interjections | <i>bother, boy, careful, Christ, damn, drats, fire, fuck, heavens, hell, help, shame, shit</i> |

Sources: e.g., Ameka (1992a); Norrick (2009: 868)

Norrick (2009: 889) treats primary interjections as a *sui generis* class, while he prefers ‘to leave secondary interjections in their proper lexical classes, describing their functions as interjections in general pragmatic terms.’ He observes that secondary interjections have their roots in religion, sex, and scatology, to express strong emotions (Norrick 2009: 868). In fact, they can be assumed to have been coopted from sentence grammar, frequently being reduced in form, like *blimey*, which goes back to *God blind me*, *gee* (<*Jesus*), *zounds* (<*Christ’s wounds*) (O’Connell and Kowal 2008: 134), or the German interjection *oje* from *oh mein Jesus* ‘oh my Jesus’ (see Section 4.3.2).

Instead of two types of interjections, Goddard (2014: 54) proposes a tripartite division. As shown in the scale of (13), he distinguishes between two categories of English primary interjections, namely phonologically aberrant ‘noise-like’ ones and phonologically normal ‘word-like’ ones, where the latter ‘have the potential to serve as base forms for normal words, such as the verb *to wow* (*someone*), the adjective *yucky*, and expressions like *wow of a time* and *Yuck factor*’ (see also Section 4.2.5).

(13) Three formal categories of English interjections according to Goddard (2014: 54, Figure 1)

|                                   |                          |                              |
|-----------------------------------|--------------------------|------------------------------|
| ‘Noise-like’ primary ones         | ‘Word-like’ primary ones | Secondary interjections      |
| <i>Mmm, Psst!, Tsk tsk!, Ugh!</i> | <i>Gee!, Wow!, Yuck!</i> | <i>Christ!, Fuck!, Shit!</i> |

In some other works, secondary interjections are not considered interjections at all but are treated rather as belonging to some other class of linguistic forms in English. For Quirk et al. (1985: 852–3), for example, ‘interjections’ are restricted to primary interjections whereas secondary interjections belong to what these authors call ‘formulae.’

Biber et al. (1999: 1094–5) also restrict the term ‘interjection’ to primary ones, referring to secondary interjections as expletives. Furthermore, they make a distinction between ‘taboo expletives,’ like *bloody hell*, *damn*, *fuck*, *God*, *Goddamn assholes*, *Goddammit*, *hell*, *Jesus*, and *shit*, and moderated (or euphemistic) expletives, like *gee*, *geez*, *good grief*, *heavens*, *heck*, and *my gosh*. We will return to these distinctions and to ‘moderated expletives’ in Section 4.3.2.

Primary and secondary interjections frequently collocate, and the former tend to precede the latter, cf. English *ah Jesus*, *oh boy*. Collocations having the opposite order are frequently disallowed. Thus, in the following collocations of German interjections, the primary interjections *pfui* ‘pew’ and *oh* ‘oh’ precede the secondary interjections *Teufel* ‘devil’ and *Gott* ‘God’ whereas the reverse order is considered to be odd (Nübling 2001: 25):

- (14) German (Nübling 2001: 25)
- a   **Pfui   Teufel!   #Teufel pfui!**  
    pew   devil  
    ‘Ugh,   yuck!’
- b   **Oh   Gott!   #Gott oh!**  
    ‘Oh   God!’

The distinction discussed in this section has been proposed only for interjections. But there are also other types of interactives making such a distinction, as shown in Table 3.14. The distinction is not confined to the five types listed in Table 3.14, it can also be found in animal directives (Section 3.11). For a discussion of the distinction in discourse markers see Heine et al. (2021: 263–5).

**Table 3.14** Types of English interactives showing a primary vs. secondary form distinction

| Type               | Primary forms                  | Secondary forms                                |
|--------------------|--------------------------------|------------------------------------------------|
| Attention signals  | <i>hey, oh, oi, pst, yo(o)</i> | <i>listen, say</i>                             |
| Directive          | <i>Shh</i>                     | <i>piss off, wait</i>                          |
| Discourse markers  | <i>uh, um</i>                  | <i>anyway, you know</i>                        |
| Ideophones         | <i>bang, boing, thud</i>       | <i>cock-a-doodle-doo, knock! knock! knock!</i> |
| Interjections      | <i>oh, ouch, wow, yuck</i>     | <i>bloody hell, God</i>                        |
| Response elicitors | <i>eh?, huh?</i>               | <i>right?, see?</i>                            |
| Response signals   | <i>mhm, uh huh, unh unh</i>    | <i>exactly, right</i>                          |

The forms *cock-a-doodle-doo* and *Knock! Knock! Knock!* are taken from Clark (2019: 241), who classifies them as an English ideophones. Note that Clark’s classification of ideophones into primary and secondary ones differs from the classification proposed here. It is based on whether being an ideophone is the primary or the secondary usage of a given form.

### 3.6.4.2 Reactive vs. pragmatic interjections

An interesting distinction is made by [Ljung \(2011: 83–93\)](#) between two kinds of what he calls ‘expletive interjections,’ that is, secondary interjections, namely between reactive (expletive) and pragmatic (expletive) ones: The latter are by far more common than the former, making up four fifths of all expletive interjections that he found in his one-million-word corpus of the British National Corpus (BNC).

Roughly speaking, the distinction concerns the question of whether the use of an interjection is triggered by extralinguistic factors (the ‘context’) or by factors recoverable from the surrounding text (the ‘co-text’). Reactive interjections describe the reaction of the speaker to a certain situation—they are, as we say here, immediately anchored to the situation of discourse, that is, the context, as *oh my god* and *shit* are in the following examples:

- (15) English (KB6 347; [Ljung 2011: 85, \(3\), \(4\)](#))

a < crash as kid falls over > **Oh my god.**

b Again, this is < knocking on the wall > it’s the same as this. **Shit!**

Pragmatic interjections, by contrast, are ‘used as *pragmatic markers* (a category also known as *pragmatic particles*, *discourse markers* and *discourse particles*)’ ([Ljung 2011: 86](#); underlining in the original). [Ljung \(2011: 87\)](#) gives examples such as the following to illustrate various pragmatic-marker usages, where the interjection occurs initially in (16), medially in (17), and finally in (18), but as a turn of its own in (19).<sup>29</sup>

- (16) English (KBR 1866; [Ljung 2011: 87, \(9\)](#))

**Jesus bloody Christ** how could he have missed that

- (17) English (KBE 1632; [Ljung 2011: 89, \(14\)](#))

I’m not **fuck** laughing am I?

- (18) English (KB7 6582; [Ljung 2011: 88, \(12\)](#))

I’m not playing this, **bugger it!**

- (19) English ([Ljung 2011: 90, \(23\)](#))

A: I’m driving, there’s this big bang, and the whole bonnet lit up.

B: **Oh God!**

Pragmatic interjections have a text-structuring function. In the framework proposed here, they take a T argument which provides their textual anchor. Most of the time, this argument is encoded in the same utterance, as in (16) through (18), but it may as well form a separate utterance, e.g., a distinct turn as *Oh God!* does in (19). On the interpretation proffered here, the difference between the two kinds of

<sup>29</sup> Furthermore, [Ljung \(2011: 88–90\)](#) discusses expletive pragmatic markers with reference to their function as gap fillers and slot fillers, a distinction that is ignored here (see Section 3.3.4 on fillers).



interjections is mainly one of valence structure (Section 3.6.3): Whereas reactive interjections are paradigm instances of secondary interjections having (S) as their argument structure, that of pragmatic interjections is typically (S, T). Pragmatic interjections can be hypothesized to mark the first step on the way to the grammaticalization from interjection to discourse marker—note that discourse markers are called ‘pragmatic markers’ by some researchers (Section 3.3). This process of grammaticalization will be described in Section 4.2.2.2.

#### 3.6.4.3 Other classifications

Among the many sub-classifications that have been proposed for interjections one may mention in particular the influential classification by Ameka (1992a), who distinguishes the following kinds of interjections on the basis of their functions:

- (a) *Expressive interjections*.<sup>30</sup> They express the speaker’s emotions, feelings, or state of mind. They are used for the basic emotions like joy (*wow*), surprise (*well, well*), sadness (*aw*), anger (*da*), disgust (*yuck*), or fear (*help*; cf. also Ekman et al. 1972).
- (b) *Phatic interjections*. They express the mental state of the speaker towards the on-going discourse (*ahm, yeah*).
- (c) *Conative interjections*. They are used to get the hearer’s attention or demand the hearer’s response. Thus, the English conative interjection *shh* is used to demand silence.

Ameka and Wilkins (2006) propose an updated version of this classification, for which see Section 5.1.1.

All of the forms discussed in this classification do, in fact, belong to the domain of interactives, but only forms of class (a) are interjections in accordance with the definition in Section 3.6.2. The forms *ahm* and *yeah* in (b), by contrast, are classified here as response signals (Section 3.8). ‘Phatic interjections’ concern what is widely known as phatic communication and include what is discussed in Section 3.9 as social formulae. The form *shh* presented in class (c) is also not treated here as an interjection but rather as a directive (Section 3.2).

And much the same applies to other classifications that have been proposed. For Dixon (2010b: 27), an interjection is ‘a conventionalized cry, typically indicating the speaker’s emotional response,’ expressing the speaker’s internal state (*ouch, whoopee*), a comment (*yuck, wow*), or something to the interlocutor (*hey, yo*). It is only forms of the first two groups (*ouch, whoopee, yuck, and wow*) that are in accordance with our definition of interjections, while the third group (*hey* and *yo*) are classified here as attention signals (Section 3.1).

<sup>30</sup> Franck Floricic (p.c. of July 22, 2021) is wondering whether all interjections are not expressive. The term ‘expressive’ is used by Ameka (1992a) in a more restricted sense.

Similarly, Jackendoff (2002: 240) distinguishes in English between an ‘*ouch* type’ and a ‘*shh* type’: The former is often used non-communicatively and is claimed to be more likely to be uttered involuntarily, while the latter calls for a hearer and is usually under conscious control.<sup>31</sup> In the convention used here it is only forms of the ‘*ouch* type’ which qualify as interjections whereas those of the ‘*shh* type’ are directives.

Goffman (1978: 800) singles out a class of interjections, like *oops*, that he calls ‘response cries.’ But again, not all instances of this class are classified here as interjections. In particular his group of fillers (*uh* and *um*) are assigned here, as well as by a number of other authors, to the type of discourse markers (Section 3.3.4) (see Heine et al. 2021: 260–6 for discussion).

Depending on their discourse functions, various specific subgroups of interjections have been identified. Mention should be made in particular of what has been described as ‘laughter interjections,’ that is, interjections specifically built on the concept of laughing (e.g., Levisen 2019; Andrason 2021b). Rather than representing the product of linguistic strategies mimicking the physiological phenomenon of laughter and being coined spontaneously, they are entrenched and socially stable. They have conventionalized expressive semantics and can be distinguished structurally from other kinds of interjections. And, as Levisen (2019) and Andrason (2021b) demonstrate, laughter interjections can, in fact, be observed cross-linguistically.

Depending on the context in which they are used, laughter interjections may assume functions such as expressing agreement or disagreement. In the following exchange between speakers of the South African Xhosa language, speaker B refuses an invitation by speaker A to smoke weed by using the laughter interjection *ha ha ha*, which in this context is equivalent to the negative response signal *hayi* ‘no’ (Andrason 2021b: 8).

(20) Xhosa (Bantu, Niger-Congo; Andrason 2021b: 8)

- A: U-funa                                      intsango?  
     2.SG.S-want                              NC9.weed  
     ‘Do you want some weed?’  
 B: **Ha ha ha!**  
     INT  
     ‘Ha ha ha!’ (= ‘no!’)

#### 3.6.4.4 Interjections vs. expletives, taboo words, and swear words

Swearing is widely viewed as a type of behavior that society regards as disrespectful, vulgar, and even offensive, and its study is strongly associated with that of interjections. Secondary interjections have been associated with a number of

<sup>31</sup> Concerning the notions ‘control’ and ‘spontaneous’ in the usage of interjections, see also Section 3.6.2.

different terms, like expletives, swear words, or taboo words. Ljung (2011: 79) calls forms like *Jesus!*, *oh my God!*, *Crikey!*, *gee!*, *gosh!*, *shit!*, and *fuck!* expletives, but these forms are also classified by him and some others as paradigm examples of secondary interjections. Gehweiler (2008: 73) refers to forms like *Jesus!*, *oh my God!*, *crikey!*, *shit!* *fuck!* as expletives but adds that they form ‘a subclass of the word-class interjections,’ and Stange (2016) as well treats expletives as emotive interjections:

With respect to their classification, I suggest to view all of them as emotive interjections because the emotive component weighs stronger than the pragmatic function. More specifically, they are emotive interjections with the additional feature of being fully integrated into discourse.

(Stange 2016: 15–16)

Are expletives, swear words, or taboo words in fact interjections and if not, is there a boundary between the two? It is hard to find a conclusive answer to this question in the relevant literature.

Taboo words are defined generally as ‘words with restrictions in their usage,’ where ‘taboo’ ‘refers to a prohibited or restricted social custom’ (Jay 2020: 37). They are said to include concepts of profanity (religious word taboos) or blasphemy, obscenity, slang, sex (for body parts and behaviors), name calling, scatology, disease, racist/gender/ethnic slurs, and vulgarity (Jay 2020: 37–8; see also Ljung 2011: 35; Nübling and Vogel 2004). Other themes figuring in taboo words or swearing include ancestors, specific animal species, diseases, prostitution, death, or terms for roles of kinship relations.

Defining expletives as taboo or semi-taboo expressions used as exclamations, ‘especially in reaction to some strongly negative experience,’ Biber et al. (1999: 1094) restrict the term ‘interjection’ to primary ones and refer to what other people call ‘secondary interjections’ as expletives. But they make a distinction between ‘taboo expletives,’ like *bloody hell*, *damn*, *fuck*, *God*, *Goddamn asshole*, *Goddammit*, *hell*, *Jesus*, and *shit*, and ‘moderated (or euphemistic) expletives,’ like *gee*, *geez*, *good grief*, *good grief*, *heavens*, *heck*, and *my gosh*. Whereas taboo expletives make reference to religion, sex, and bodily excretion, moderated (or euphemistic) expletives differ from the former in that they ‘camouflage their taboo origin by various phonetic modifications (e.g. *gosh* for *God*) or by substitution of different but related words (e.g., *goodness* for *God*)’ (Biber et al. 1999: 1094). The authors furthermore assume that moderated expletives are ‘socially acceptable in many situations where some taboo expletives are ruled out’ and are similar in function to interjections such as *oh* and *wow*, in that they may express a generalized reaction of surprise or emotional involvement (Biber et al. 1999: 1095; see Gehweiler 2008 for detailed discussion and analysis). We will return to this issue in Section 4.3.2.

The concern of this book is with interjections as a type of interactives, as defined in Section 3.6.2, and other terms are used only when referring to studies using these terms. Accordingly, what has been called ‘expletives,’ ‘swear words,’ or ‘taboo words’ are included here only to the extent that they conform to that definition.

### 3.6.5 Discussion

While interjections are usually understood to form a closed class, Norrick (2009: 868–9) argues that English interjections are an open-class category like nouns and verbs, and that this applies to both primary and secondary interjections: They ‘represent a large, potentially infinitely extendable class of items’ (Norrick 2009: 866). In some languages, the number of interjections is in fact large; Andrason and Dlaki (2020) counted nearly 350 of them in the Xhosa language of South Africa. Nevertheless, it would seem that the number of interjections is cross-linguistically distinctly smaller than that reported for ideophones in some languages (Section 3.5).

## 3.7 Response elicitors

Usually subsumed under the label of ‘interjections,’ sometimes also under ‘discourse markers,’ response elicitors are classified here as a type of interactives of their own. We follow Biber et al. (1999: 1089) in assuming that they express a discourse function distinct from interjections and other types of interactives, adopting both the term and the concept from these authors.

### 3.7.1 Introduction

Biber et al. (1999: 139, 1089) characterize response elicitors ‘as generalized question tags,’ which include *alright?*, *okay?*, *right?*, *see?*, as well as *huh?* (especially in American English) and *eh?* in British English, as in (1). The last two are illustrated in (1). Differing from clausal question tags (*isn’t it*), which invite agreement or confirmation from the hearer, response elicitors (*eh?*, *huh?*, *okay?*, *right?*, *see?*) seek a signal that the message has been understood and accepted (Biber et al. 1999: 1089).

- (1) English (Biber et al. 1999: 1089)
- a Jordan what’s the matter? What’s the matter **eh?** (BrE)
  - b **Oh hi**, you’re Brent’s, you’re Brent’s older sister, **huh?** Your brother’s so cool (AmE)

This distinction is not made here; rather, we will be subsuming both under the same, more general type of response elicitors, for the following reasons. First, we did not find clear cross-linguistic support for the distinction; in the languages that we were able to survey, both appear to be treated essentially the same way. And second, the extent to which response elicitors have a role of inviting agreement or confirmation from the hearer or seek ‘a signal that the message has been understood and accepted’ varies in most forms depending on the context in which they are used and it seems to be a matter of degree rather than of kind, and so appears to be the difference between clausal question tags and other response elicitors.

Nevertheless, English question tags, such as *doesn’t he?*, *won’t you?*, *have they?*, etc. are not a prototypical representative of the type because they do not conform to our definition of interactives in Section 1.2: Rather than being invariable forms, they exhibit a range of morphological variants and thus contrast with prototypical instances of the type.

### 3.7.2 Working definition

We propose the following definition for response elicitors:

- (2) A response elicitor is an interactive used to seek the hearer’s agreement or confirmation of a remark just made.

Examples of English forms serving at least in many of their usages as response elicitors are *eh?*, *huh?*, *isn’t it?*, *okay?*, *right?*, *see?*. Biber et al. (1999: 1089) single some of these (*huh?*, *okay?*, *eh?*, and *see?*) out as being ‘somewhat peremptory: they are typically used in familiar, casual exchanges, and would seem rude if they occurred in a formal situation.’

A detailed analysis of English response elicitors such as *eh?* and *huh?*, referred to by the author as confirmationals, is found in Wiltschko (2021, Chapter 5). For a list of functions that have been proposed for *eh?*, see Wiltschko (2021: 95).

As a rule, the speaker is seeking the hearer’s agreement or confirmation of her own remark. But we will also include expressions reacting to remarks made by the previous speaker, where the speaker may signal doubt, disbelief, or surprise on what the previous speaker just said. A paradigm instance of an English response elicitor of this kind is *really?*, as in the following example.

- (3) English (Biber et al. 1999: 1071)
  - A: **So** do you think an alligator would like salt water?
  - B: It would probably kill him **wouldn’t it?**
  - A: That one on the news has been out in the ocean for a while.
  - B: **Really?**

Like interjections, response elicitors also exhibit a distinction between primary and secondary forms in English, see Section 3.6.4.1.

### 3.7.3 Grammar

In the following example, use of the response elicitor *eh?* presupposes a speaker asking a hearer, and the usage of *eh?* would be incomplete without the preceding T argument *Might as well get rid of it*. Thus, the descriptive content of a response elicitor like *eh?* in (4) can be paraphrased as in (5a), suggesting that there are three arguments, cf. (5b).

- (4) British English (Biber et al. 1999: 1089)

Might as well get rid of it **eh?**

- (5) a Paraphrase: Do you (H) agree with what I (S) have just said (T)?

b Argument structure: *eh?* (S,H,T)

Wiltchko (2021: 101) observes that the core functions of confirmationals, that is, response elicitors like *eh?* and *right?*, are (a) to request a response and (b) to introduce a propositional attitude. On the perspective proposed here, (a) relates to the descriptive content of response signals as captured by the paraphrase in (5a) whereas (b) concerns their T argument, which usually, though not necessarily, has a propositional format, as *Might as well get rid of it* in (4) in fact does.

Response elicitors differ from many other types of interactives in three important ways. First, they have the format of questions; accordingly, they are typically pronounced with an interrogative intonation. Second, when combining with text pieces of sentence grammar their position is typically utterance final or at the right periphery of an utterance. In this respect they resemble ideophones, which are also likely to appear utterance finally (Section 3.5).

And, their argument structure differs from that of many other interactives in containing three arguments: For one thing, they share with other interactives in having an obligatory but covert S (speaker) argument. For another, there is also a covert H (hearer) argument referring to the hearer or hearers from whom agreement or confirmation is sought. And finally, there is an obligatorily expressed T argument, standing for the content of the preceding text piece to which the response elicitor refers. For example, the T argument of the response elicitor *see?* in the English text piece *It's like a magnet obviously see?* (Biber et al. 1999: 1089) is provided by the utterance *It's like a magnet obviously* preceding *see?*. But T can be instantiated by text pieces of various size and complexity, extending from phrases (e.g., *On Sunday, right?*) to whole paragraphs or even longer pieces of text.

Response elicitors usually carry the intonation contour characteristic of question formation. But this need not be the case. For one thing, their intonation contour may gradually be lost due to frequent use. And for another, there may be a special intonation contour reserved for response elicitors. In the Paamese language of Vanuatu, for example, the response elicitor (question tag) *aa* has a sharply rising intonation apparently not found in other interrogative constructions, cf. (6).

- (6) Paamese (Oceanic, Austronesian; Crowley 1982: 241; intonation not marked in the original)

kaiko ko+tāhosi      **aá?**  
 2.SG 2.SG+REAL.good Q  
 ‘You’re good, aren’t you?’

In Standard Chinese (Mandarin) there is a family of X-NEG-X structures (or A-not-A questions; Li and Thompson 1981: 546) serving as response elicitors. The following example illustrates the form *shì bu shì* (be not be) ‘right?’, where S and H are also implied while the T argument (*taā zài gēng - tián*) is overtly expressed:

- (7) Standard Chinese (Sinitic, Sino-Tibetan; Li and Thompson 1981: 546)

tā zài gēng - tián, **shì bu shì?**  
 3:SG DUR plow field be not be  
 ‘S/he is plowing the field, right?’

In accordance with its function as a ‘response elicitor,’ T is almost invariably placed before the interactive eliciting the response. The reverse order is, however, not entirely ruled out. There are examples showing that the response elicitor can precede its T argument. Thus, one may occasionally hear text pieces like *Right? you’re Swiss?* in English, and similar observations can be made in other languages. In the Akie language of north-central Tanzania, the response elicitor *máleena* ‘isn’t it?’ (lit. ‘doesn’t it say this?’) commonly occurs at the right periphery of an utterance, cf. (8a). But it may as well appear at the left periphery, as in (8b), in which case T must be an interrogative clause introduced by the negation particle *ma*, and no copula (*ko*) may be used:

- (8) Akie (Southern Nilotic, Nilo-Saharan; König, Heine, and Legère 2015a: 120)

a ko chicheε púuní, **máleena.**  
 COP 3.PL.A Maasai.N isn’t it?  
 ‘They are Maasai, aren’t they?’

b **máleena,** má (\*ko) chicheε púuní?  
 isn’t.it NEG (\*COP) 3.PL.A Maasai.N.Q  
 ‘Isn’t it (true), they are Maasai?’

Furthermore, response elicitors are also found occasionally utterance-internally, as in the following English example, where the question tag *isn't it* is inserted within a noun phrase.

- (9) English (Dehé and Braun 2013: 131)

It's a mixture **isn't it** of original instruments.

In narrative discourse, one of the most common sources of response elicitors is provided by perception verbs for 'hear' and 'see,' employed by the narrator to keep the audience involved in the unfolding text. Examples are provided by the interactives *aríí kas ira* (have you heard) 'you understand?' and *arii sowé íra* (have you seen) 'you see?' in the Akie language of north-central Tanzania (König, Heine, and Legère 2015b) or *à tùéí* 'you hear' in the West African language Kisi (Childs 1995: 268).

### 3.7.4 Functional space

The term 'response elicitor' implies that instances of this type of interactives are expected to elicit responses from the hearer. However, these forms tend to show a range of context-dependent usages, not all of which may be in accordance with the definition in Section 3.7.2. A look at the functions of English question tags may illustrate this claim. Building on a wide range of studies that have been carried out on the colloquial usages of these question tags both in American and British English, Tottie and Hoffmann (2006) identify six main functions expressed by them, namely the ones in (10). The authors found nine times as many tag questions in colloquial British English as in colloquial American English.

- (10) Functions of question tags in colloquial American and British English (Tottie and Hoffmann 2006: 300–1)

- a Aggressive
- b Attitudinal
- c Confirmatory
- d Facilitating
- e Informational
- f Peremptory

Perhaps interestingly, the informational function, which concerns genuine requests for information, apparently is of a more marginal nature, covering only 3% of all uses. Clearly most central are the confirmatory, the attitudinal, and the facilitating functions, which together account for over 90% of all uses.

In confirmatory uses the speaker is not sure of what s/he says, and therefore wants confirmation, as in (11).



- (11) American English ((LSAC); [Tottie and Hoffmann 2006](#): 300)

A: I'm gonna try to go walking for a little bit. I don't need a jacket, **do I**?

B: **No**, it's still pleasant.

Attitudinal uses emphasize what the speaker says and do not expect involvement, or a reply by the hearer. This is different with facilitating uses, where the speaker is sure of the truth of what s/he says but nevertheless wants to involve the hearer, as in (12).

- (12) American English (BNC-SDM; [Tottie and Hoffmann 2006](#): 300)

Teacher: Right, it's two **isn't it**?

Pupil: **Mm**.

With a peremptory use, a speaker is assumed to intend to close off a debate, while aggressive uses function as an insult or a provocation.

To our knowledge there are no detailed comparative studies on the functions of response elicitors. It is therefore hard to know whether or to what extent the range of functions covered by English question tags reflects a cross-linguistically more widespread pattern.

Uses of response elicitors differ greatly from one another in the extent to which the hearer is expected to react to the use of the interactive forms. At one end of the scale there are forms where the speaker is highly likely to expect a response from the hearer—that is, where the form is interpreted as a question. At the other end there are forms where the speaker does not expect, or even welcome a response. The latter concerns response elicitors or specific usages of them where in the model of [Wiltschko \(2021: 104\)](#) no response layer is involved. We are not aware of any substantial information on where a given use is located along such a scale. With regard to response elicitors derived from alternative connectives for 'or?', [Heine, Kaltenböck, and Kuteva \(2015: 124\)](#) propose the three-stage scenario in (13).

- (13) A three-stage scenario of grammaticalization of 'or?' as a response elicitor (REE; based on [Heine, Kaltenböck, and Kuteva 2015: 124](#))

I: REE serves to invite the hearer to take the floor.

II: REE is no longer clearly inviting the hearer to take the floor; rather, it now serves mainly functions relating to the situation of discourse.

III: The hearer is no longer expected to take the floor. REE now serves text-structuring functions, e.g., as a filling device for online planning, as a boundary marker, etc.

Differences in the functions of response elicitors appear in fact to be mainly due to grammaticalization, and the three stages in (13) can be taken to reflect different degrees of grammaticalization, with stage III being suggestive of the final stage of the process. The more a response elicitor is grammaticalized the wider the range of contexts is with which it is associated, the more it loses its interrogative force, and the less it will signal that the speaker expects a verbal response. For example, the

English response elicitor *right?* in (14) is suggestive of Stage I, inviting the hearer to take the floor.

(14) American English (Biber et al. 1999: 1089)

A: You know who Stan is, **right?**

B: I've heard his name.

In other contexts, response elicitors are grammaticalized to the extent that a response by the hearer is not expected, or even not welcomed. The example in (15a) can be taken to represent a grammaticalized Stage-II use of the response elicitor *right?*. The example is taken from the Howison Lecture in Philosophy that the philosopher David Kaplan delivered on semantic theory at the University of California at Berkeley in 1999. In the course of the lecture, Kaplan (1999) used *right?* eighteen times, and considering the formal context in which this discourse took place, the audience does not appear to have been invited to respond. A usage close to Stage III can be seen in (15b), where the main function of *right?* appears to be to signal the end of a paragraph. While still involving the hearer, *right?* jointly with the discourse marker *OK* marks a boundary, preparing the hearer for a new topic of discourse.

(15) English (Kaplan 1999: 16; boldening ours, italics in the original)

a I mean, the question is, could we eliminate expressives? Why do we need expressives? Do we need expressives? There's a similar, very interesting question, can we eliminate indexicals from our language, **right?**

b But when you meet the queen, you can say, *Nice day, Your hat's on funny, you know*, but you can't say *Who does your hair?* **right?** OK.

Still, an interrogative function cannot entirely be ruled out in (15b). But this is no longer the case in the Stage-III example of (16), where the hearer could not possibly be able to volunteer an answer, having just met the speaker.

(16) English (Brown and Levinson 1987: 119 (78))

I had a really hard time learning to drive, **didn't I**.

Strongly grammaticalized response elicitors suggestive of Stage III, in fact, have the appearance of discourse markers, and there is remarkable overlap between response elicitors and discourse markers (see Section 4.2).

A more marginal kind of a response elicitor is provided by expressions used by an interlocutor as a response to an address call made by the previous speaker, inquiring about the reason for being called, translatable roughly as 'Yeah? What's the matter?'. Some languages have a marker dedicated to this function. The following example is taken from the Tigemaxo language in Mali, where the expression is *ɔhʒ*, illustrated in the following text piece:

(17) Tigemaxo (Mande, Niger-Congo; [Blecke 1996](#): 189, B.217)

... a wo n juʃaalu a wo ɕhɔ̃?  
3.SG say 1.SG girl 3.SG say REE  
‘He said: My girl. She said: Yeah (what’s the matter)?’

Another somehow marginal type of a response elicitor is discussed in the next sections.

3.7.5 Other-initiated repair markers

The form *huh?* is an instance of an English response elicitor, as can be seen in (18). In accordance with what appears to be its most salient function it can loosely be called a ‘confirmation seeker.’

(18) American English ([Biber et al. 1999](#): 1089)  
**Oh hi**, you’re Brent’s, you’re Brent’s older sister, **huh?**

But in addition, it also has an important alternative usage as what is described by [Dingemanse, Torreira, and Enfield \(2013](#): 7) as an ‘other-initiated repair marker,’ occurring for example when one has not clearly heard what the other speaker just said (see also [Enfield 2017](#), Chapter 8). In this usage, one participant produces a turn at talk, the other then signals some trouble with this turn, and finally the first speaker produces a next turn which aims to resolve the trouble, usually by means of repetition and/or slight modification; [Table 3.15](#) depicts the structural schema of discourse organization concerned, an example is found in (19).

**Table 3.15** Structural schema imposed by other-initiated repair (OIR) markers in conversations

| Turn | Speaker | Content of text                                  |
|------|---------|--------------------------------------------------|
| 1    | A       | Text piece believed by B to be in need of repair |
| 2    | B       | <b>OIR marker</b>                                |
| 3    | A       | Text piece proposing repair                      |

Source: Based on [Dingemanse, Torreira, and Enfield \(2013\)](#)

(19) American English ([Biber et al. 1999](#): 1092)  
A: And then what brought them here, **I wonder**.  
B: **Huh?**  
A: What brought them here?

The other-initiated repair usage, as in (19), differs from the confirmation-seeking usage of response elicitors in three ways. First, its functions are quite different: Whereas the former can be said to have an ‘information seeking’ function, that

of the latter is primarily confirmation seeking, at times geared to strengthen the relationship between speaker and hearer (e.g., ‘are you still with me?’). Second, the expected response by the hearer is different: Rather than simply triggering a polar response form, like *yes* or *no*, other-initiated repair signals are expected to elicit a more complex response. And third, the T argument (*And then what brought them here, I wonder*) and *huh?* in (19) are not produced by the same speaker but by different speakers (but see also Section 3.7.2). Thus, *huh?* as an information seeker constitutes a reaction of speaker B to what the previous speaker A had said.

Note that *huh?* is not the only English other-initiated repair marker, another common instance is provided by *pardon?*, as in example (20). Cross-linguistically, the two functions, information seeking and confirmation seeking, tend to be expressed by different forms of response elicitors.

(20) English (Biber et al. 1999: 21)

<Watching television>

[...]

A: Who wins?

B: **Pardon?**

A: Who wins?

B: I don’t know (CONV)

All evidence that there is suggests that *huh?* as an other-initiated repair marker conforms to the definition of interactives in Section 1.2. It is treated by Dingemanse, Torreira, and Enfield (2013) as an ‘interjection,’ and it is part of a larger paradigm of English expressions for other-initiated repair, including *what?* and *pardon?*. And there could be justification to conceive this set as a separate type of interactives distinct from confirmation seeking response elicitors. However, on account of the fact that the central function of both is to elicit a response from the hearer, and that the two have the same argument structure, namely that in (5b), we tentatively classify other-initiated repair forms as a case of response elicitors. Thus, one way of paraphrasing *huh?* as an other-initiated repair marker would be as in (21a), having the same argument structure as confirmation seeking *huh?*, cf. (21b).

(21) a Paraphrase: I (S) have not clearly heard what you (H) just said (T).

b Argument structure: *huh?* (S,H,T)

Dingemanse, Torreira, and Enfield (2013: 7) say that the form is ‘a question word devoid of semantic content.’ This is not the interpretation proffered here. If in fact the question word were devoid of semantic content it would be hard to understand why the hearer reacts exactly the way she does—that is, one would not expect *huh?* to necessarily trigger exactly the kind of response requested. For example, if *huh?* in (19) were replaced by response elicitors like *isn’t it?*, *okay?*, *right?* or

*see?*, or by some other kind of interactive for that matter, the response would presumably not be the same, or there would be no response at all. Rather, we argue that in addition to its expressive meaning, *huh?* has descriptive content roughly as sketched in the paraphrase of (21a) (Section 2.1). In more abstract terms, this content can be described thus: With the use of *huh?* in (19) the speaker expresses (a) that she has a problem and (b) that she invites the hearer to solve that problem.

Importantly, [Dingemanse, Torreira, and Enfield \(2013: 8\)](#) classify other-initiated repair markers as a universal word, which is likely to be attested in similar form across languages:

All interjection tokens in all languages in our sample are syllable-like utterances consisting of one vowel-like sound optionally preceded by a consonant-like sound. We refer to these elements as syllable, vowel, and onset. We never found forms longer than one syllable and we never found any final consonantlike sounds.

([Dingemanse, Torreira, and Enfield 2013: 7](#))

The authors found such ‘interjections’ in fact in 31 languages all over the world. The forms listed in [Table 3.16](#) represent a selection of these forms.

**Table 3.16** A sample of ‘other-initiated repair markers’

| Language         | Where spoken<br>(main country) | Genetic affiliation         | Form  |
|------------------|--------------------------------|-----------------------------|-------|
| Cha'palaa        | Ecuador                        | Barbacoan                   | ʔà:   |
| French           | France                         | Romance, Indo-European      | h[é̃] |
| Icelandic        | Iceland                        | Germanic, Indo-European     | há    |
| Kikongo          | D.R. Congo                     | Bantu, Niger-Congo          | é     |
| Korean           | S. Korea, N. Korea             | Isolate                     | é     |
| Lahu             | China, Thailand                | Tibeto-Burman, Sino-Tibetan | hâi   |
| Lao              | Laos                           | Tai, Kra-Dai                | hâi:  |
| Mandarin Chinese | China                          | Sinitic, Sino-Tibetan       | á:    |
| Murrinh Patha    | N. Australia                   | Southern Daly?              | á:    |
| Siwu             | Ghana                          | Kwa, Niger-Congo            | á:    |
| Spanish          | Spain                          | Romance, Indo-European      | é     |
| Tzotzil          | Mexico                         | Tzeltal-Tzotzil, Mayan      | é     |
| ǀAkhoe ǁaillom   | Namibia                        | Khoe (‘Central Khoisan’)    | hé    |

Source: [Dingemanse, Torreira, and Enfield \(2013: 3\)](#)

Note:  $\acute{x}$  = rising pitch,  $\grave{x}$  = falling pitch)

The findings on other-initiated repair markers presented by [Dingemanse, Torreira, and Enfield \(2013\)](#) are noteworthy. They presumably do not constitute an isolated case, as we saw in Section 3.1.5 with reference to the attention signal \**hey*.

What such observations might suggest is that there is some cross-linguistic base underlying the structure of primary interactives. The exact nature of such a base is in need of further investigation.

### 3.8 Response signals

Let what you say be simply ‘Yes’ or ‘No’; anything more than this comes from evil.

(Matthew 5:37)

This citation from the New Testament may be taken to be an early illustration of the significance that interactives have in social interaction: ‘Yes’ and ‘no,’ which are the main but not the only subject of the present section, are paradigm instances of response signals. In formal linguistics, they tend to be described as assigning a truth value to the propositional content of a preceding question. Response signals are not only useful, they are also efficient, and also highly frequent tools of social interaction. Thus, [Enfield \(2017\)](#) writes:

The average time that people take to respond to a question is about the same time that it takes to blink the eye: 200 milliseconds.

([Enfield 2017](#): 1)

Like in the case of response elicitors, the concept of response signals is taken from [Biber et al. \(1999: 1089\)](#), who refer to them as ‘response forms.’ Response signals have received quite some coverage in connection with research on question-response sequences (e.g., [Stivers and Enfield 2010](#); [Enfield, Stivers, and Levinson 2010](#); [Enfield et al. 2019](#)). An especially detailed account is found in [Wiltschko \(2021, chapter 6\)](#). While her concern is with the syntax of response signals, called by her response markers, she provides an insightful analysis of their functions and constraints of usage.

Response signals are widely classified as interjections (e.g., [Bloomfield \(\[1933\] 1962; Ameka 1992a; Childs 1995: 145–7; Ameka and Wilkins 2006\)](#), less commonly also as discourse markers (e.g., [Lee-Goldman 2011](#)). But in accordance with the framework and the definitions proposed here they need to be distinguished from both interjections and discourse markers (Section 3.8.2).

English response signals such as *mhm*, *uh-huh*, and *yeah* are treated by [Ameka \(1992a: 114\)](#) as ‘phatic interjections,’ used for ‘backchanneling or feedback signalling vocalisations,’ and in fact, backchanneling is a common function of response signals. Backchannels, or ‘continuers,’ signal that some part of a telling has been received and that more is awaited, thus encouraging the previous speaker to continue (see also [Stange 2016: 12; Wiltschko 2021: 46–7](#)).

### 3.8.1 Introduction

Response signals are paradigm examples of stance markers. According to the framework of stancetaking of [Du Bois \(2007\)](#), they instantiate the concept of alignment: ‘More commonly, speakers show alignment by stance markers like *yes* or *no*, or gestures like a nod or a headshake, or any number of other forms that index some degree of alignment’ ([Du Bois 2007](#): 144).

Following [Biber et al. \(1999: 1089–92\)](#), the term ‘response signal’ is used here in a restricted sense for forms typically expressing confirmation or agreement or their negative counterparts. The main instances of response signals in the languages of the world are expressions for ‘yes’ and ‘no’. The two do not only constitute contrastive responses, their structure also differs in a number of ways. For example, [Enfield \(2017: 74\)](#) found that in English conversations, ‘yes’ answers had an average delay of just 35 milliseconds while that of ‘no’ answers was almost twice as long to arrive.

Some responses to greetings (e.g., *fine* in *How are you?—Fine.*) can in a wider sense be said to also belong to the type of ‘response signals.’ On account of their specific restriction to social rituals, however, greeting formulae are not included here but are discussed as parts of social formulae in Section 3.9.

And some forms for response signals can also be extended to serve as response elicitors, in which case they change from a declarative to an interrogative intonation. Thus, in the following text example, the interactive *right* occurs both as a response elicitor (1, A) and as a response signal (1, B):

- (1) English (Frown L05: 35; [Méndez-Naya 2006](#): 146, (13); [Heine et al. 2021](#): 127)
- A: Just a small point to be verified. No reading glasses were found, **right**?
- B: **Right.** Just sunglasses.

While forms for ‘yes’ and ‘no’ have contrasting functions, they exhibit cross-linguistically a wide range of functions far beyond those of signaling confirmation and disconfirmation (e.g., [Lee-Goldman 2011](#)). Accordingly, [Wiltschko \(2021\)](#) concludes:

Most analyses of response markers focus on their use as answers to yes/no-questions. In this context, they seem to be intrinsically associated with polarity values. But this is not their only use by far. Across many unrelated languages, response markers, which may be used to answer a polar question, can also be used to react to other initiating moves, including declaratives, wh-questions, imperatives, and exclamations, as well as non-verbal situations. When responding to such moves, *yes* and *no* clearly do not denote the polarity of a proposition. Rather they may serve as agreement markers, markers of compliance, or simply markers of listening.

([Wiltschko 2021](#): 156)

Furthermore, the two can combine into one signal for specific discourse functions in some languages. An especially interesting case is provided by the combination *yeah-no* in Australian English (Burridge and Florey 2002; Moore 2007; Wiltchko 2021: 174–5), and a similar combination, *ʔäre-na* ‘yeah-no,’ is found in Persian (Ghaderi 2021). In colloquial German, the combination of *ja* ‘yes’ and *nein* ‘no’ results in a blended form *jein* expressing doubt or a ‘middle position’ in one’s standpoint to a particular state of affairs (Tania Kuteva, p.c. of February 7, 2022).

### 3.8.2 Working definition

Biber et al. (1999) describe response signals (‘response forms’) in the following way:

Response forms are inserts used as brief and routinized responses to a previous remark by a different speaker. These include responses to questions (typically *yes*, *no*, and their variants), responses to directives (e.g. *okay*), and responses to assertions (e.g. **backchannels** such as *uh huh*, *mhm*).

(Biber et al. 1999: 1089; italics and bold print in the original)

Based on this characterization, we propose the following definition for response signals:

- (2) A response signal is an interactive used as a concise response to a previous remark by a different speaker, typically expressing confirmation or disconfirmation.

This definition takes care of prototypical instances of response signals but is at variance with many of their context-dependent usages (see Wiltchko 2021, chapter 6 for detailed discussion). Instead of or in addition to confirmation, acceptance, acquiescence, agreement, assent, etc., a response signal may express a range of other functions, such as responses to vocatives (e.g., *Waiter!—Yes, Sir!*). Examples of English forms serving in many of their usages as response signals are:

Assent, agreement: *absolutely, alright, aye, certainly, exactly, I see, indeed, mhm, mm, of course, okay, okie-dokie, right, sure, uh huh, yeah, yep, yes, you bet*

Dissent, denial, disagreement: *certainly not, definitely not, huh uh, m'm, no, nope, not at all, not likely, not yet, unh unh*

As this listing suggests, response signals for confirmation outnumber those expressing disconfirmation. Choice of response signals tends to be contingent



on the kind of challenge expressed by the previous speaker (see Biber et al. 1999: 1089 on English). According to Brown and Levinson (1987: 271), another parameter determining the choice between functionally related forms is provided by distinctions of politeness. Thus, English *yes*, *no*, and *all right* are said to be associated with negative politeness and *yup*, *uh-huh*, *nope*, and *sure* with positive politeness.

Some languages have rich inventories of response signals, distinguishing degrees of emphasis, politeness, or other functions. In the Mayan language Tzeltal of Mexico, for instance, the inventory includes the following expressions: *haʔuk* ‘so it may be,’ *hičuk* ‘thus it may be; so let it be; as if it were so,’ *mančuk* ‘OK, it’s all right,’ *yak* ‘yes,’ and *yakuk* ‘OK, I agree’ (Brown and Levinson 1987: 158).

Response signals may be either solicited or unsolicited. In the former case they typically constitute the second part of an adjacency pair in a question-answer sequence (Schegloff and Sacks 1973; Schegloff 2007).<sup>32</sup> In the latter case they are not required, nor may they be expected, as in the following case.

- (3) English (Biber et al. 1999: 77)

“Oh, how good you are to me!” Yes, I am, she told herself; though she stoutly denied it to his face. (FICT)

Like other interactives, response signals are as a rule short, invariable forms. In many African languages, ‘yes’ is or can be expressed by a simple low or mid-front vowel and ‘no’ by a repetition of that vowel with an interspersed glottal stop. Thus, in the Akie language of north-central Tanzania, the main response signals are *àà* ‘yes’ and *áʔà* ‘no’ (König, Heine, and Legère 2015a: 117–18), in the Dhaasanac language of southern Ethiopia *ʔěě* ‘yes’ and *ʔěʔě* ‘no’ (Tosco 2001: 249), and in the unrelated West African language Kisi it is *ěě* ‘yes’ and *(ʔ)ěʔě* ‘no’ (Childs 1995: 146). Alternatively, there may be a bilabial nasal instead of a vowel expressing the same distinction, as is the case in the Kanyok language of the Democratic Republic of the Congo: *m̩m̩* ‘yes’ vs. *m̩-m̩* ‘no’ (Stappers 1986: 68).

Some languages exhibit sets of forms distinguished on the basis of sociolinguistic parameters. Korean response signals, for example, are sensitive to degrees of politeness, as Table 3.17 shows.

Response signals belong to the types of interactives that are strongly associated with body movements—in this case almost invariably with head movements (e.g., nods; see Enfield et al. 2019).

Like interjections, response signals also exhibit a distinction between primary and secondary forms in English (3.6.4.1).

<sup>32</sup> ‘The structure of an adjacency pair spans utterances that are produced by two people in interaction. An adjacency pair has two PAIR-PARTS. The first pair-part is spoken by Person A, the second by Person B. Examples include sequences in which a greeting elicits a return greeting, or in which a question elicits an answer’ (Enfield et al. 2019: 282); see also Section 7.3.2.

**Table 3.17** Degrees of politeness in Korean response signals based on speech levels

| Degree of politeness | ‘Yes’          | ‘Yes, that’s right’    | ‘No’               | ‘Not at all’         |
|----------------------|----------------|------------------------|--------------------|----------------------|
| Plain                | <i>onya</i>    | <i>kuleh-ta</i>        | <i>ani-ta</i>      |                      |
| intimate             | <i>ung</i>     | <i>kulay</i>           | <i>ani(-ya)</i>    | <i>chenman-ey</i>    |
| Familiar             |                | <i>kuleh-ney</i>       | <i>ani-ney</i>     |                      |
| Polite               | <i>yey/ney</i> | <i>kulay-yo</i>        | <i>ani-eyyo</i>    | <i>chenman-ey-yo</i> |
| Deferential          |                | <i>kuleh-sup-ni-ta</i> | <i>ani-p-ni-ta</i> |                      |

Source: Sohn (1999: 399)

### 3.8.3 Grammar

Response signals appear to occur cross-linguistically quite commonly as stand-alones, forming utterances of their own. For example, Ghaderi and Amouzadeh (2020: 15) found the Persian response signal *bale* ‘yes’ to show the following distribution of relative frequency in telephone conversations: 67.9% as stand-alones (‘detached’), 20.5% at the left periphery, 9.5% at the right periphery, and 2.1% in medial position.

Response signals have a complex argument structure. Take the following example:

- (4) American English (Biber et al. 1999: 1089–90)

A: Does someone have a pencil?

B: **Yeah**, here.

The use of the response signal *yeah* implies a speaker (S) argument, who is B in (4), and a hearer (H) argument, who is A in (4) to whom *yeah* is directed. And there is also a T argument standing for the utterance made by A, namely *Does someone have a pencil?*. That H and T are obligatory participants of *yeah* is suggested by the following: If the utterance were made by B in soliloquy, that is, in the absence of speaker A (= H), it would not normally be meaningful. And being a response to the question raised by speaker A, *yeah* would also not be meaningful if T were missing, that is, if there were no preceding question.

The usage of the response signal *yeah* in (4) can therefore be described as in (5), where (5a) proposes a paraphrase and (5b) the argument structure for *yeah*.

- (5) a Paraphrase: I (S) answer your (H) question (T) in the affirmative.  
 b Argument structure: *yeah* (S,H,T)

Mention may also be made that the response by B in (4) contains another participant, namely an adjunct (see Section 2.2.5): The adverb *here* is an optional participant which is not required by the valence of *yeah*: If it were omitted the utterance would nevertheless remain well-formed. As was observed in Section 2.2.5, English response signals, like *yes* and *no* and their variants, are particularly prone to take such adjuncts. Adjuncts, referred to by Wiltschko (2021: 147) as reacting moves typically following the response marker, present in most cases information elaborating on the usage of the response signal.

While T is an obligatory participant of discourse, in some cases it may be provided by the extra-linguistic context. In example (6) from Persian it is the particular situation where the response signal *bale* ‘yes’ on its own suffices to provide the information normally expressed by some text piece instantiating T:

- (6) Persian (Ghaderi and Amouzadeh 2020: 13, (25))  
 [Pilot assistant addressing a cowardly amateur paratrooper who refuses to jump]  
**bale! bale!**  
 yes yes  
 ‘Yes! Yes!’

The T argument is most commonly, but by no means necessarily, provided by a question. In (7a) it is a request (though phrased in terms of a question) and in (7b) it is an assertion, and the latter also applies to (8).

- (7) American English (Biber et al. 1999: 1090)  
 a A: Can you reach me three of those small tins – at the top – the processed ones.  
 B: **Yeah.**  
 b A: People don’t like to sit in banquets for a long time.  
 B: **Unh unh.**
- (8) English (Biber et al. 1999: 223)  
 A: Well – the good news for the environmentalists is the bike runs on unleaded.  
 B: **Mhm.** (CONV)

*Yeah* is most frequently used to acknowledge the receipt of information that is new to the discourse but consistent with current active information. Paradigm examples are provided by response signals serving as responses to questions. However, this function accounts for only one part of its usage. Functioning as backchannel is another important area of use, where the response signal presents feedback to the speaker that the message is being understood and accepted (Yngve 1970; Wiltschko 2021: 180–1). Accordingly, Biber et al. (1999) observe:

Given the interactive nature of conversation, backchannels are important in indicating that speaker and hearer are keeping in touch with one another, and that communication is still in progress.

(Biber et al. 1999: 1091)

Response signals serving backchannel functions are found in languages across the world (e.g., Clancy et al. 1996; Cecil 2010). English examples are *mhm*, *mm*, *uh huh*, and *yeah* and their negative counterpart *huh uh*, consider the following exchange (for a discussion of English *mm-hmm* and *uh-huh*, see Enfield 2017: 107–12):

(9) British English (Biber et al. 1999: 1091)

A: and their lawn is high up, up to the road, [you know]

B: **Mm**

C: **Yeah**

A: so he dug up the lawn and put a little garage in.

Furthermore, signaling acquiescence to a suggestion or proposal made by the previous speaker is also a common function of specific response signals, like *ngoi* ‘O.K.’ in the Mayali language of Western Arnhem Land, Australia:

(10) Mayali (Gunwinyguan, Arnhem; Evans 1992: 233, (10))

A: **Mah**, yi-wayini!

let.it.happen.now 2.SG-sing

‘Come on, sing that song!’

B: **Ngoi**, galuk nga-wayini.

O.K. FUT 1.SG-sing

‘O.K., I’ll sing it.’

Prototypically, response signals are reactions to polar (‘yes/no’) questions but depending on the perspective adopted by the hearer on the preceding T argument, they may even express reactions to wh-questions, as pointed out by Wiltschko (2021: 161):

(11) American English (BB-2012–05-23, BB-2012–03-27; Wiltschko 2021: 161–2, (29), (32))

a Katie: Why would he do something like that?

Brooke: **Yes**, I know. That is the question.

b Bill: What do you want to bet?

Liam: **No**, I am not playing this game with you.

While response signals typically express confirmation or disconfirmation, or agreement or disagreement, they may also signal more qualified and complex reactions. For example, the German response signal *hm* expresses a number of

different functions depending on the tone pattern it adopts, as shown in Table 3.18 (see also Table 1.3 of Section 1.3.1.4).

**Table 3.18** Tonological variants of the German response signal *hm*

| Form         | Tone pattern   | Approximate translation | General meaning         |
|--------------|----------------|-------------------------|-------------------------|
| <i>hm̃</i>   | Falling        | 'that's strange'        | Perplexity              |
| <i>hm̂</i>   | Rising         | 'why that?,' 'really?'  | Doubt, divergence       |
| <i>hm̃m̂</i> | Falling-rising | 'I agree,' 'correct!'   | Agreement, confirmation |

Source: based on Ehlich (1986)

Note: Tone marking is in accordance with that commonly employed for tone languages

In accordance with the definition in Section 3.8.2, expressing confirmation or disconfirmation are the most prototypical functions of response signals. But they are not the only ones; indicating one's interest, surprise, or doubt on what the previous speaker had to say are also functions signaled by response signals, such as *really* in (12).

(12) American English (Biber et al. 1999: 1092)

A: She's real good, the people like her real well over at the hospital.

B: **Really?**

Biber et al. (1999: 1092) also include *huh?*, *eh?*, and *what?* as having response signal functions:

A more negative type of feedback is supplied by question forms such as *huh?* and *eh?* (as well as the question word *what?*), which can act as (somewhat impolite) response forms as well as response elicitors ... As response forms, they signal lack of understanding and wish to have the message repeated.

(Biber et al. 1999: 1092)

It would seem that *huh?* and *eh?* in the three examples provided by the authors, including (13) below, are suggestive of straightforward instances of a specific kind of response elicitors called 'other-initiated repair markers' that was discussed in Section 3.7.5 rather than of response signals.

(13) British English (Biber et al. 1999: 1092)

A: Where are we at?

B: **Eh?**

A: Where are we going?

Response signals are not uncommonly repeated to express emphatic agreement (*yes yes!*) or disagreement (*no no!*). The Mparntwe Arrernte language of central Australia has three variants of the signal for 'yes': The simple form *ye* is most

commonly used for backchanneling, simply affirming that one is attending to what has been said. When there is strong agreement with the force of a proposition, the form *yewe* or *yawe* is used, while the reduplicated form *yewe-yewe* is used mainly when affirming that a proposition is definitely correct (Wilkins 1989: 375).

The repeated form need not be the same token but may be a different variant of the same type, as the sequence of the two response signals *yeah* and *uh-huh* below shows.

- (14) English (LSWEC-AC (111401; Norrick 2009: 874)

[...]

Nancy: **yeah** she's just fascinated by that.

Margaret: **yeah uh-huh.**

An interesting case of repetition of response signals can be observed in English, where repeating the positive response signal *yeah* may be interpreted as a sarcastic disconfirmation marker if the first use of *yeah* in *yeah yeah* is pronounced with a rising intonation, and the second one with a falling intonation (Tania Kuteva, p.c. of February 7, 2022).

But at least in English, they may also be emphasized by means of interjections. Thus, Norrick (2009: 880) reports that when placed before the bare response signals *yeah* and *no* in initial turn position, English secondary interjections such as *hell*, *fuck*, and *shit* can serve as intensifiers of the following response signals, as *hell* seems to do in (15a) and *shit* in (15b).

- (15) English (SCoSE (Football); WSC DPC238; Norrick 2009: 880)

a Lana: did you burn anything?

Kevin: **hell yeah.**

Lana: **oh good.**

((laughs))

b DN> any, any **er** improvement on seeing client base.

CE> °**shit no**°.

DN> **no**?=

CE> =he's all up and down.

As we will see in Section 4.2.2.1.3, interjections can gradually develop into response signals for disconfirmation in contexts such as (15b).

Frequently used response signals can undergo phonetic modification for some reason or other. Thus, the English response signals *yes* and *no* have a number of variants, reflected in spellings like *yeah*, *yep*, and *nope*, respectively (Biber et al. 1999: 1083).

In a number of languages, response signals have been grammaticalized to constructions where they are introduced by a verb for 'say,' cf. English *She said no*.

The constructional schema employed is typically that of direct speech, as in the following example from the Sepik language Iatmul of Papua New Guinea, where the verb *wa(a)* ‘say’ introduces the response signal *ka’i* ‘no’, and the latter takes the complement *gu yaaku-vak-kak*:

- (16) Iatmul (Ndu, Sepik; Jendraschek 2012: 220, 9.46)  
       gu     yaaku-vak-kak     ka’i   wa-wun  
       water wash-NML-DAT no   say-1.SG  
       ‘I don’t want to wash’ (lit.: ‘I say no to washing’)

In French, the schema appears to involve indirect rather than direct speech, as in *Elle a dit que non* (she has said that no) ‘She said no’, where the response signal *non* is introduced by the subordinating conjunction *que* ‘that’.

Response signals are usually accompanied by pronounced facial gestures in some languages. In the Ambrym language of the New Hebrides in the western Pacific Ocean, the form *lɔŋ* ‘yes’ is accompanied by a slight lifting of the eyebrows, and the form *ɛhɛ*: ‘no’ by tightening the lips with a pulling-down of their corners (Paton 1971: 93).

### 3.8.4 Response signals expressed by clicks

Response signals are among the types of interactives that tend to exhibit cross-linguistic regularities in their phonetic shape. One example is provided by forms for ‘yes’, as has been established by Parker (2021). Based on a survey of 604 attested forms for ‘yes’ taken from 512 languages of world-wide distribution, Parker shows that there is a marked tendency in these forms to contain the laryngeal phonemes /ʔ/ or /h/.

An equally pervasive example can be seen in the world-wide distribution of ‘naked clicks’, that is, clicks without distinctive phonetic accompaniments or other phonological material such as vowels or consonants, as we saw in Section 1.3.2.8. Using clicks as an alternative to expressions for affirmation (‘yes’) and/or negation (‘no’) is a fairly widespread feature in the languages of the world. In his world-wide study of para-linguistic usages of clicks, Gil (2013) distinguishes the following types of click usage: (a) Clicks that may express ‘logical’ meanings (‘yes’ or ‘no’), (b) clicks that may express ‘affective’ but not logical meanings, and (c) clicks that may express neither affective nor logical meanings. Our main concern here is with (a), that is, with clicks used as response signals. For the most part there is no information on which click types are involved.

The quantitative survey carried out by Gil (2013) shows that 32.9% (43) of the 143 languages analyzed belong to type (a), as opposed to 49.6% (71) to type (b) and 17.5% (25) to type (c). While it does not become entirely clear to what extent this sample is representative of the world’s languages, the fact

that almost one-third of all languages draws on clicks as response signals is noteworthy.<sup>33</sup>

There does not appear to be any inherent relationship between function and click type. For example, a lateral click ([ɬ]) is used in the Avar language of Dagestan for negation but for affirmation in the Somali language of eastern Africa. Similarly, Hebrew and many Arabic dialects use a dental click ([ɮ]) for ‘no,’ but in the San’ani dialect of Arabic, the dental click is used for ‘yes’ rather than for ‘no’ (Gil 2013).

An observation that we made during our field research in West African languages is the following: There is a widespread contrast between confirmation and disconfirmation expressed by means of a single click for the former, and a repeated click of the same type for the latter. Depending on the language concerned, this may involve different click types (Lionnet 2020; see also Section 1.3.2.8). As noted already by Gil (2013), languages exhibiting such a contrast are limited to Africa.

### 3.8.5 Discussion

Do all languages have dedicated response signals for ‘yes’ and ‘no’? There is so far no conclusive answer. To be sure, many languages have typically short forms which may have no function other than signaling confirmation or disconfirmation and are etymologically opaque, like English *mhm* or *uh huh*, that is, they are primary interactives (Section 3.6.4.1). But there are also languages where such response signals are historically derived from other text pieces. In Swahili, for example, the main response signals *ndiyo* ‘yes’ and *hapana* ‘no’ can be traced back to the verbal expressions *ndi-yo* ‘it is that’ and *ha-pa-na* ‘there is none,’ respectively.<sup>34</sup> A seemingly unusual situation is found in Korean where there are response signals derived from interrogative constructions via rhetorical questions, that is, questions that do not seek information or require an answer (Rhee 2004: 413). For example, the question words *way* ‘why?’ and *eti* ‘where?’ have acquired uses as markers for emphatic confirmation (‘yes’) and disconfirmation (‘no’), respectively.

Note further that in Indian English one of the conventionalized ways to express confirmation or agreement, *why not*, meaning ‘yes’ can be traced to a rhetorical question (Tania Kuteva, p.c. of February 7, 2022) and similarly, German *warum nicht?* (‘why not?’) and *und ob* (‘and if’) can in certain situations express ‘yes.’ The latter expresses an emphatic form of confirmation, cf. *Liebst du sie?* ‘Do you love her?’—*Und ob!* ‘Absolutely!’

Furthermore, social formulae for ‘please’ or ‘thank you’ have in some European languages assumed functions of response signals in specific contexts, either

<sup>33</sup> Floricic (2015: 101) notes, however, that clicks used for ‘yes’ and ‘no’ are not unconstrained: their cooccurrence with other markers can produce ill-formed configurations in the same way as other markers.

<sup>34</sup> The negation form *la* ‘no’ of Swahili, though, has been borrowed from Arabic.



in combination with extant markers (cf. English *Yes please!*, *No thanks!*) or on their own, for example, in an exchange such as *Would you like to have another cup of tea?—Please!*.

Furthermore, there are also languages that lack paradigm response signals. In the Arawá language Jarawara of Brazil, for example, there is no form for 'no'; rather, speakers have to use an appropriate verb with the negative suffix *-ra* for this purpose (Dixon 2004: 389).

The communicative significance that response signals have can be seen, for example, in the role they play in situations of intense language contact, being easily transferred from one language to another. For example, in the English discourse of Hebrew–English bilinguals in Jerusalem there are also Hebrew expressions confirming a comment made by the previous speaker, like *ken* 'yes', *bidyuk* 'exactly', *naxon* 'true, right', *betax* 'sure', and *beseder* 'all right', also for disconfirming (*lo* 'no'), thus facilitating 'the negotiation of interpersonal closeness vs distance between participants' (Maschler 1994: 341, 350).

### 3.9 Social formulae

Unlike other types of interactives, social formulae tend to enjoy a privileged social status. Parents or teachers may discourage children to use interjections like *damn*, *shit*, etc., or discourse markers like *um*, *uh*, etc. This is strikingly different with social formulae, as aptly pointed out by Davies (1987):

It is likely that in any speech community, certain formulas are felt to be a basic part of the everyday politeness that even quite small children are expected to show, whereas others may be acquired much later and without the insistent prompting by parents which so often seems to be used to elicit the first formulas a child uses. British parents, for instance, seem to lay great stress on the need for three-, four-, or five-year-old children to use *please* and *thank you* appropriately.

(Davies 1987: 86)

Social formulae perform functions that are part of the social institutions of the cultures in which they are used, reflecting and reinforcing the institutions and attitudes of those cultures (Tannen and Öztekin 1981: 44). They are routines in the sense of Coulmas (1981: 2–3), being 'highly conventionalised prepatterned expressions whose occurrence is tied to more or less standard communication situations.' In the framework of Functional Discourse Grammar, social formulae are, in a narrow sense, classified as 'interactives,' differing from 'expressives' in that they are 'clearly directed to the Addressee' (Hengeveld and Mackenzie 2008: 77), that is, they have—in the framework proposed here—an inherent H (hearer) argument as part of their valence. In the wording of Wilkins (1992: 142), social formulae are

‘one-word formulae’ that ‘are tied to specific, and very common, situations, and their function is mainly to acknowledge, promote, and/or maintain social relations in accordance with cultural conventions.’

Arguably the most detailed account of social formulae is that by [Coulmas \(1979\)](#). His ‘routine formulae’ represent ‘petrified forms of typical behavior for handling certain types of situations.’ Their chief function is ‘to reduce the complexity of social interaction’ ([Coulmas 1979](#): 239, 251). Routine formulae include greeting, farewell-taking, congratulating, thanking, toasting—that is, they fall squarely within our type of social formulae.

Another related notion is that of ‘politeness formulas’ of [Ferguson \(1981](#): 30; see also [Davies 1987](#)), which include greetings, farewells, thank yous, apologies, pardons, wishes for health, wishes for recovery from sickness, condolences, topical blessings, and curses. [Ferguson \(1981\)](#) argues that the use of politeness formulas ‘is universal in human societies.’ These formulas for ritual interpersonal exchange are included in but do not coincide with the type of social formulae proposed here. [Ferguson \(1976](#): 22) speculates that they are part of ‘the biological substrate of the universal human exchange.’ Whether social formulae really belong to some ‘biological substrate’ need not concern us here as long as it remains unclear how such a substrate is to be conceived and what exactly is ‘biological’ about politeness formulas. But what is obvious is that such formulas are vital to human communication, and can be expected to be found in languages across the globe.

A third notion showing significant overlaps with the type of social formulae is that of ‘polite speech act formulas’ of [Biber et al. \(1999](#): 1093–4), which includes thanking, apologizing, requesting, and congratulating. A characteristic of this group is, as the authors say, that it frequently elicits a polite reply.

Social formulae are part of what [Quirk et al. \(1985](#): 852–3) call ‘formulae,’ but the latter is conceived as a wider category which also includes response signals (e.g., *absolutely, certainly not, yes*), secondary interjections (e.g., *My Gosh, Oh hell*), alarm calls (*Fire!, Help!*), as well as a number of other kinds of expressions. Another kind of wider category is that of the ‘speech act formulas’ proposed by [Pawley \(2009](#): 7–8). Many of his formulas, like English *hello, sorry, congratulations*, do in fact qualify as social formulae, but others do not, being suggestive of formulaic collocations or idioms of sentence grammar (e.g., *You can say that again, That’s another story*; [Pawley 2009](#): 6).

### 3.9.1 Introduction

Social formulae are arguably the most important means of shaping our daily interaction, regulating how we interact with other people. What Eastman observed on the significance of ‘interjections’ in the Swahili language of East Africa, applies in a similar way to many other kinds of interactives and to languages other than English:

To be communicatively competent in Swahili, in addition to knowing how to speak, one must know how to greet, exhort, warn, beseech, surprise, call, wish, persuade, cooperate, disapprove, challenge, and entertain others.

(Eastman 1992: 282)

Prototypically, usage of social formulae presupposes a situation of discourse where the speech participants involved are close to and face one another. However, this is not necessarily the case. In the Ewe language of Ghana and Togo, the expressive *ú:ru!* (or *u:wúí!*), used to call people, specifically serves to address people across a distance in the bush or on the farm where the hearer is not visible to the speaker (Ameka 1992b: 257).

Knowing 'how to speak,' or to write, covers but one component making up the competence that a speaker needs to acquire to be at home in a language. The functions listed by Eastman are exactly the ones that are central to the domain of social formulae discussed in this section.

Social formulae relate to the *Appell* ('appeal, plea') function of Bühler ([1934] 1965) and the phatic function of Jakobson (1960), in that they contribute to the establishment and maintenance of communicative relations. More centrally, they concern what Lyons (1977: 50) refers to as the social function of language.

Ameka (1992a: 114) classifies social formulae, like *goodbye*, *welcome*, *sorry*, and *thank you*, on the one hand, as 'phatic' interjections, which 'are used in the establishment and maintenance of communicative contact.' On the other hand, he notes that they differ from other interjections in that they 'are interactional and are speech acts' (Ameka 1992b: 269). In spite of this dramatic difference, however, he argues that the two categories belong to one and the same continuum or scale, where one end consists of interjections like *brrr!* 'I feel cold,' while the other end is made up of the social formulae mentioned.

Social formulae usually have a high emblematic value, they are sometimes even retained when speakers switch from one language to another. For example, among groups of German Americans, some barely knowing German, many German social formulae have been retained. German interpolations in English discourse serve to mark speakers as Texas Germans or Indiana Germans, but in addition to speaker attitudes, they also involve speaker-hearer interaction. It is mainly German formulae of greeting, farewell taking, thanking, or welcoming that are used in their English discourses (Salmons 1990: 471).

### 3.9.2 Working definition

The following definition is proposed for social formulae:

- (1) A social formula is an interactive used to foster social relations with the hearer.

This definition is highly general, it is meant to distinguish social formulae from other types of interactives. Social formulae relate to but must not be identified with what Ameka (1992a; 1992b) calls ‘phatic interjections,’ which ‘are used in the maintenance of social and communicative contact’ (Ameka 1992b: 245).<sup>35</sup> More than other types of interactives, social formulae tend to involve adjacency pairs (Schegloff and Sacks 1973; Schegloff 2007; Enfield et al. 2019), that is, schemas of conversational routines consisting of a sequence of two parts, where the first part, produced by speaker A, entails that speaker B is expected to respond in a conventionalized way in the second part (see also Section 7.3.2). Responses need not consist of linguistic expressions, they may as well consist, for example, of a gesture.

The following are examples of cross-linguistically widespread functions of social formulae that may involve adjacency pairs:

- Initiating social contacts
- Apologies
- Blessings
- Condolences
- Congratulations
- Farewells
- Good wishes
- Greetings
- Introductions
- Offerings
- Thanking
- Toasts
- Wishes for health
- Wishes for recovery from sickness

However, not all types of social formulae are necessarily organized in the form of adjacency pairs.

For English examples of social formulae, see Table 1.6 of Section 1.5.

### 3.9.3 Grammar

One difference between (primary) interjections and social formulae is described by Ameka in the following way: ‘primary interjections do not have addressees,

<sup>35</sup> Ameka (1992a) includes the English forms *pst!* and *shh!* in his category of ‘phatic interjections.’ On account of their function, these are treated here as directives (Section 3.2).

but formulae have addressees' (Ameka 1992a: 109)—that is, unlike interjections, social formulae have at least two arguments, namely S (speaker) and H (hearer).

The conceptual content of social formulae is typically of two kinds. Either it involves adjacency pairs, being restricted to a message of the speaker to the hearer, as in greeting rituals of the kind appearing in the exchange of (2), where there are two greeting formulae: *hello* and *good morning*. The latter, for example, can be paraphrased as in (3a), and its argument structure is provided in (3b).

- (2) British English (Biber et al. 1999: 1085)

A: **Hello, Joyce.**

B: **Good morning, Bob.**

- (3) a Paraphrase: I (S) wish you (H) a good morning.  
b Argument structure: *good morning* (S,H)

With regard to their valence, two sub-types of social formulae need to be distinguished. One sub-type, illustrated by the paraphrase in (3), has S and H as its only arguments, paradigm instances being provided by greetings and leave-takings. Thus, *Hi!* or *Goodbye!* constitute independent utterances.

This is different with the second sub-type. If one says *Thanks!* or *Excuse me!* then this is likely to trigger a response like *What for?*, that is, a quest for information on the reason for using the formula concerned—in other words, the utterances would be incomplete without the speaker providing a justification for using them. Take the example of the English form *sorry*, illustrated in (4), one reading of which can be paraphrased as in (5a). The usage of *sorry* does not seem to be complete unless the speaker makes it clear to the hearer why she uses this interactive. This means that in addition to S and H, *sorry* presupposes a T argument, which gives the reason for why the interactive is used—hence, there are three arguments, cf. (5b). In example (4), this reason, that is T, is provided by the text piece *to keep bothering you*.

- (4) English (Biber et al. 1999: 1093)

**Sorry to keep bothering you.**

- (5) a Paraphrase: I (S) apologize to you (H) for having done T.  
b Argument structure: *sorry* (S,H,T)

Social formulae belonging to this sub-type were called T-oriented in Section 2.2.3.1. They include most of all apologies (*Pardon me!*), forms for thanking (*Thank you!*), praising (*Well done!*), etc. All of them imply a T argument providing the reason for the use of the social formula, even if T may be recoverable from the context rather than being explicitly encoded.

There are also other ways in which the T argument of *sorry* can be encoded. Quite commonly it appears in the immediately preceding discourse. Thus, in

the following example, A apologizes to B and the reason can be sought in the predication *it's a little messy* produced by B in the previous turn.

- (6) American English (Biber et al. 1999: 1084)

B: It picks up good, but it's a little messy

A: **Oh yeah**, it's <unclear> messy - **oops, sorry**. <while eating>

But T need not be formally encoded, it can as well be provided by the situation of discourse rather than by some text piece. Suppose I step on someone's foot and apologize by simply saying *Sorry!*. In such a case the T argument is not provided by text material but rather by a specific event in the situation of discourse. An example concerning apologies was discussed in Section 2.2.3.1; (7) is another example involving thanking. Use of an utterance like (7) would be incomplete without the hearer giving a reason for producing the utterance. This reason is provided not by a text piece but by the contextual information added in square brackets. Thus, *thanks* has the same argument structure as *sorry*, cf. (8).

- (7) English

[Noticing that Fred needs sugar for his tea, Jay passes the sugar bowl to him]

Fred: **Thanks, Jay!**

- (8) a Paraphrase: I (S) express my gratitude to you (H) for having done T.  
b Argument structure: *thanks* (S,H,T)

Responses to greetings (e.g., *How are you?*—*Fine.*) can be said to belong in a wider sense to the type of response signals that were discussed in Section 3.8. However, on account of the conventionalized function they have in greeting rituals they are included here. Such responses can be more complex, they may show variations of several kinds. In the following greeting exchange in Yiddish in (9), the hearer uses the response form *Got tsu danken* 'Thank God' (lit. 'God to thank') etymologically derived from a verbal infinitive form:

- (9) Yiddish (Matisoff [1979] 2000: 13, (6a-b))

A: **Nu**, vos makht Ir epes?  
(now what make you ?)  
'So, how are you?'

B: **Got tsu danken.**  
(God to thank)  
'Thank God.'

The three-place forms *please*, *sorry*, and *thank you* are highly flexible social formulae of English: They frequently enter into collocation with a preceding response signal, like *yes* in (10a), and they can occur in various slots of an utterance, or as stand-alones, cf. *thank you* in (10b).

- (10) British English (Biber et al. 1999: 1093)  
 a A: Would you like to have another drink **Adam**?  
 B: **Yes please.**  
 b Can I have a – another two Diet Cokes **please? Thank you.**

In combination with imperatives, *please* has a softening effect, giving the imperative the form of a request (Biber et al. 1999: 220).

- (11) English (Biber et al. 1999: 220)  
 “**Please** do come over,” she invited. (FICT)

There is a common collocation pattern where a formula of social exchange is followed by a vocative (Biber et al. 1999: 1086):

- (12) English (Biber et al. 1999: 1086)  
 A: **Okay. Bye Butch.** DM - SOF - VOC  
 B: **Bye Butch, bye Marc.** SOF - VOC

### 3.9.4 Grammaticalization

Social formulae, and especially greetings, are highly frequently used interactives that tend to undergo massive grammaticalization—to the extent that neither their form nor their meaning may have much in common with their original shape. Take the following examples of greetings in English (13) and Swahili (14).

- (13) The development of the greeting *hi* of American English (Ferguson 1981: 31–2)  
*Orthography*: How are you? > How are you?, Hiya! > Hiya!, Hi! > **Hi!**  
*Phonetic form*: [haway<sup>a</sup>] [haway<sup>a</sup>], [hay<sup>a</sup>] [hay<sup>a</sup>], [hay] [hay]
- (14) The development of the greeting *shikamoo* of Swahili (Bantu, Niger-Congo)  
 Na-shika mi-guu y-ako! > **shikamoo!**  
 1.SG.PRES-seize NC4-foot NC4-2.SG.POSS Hi!  
 ‘I grab your feet!’ ‘Hi!’ (Respectful form)

English *hi* and Swahili *shikamoo* have a similar function, but the latter is a respectful greeting used by an inferior to a superior, or a younger to an older person. But both expressions have undergone the same kind of development, to be described in Section 4.2 by means of parameters of grammaticalization; suffice it here to provide a brief outline of their development. By means of desemanticization, English *How are you?* lost its character of a question, whereas the Swahili expression *Na-shika mi-guu y-ako!* lost its literal meaning of grabbing someone’s feet in favor of its new function as a greeting formula. Via decategorialization, the two expressions lost their internal word and morpheme boundaries, merging into unanalyzable

forms, and via the parameter of erosion the expressions lost much of their phonetic substance, being reduced to the short forms *hi* and *shikamoo*, respectively.

The English example in (13) illustrates another salient feature of the process just sketched: In accordance with the overlapping structure of grammaticalization, described by means of the overlap model of Heine (1993: 48–53), a change from A to B is not a discrete step but typically involves an intermediate stage where A and B co-occur, being used side by side, before in the final stage A may be lost, hence:  $A > A/B > B$ . As the case of English *hi* in (13) demonstrates, however, earlier stages need not be lost but may continue to co-exist in the modern usage of the language concerned.

Similar developments can be observed in other kinds of social formulae, like English *goodbye* (<*God be with you*) or the Swahili salutation to mourners *makiwa* (Section 4.2, (6)); we will deal with such developments in more detail in Section 4.2.

### 3.9.5 The English formula *please*

The social formula *please* may serve as an example of another effect of grammaticalization, suggesting that not uncommonly interactives exhibit two contrasting patterns of usage, where one is suggestive of a prototypical structure of an interactive, whereas the other pattern is characterized by more or less pronounced features of structural integration in sentence grammar; we will return to this issue in Sections 4.2 and 4.4.

For good reasons, *please* has been analyzed variously as an adverbial, an interjection, a discourse marker (e.g., Biber et al. 1999), a politeness marker, or a pragmatic marker (Wichmann 2004; 2005). To the extent that its usages meet the definition in Section 3.9.2 it is treated here as a social formula. According to Wichmann (2005: 229, 248), *please* is typically a routine expression of courtesy ‘but depending on its intonational realization, can convey meanings at the more emotional end of the interactional spectrum.’ The function of *please* is strongly associated with requests. For example, Woods (2021: 123) shows that the utterance of (15a) can have either an information-seeking or a request function. In the utterance containing *please* in (15b), by contrast, there is only a request function.

- (15) English (Woods 2021: 123)
- a Can you open the window?
  - b **Please** can you open the window?

*Please* is highly versatile in its placement potential, occurring initially, medially, and finally, and also in subordinate clauses, cf. (16). Final *please* differs from initial and medial *please* in that it ‘does not mark requests per se but “bonds” a



speaker and addressee, reinforcing their relationship as requester and requestee' (Woods 2021: 121–2, 125).

- (16) English (BNC, CL2 2285; Woods 2021: 160)

I asked if they would **please** call me Richard – Dick, I said, made me feel like a symbol of some kind.

As the detailed analysis by Woods (2021) shows, two main kinds of usage patterns can be distinguished, which we propose to refer to as *please*<sub>1</sub> and *please*<sub>2</sub>.

*Please*<sub>1</sub> is described by Woods (2021) as 'the overt realisation of a high functional head,' it is a syntactically unintegrated stand-alone, even if it may, and frequently does combine with other text pieces. It exhibits the structure of a paradigm instance of an interactive in accordance with the definition in Section 1.2. *Please*<sub>2</sub>, by contrast, is called by Woods (2021) 'syntactically integrated' *please*. As demonstrated by Woods, *please*<sub>2</sub> does enter into a wide range of collocations both with other interactives and with pieces of sentence grammar, to the extent that it can be said to have a 'syntax.'

In more general terms, the difference between the two is one of grammaticalization: Whereas *please*<sub>1</sub> shows essentially no traces of grammaticalization, *please*<sub>2</sub> is a grammaticalized version of *please*<sub>1</sub> and, accordingly, the differences can be accounted for with reference to parameters of grammaticalization (Table 4.3 of Section 4.2.1).

The two patterns thus differ from one another in a number of ways. First, they differ in their frequency of occurrence. *Please*<sub>2</sub> is clearly more widely used than *please*<sub>1</sub>. *Please*<sub>1</sub> is most commonly associated with imperatives in both spoken and written texts, followed by interrogatives, while stand-alone *please*<sub>1</sub> is generally rare (Woods 2021: 126). Second, whereas *please*<sub>1</sub> is or may be set off by pronounced prosodic marking, *please*<sub>2</sub> is largely prosodically integrated. And third, the two differ also in the way they combine with other text material: *please*<sub>2</sub> has lost much of the freedom of placement that *please*<sub>1</sub> has. For example, whereas syntactically unintegrated *please*<sub>1</sub> can occur in the initial position of declarative clauses this does not appear to apply to *please*<sub>2</sub>. Furthermore, and relatedly, initial *please*<sub>2</sub> is only compatible with imperatives and polar questions having modal meaning (Woods 2021: 132). Information-seeking questions that cannot be construed as requests are incompatible with initial *please*<sub>2</sub>, cf. (17a), while *please*<sub>1</sub> does not show such a constraint, cf. (17b).

- (17) English (Woods 2021: 131, (29c), (31))

a \**Please*<sub>2</sub> did you get that dress at Macy's?

b ***Please***<sub>1</sub> – did you get that dress at Macy's?

Nevertheless, both patterns exhibit salient features of interactives. For example, both are invariable forms that can neither be participants of a sentence nor may they be modified or coordinated, as (18) illustrates.

- (18) English (Woods 2021: 131, (25))  
 a \*Very **please**. (Cf. very kindly)  
 b \*Many **please**. (Cf. many thanks)

Furthermore, both have a three-place argument structure, distinguishing S, H, and T, as the example of *please*<sub>2</sub> in (19) and its argument structure in (20b) may show.

- (19) English (Woods 2021: 142, (66))  
 You should seriously **please**<sub>2</sub> stop replying to those messages.

- (20) a Paraphrase: [With seriousness [I (S) request [you (H) stop replying to those messages (T)]]]<sup>36</sup>  
 b Argument structure: *please*<sub>2</sub>(S,H,T)

That *please*<sub>1</sub> also has a three-place argument structure may not immediately be obvious when looking at stand-alone examples, like (21). But in addition to the implicit arguments S and H, *please*<sub>1</sub> can be claimed to have a third implicit argument, namely T. This argument can be provided by the preceding text. For example, there may have been a text piece like *May I come in?* preceding (21), uttered by a different speaker. Alternatively, T might be provided by the situation of discourse, for example, when a mother wants her child to stop doing what it is not supposed to do<sup>37</sup>—in short, without an overtly or covertly present T, (21) would not seem to be used felicitously. Thus, the argument structure of (21) can be rendered as in (22).

- (21) English (Woods 2021: 126)  
**Please**<sub>1</sub>!  
 (22) a Paraphrase: I (S) request you (H) to do X (T)  
 b Argument structure: *please*<sub>1</sub>(S,H,T)

To conclude, there is reason to assume that many interactives are characterized by a distinction between non-grammaticalized and grammaticalized usages of the kind sketched here with reference to the difference between *please*<sub>1</sub> and *please*<sub>2</sub>, even if there is so far no information for most of the interactives. The impression conveyed in the course of this section was that the two patterns of usage are neatly separated from one another. This, however, is presumably a simplified understanding of the actual situation. On the basis of observations generally made in studies of grammaticalization, it is distinctly more likely that there is a gradual transition from one pattern to another, as will be shown in Section 4.4 with reference to ideophones. But so far there is no appropriate information on this issue—neither for *please* nor for most other interactives.

<sup>36</sup> The brackets in the paraphrase are adopted from Woods (2021: 142, (66)).

<sup>37</sup> From a different theoretical perspective, Woods (2021: 126) suggests treating stand-alone *please*<sub>1</sub> ‘as an imperative in which a structure such as “spare me” or “stop now” has been elided.’

### 3.9.6 How to use social formulae

Being immediately anchored to the situation of discourse, interactives are highly sensitive to the contexts in which they are or can be used. This applies presumably more to social formulae than to other types of interactives: To know where, when, and how a given formula is to be used presupposes an advanced sociolinguistic competence in the language concerned. We may illustrate this by looking at a somewhat trivial situation that speakers of English are likely to be familiar with:

To take one example, in order to decide whether or not it is appropriate to utter the formula *good evening* in greeting, a speaker has to take account of the time of day at which the encounter occurs, the circumstances of the meeting (it would not normally be used, for instance, on meeting a person for the second time in a single evening), the formality of the situation, the nature of the addressee, and the relationship between speaker and addressee (one is unlikely to use it in very informal situations, to a small child, or to one's spouse).

(Davies 1987: 79)

A speaker of Moroccan Arabic would need to rely on a different world of social competence to say the same thing. 'Good evening' can be expressed by *msakum* ('your afternoon'), but only if the speaker is a woman. Otherwise, *msalxir* ('afternoon of goodness') would be used, which is acceptable for both men and women. But the timespan covered by the two expressions is different from that of English *good evening*: Both *msakum* and *msalxir* can be used at any time between early afternoon (after the midday meal) and late evening, thus, extending over a timespan covered by both *good afternoon* and *good evening* in English (Davies 1987: 84).

Constraints on the modalities of usage of social formulae are arguably more pronounced in some languages other than in English. Coulmas (1981: 82–9) characterizes the usage of the Japanese formula *sumimasen* and its variants (*sumimasen doōmo*, *doōmo sumimasen deshita*, etc.) in the following way. Depending on the context, the formula can be used as general conversation opener, attention getter, to take leave, to apologize, or to express one's gratitude. It is frequently used when English speakers would say *thank you*, but depending on the context it may mean not only 'thank you' but also 'I'm sorry'. It may also be used to indicate a general feeling of gratitude and obligation without there being any specific object of gratitude or obligation.

*Sumimasen* thus covers a spectrum of contexts. But in more dramatic situations, where in fact something wrong has happened and one wishes to voice one's empathy to the person concerned, one is likely to draw on a range of other forms, like *O-kinodoku-sama* (*degozaimashita*) 'I'm terribly sorry', which literally means 'this must have been poison to your soul'. In an intimate setting, by contrast, one

might pick the very casual *gomen ne* instead to apologize. And in a similar way there are situations where one would not use *sumimasen* to say ‘thank you’ but rather choose from a pool of alternative forms, such as *arigatō(gozaimashita)* or *sore-wa kyōshuku desu*, the latter meaning literally ‘I feel ashamed.’

To be able to use an expression like *sumimasen* in all its contexts, embeddedness in Japanese culture and history requires a pronounced competence of the language far beyond that of Japanese sentence grammar. The literal meaning of *sumimasen* is somehow negative, paraphrasable as ‘that does not sort itself out’ or ‘that is not enough.’ The expression is first attested around 1800, apparently signaling dissatisfaction of the speaker with himself or herself, and as such it could be used for both apologizing and thanking (Heiko Narrog, p.c. of November 7, 2021).

*Sumimasen* is but one out of a plethora of social formulae that speakers of Japanese need to control. Social formulae have been studied in their social context (see, e.g., the contributions in [Coulmas 1981](#)), and they suggest that the nature of social embeddedness of interactives in general and of social formulae in particular is complex. But so far our knowledge of this embeddedness is too scant to allow for any significant generalizations.

### 3.10 Vocatives

‘God did tempt Abraham and said unto him, Abraham: and he said, Behold, here I am.’

(Genesis 22:1)

The above citation from the bible represents an early attestation of the role played by vocatives in linguistic communication. But among the types of interactives examined in this chapter, vocatives constitute a more marginal one. For one thing, they may take on a fairly complex format, especially but not only as address forms in letter writing. Consider the following somewhat extreme format of a vocative that was found in a letter written in 1623 by Sir Thomas Wentworth to his friend Christopher Wandesford:

- (1) Seventeenth-century English ([Nevalainen and Raumolin-Brunberg 1995: 560](#))

‘My diligent and expedite Ambassadors resident with his Maiestie of greate Brittainē’

And for another, and more importantly, vocatives differ from most other types in having an essentially open-ended membership that is hard to delimit—for

example, any English first name, second name, or other proper name may function as a vocative.

It therefore comes as no surprise that vocatives tend to be treated separately from other types in grammatical accounts. Thus, the category of English inserts proposed by Biber et al. (1999) contains instances of most types of interactives but not of vocatives, which are treated as a category of their own (Biber et al. 1999: 1103, 1108–13). Similarly, Hengeveld and Mackenzie (2008: 77–8), while noting that vocatives ‘constitute a special class of Interactives,’ do not discuss them within their category of ‘interactives,’ which is restricted essentially to social formulae (Section 3.9). Vocatives, in the sense of the term as used here, correspond to what are called ‘vocative interjectives’ in grammars of South African Bantu languages (e.g., Cole [1955] 1987: 396; Doke and Mofokeng [1957] 1985: 367).

Vocatives are included here to the extent that they conform to the definition of interactives in Section 1.2—that is, we will not be concerned with vocatives of the kind presented in (1). First, vocatives express discourse functions similar to some other types of interactives, especially to attention signals (Section 3.1). Second, they share a paradigmatic relationship with other types of interactives (see, e.g., Section 1.3.1.3, Table 1.2). And finally, they are also related to other kinds of interactives in their patterns of grammaticalization (Section 4.2).

In a number of languages, ‘vocative’ is encoded in the nominal case system, forming one of the morphological case categories distinguished. Well-known examples are provided by Indo-European languages having morphological vocative endings, like Latin, Greek, Latvian, Lithuanian, and most Slavic languages (see Anstatt 2011), in accordance with the following note by Haiman (1980):

The vocative has been traditionally named a case on paradigmatic grounds, although it differs syntactically from all other cases: while true cases define the role a noun plays as part of a sentence, the vocative form of a noun is in itself a complete sentence.

(Haiman 1980: 240)

Rather than referring to a morphological category of case marking, the term ‘vocative’ as used here stands for a functional type of interactives, differing from the former especially in the following (see Floricic 2011: 106): First, rather than the function of *Darstellung* (representation) of Bühler ([1934] 1965), vocatives concern Bühler’s function of *Appell* (appeal, plea) (Kurylowicz [1949] 1973: 146–7; Floricic 2011: 106). Second, unlike the former, they are not a syntactic or prosodic part of a sentence, they do not contribute to the asserted content of an utterance and cannot be used in indirect discourse (Eckardt 2014: 223–4, 246); but see also Espinal (2013). And third, their function has more in common with that of imperatives or interjections than with noun phrases, as was already pointed out by Schleicher ([1862] 1877: 162).

### 3.10.1 Introduction

Vocatives are an important tool for creating, maintaining, and reinforcing social relationships (cf. Biber et al. 1999: 1108, 1110): They may signal a bond of affection by means of forms of endearment (*darling, sweetie pie*, etc.) or informal relations with familiarizers (*folks, guys*), strengthen group relationship by means of family terms (*mom, dad*), or mark social distance by using surnames (*Mr Woodburn*), or honorifics (*sir, madam*). Thus, they express discourse functions that fall squarely within the category of interactives. Nevertheless, on account of what was observed in the introduction to this chapter, including vocatives within the pool of interactives may seem surprising and not everybody might welcome this decision.

Like attention signals, vocatives have been classified within a higher category of ‘appellatives’ (Sonnenhauser and Noel Aziz Hanna 2013: 12). For Fraser (1996: 169, 176), vocatives are part of what he calls ‘parallel markers’ since they ‘signal an entire message in addition to the basic message.’ Thus, they are said to differ from ‘basic markers,’ which signal only ‘the force of the basic message.’

Vocatives share with directives some of the hallmarks of the sentence grammar categories from which they are derived. But whereas directives show verbal properties (Section 3.2), vocatives tend to exhibit nominal properties. At the same time, the two differ considerably from the categories from which they are derived, as has been pointed out by a number of authors (Jakobson 1971: 10; Sonnenhauser and Noel Aziz Hanna 2013: 11; see Section 3.2.1).

Vocatives resemble nominal appositions in certain contexts. For example, *my brother* in (2) can be interpreted both as a vocative and as an apposition, both being theticals in the sense of Heine et al. (2013; see Section 5.2.3). The two can, however, be distinguished on account of their structural behavior and contrastive functions. With regard to (2), *my brother* is an apposition when paraphrasable by *being my brother* but a vocative when paraphrased by *brother!*. In the latter case, *my brother* differs from appositions in serving speaker–hearer interaction.

(2) English

What did you, my brother, tell them?

### 3.10.2 Working definition

The definition of vocatives proposed here is as follows, where ‘hearer’ stands for one or more addressees:

- (3) A vocative is an interactive used to call or address the hearer, asking for their attention.

This definition contains, on the one hand, features that are highlighted in other definitions that have been proposed. On the other hand, it differs from other definitions which focus on the nominal character of vocatives. It is especially hard to reconcile with the claim that vocatives are noun phrases (e.g., [Levinson 1983: 71](#)).<sup>38</sup> For one thing, unlike noun phrases, they conform to the definition of interactives in Section 1.2, and for another, vocative expressions need not consist of nominal elements, as the following examples show.

- (4) English ([Zwicky 1974](#); [Schaden 2010: 177](#), (5))
  - a **You with the sweater on**, move about a foot to the left.
  - b **Everyone who has a dog**, you need a dog permit!

While most vocative expressions in fact resemble and can historically be traced back to noun phrases, neither their usage nor their discourse functions have much in common with those of the syntactic category of NPs (Section 3.10.1; see also below). Our definition also differs from the one in (5) since vocatives need not involve a sentence; rather, they can occur as stand-alones without involving or implying a sentence.

- (5) By definition, a vocative is a nominal element referring to the addressee(s) of a sentence. ([Schaden 2010: 176](#))

Furthermore, the definition in (5) ignores features such as the following: Vocatives are said to establish or maintain a relationship between the speaker and some proposition ([Lambrecht 1996: 267](#)).<sup>39</sup> This assertion holds true for most but not for all usages of vocatives in that vocatives are not required to co-occur with or imply some proposition. Finally, the definition in (5) does not mention the functions expressed by vocatives. To be sure, they ‘are used, for example, to summon, to rebuke, to question whether the person addressed is present, or to remind of an order or a request’ ([Quirk et al. 1985: 851](#)), but this by no means exhausts the range of all the functions they may serve.

[Zwicky \(1974: 796\)](#) argues that vocatives are not emotionally unmarked, in that ‘vocative NPs in English are almost never neutral: they express attitude, politeness, formality, status, intimacy, or role relationship, and most of them mark the speaker ..., there is virtually no affectively neutral vocative.’ In this respect, vocatives resemble imperatives (Franck Floricic, p.c. of August 5, 2021); see Section 3.2.

Our concern here is only with formulaic, conventionalized vocatives, that is, with forms corresponding to our definition in 1.2. But in addition, we will not

<sup>38</sup> ‘Vocatives are noun phrases that refer to the addressee, but are not syntactically or semantically incorporated as the arguments of a predicate; they are rather set apart prosodically from the body of a sentence that may accompany them’ ([Levinson 1983: 71](#)). Franck Floricic (p.c. of August 5, 2021) reminds us that ‘referring’ of vocatives is not of the same kind as it is in canonical cases of reference.

<sup>39</sup> ‘Vocatives serve to call the attention of an addressee, in order to establish or maintain a relationship between this addressee and some proposition’ ([Lambrecht 1996: 267](#)).

exclude a few items that do not correspond to the notion of ‘plain vocatives’ but exhibit a paradigmatic relationship with them, such as English *everyone* or *you* in specific uses of them.

### 3.10.3 Grammar

Being interactives, vocatives are syntactically detached, and they cannot be modified by restrictive, additive, or focus particles, as demonstrated by Schaden (2010: 178).<sup>40</sup> They can, and not uncommonly do, occur as stand-alones, that is, as independent speech acts without implied T argument, as in (6). In this capacity, vocatives are argued to have two arguments, namely S and H, both unexpressed, where the speaker calls the hearer, as is suggested by the paraphrase of (6) in (7a). According to Biber et al. (1999: 1111), such stand-alone occurrences, which are treated here as ‘prototypical vocatives,’ account for about 10% of all uses of English vocatives.

- (6) English (Biber et al. 1999: 1108, 1111)

**Mom!**

- (7) a Paraphrase: I (S) am calling you (*Mom*, H)  
 b Argument structure: *Mom* (S,H)

In addition to stand-alone vocatives, there are also what one may loosely call T-oriented vocatives (Section 2.2.3.1), exemplified in (8), which clearly constitute the majority of usages in English: They have not only an inherent speaker (S) and hearer role (H), both formally unexpressed, but also an overtly marked T argument. For example, (8) can be paraphrased roughly as in (9a) and its argument structure in (9b). Thus, there are two argument structures associated with vocatives (see Section 2.2.4).

- (8) American English (Biber et al. 1999: 1109, 1111)

**Honey**, can I use that ashtray **please**?

- (9) a Paraphrase: I (S) am calling you (*honey*, H) to ask if I can do T  
 b Argument structure: *honey* (S,H,T)

There is a fairly frequent collocation pattern where vocatives combine with other types of interactives. In (10a), this is an attention signal (*hey*) and in (10b) a social formula (*please*), both preceding the vocative.

<sup>40</sup> It goes without saying that this applies only to vocatives used as interactives. Otherwise, vocative expressions can have highly complex internal structures; Franck Floricic (p.c. of August 5, 2021) reminds us of French examples like the following: *Les enfants qui sont à ma droite, suivez-moi!* ‘The children who are to my right, follow me!’ See Espinal (2013) for ways in which vocatives can combine with personal pronouns and particles into complex constructions. Furthermore, Espinal (2013) proposes to distinguish between ‘true’ vocatives and ‘fake’ vocatives.



- (10) English (Biber et al. 1999: 140, 1088)
- a **Hey Rob** do you want to lend me fifty p? (BrE)  
ATT - VOC
  - b **Please, Dad,** I know you're busy and that, but I must know. (FICT)  
SOF - VOC

Vocatives are fairly flexible in their placement. The following examples from the Akie language in north-central Tanzania show that they can occur as stand-alones (11a), utterance-initially (11b), utterance-finally (11c), as well as inserted in an utterance, as *ánanú* 'folks' is in (11d).

- (11) Akie (Southern Nilotic, Nilo-Saharan; König, Heine and Legère 2015a: 125)
- a **paapâ!**  
'Father!'
  - b **paapâ,** pányee ko anó nte?  
father.A meat.A COP where be.at  
'Father, where is the meat?'
  - c ko anó nte pányee, **paapâ?**  
COP where be.at meat.A father.A  
'Where is the meat, father?'
  - d ko má pet-u-un **ánanú** ámti!  
COP NEG miss-VEN-2.SG.O folks food.N  
'(Even) food, you won't miss it, folks!'

Biber et al.'s (1999: 1111–12) text analysis of English grammar suggests that vocatives appear most frequently in final position (about 70% of their usages) and distinctly less frequently in initial (about 10%), and in medial position (also about 10%), cf. (12).

- (12) English (Biber et al. 1999: 1109, 1111)
- a **Huh,** you get to do this next year, **Jason.** (AmE)
  - b I'm sorry **Asia** that's just not my identity.

Initial vocatives tend to be associated with longer text units, whereas final vocatives are associated with shorter units. This, however, need not be the case, as the following text piece from the West African Ewe language shows:

- (13) Ewe (Kwa, Niger-Congo; Ameka 1992b: 252, (7))
- dadaví, dzáà! dzáà! wo-é-zɔ.**  
madam welcome welcome 2.SG-FOC-walk  
'Madam, Welcome! Welcome! You have travelled.'

In the following example, the vocative *man* occurs both initially and finally. Note, however, that *man* is also used as an interjection (see Table 1.6 of Section 1.5), and it is not entirely clear which of the two types is involved in (14).

(14) American English (Biber et al. 1999: 1109)

**Hey, man.** I'll make this real short. What's happening, **man**?

Based on their functions, the two kinds of vocatives in (15) can be distinguished (Schegloff 1968), each associated with different positions vis-à-vis their host (Zwicky 1974: 787; Levinson 1983: 71; Biber et al. 1999: 1112):

(15) Vocatives as calls vs. addresses

- a Calls (or summonses), which are designed to get the hearer's attention. They are utterance- and conversation-initial, forming independent speech acts, cf. (16a). They include not only all stand-alone vocatives but also host-initial T-oriented vocatives (Section 2.2.3.1). The latter tend to impose a topic-comment schema on the usage of the vocative, where the vocative acts as the topic for the ensuing sentence grammar unit, which provides the comment.
- b Addresses, which serve to maintain or emphasize the contact between speaker and hearer (Zwicky 1974: 787), or to identify someone as a hearer (Biber et al. 1999: 1112) and occur in positions other than the host-initial one, cf. (16b). They are restricted to T-oriented vocatives, that is, the sentence grammar unit is an obligatory part of the use of the vocative in this kind of usage.

(16) English (Zwicky 1974: 787)

- a **Hey lady**, you dropped your piano.
- b I'm afraid, **sir**, that my coyote is nibbling on your leg.

These observations suggest that placement at the left periphery of its T argument ('initial vocative') or at its right ('final vocative') each has a bearing on the usage characteristics of vocatives:

For example, an initial vocative can serve as an attention-getter and can also have the function of clearing space for a lengthy turn. The final vocative, on the other hand, is more likely to occur after a short remark, where attracting attention is not a problem, and where the social role of the vocative can combine with that of singling out the addressee.

(Biber et al. 1999: 1113)

A different kind of classification is proposed by Schaden (2010: 181–3), who argues that vocatives have three basic functions, namely (a) to identify the addressee,

(b) to predicate a property of the addressee, and (c) to activate the addressee. Accordingly, three kinds of vocatives need to be distinguished, illustrated with the following examples.

- (17) English (Schaden 2010: (22), (28))
- a **George**, could you pass me the salt, **please**? Identification
  - b **Dear friends**, let us go inside! Predication
  - c [...] **Grandma**, what big teeth you have! Activation

In identificational vocatives, the addressee(s) still must be established as such, whereas in predicational and activational vocatives it is presupposed, that is, it has already been established (Schaden 2010: 184). A modified classification is proposed by Espinal (2013), distinguishing between an identity, an identificational, and a predicational type of vocatives.

3.10.4 Sub-types

A number of classifications of vocatives have been proposed. For a common typology, see Fraser (1990: 391–2). Biber et al. (1999: 1108–11) propose the eight types of vocatives in English in Table 3.19, arranged in their order of frequency of occurrence in texts of American and British English. These types are illustrated in (18).

Table 3.19 Types of English vocatives

| Category                                                      | English examples                                                       |
|---------------------------------------------------------------|------------------------------------------------------------------------|
| a First names (in full)                                       | <i>Jennifer, Marjorie, Paul, Thomas</i>                                |
| b Familiarized first names (shortened, or with suffix -y/-ie) | <i>Jackie, Marj, Paulie, Tom</i>                                       |
| c Familiarizers                                               | <i>bud, dude, guys, man</i>                                            |
| d Kinship forms                                               | <i>da, dad, daddy, ma, mom, mommie, mum, pa, pop</i>                   |
| e Endearments                                                 | <i>baby, (my) darling), (my) dear, hon, honey, love, sweetie (pie)</i> |
| f Others (including nicknames)                                | <i>boy, lazy!, everyone, red dog, Uncle Joe, you</i>                   |
| g Titles and surnames                                         | <i>Mr Graham, Mrs Jones</i>                                            |
| h Honorifics                                                  | <i>madam, sir</i>                                                      |

Source: Biber et al. (1999: 1108–11)

- (18) Examples of the types of English vocatives distinguished by Biber et al. (1999: 1108–11; boldening adopted from the original)
- a Morning **Diane**.
  - b Hey, **Mike**, grab your dominoes! (AmE)
  - c Got a ticket **mate**?
  - d Thanks **Mom** – okay – talk to you later – see you soon – bye. [On the telephone] (AmE)
  - e Is that you **darling** come here **sweetie pie**. (AmE)
  - f Hi, **Aunt Margaret**! (AmE)
  - g Hello **Dr. Denton**. How do you do?
  - h Okay yes **sir**. Stuffed mushrooms. [From a restaurant meal] (BrE)

According to the text analysis by Biber et al. (1999: 1111), first names, as in (a) and (b) of Table 3.19, are by far the most common, accounting for more occurrences than the remaining six types taken together. (c) and (e) are commonly used to claim in-group membership. They may serve, for example, to soften a preceding directive or, when used to children, to turn a command into a request (Brown and Levinson 1987: 107–8). In addition, vocatives may also consist of second person pronouns (e.g., *You, hand me those forceps*; Zwicky 1974: 787). In some languages there are specific personal pronouns used as vocatives; the Maale language of Ethiopia distinguishes the vocative pronouns *ʔéézzay* and *ʔenááre* to call or address male and female persons, respectively (Amha 2001: 261).

As a rule, vocatives address human beings. But in certain text genres, such as tales or stories, the addressee can be inanimate. In the following example from a fable in the Akie language of north-central Tanzania, the addressee is a homestead (*kaav* ‘homestead,’ here reduced to *kaa*):

- (19) Akie (Southern Nilotic, Nilo-Saharan; König, Heine, and Legère 2015a: 126)
- |             |    |      |                 |
|-------------|----|------|-----------------|
| <b>kaa</b>  | í  | ó    | yaat-u-aan!     |
| homestead.A | DM | 2.PL | open-VEN-1.SG.O |
- ‘Homestead! You open me!’

Other examples are not hard to come by in songs and poems. Mention may be made of the following text piece from the national anthem of Canada: *O Canada! Our home and native land!* (Neal Norrick, p.c. of July 10, 2021).

In a number of accounts, attention signals (Section 3.1) are treated as part of vocatives. For example, Hengeveld and Mackenzie (2008: 81) classify the English form *hey* as a paradigm instance of a vocative. This is not the convention proposed here; see Section 3.1.5 for discussion.

At the same time, we follow Hengeveld and Mackenzie (2008: 78–9) in treating epistolary salutations like *Dear John* as vocative forms—even if ‘peripheral’ ones

because such forms constitute fixed formulae of written discourse that can equally well be classified as instances of the type of social formulae (Section 3.9).

### 3.10.5 Typological variation

In a number of languages, vocatives have the same form as corresponding noun phrases of sentence grammar. However, there are also quite a number of languages where vocatives present special forms. For one thing, sentence grammar and interactive grammar may employ different terms for one and the same referent. In the Muna language of Sulawesi in Indonesia, the terms *ghane* ‘boy’ and *bhela* ‘friend’ do not occur as nouns, they are restricted to vocatives (van den Berg 1989: 187). And for another, terms in the two grammars may be similar but nevertheless be distinguished by specific grammatical features.

Features that may set vocatives off from noun phrases are in particular the following:

- (a) Special vocative markers
- (b) Truncation
- (c) Case marking
- (d) Sex gender marking
- (e) Politeness marking

These features are now looked at in turn.

#### (a) Special vocative markers

In order to mark a nominal form as a vocative, speakers of some languages use special markers whose main or only function it is to signal an interactive function, or as Hengeveld and Mackenzie (2008: 79) say, an ‘interpellative illocution.’ Such a marker appears to be *A* in Scottish Gaelic, as in *A Sheumais!* (*A James.VOC*) ‘James!’.

#### (b) Truncation

Another way in which vocatives are distinguished from noun phrases of sentence grammar in a number of languages involves reduction in their form. This will be the subject of Section 3.10.6.1.

#### (c) Case marking

There are a number of languages having ‘vocative’ as part of their paradigms of the nominal case system, where the vocative forms one of the cases distinguished. For example, Latin masculine singular nouns like *dominus* ‘master’ include vocative *-ē* in their paradigm of case inflections:

|                   |          |
|-------------------|----------|
| <i>Nominative</i> | domin-us |
| <i>Genitive</i>   | domin-ī  |
| <i>Dative</i>     | domin-ō  |
| <i>Accusative</i> | domin-um |
| <i>Ablative</i>   | domin-ō  |
| <i>Vocative</i>   | domin-ē  |

In addition to Latin, examples are provided by a number of other Indo-European languages having morphological vocative inflections, like Greek, Latvian, Lithuanian, and Slavic languages like Polish, Upper Sorbian, Czech, Ukrainian, Croatian, Serbian, Bulgarian, and Macedonian (Anstatt 2011; Noel Aziz Hanna and Sonnenhauser 2013: 293). In the Atlantic language Noon of Senegal, vocative expressions serving as calls in utterances of their own take the vocative suffix *-óo*, cf. (20a). But this suffix is usually absent when the vocative expression takes a T argument in the immediately following clause, cf. (20b) (Soukka 2000: 266–7).

(20) Noon (Atlantic, Niger-Congo; Soukka 2000: 266–67)

- a **Yiyóo!**                      **Yiyóo!**                      (<*yiyaa* + *-óo*)  
 mother.VOC                      mother.VOC  
 ‘Mother! Mother!’
- b **Yiyaa!**    mi    heel-in-daa.                      u    tum-ee    yaa?  
 mother    1.SG    search-PERF-O.2.SG    2.SG    do-PAST    what  
 ‘Mother! I have been looking for you. What did you do?’

#### (d) Sex gender marking

In the Punjabi language of northwest India and eastern Pakistan, the particles *óé* and *nii* are used for calling male and female referents, respectively. Alternatively, these particles are replaced, respectively, by *saalaa* (‘brother-in-law’) and *saalii* (‘sister-in-law’) to add ‘rudeness’ (Hengeveld and Mackenzie 2008: 79).

#### (e) Politeness marking

Another group of languages encodes distinctions in social relations in their vocative forms. Such languages include Kashmiri of northern India and Pakistan, where polite and impolite vocatives are distinguished (Hengeveld and Mackenzie 2008: 79).

### 3.10.6 Special features of vocatives

Vocatives are commonly, though not necessarily, derived historically from noun phrases of sentence grammar, but they may acquire grammatical features that set them apart from their counterparts of sentence grammar: On the one hand, they

may be shorter, and on the other hand, they may have a more extended form. These features are now looked at in turn.

3.10.6.1 Truncation

Being frequently derived from noun phrases and separated from the latter by their functions, vocatives tend to become formally different with increasing usage. One common effect is that they may lose morphological and/or phonological substance.<sup>41</sup> English nouns, for example, require the use of articles in many contexts (e.g., *the doctor*), but when coopted as vocatives, they occur without articles (*Doctor!*). And in European Portuguese, proper names used referentially take the definite article, e.g., *o João* (DEF.SG John) ‘John.’ But the article disappears in vocative uses: *João!* ‘John!’ (Hengeveld and Mackenzie 2008: 80).

Truncations of form are particularly common when involving phonological rather than morphological substance. For example, English first names are frequently reduced to monosyllabic forms, e.g., *Elizabeth* > *Liz* or *Beth*, *Frederick* > *Fred*, *Joseph* > *Joe*, *Patricia* > *Pat*, etc. The following Romance examples in Tables 3.20 and 3.21 illustrate two common hypocoristic patterns of reduction, one involving final (Italian) and the other initial reduction (Catalan); see also Floricic (2011) and Floricic and Molinu (2012) for discussion and more examples.

Table 3.20 Reduction of  
some Italian first names

| Full form         | Hypocorism   |
|-------------------|--------------|
| <i>Alessandro</i> | <i>Ale</i>   |
| <i>Antonella</i>  | <i>Anto</i>  |
| <i>Patricia</i>   | <i>Patri</i> |
| <i>Teodoro</i>    | <i>Teo</i>   |

Source: Floricic and Molinu  
(2012: 158)

Such examples are far from exceptional: Vocative forms tend to be shortened versions of the noun phrases from which they are historically derived. Thus, Jakobson (1971: 10–11) notes that vocatives in Russian are reduced to the word stem, cf. *Maša* (nominative) vs. *Maš!* (vocative). In the Runyankore language of western Uganda, the initial vowel of a noun is omitted in the vocative, that is, when someone is addressed directly (Morris and Kirwan 1957: 150), and much the same

<sup>41</sup> Rather than reducing material, however, there are also some languages that use extra morphological material for vocatives, such as vocative case forms (see above).

**Table 3.21** Reduction of some Catalan first names

| Full form        | Hypocorism   |
|------------------|--------------|
| <i>Clotilde</i>  | <i>Tilde</i> |
| <i>Enriqueta</i> | <i>Queta</i> |
| <i>Manolita</i>  | <i>Lita</i>  |
| <i>Remei</i>     | <i>Mei</i>   |

applies to vocatives in a number of other Bantu languages which are characterized by the omission of the initial vowel of nouns when used as vocatives, cf. the examples from Zulu of South Africa in Table 3.22.<sup>42</sup>

**Table 3.22** Reduction of some Zulu nouns (Bantu, Niger-Congo)

| Noun           | Meaning  | Vocative      | Meaning     |
|----------------|----------|---------------|-------------|
| <i>abafana</i> | ‘boys’   | <i>Bafana</i> | ‘boys!’     |
| <i>inkosi</i>  | ‘chief’  | <i>Nkosi</i>  | ‘O chief!’  |
| <i>umuntu</i>  | ‘person’ | <i>Muntu</i>  | ‘O person!’ |

Source: [Doke \(\[1927\] 1988: 280\)](#)

Examples from North American languages include the following: In the Thompson language (Nlaka’pamuctsin) of British Columbia, there are *kík!* ‘mother!’ vs. *skíx-ze?* ‘mother’ and *sín!* ‘younger brother!’ vs. *sínci?* ‘younger brother’ ([Thompson and Thompson 1992: 186](#)). In the Yurok language of North-western California there are also vocative forms for kin terms that are less marked than nouns are in other functions, e.g., *kók!* ‘mother!’ vs. *kókos* ‘mother,’ or *pic* ‘grandfather!’ vs. *picowos* ‘grandfather’ ([Robins 1958: 23](#)), or Fox *ifkwε* ‘woman!’ vs. *ifkwε:wa* ‘woman’ ([Bloomfield \[1933\] 1962: 177](#)).

Similarly, in Ma Manda, a non-Austronesian language of Papua New Guinea, vocatives are for the most part the shortest and the most unmarked forms of kinship terms, commonly formed by omitting the possessive suffix, e.g., *bepma* ‘my father, father’s brother’ vs. *bep* ‘father!’, *taam(n)a* ‘my wife’ vs. *taam* ‘wife!’ ([Pennington 2016: 99–100](#)).

Vocatives share this feature with some other types of interactives like imperative directives, which may be marked off phonologically in some way (Section 3.2).

<sup>42</sup> Omission of the initial vowel of nouns may ultimately have involved a morphological rather than a phonological mechanism since the vowels are historically pre-prefixes.



The Logudorese dialect of Sardinian has what Floricic (2011: 114) calls a ‘minimality’ constraint: Autonomous information units or content words must have at least two syllables. The monosyllabic vocative and the monosyllabic (canonical) imperative, however, violate the essential constraint on disallowing monosyllabic autonomous forms and content words (Floricic 2011: 115).

In some languages, vocatives are formed by omitting some morphological element required on nouns, and this omission may result in ‘unusual’ forms. This appears to apply to the Palula language of northern Pakistan, where some kinship terms form vocatives by reduction of the corresponding nominative forms, appearing without the gender-signaling endings *-u* or *-i* (Liljegen 2016: 110), e.g.,

phéepi (father’s.sister.N.SG)  
vs. **phéep** (father’s.sister.VOC)

Similarly, in some Bantu languages, noun class pre-prefixes are obligatory parts of nouns, but these pre-prefixes are dropped to form vocatives, thereby creating ‘deviant noun forms.’ The Zulu example that was mentioned in Table 3.22 illustrates this pattern: A noun like *abafana* ‘boys’ loses its pre-prefix *a-* to give rise to the vocative form *bafana* ‘boys!’.

Finally, in the Hausa language of Northern Nigeria, ‘short-form names’ are created by chopping off the end of the word in vocatives, e.g., *Hànnatù* > *Hànnè*. Common nouns, except for many loanwords, normally end in a long final vowel; proper nouns including vocatives, by contrast, very often have a short final vowel (Newman 2000: 338).

However, there are also a few exceptions in that in some languages, vocatives require more material for their phonological expression than corresponding forms of sentence grammar. For example, in the Wampis language of the Peruvian Amazon, vocatives differ from the nominative form of nouns by adding a high tone on the last syllable or a high-toned suffix:

- (20) Wampis (Jivaroan family; Peña 2015: 492)  
nua-ru (wife-my) ‘my wife (nominative)’  
**nua-rú** (wife-my.VOC) ‘my wife (vocative)’

Similarly, in the Manambu language of the Ndu family of Papua New Guinea, vocatives are longer than corresponding non-vocative kinship nouns and personal names, e.g., *Máli*, vocative *Maliyé:(y)* ‘Mali!’ (Aikhenvald 2008: 44). In Sardinian vocatives, lengthening of the final vowel is attested for emphasis (Floricic and Molinu 2018).

There is no necessary contradiction between removing and adding phonological material in vocative forms since the two processes may serve different functions. While this issue is in need of more information one may wish to draw attention to cases such as the following. When vocatives are coopted from nouns in the Northern Wakashan language Southern Kwakiutl of British Columbia, they

lose final [a] or [ʔ], but when vocatives are shouted, -é is added to the form (Jacobsen 1994: 32).

### 3.10.6.2 Semantic extension

The following examples may show that the semantic space of social interaction is not necessarily the same as that of propositional sentence grammar: For example, when coopted from sentence grammar to usage as vocatives, kinship terms like ‘father’ or ‘mother’ lose the referential function they had within the sentence. Their referentiality is now defined by the situation of discourse (the context) and their semantic features can be redefined in accordance with the special social needs characterizing their use as interactives. The result is that they may acquire an extended usage, such as being associated with larger sets of referents defined by the relevant situation of discourse (Section 4.1).

The Spanish noun *hombre* ‘man, person’ has been coopted as a generic vocative used to establish contact with or get the attention of the hearer. And it has been grammaticalized to an interjection of surprise. Grammaticalization entailed desemanticization: On the one hand, *hombre* lost its appellative function in certain contexts; on the other hand, it largely lost its gender and number features: its use is no longer restricted to male and singular referents; rather, it can also be found to refer to female or plural referents, and even to animals (Kleinknecht 2019: 299).<sup>43</sup>

In the Bantu language Ha of western Tanzania, kinship terms tend to be used with a meaning that is ‘literally not true,’ as Harjula (2004: 151) says: When addressing people, kinship terms are often extended to a wider range of referents: Any older or respectable man can be called ‘my father,’ a woman can be addressed as ‘my mother’ and people of the same age and status as ‘siblings’. In a number of Bantu languages, the term for ‘mother’ is used to address female non-kin elders (Mitchell 2017: 291). Similarly, speakers of the Ulithian language of Micronesia have extended their kinship terms *tama-yi* ‘my father’ and *sila-yi* ‘my mother’ more generally to uses as polite vocatives for ‘Sir!’ and ‘Madam!’, respectively (Sohn and Bender 1973: 85).

Depending on the situation of discourse concerned, semantic extension tends to exhibit specific social constraints. Among the Nyakyusa of Malawi and western Tanzania, the situation is one of formal traditional meetings (Lusekelo 2021: 72). In such meetings, the mass attending the assembly uses the kinship terms *tata*, plural *batata* ‘father’ and *juba*, plural *bajuba* ‘mother’ in the following, hierarchically structured senses as address forms:

- The term *batata* is not only used refers to fathers but also to parents and all other participants of the meeting, both male and female, except young members of the community.

<sup>43</sup> In Asturian, *hombre* occurs in the reduced form *ho* (Franck Floricic, p.c. of August 5, 2021).

- The term *bajuba* not only includes mothers but all female participants except children.

In the Dagaare language of northwestern Ghana, some vocatives ‘show culturally conditioned respect and endearment. For instance, a woman can address a brother-in-law, a cousin-in-law, and a nephew-in-law as *ɪ sɪr* (‘my husband’) and in return receive the address *ĩpɔw* (‘my wife’). Also, a married woman and a male friend of her husband’s close friend will normally address each other as *ĩ sɛn* (‘my boyfriend/girlfriend’), while children address their paternal uncles as *Baba* (‘Dad’) (Mwinlaaru 2017: 135).

In the Palula language of northern Pakistan, some kinship terms used as vocatives also acquire a socially extended meaning in that the noun *phéepi* ‘father’s sister’ has the wider meaning of ‘middle-aged woman’ when used in the vocative form *phéep* (Liljegren 2016: 110):

- (21) Palula (Indo-Iranian, Indo-European; Liljegren 2016: 110)

phéep                      séer-i-ee

father's.sister.VOC fine-F-Q

‘How are you auntie [politely addressing any middle-aged woman]?’

Similarly, speakers of the Punjabi (Panjabi) language of northwest India and eastern Pakistan appear to have extended the kinship terms *saalaa* ‘brother-in-law’ and *saalii* ‘sister-in-law’ for rude address to call male and female referents, respectively (Hengeveld and Mackenzie (2008: 79). And speakers of the Wampis language of Peruvian Amazon have extended nuclear family terms to other relatives in vocative usages (Peña 2015: 495):

nuku 'mother' > 'woman of status within the family or social group'

iatsu- 'brother' > 'relative or close friend'

In the Rhineland area of western Germany, there used to be a usage before the turn of the millennium whereby children would address any unknown woman as *Tante* ‘aunt’ and an unknown man as *Onkel* ‘uncle.’ While this usage appears to have been discontinued, the following is still there: When shopping in a market place, a customer may be addressed as *Junger Mann!* ‘young man!’, and a female customer as *Junge Frau!* ‘young woman!’ even if the vendor may well be aware that the customer has passed their fifties or sixties. In a similar way, Bulgarian children would use the form *lelja* ‘aunt,’ usually in its diminutive form *lelichka*, and *chicho* ‘uncle,’ also usually in the diminutive form *chichko*, as terms of address for an unknown woman and an unknown man, respectively, before 1989 (Tania Kuteva, p.c. of March 9, 2022).

The shift from noun phrase to vocative may also involve what appear to be metaphorical strategies. Among the Wampis in Peruvian Amazon, fathers may use the noun *suki* ‘testicle’ to address their sons in a familiar way, and

the noun *juni* ‘type of larva’ is also used to refer to or address a child (Peña 2015: 494).

An example of extreme semantic context extension of a kin term used as a vocative address form has been reported from the Datooga language of north-central Tanzania, where the noun *îiyá* ‘mother’ is frequently transferred metaphorically from a child–mother relation to a range of other relations. Table 3.23 shows the kind of functions that *îiyá* may assume in such contexts. The functions of *îiyá* include establishing a fictive child–mother kinship relationship, or to express respectful affiliation between non-intimates. Most often, *îiyá* functions as a term of endearment or it ‘helps to mitigate potential impositions on the addressee’ (Mitchell 2017: 289–90).

The only clear constraint appears to be that adult male persons may not receive the vocative *îiyá*. Note that semantic extension also includes role inversion, in that a mother can call her child *îiyá* ‘mother.’ Thus, Mitchell (2017: 293–4) found a male toddler being told the following by his stepmother while she was feeding him:

- (22) Datooga (Southern Nilotic, Nilo-Saharan; Mitchell 2017: 294)  
léah      èan-óo-ga      **îiyá**  
drink.IMP   milk-PS-MR   mother  
‘Drink the milk, dear’

Role inversion in vocatives is a well documented phenomenon, for example, in Sicilian and Abbruzzese (Franck Floricic, p.c. of August 5, 2021).

**Table 3.23** Extensions of the term *îiyá* ‘mother’ in Datooga (Southern Nilotic, Nilo-Saharan)

| Nature of addressee                              | Main function signaled     |
|--------------------------------------------------|----------------------------|
| ‘Biological mother’ (default)                    | ‘Mother’                   |
| Female kin of ascending generation               | Respectful affiliation     |
| Female non-intimate peer (equals)                | Positive affective meaning |
| Intimates of the same generation                 | Affiliation and closeness  |
| Younger sister                                   | Endearment                 |
| Female children and small children of both sexes | Affiliation, affection     |
| Role inversion (e.g., mother to child)           | Endearment                 |
| After imperatives, disagreement, etc.            | Mitigation                 |

Source: Mitchell (2017: 289–97)

Creative innovations associated with the usage of vocatives can be observed, last but not least, in a situation of discourse involving interactions among teenagers.

Stenström (2020: 282) concludes that ‘some of the most frequent rude, and potentially insulting, vocatives are used as intimate address terms by English- and Spanish-speaking teenagers alike in their everyday conversational encounters.’ In the right context an utterance such as *You fucking bastard!* could mean ‘You are a great guy!’ among some teenagers in London, and the noun *cunt* is not restricted to refer to females; rather, it is sometimes used by a boy to address another boy, e.g., *fuck off you cunt*. And in the Spanish of Madrid a form like *cabrón*, presumably derived from the noun *cabrón* ‘billy goat,’ means ‘a person who tolerates husband’s or wife’s adultery’ or ‘asshole, fucker’ but can well be used also to address a boy (Stenström 2020: 292, 294).

### 3.11 Interaction with animals

The present section differs from all the preceding ones in not being restricted to human communication. It is concerned with what is referred to variously as ‘animal calls’ by Bynon (1976: 40), ‘directives’ by Amha (2013), ‘conative animal calls’ by Andrason and Karani (2021b: 3), and what we propose to call ‘animal directives,’ that is, with interactives directed to animals, classified by Ameka (1992a: 113–14) as ‘conative interjections.’ They form a subgroup of the directives discussed in Section 3.2 but differ from the latter in a number of ways and, hence, need to be looked at in their own right.

Interaction with animals involves on the one hand domestic animals, and on the other hand wild animals. The concern of the present section is primarily with the former. Communication with wild animals is said to be, as Bynon (1976: 41) puts it, more of the ‘mimicry type,’ frequently involving imitations of naturally occurring cries. It is functionally more restricted, involving mainly directives for dispersal (see Andrason and Karani 2021b: 10 for discussion).

#### 3.11.1 Introduction

Addressees of directives are normally human beings, but they may as well be animals, and the present section is devoted to the latter. Expressions directed to animals, both domestic and wild animals, in fact appear to be legion in the languages of the world.

What distinguishes animal directives from other kinds of interactives are most of all the interrelated features in (1) (cf. Ameka 1992b: 268; Andrason and Karani 2021b):

- (1) Features characterizing the use of animal directives
  - a There is an asymmetric social relationship between speaker and hearer in that the former is in a position of power over the latter.
  - b The hearer is non-human, and hence, not in command of a human language.
  - c Nevertheless, the hearer is expected to understand the content of the message and react to it accordingly.
  - d Communication is typically non-reciprocal: The speaker is always human and the hearer always an animal.
  - e. Communication is restricted to the spoken medium.

The generalizations proposed in (1) are presumably biased in favor of the human part in communication with animals. There is reason to suggest that exchanges between humans and animals can be reciprocal, considering that at least some animal species, including dairy cows, participate for example in greeting rituals (e.g., [Cornips 2022](#)).<sup>44</sup>

[Meinard \(2015: 153\)](#) claims that animal directives are ‘semantically empty calls.’ This is not the conclusion drawn here, as can be shown with the following example. In a number of languages there are different calls depending on whether one summons or disperses an animal. So for speakers of the Zargulla language in southwestern Ethiopia it makes quite some difference if they address their goats with *miŋo* ‘come here!’ or with *ʔiʃʃá* ‘go away!’ and the goats also seem to be aware of the semantic difference, usually reacting accordingly (cf. [Amha 2013: 224](#)). As was argued in Section 2.1, interactives, including animal directives, are meaningful, having both descriptive and expressive meaning content, and in the present example the descriptive meaning is one of deictic motion, namely motion either towards or away from the speaker.

For a world-wide survey of animal directives, see [Andrason and Karani \(2021b\)](#).

### 3.11.2 Working definition

Animal directives are directed to specific animals, requesting well defined types of action. The following definition is in the spirit of [Andrason and Karani \(2021b\)](#), who refer to animal directives as conative animal calls (CAC).

- (2) An animal directive is an interactive used to express a request or command directed to an animal.

The definition resembles that by [Andrason and Karani \(2021b: 33\)](#); see also [Andrason 2022](#)), according to whom animal directives (‘conative animal calls’) convey requests, wishes, desires, demands, or orders addressed to animals. And it is also

<sup>44</sup> We are grateful to Ad Foolen for having drawn our attention to this point.

similar to but not the same as that by Bynon (1976: 40): ‘Animal calling may, then, be defined as the transfer by acoustic means of information from a human being to an animal for the purpose of influencing its behavior.’ The latter definition is more general, it also covers acoustic means produced by the use of instruments such as whistles, drums, rattles, etc. The definition proposed here, by contrast, is essentially restricted to linguistic and paralinguistic means. Even if we will not entirely ignore extra-linguistic phenomena, our concern here is with interactives, that is, a linguistic category. Except for the fact that it refers to animals, (2) is in accordance with the definition that was proposed in Section 3.2.2 for directives in general. Examples of animal directives in various languages are provided in the remainder of this section.

### 3.11.3 Grammar

As just observed, animal directives can be classified as a subgroup of directives (Section 3.2), and they have the same argument structure, even if their H argument is an animal, or several animals, rather than a human. (3) is an example of the animal directive *ss* ‘move away (of cows)’ of the Arusa Maasai language of northern Tanzania. Its paraphrase and argument structure is found in (4). Animal directives in this language frequently combine with vocatives, as in (3), where the vocative form *na kiteŋ* ‘o cow!’ of the noun *en-kiteŋ* ‘cow’ is an adjunct referring to the H argument.

- (3) Arusa Maasai (Eastern Nilotic, Nilo-Saharan; Andrason and Karani 2021b: 30, (5a))

|                                   |              |                   |
|-----------------------------------|--------------|-------------------|
| <b>Na</b>                         | <b>kiteŋ</b> | <b>ss!</b>        |
| VOC:SG:F                          | cow          | DIR <sup>45</sup> |
| ‘Hey cow, <i>ss</i> [move away]!’ |              |                   |

- (4) a Paraphrase: I (S) ask you (cow, H) to move away  
 b Argument structure: Arusa Maasai *ss* (S,H)

As the descriptions that there are suggest, animal directives have all the grammatical hallmarks of directives and of interactives in general. For example, they can and frequently do occur as stand-alones and are normally syntactically and prosodically independent indexicals (see Andrason and Karani 2021b). However, they may also show features of syntactic integration. In the following example, the directive *wuu wuu* of Arusa Maasai takes what appears to be an object complement:

<sup>45</sup> The directive *ss* is glossed ‘CAC’ (conative animal call) by Andrason and Karani (2021b: 3), while here we use the general gloss for directives (DIR).

- (5) Arusa Maasai (Eastern Nilotic, Nilo-Saharan; [Andrason and Karani 2021b: 28, \(1\)](#))  
 [Directed to a dog]  
**Wuu Wuu** e-n-kitejo!  
 DIR DIR SG-F-rabbit  
 'Chase [*wuu wuu*] the rabbit!'

We observed at the beginning of this section that animal directives are first and above all 'directives,' expressing most of all directed motion. But there are exceptions, as can be illustrated with the examples from Georgian in (6): This language has an extremely rich inventory of animal directives, and some of them express emotional concepts rather than commands.

- (6) Animal directives in Georgian relating to emotional states ([Hewitt 1995: 99](#))
- |                                    |                               |
|------------------------------------|-------------------------------|
| <b>aga-gag-aa</b> or <b>azgamo</b> | 'to frighten dogs'            |
| <b>at'</b> or <b>čpveoo</b>        | 'sign of wrath to cattle'     |
| <b>cxa=cxa</b>                     | 'to frighten mice'            |
| <b>diido</b>                       | 'to settle a cow'             |
| <b>hae(e)</b>                      | 'to encourage bulls'          |
| <b>maamoh</b>                      | 'to show affection to cattle' |
| <b>mc'a</b>                        | 'to quieten a dog'            |
| <b>nč'a nč'a</b>                   | 'to show anger to dogs'       |
| <b>tpee oho</b> or <b>ogečo</b>    | 'to show anger to cattle'     |

#### 3.11.4 Usage and functions

Presumably the main function of animal directives is to express directed motion, telling an animal either to move towards the speaker ('venitives'), or to move away from the speaker or some other location ('andatives'); concerning the latter, referred to as dispersal calls, see also Section 3.11.6. The two kinds of motion are usually accompanied by contrasting patterns of prosody and gestures. Speakers of the Muna language of Sulawesi in Indonesia may call a cat by means of an alveolar click or by saying *mamamama* usually accompanied by tapping on the ground ([van den Berg 1989: 189](#)).

[Table 3.24](#) illustrates this distinction with examples from the Zargulla language of southwestern Ethiopia (for a similar inventory of directives in the Maale language, see [Amha 2001: 262](#)). A third, but clearly less common type of directives is that for 'Halt!'

An interesting observation on the distinction between venitives and andatives is reported by [Enfield \(2007: 315\)](#) from the Lao language of Laos: Whereas directives for movement away from the speaker involve forms as single units, forms for movements towards the speaker are usually repeated over and over.



**Table 3.24** Animal directives for deictic motion in Zargulla (Omotic, Afroasiatic)

| Animal sp. | Venitives                                           | Andatives                |
|------------|-----------------------------------------------------|--------------------------|
| Cows       | <i>ʔaa</i> (also ‘to stop/calm down when agitated’) | <i>káis</i>              |
| Oxen       | <i>ʔāā</i> (‘to lead to mating location’)           | <i>wāā</i>               |
| Goats      | <i>miʔo</i>                                         | <i>ʔiššá</i>             |
| Sheep      | <i>baáʔ</i>                                         | <i>ʔúss</i>              |
| Chicken    | <i>lúk, bíkʼ</i>                                    | <i>čúk</i>               |
| Cats       | <i>wurruu</i>                                       | <i>kíp</i>               |
| Dogs       | <i>háč(i)</i>                                       | <i>káč</i> or <i>kéč</i> |

Source: Amha (2013: 224)

Which function is ‘lexically’ distinguished is contingent on the nature of the kind of regular interaction that humans have with a given animal species. For example, Amha (2013: 224) found that the Zargulla have only an andative form for birds (*caáy*), to disperse birds, and only a venitive form for bees (*búl búl táná bir-á, kawó ʔamaat*), to attract bees, the latter serving to summon a queen bee and its colony.

But there is a range of additional functions, such as to initiate, continue, terminate, or modify the motion of the animal, to get its attention, to calm it, to make it feel more relaxed, to incite an animal to fight or to stop. Xhosa people in South Africa may say *hanewu!* to stop their cattle, *iyaphi!* to slow them down, *hoyi!* to line them up, and *truyi!* to let them move back (Andrason and Dlali 2020: 187).

In Ayt Hadiddu of Central Morocco there are up to ten different directives for each of the domestic animals, distinguishing various actions that the animal is expected to perform. Table 3.25 gives an impression of the kinds of action distinguished, using mules as an example. Of the ten directives distinguished, five are also used for horses and/or donkeys, while the remaining directives appear to be restricted to mules. Note that clicks and whistles are part of the inventory of directive forms.

On the basis of the data available, the addressees of animal directives are most of all domestic animals (cattle, sheep, goats, fowl, horses, donkeys), followed by pets (dogs, cats), while wild animals are the least common addressees, and for the last group, directives serve most of all as dispersals. In general, directives for wild animals are conceptually less differentiated than directives for domestic animals. Nevertheless, even the former may exhibit various conceptual distinctions. Speakers of the Southern Sotho language of Lesotho, for example, employ different forms for dispersal depending on which animal species is concerned, as illustrated in (7).

**Table 3.25** Functions of animal directives for mules in Ayt Hadiddu (Berber, Afroasiatic)

| Function                           | Form                              |
|------------------------------------|-----------------------------------|
| a To make it advance               | <i>ušt!</i>                       |
| b To make it come to the caller    | <i>ttzay!</i> (repeated)          |
| c To make it slow down or stop     | <i>ššt(a)<sup>n</sup>!</i>        |
| d To make it get up                | <i>šta-nšr!</i>                   |
| e To make it put itself straight   | <i>šta-ssšr!</i>                  |
| f To make it advance to be tied up | <i>šta-qqn!</i>                   |
| g To make it drink                 | Series of short bilabial whistles |
| h To make it roll                  | <i>kkšš!</i>                      |
| i To warn it before handling       | Unilateral velaric clicks         |
| j To chase or send it off          | <i>šta-ddu!</i>                   |

Source: Bynon (1976: 45)

- (7) Animal directives in Southern Sotho used for driving wild birds away from the fields (Guma 1971: 247)

|                        |                  |
|------------------------|------------------|
| <i>hobei, hoboi</i>    | Doves            |
| <i>holia</i>           | Sparrows         |
| <i>hukubi</i>          | Crows            |
| <i>khube, thakhube</i> | Red-bishop birds |

Ideally, there is a distinct form or set of forms for each of the domestic animals concerned. When speakers of the Zargulla language of southwestern Ethiopia want to stop the movement of an animal they use the directive *?aa?aa* for cows, *tooššé* for donkeys, and *noóc* for horses and mules (Amha 2013: 226). In addition, further terminological distinctions may be made according to the age or the sex of an animal. For example, the Berber-speaking Ayt Hadiddu of Central Morocco use the directive *mšši!* to chase away a grown-up dog, but *xišš!* to chase away a puppy (Bynon 1976: 48).

But not uncommonly, more than one animal species is addressed with the same term. Clusters of different animal species that may receive the same directive are, for example, sheep and goats, or horses, mules, and donkeys, or different species of birds. Thus, to send an animal away, speakers of the Ewe language of West Africa use *kái!* for sheep and goats, *sã!* for dogs and cats, and *súi!* for domestic birds (Ameka 1992b: 267). And among the Ayt Hadiddu of Central Morocco, horses, mules, and donkeys share the same directives for ‘Drink!’ and also for ‘Roll!’ and in a similar fashion, identical directives are used for sheep and goats for ‘Drink!’ and there are also the same terms for ‘Lick salt!’ (Bynon 1976: 52).

The directive *Aca!* of the Atlantic language Noon in Senegal is employed to chase away bigger animals, but it is also used to chase away children (Soukka 2000: 266).

In the Awtuw language of Papua New Guinea, dogs and pigs each have their individual names, and different sets of names are employed for male and female animals. Pigs are called by interspersing their name with a prenasalized, lax, voiced, protracted bilateral trill followed by a nasalized low front vowel. Chickens are called with a protracted voiced, apical-alveolar trill uttered with a high pitch followed by a series of very tense bilabial fricatives with expressive velar air (Feldman 1986: 196).

Animal directives can be combined within one and the same utterance. In the following example from Arusa Maasai, the directive *popopopo*, dedicated to chasing monkeys away, is followed by *fomo*, having a similar function but directed also to other animal species.

- (7) Arusa Maasai (Eastern Nilotic, Nilo-Saharan; [Andrason and Karani 2021b](#): 28, (2b))

|                 |             |          |
|-----------------|-------------|----------|
| <b>Popopopo</b> | <b>fomo</b> | naleŋ!   |
| DIR             | DIR         | very.far |

‘*Popopopo* [go away (to a monkey)] very far!’

As a rule, animal directives are restricted to interaction with animals. They are, however, occasionally also extended to humans, typically in a degrading sense. For example, [Bynon \(1976: 48\)](#) reports that an Ayt Hadiddu husband may be found using *mššimt!*, a feminine plural directive for chasing dogs away, but also to throw his wife and daughter out of the house.

### 3.11.5 Form

While animal directives have overall the structure of other interactives, they deviate in their phonological structure from sentence grammar more pronouncedly than any other types of interactives do. Directives may consist of a single consonant or contain phonetic features unheard of in other forms of language use. In Zargulla, for example, vowel nasalization is found in animal directives but not in sentence grammar ([Amha 2013: 223](#)). Furthermore, there are some other peculiarities setting animal directives off from other language material. In Arusa Maasai of northern Tanzania, for example, quadruplication of forms is a common phenomenon in animal directives (e.g., *OOOO* ‘to summon dogs’) but not elsewhere in the language.<sup>46</sup> [Andrason and Karani \(2021b: 26\)](#), in fact, observe that ‘triplicated and quadruplicated sequences, tend to be particularly common of CACs [conative animal calls; a.n.] across languages – perhaps, being nearly exclusive to this class of lexemes.’

<sup>46</sup> ‘O’ is a bilabial click.

Two particular features seem to be fairly common in animal directives, namely clicks and whistles. While click consonants are not uncommonly encountered in interjections and response signals (Section 3.8.4; but see also Section 3.3.5), in animal directives they are of especially high frequency. For example, to summon a dog, speakers of the Manambu language of Papua New Guinea use an alveolar click ([!]), cf. Table 3.26.

**Table 3.26** Animal directives in Manambu (Ndu family)

| Form                                  | Function                  |
|---------------------------------------|---------------------------|
| Alveolar click (repeated three times) | To summon a dog           |
| <i>diu diu diu diu</i>                | To call puppies           |
| <i>ps ps ps</i>                       | To call a cat (or a baby) |
| <i>ti ti ti</i>                       | To call chickens or ducks |

Source: Aikhenvald (2008: 44)

In the Tjwao language of Zimbabwe, itself a click language of the Khoisan Sprachbund in southern Africa, fowls are addressed with alveolar clicks ([!-!-!]) and dogs with dental clicks ([!/-!/-]) (Andrason, Fehn, and Phiri 2020: 14).<sup>47</sup> Similarly, speakers of the Tswana language of Botswana, also a click language, use a series of four alveolar ('prepalatal') clicks ([!-!-!-!]) in calling fowls (Cole [1955] 1987: 396).

Andrason and Karani (2021b) found three click types in use among the Arusa Maasai of northern Tanzania, namely a lateral (//), a bilabial (⊙), and a palatal (ʃ) one, all three occurring as stand-alone interactives.<sup>48</sup> In this language there is a common combination of four clicks of the same type, where // // // // serves, on the one hand, as a directive to donkeys and, on the other hand, to summon goats, ⊙ ⊙ ⊙ ⊙ to summon dogs, and ʃ ʃ ʃ to summon cats (Andrason and Karani 2021b: 13). In Ayt Hadiddu of Central Morocco, Bynon (1967: 53) counted six different clicks used in animal directives, even if the exact nature of the clicks is not described in detail. And speakers of Muna in Sulawesi, Indonesia, may use an alveolar click to call a cat (van den Berg 1989: 189).

Clicks in animal directives are for the most part used on their own, even if they are almost invariably reduplicated or repeated. They thus differ from the clicks in the sentence grammar of Khoisan languages and some South African Bantu languages in not combining with vowels and consonants (Section 1.3.2.8).

<sup>47</sup> For good reasons, 'Khoisan' is nowadays avoided as a term of genetic linguistics. Rather than as a diachronic term, it is used here strictly as a *referential* term for all the languages in southern and eastern Africa that have traditionally been subsumed under the label 'Khoisan.' Nowadays, these languages tend to be referred to as 'Khoe and San,' but whether 'Khoe and San' really constitutes a linguistically significant unit remains debatable.

<sup>48</sup> Andrason and Karani (2021b) present the three phonetic click symbols presented here but do not use the articulatory labels proposed here.

Furthermore, [Bynon \(1976: 54\)](#) found click interactives in Ayt Hadiddu to be more strongly polysemous than other animal directives. For example, his ‘Click 1’ (whose exact phonetic identity is not entirely clear) is not only used to make a horse, mule, or donkey come to one but also to make a domestic rabbit come to be fed, as well as to encourage sheep and goats to lick salt.

More than other interactives, and distinctly more than forms of sentence grammar, animal directives draw on extra-linguistic means, and the boundary between linguistic and extra-linguistic means of expression is at times fluid.<sup>49</sup> For example, speakers of Arusa Maasai were found to include five whistles in their inventory of animal directives, described by Andrason and Karani thus:

The meaning of the five whistles is also extremely specific. Whistle 1 is used to summon a dog that is not far away; whistle 2 is used to call a dog from a distance; whistle 3 is used to summon flocks; whistle 4 is produced to calm flocks when pasturing; and whistle 5 is used to calm cattle, for example cows when milking them or an ox when injecting medicine or putting a yoke on its neck.

([Andrason and Karani 2021b: 13](#))

Consider the following example, which consists entirely of interactives: In addition to the vocative (VOC) *na kineji* ‘o goats!’ and four animal directives (DIR), called ‘conative animal calls’ by [Andrason and Karani \(2021a\)](#), and one of them is a whistle used to interact with goats (WHISTLE.3):

- (8) Arusa Maasai (Eastern Nilotic, Nilo-Saharan; [Andrason and Karani 2021a: 31, \(8a\)](#))

|                                                                   |               |                     |                  |           |            |
|-------------------------------------------------------------------|---------------|---------------------|------------------|-----------|------------|
| <b>Na</b>                                                         | <b>kineji</b> | <b>prupruprupru</b> | <b>WHISTLE.3</b> | <b>ae</b> | <b>ae!</b> |
| VOC:F                                                             | goat.PL       | DIR                 | DIR              | DIR       | DIR        |
| ‘O goats <i>prupruprupru</i> <i>whistle.3 ae ae</i> [come here]!’ |               |                     |                  |           |            |

Animal directives can be divided into three classes (cf. [Andrason and Karani 2021b: 21](#)):

- (9) Three classes of animal directives
- a Primary forms
  - b Secondary forms
  - c Borrowings

The ordering in (9) presumably reflects the relative frequency of occurrence in which animal directives occur cross-linguistically. The distinction between primary and secondary forms was discussed in Section 3.6.4.1.

Primary directives differ from secondary ones in that they are not transparently related to any forms of sentence grammar, they are etymologically opaque, even if

<sup>49</sup> Since this study is confined to linguistic expressions, extra-linguistic means of expression can only be mentioned in passing, requiring separate analysis.

they may have an onomatopoeic source. Secondary interactives are, as is argued here, mostly derived from forms of sentence grammar via cooptation (Section 4.1), hence they resemble their etymological sources in form and meaning. Table 3.27 gives a few examples of the three classes of animal directives from the Arusa Maasai language of northern Tanzania, where primary directives form the clear majority of forms. Note that the four primary directives mentioned in the table, *bee*, *kurukuru*, *mee*, and *naau*, appear to have an onomatopoeic base (Andrason and Karani 2021b: 21–2).

**Table 3.27** Classes of animal directives in Arusa Maasai (Eastern Nilotic, Nilo-Saharan)

| Class of interactive | Animal directive | Meaning    | Sentence grammar source | Meaning                       |
|----------------------|------------------|------------|-------------------------|-------------------------------|
| Primary              | <i>bee</i>       | ‘Goat!’    |                         |                               |
|                      | <i>kurukuru</i>  | ‘Chicken!’ |                         |                               |
|                      | <i>mee</i>       | ‘Sheep!’   |                         |                               |
|                      | <i>naau</i>      | ‘Cat!’     |                         |                               |
| Secondary            | <i>fomo</i>      | ‘Go!’      | < <i>fomo</i>           | ‘Go!’ (imperative)            |
|                      | <i>wou</i>       | ‘Come!’    | < <i>wou</i>            | ‘Come!’ (imperative)          |
|                      | <i>kine</i>      | ‘Goat!’    | < <i>ol-kine</i>        | ‘goat’                        |
| Borrowing            | <i>kamata</i>    | ‘Take it!’ | < <i>kamata!</i>        | ‘Take it!’ (Swahili loanword) |
|                      | <i>toka</i>      | ‘Get out!’ | < <i>toka!</i>          | ‘Get out!’ (Swahili loanword) |

Source: Andrason and Karani (2021b: 21–2)

Being directives, secondary animal directives are often coopted from imperative forms of verbs. These forms tend to be gradually grammaticalized into invariable interactives, whereby the forms lose their internal compositionality. In their early development, the forms are still indistinguishable from verb forms. In their transitional stages from inflected verb form to unanalyzable interactive there may be traces of the earlier situation in that animal directives still exhibit morphological relics, such as optional imperative or number affixes.

While occurring frequently as stand-alones, animal directives can combine in various ways with segments of sentence grammar. In Ayt Hadiddu, for example, a directive is frequently followed by a sentence grammar comment. Thus, in the following example the speaker shouts a text piece at a mule that is going too fast for his comfort, where *šštaaa!* is a mule directive (cf. *ššt(a)*<sup>n</sup> ‘to make a mule slow down, stop, remain stationary’), followed by an elaboration in sentence grammar.

- (10) Ayt Hadiddu (Berber, Afroasiatic; [Bynon 1976](#): 44, 56–7)<sup>50</sup>

šštaaa! ddu ss-ttawil! matta twada-nnay?

‘Slow! Go gently! What’s that gait?’

### 3.11.6 Voiceless sibilants in animal dispersal calls

We saw in earlier sections that at least some types of interactives exhibit a sound symbolic behavior. Attention was drawn in particular to attention signals (*\*hey*; Section 3.1.5) and a subtype of response elicitors, called other-initiated repair markers (*huh?*; Section 3.7.5). Mention was also made of the predilection of response signals to be expressed by clicks, and of forms for ‘yes’ to contain the laryngeal phonemes /ʔ/ or /h/ (Section 3.8.4).

Similar observations have been made by [Andrason \(2022\)](#) on one subtype of animal directives, namely dispersal calls. Dispersal calls, that is, animal directives used for repelling animals, are likely to be built around voiceless sibilants, as has been established in a worldwide survey by [Andrason \(2022\)](#).

The most relevant types of sibilants recruited to this end are alveolar ([s]), partially palatalized post-alveolar or palato-alveolar ([ʃ]), alveolo-palatal ([ɕ]), and retroflex ([ʂ]) fricative sounds as well as corresponding affricates. A few examples are provided in (11), for a comprehensive treatment see [Andrason \(2022\)](#), on which the present discussion rests.

- (11) Examples of voiceless sibilants in animal dispersal calls (Source: [Andrason 2022](#), Table 1)

|            |                                   |              |        |
|------------|-----------------------------------|--------------|--------|
| Basque     | (Isolate)                         | <b>sapi</b>  | [sapi] |
| German     | (Germanic, Indo-European)         | <b>ksch</b>  | [kʃ]   |
| Hausa      | (Chadic, Afroasiatic)             | <b>uss</b>   | [us:]  |
| Indonesian | (Malayo-Polynesian, Austronesian) | <b>cik</b>   | [tʃiʔ] |
| Lushai     | (Tibeto-Burman, Sino-Tibetan)     | <b>sik</b>   | [sik]  |
| Ndebele    | (Bantu, Niger-Congo)              | <b>shi</b>   | [ʃi]   |
| Vietnamese | (Vietic, Austroasiatic)           | <b>[sʰi]</b> | [sʰ]   |

[Andrason’s \(2022\)](#) survey comprises animal directives from 79 languages, yielding overall 247 voiceless-sibilant dispersal calls, that is, more than three calls per language. His analysis results in a range of generalizations, including the following:

- (a) All 79 languages of his survey contain more or less common, entrenched, and/or conventionalized dispersal calls built around voiceless sibilants.
- (b) 55 of these languages make use of more than one voiceless-sibilant dispersal call.

<sup>50</sup> No glosses are provided by the author.

- (c) In the case of 24 languages showing only one voiceless-sibilant dispersal call, that call tends to be by far the most typical and salient among all dispersal calls available.
- (d) Voiceless sibilant dispersal calls tend to be pronounced with a powerful blow of airstream.
- (e) They frequently contain vowels and/or consonants but no clicks or whistles. As we saw in Section 3.11.5, clicks and whistles are fairly common in animal directives but are absent in voiceless sibilant animal dispersal calls.

As in the case of the other interactives mentioned at the beginning of this section, the physiological base underlying sound symbolic behavior is essentially unclear and in need of more research.

### 3.11.7 Discussion

In his survey of animal directives in the Berber language Ayt Hadiddu of Central Morocco, [Bynon \(1976: 57\)](#) notes that it would seem simplest to treat domestic animal calling as a register of language specifically reserved for addressing domestic animals and characterized by certain special features and restrictions. But then he concludes:

Such a description, however, even though it may account adequately for the caller's utterances, may not be the best from the point of view of man-animal communication. To account for the way in which this actually takes place—WHEN in fact it DOES take place—a description which treats domestic animal calling as operating through a restricted set of discrete acoustic signals (optionally supplemented by some elements of normal language) may prove more effective.  
([Bynon 1976: 57](#); upper-case print in the original)

This conclusion comes close to the one proposed here. Animal directives are in fact part of language, but as is argued here, not part of sentence grammar: They appear to be indexicals in accordance with our definition of interactives in Section 1.2, being a sub-type of directives (Section 3.2) characterized by the function-specific features mentioned in the preceding paragraphs.

For good reasons, animal directives are described by [Andrason and Karani \(2021b\)](#) as a prototype, and in fact these directives do not form a discrete category clearly set off from sentence grammar, and the boundary between the two is generally fluid. Less prototypical instances of animal directives show various kinds and degrees of integration in sentence grammar via grammaticalization or lexicalization.



The Pitjantjatjara language of Central Australia illustrates an instance of a completely grammaticalized animal directive: The expression (‘interjection’) *pai!*, directed to dogs and meaning ‘clear off!, be off!, piss off!’, is the source of the transitive verb *paiṇi*, which means ‘to drive someone or something away; to shoo off, to hunt away’ (Wilkins 1992: 131). An example of a less grammaticalized instance of an animal directive but of a kind that can quite commonly be observed is the following: In the Ewe language of Togo and Ghana, the animal directive *súi!*, directed to domestic birds, appears to have acquired the status of a complement of the verb *dó* ‘say’, as in (12).

- (12) Ewe (Kwa, Niger-Congo; Ameka 1992b: 268)  
dó **súi** ná koklo-á.  
say *súi* (INT) to hen-DEF  
‘Say ‘*súi*’ to the hen.’ [Sending the hen away]

Integration of an animal directive may give rise to lexicalized or idiomatic phrases. For example, in the Xhosa language of South Africa, the verb *ukuthi* ‘to say’ is combined with the directive (‘interjection’) *hekh*, used to urge oxen to advance, in order to form the fixed phrase *ukuthi hekh* (lit. ‘to.say *hekh*’) ‘start working with oxen’ (Andrason and Dlali 2020: 199).

Finally, there is a pathway leading from animal directives to sentence grammar whereby the usage of the former is extended, turning into referential lexical items (Section 4.2). In Table 3.28 there are a few examples from the Muna language of Sulawesi in Indonesia (van den Berg 1989: 189), showing how the interactives concerned acquire uses as verbs.

This pathway seems to be particularly widespread in the domain of nursery language. Table 3.29 provides a list of selected directives that provided the source of nursery terms for animals in Ayt Hadiddu of Central Morocco, the nursery terms being phonologically adapted versions of the directives (for more detailed discussion, see Bynon 1976: 61).

Concerning the role played by interactives in the exchange with small children, see also Section 3.2.6.

**Table 3.28** From animal directive to verb in Muna (Malayo-Polynesian, Austronesian)

| Animal directive         | Meaning                      | Verbal usage          | Meaning                                                        |
|--------------------------|------------------------------|-----------------------|----------------------------------------------------------------|
| <i>koro-koro; krrrrr</i> | ‘to call chickens’           | <i>ae-koro-koro</i>   | ‘I call chickens’                                              |
| <i>sigha</i>             | ‘to chase away cats or dogs’ | <i>ae-sigha bheka</i> | ‘I chase a cat ( <i>bheka</i> ) away by calling <i>sigha</i> ’ |
| <i>sio, siu, sy(o)</i>   | ‘to chase away chickens’     | <i>ae-sio</i>         | ‘I chase chickens away by calling <i>sio</i> ’                 |

Source: van den Berg (1989: 189)

**Table 3.29** Nursery terms derived from animal directives in Ayt Hadiddu (Berber, Afroasiatic)

| Animal directive           | Meaning                         | Nursery language    | Meaning   |
|----------------------------|---------------------------------|---------------------|-----------|
| <i>h(zi)<sup>n</sup>!</i>  | ‘To make a goat kid come’       | <i>zizi, hzizi</i>  | ‘goat’    |
| <i>khullu!</i>             | ‘To make a chicken come’        | <i>khullu</i>       | ‘chicken’ |
| <i>rra!</i>                | ‘To make a donkey advance’      | <i>rrarra</i>       | ‘donkey’  |
| <i>sipsil!</i>             | ‘To make a cat come’            | <i>sibsi, sipsi</i> | ‘cat’     |
| <i>ššt(a)<sup>n</sup>!</i> | ‘To make a mule or donkey stop’ | <i>štašta</i>       | ‘mule’    |
| <i>zaw!</i>                | ‘To make a cow advance’         | <i>zawa</i>         | ‘cow’     |

Source: Bynon (1976: 61)

Rather than grammaticalizing animal directives to words of sentence grammar, the process can be restricted to a shift in function from communication with animals to communication with humans. For example, in the Shona language of Zimbabwe, usage of the animal directive *fútsèkì!* ‘Stop barking!’ was extended from dogs to people with the meaning ‘Get lost!’ (Brauner 1993: 228)—thus, *fútsèkì!* retained its status as a directive (Section 3.2).

### 3.11.8 Conclusions

The set of conventionalized expressions of animal directives is usually small in western industrial societies but can be fairly extensive in other societies; Andrason and Karani (2021b: 9) found 46 of them in the Arusa Maasai language of northern Tanzania, and an even larger set exists in Ayt Hadiddu of Central Morocco. The directives are especially common in societies having an economic focus on animal husbandry and among people dealing professionally with animals. And it can be said to have features of a ‘mini-grammar.’

Animal directives can presumably be created spontaneously, but there is usually a core set that is stored and retrieved whole, having to be learned and being transmitted from one generation to another.

As aptly pointed out by Andrason and Karani (2021b), animal directives allow for a number of generalizations.<sup>51</sup> It is in particular the features in (13) that can be expected to be found in different languages.

- (13) Features to be expected in animal directives
- a They constitute a set of conventionalized expressions.
  - b Being conventionalized, they have to be learned by speakers.

<sup>51</sup> Proposing a cross-linguistic prototype of what they refer to as conative animal calls, Andrason and Karani (2021b) present eighteen semantic-pragmatic, phonetic, morphological, and syntactic features characterizing this prototype; see there for more details.

- c Like other directives they typically express a request for action.
- d The action requested most commonly concerns motion.
- e If the hearer is a wild animal the request relates most of the time to dispersal.
- f They have a consonant bias, not seldom consisting of a single consonant.
- g They make extensive use of click consonants.
- h They also make extensive use of whistles and other extra-linguistic means.
- i They are commonly reduplicated, triplicated, or quadruplicated.

In concluding, it should be emphasized that the account presented in Section 3.11 is not entirely impartial. As pointed out in the introduction to the section, communicative interaction is not restricted to the human part; rather, animals as well have been found to take part in such interactions.

### 3.12 An overview

After having provided a survey of the ten types of interactives examined in the preceding sections of this chapter, salient features distinguishing the types of interactives are now summarized in Table 3.30. Interactives differ from one another, on the one hand, in the functions they express. Since these functions were already discussed at the beginning of this chapter in Table 3.1, they are ignored here. On the other hand, they also differ in their grammatical features, namely the ones listed in (1).

- (1) Features distinguishing interactives
  - a Argument structure
  - b Forming utterances of their own
  - c Special phonological features
  - d Reduplication

#### (a) Argument structure

As we saw in Chapter 2, the following argument configurations have been observed in prototypical usages of interactives:

|       |                           |
|-------|---------------------------|
| S     | (speaker),                |
| T     | (theme),                  |
| S,H   | (speaker, hearer),        |
| S,T   | (speaker, theme), and     |
| S,H,T | (speaker, hearer, theme). |

Where arguments are placed relative to the interactive they determine depends on the functions they express, as discussed in the various sections of this chapter.

#### (b) Forming utterances of their own

While interactives commonly combine with surrounding text pieces, one of their most salient characteristics is the ability to occur as ‘stand-alones,’ that is, to form utterances of their own.

A notable exception is provided by text-related interactives, and most of all by discourse markers, which usually do not stand alone but co-occur with their textual arguments. However even discourse markers, such as *anyway*, *however*, *okay*, or *so*, may appear as stand-alones, forming the only linguistic form of a conversational turn (e.g., [Raymond 2004](#)).

#### (c) Special phonological features

‘Unusual’ phonological features have been reported in a number of studies; they were discussed in Section 1.3.2.2 of the first chapter. In fact, a number, though not all types of interactives are at variance with some phonetic and phonotactic constraints observed in sentence grammar.

#### (d) Reduplication

Reduplication is overall not a big issue in interactives. To be sure, it is a conspicuous mechanism in ideophones, which quite commonly reduplicate ([Dingemanse 2015](#); [Haiman 2018](#): 140–6), as in Korean *c'onc'on* ‘woven tightly,’ *colcol* ‘flowing liquid,’ or Siwu *mukumuku* ‘mumbling mouth movements,’ *gelegele* ‘glittery appearance,’ *gbadara-gbadara* ‘a drunkard’s wobbling gait’ ([Dingemanse 2018](#): 2). And much the same applies to animal directives (Section 3.11.5), which—like ideophones—not only tend to be reduplicated but also triplicated and quadruplicated, as shown by [Andrason and Karani \(2021b](#): 26).

In interactives other than ideophones and animal directives, however, repetition does not seem to differ dramatically from that found in sentence grammar, and this applies even to ‘response cries’ like *ouch* or four-letter words like *shit*, that is, interjections which have often been lumped together with ideophones.<sup>52</sup> Consider the following typological observations made by [Haiman \(2018](#): 137):

- In Quechua of Ecuador, 37 out of 56 ideophones reduplicate, but only one out of 10 interjections do.
- In Nkore-Kiga of southwestern Uganda, 11 out of 12 ideophones reduplicate, but only one out of 11 interjections does.
- In Kannada of southwestern India, 25 out of 43 ideophones reduplicate but only 2 out of 12 interjections do.

<sup>52</sup> Drawing attention to the English expressions *my my* or *now-now*, however, [Ameka and Wilkins \(2006](#): 7) argue that ‘reduplication tends to be an input process to the formation of interjections.’

- In Tamil of southern India and Sri Lanka, 19 out of 29 ideophones reduplicate, while only 2 out of 15 interjections do.
- In the Totonacan language Upper Necaxa Totonac of Central Mexico reduplication does not occur for most morphemes, but it is more or less the norm for ideophones.

The features examined in [Table 3.30](#) are restricted to rough generalizations and to prototypical tokens of a given type, but most of the ten types also exhibit additional structures associated either with specific tokens of interactives or with specific contexts in which a given interactive is used. More fine-grained discussions of the variables were presented in the survey of interactive types in the preceding sections of this chapter.

No claim is made to the effect that the total numbers of features listed in [Table 3.30](#) have any grammatical significance. Whether presence vs. absence of a particular argument is in any way relevant for determining the structural significance of a given type of interactives must remain debatable. Note that the highest figure of 5.0 is that of animal directives, which in some sense can be considered to be the most peripheral of all interactives since they are the only one taking a non-human argument.

**Table 3.30** Distinguishing grammatical features of interactives

| Type               | Arguments typically expressed (optional arguments in brackets) |          |          | They can form utterances of their own | They may show special phonological features | They are frequently reduplicated | Total number of features |
|--------------------|----------------------------------------------------------------|----------|----------|---------------------------------------|---------------------------------------------|----------------------------------|--------------------------|
|                    | <i>S</i>                                                       | <i>H</i> | <i>T</i> |                                       |                                             |                                  |                          |
| Attention signals  | +                                                              | +        | (+)      | +                                     | -                                           | -                                | 3.5                      |
| Directives         | +                                                              | +        | -        | +                                     | -                                           | -                                | 3.0                      |
| Animal directives  | +                                                              | +        | -        | +                                     | +                                           | +                                | 5.0                      |
| Discourse markers  | +                                                              | (+)      | +        | (+)                                   | (+)                                         | -                                | 3.5                      |
| Evaluatives        | +                                                              | +        | +        | +                                     | -                                           | -                                | 4.0                      |
| Ideophones         | -                                                              | -        | +        | +                                     | +                                           | +                                | 4.0                      |
| Interjections      | +                                                              | -        | -        | +                                     | +                                           | -                                | 3.0                      |
| Response elicitors | +                                                              | +        | +        | +                                     | (+)                                         | -                                | 4.5                      |
| Response signals   | +                                                              | +        | +        | +                                     | (+)                                         | -                                | 4.5                      |
| Social formulae    | +                                                              | +        | (+)      | +                                     | -                                           | -                                | 3.5                      |
| Vocatives          | +                                                              | +        | (+)      | +                                     | (+)                                         | -                                | 4.0                      |

*Note:* In the calculation of the total number of features, obligatory arguments are given the numerical value '1.0' and optional ones '0.5'.

### 3.13 Conclusions

In the first two chapters we were aiming to describe interactives as a linguistic category of its own. The question that was central to the present chapter then was how interactives differ from one another, and how differences between them can be described in terms of the framework established in the preceding chapters.

The catalog of the ten types distinguished here by no means exhausts the range of expressions that qualify as interactives—an issue that is to be left for future research to come up with a more comprehensive account.

# 4

## Development

In a nutshell, the problem to be dealt with in this chapter is the following: While sentence grammar and interactive grammar can clearly be distinguished on account of their grammatical structure (Section 1.2), there are interactives that are transparently related or similar to corresponding units of sentence grammar and vice versa. The question therefore is why this is so, and in this chapter an attempt is made to answer this question.

The chapter differs from the preceding ones in that it is devoted to language change. The goal of the chapter is to understand why interactives are structured the way they are. The goal thus is an explanatory one, even if we are restricted to diachronic explanations—that is, to understanding how and why phenomena to be observed today came to be what they now are.

The reconstructions to be presented rest in a few cases on historical text data. But in most cases such data are not readily available, that is, there is no or no appropriate diachronic evidence. In such cases, reconstruction work is confined to synchronic comparisons based on general observations on grammatical change. The hypotheses on the grammatical changes proposed must therefore be taken with care, being subject to verification by means of more detailed historical evidence.

The methodology employed for reconstruction rests most of all on grammaticalization theory. But, as we will see in the sections to follow, this theory does not take care of all developments of interactives, and there will also be alternative tools of reconstruction to be recruited.

It is two mechanisms that account for most of the grammatical properties of interactives, namely cooptation, which is to be discussed in Section 4.1, and grammaticalization, which is the subject of Section 4.2. While grammaticalization and cooptation are general mechanisms, there are also two more special mechanisms shaping the form of interactives, which are discussed in Section 4.3. Finally, in Section 4.4 a line of development is discussed whereby some forms of interactive grammar may gradually be integrated in sentence grammar.

We saw in Chapter 3 that interactives can be classified in various ways. One classification, discussed especially in Section 3.6.4.1, concerns the distinction between primary and secondary interactives: The former differ from the latter in that they are not transparently related to any forms of sentence grammar. Unlike secondary interactives, primary ones, such as English *eh?*, *hey*, *mhm*, *uh*, *um*, *wow*,

or *yuck*, are as a rule etymologically opaque, that is, their history is unknown. This chapter, which will be dealing with the diachrony of interactives, is therefore mainly concerned with secondary interactives. As we will see, however, even primary interactives can conform to the principles of grammaticalization discussed in this chapter (e.g., Section 4.2.2.2).

## 4.1 Cooptation: How interactives arise

Take the following pair of text examples illustrating different usages of the English form *boy*:

- (1) English (Biber et al. 1999: 389, 1103)
  - a Later, after dark, a boy brought him a plate of food. (FICT)
  - b **Boy**, there's a lot of rocks, **huh?** (AmE)

The two usages of *boy* in (1), let us call them *boy*<sub>1</sub> (1a) and *boy*<sub>2</sub> (1b), differ from one another to the extent that they are suggestive of two contrasting grammatical and semantic spaces: Whereas *boy*<sub>1</sub> is a paradigm instance of an English noun, *boy*<sub>2</sub> is of a different kind. For one thing, there is a grammatical difference: *boy*<sub>1</sub> takes an article, it can be pluralized (*boys*), it has a syntactic function as the subject of the sentence in (1a), it is prosodically integrated in the sentence, and it can only occur in specific positions of a sentence. In contrast, *boy*<sub>2</sub> has none of these properties. For another, there is a semantic difference: *boy*<sub>1</sub> is a noun typically denoting a young male person, and it is part of the meaning of the sentence. Again, *boy*<sub>2</sub> lacks these properties: Neither does *boy*<sub>2</sub> necessarily refer to a young male person nor is it part of the meaning of the sentence following it, and unlike the meaning of *boy*<sub>1</sub>, it is what was called in Section 2.1.3 an 'expressive depiction'.

The two usages of *boy* can be said to be diachronically related in that *boy*<sub>2</sub> is assumed to be historically derived from *boy*<sub>1</sub>. There must have been a period in the history of English when there was *boy*<sub>1</sub> but not yet *boy*<sub>2</sub>, and the shift from the former to the latter cannot be accounted for by means of grammaticalization theory, or of any other established method of historical linguistics.

Examples like this are legion in English, or in any other language for that matter. Take another English noun, *fire*, to illustrate the effects that this kind of shift may have: *Fire*<sub>1</sub> is a good noun of sentence grammar, whereas *fire*<sub>2</sub> is an interactive derived from the former, and as such it has quite different usages, as can be seen, for example, when someone is found shouting *Fire!* in a crowded movie theater.

Recent research suggests that the mechanism responsible for the change from the former to the latter was cooptation, and this mechanism is the subject of Section 4.1.1.



### 4.1.1 The mechanism

A question raised in much of the work of Foolen (1997, 2012, 2016) is the following: What accounts for semantic shifts from the conceptual or cognitive domain to the domain of expressive forms? A partial answer is provided in the present section: Forms having descriptive or conceptual meaning may be coopted for specific discourse functions, thereby acquiring expressive meanings of the kind observed in interactive grammar.

Many, though not all interactives are hypothesized to be the result of cooptation,<sup>1</sup> that is, a cognitive-communicative operation which enables speakers to switch their perspective from the level of reasoning anchored to the meaning of sentences to a meta-level of reasoning immediately anchored to and shaped by the situation of discourse. By means of this operation, speakers transfer expressions available on the former level to express functions on the latter level for structuring their texts (cf. Heine et al. 2017, 2021: 65–72). Cooptation is defined in the following way:

Cooptation is a fully productive operation whereby a chunk of sentence grammar, such as a word, a phrase, a reduced clause, a full clause, or some other piece of text, is deployed for use on the metatextual level of discourse processing, thereby turning into a thetical. Its functions are determined by the situation of discourse, serving (a) to overcome constraints imposed by linearization in structuring texts, (b) to provide the source of information, (c) to place a text in a wider perspective, e.g., by elaborating, proffering an explanation, a comment or supplementary information, (d) to describe the attitudes of the speaker, and/or (e) to interact with the hearer.

(Heine et al. 2021: 67, (8))

Cooptation can be viewed as a kind of *conversion* whereby one thing (e.g., *boy*<sub>1</sub> in example (1)) is converted into another thing (*boy*<sub>2</sub>). In linguistics, conversion, also known as zero derivation, is usually treated as a process of word formation leading to the creation of new words via transfer from one word class to another without involving any morphological material. Converted units may be nonce formations, but they may as well be conventionalized into regularly used forms, as in the noun-to-verb conversion of *chair* (*a chair* > *to chair a meeting*) or the verb-to-noun conversion of *catch* (*to catch* > *a good catch*).

Cooptation shares both features just mentioned with conversion of the kind just mentioned but there is a crucial difference: Conversion of the latter kind takes place within sentence grammar: When the noun *chair* was converted to a verb it

<sup>1</sup> The phrase ‘though not all interactives’ refers to the fact that in the case of primary interactives (Section 3.6.4.1) no cooptation is involved since they do not originate in pieces of sentence grammar.

remained an element of sentence structure. In contrast, when *boy*<sub>1</sub> was coopted as *boy*<sub>2</sub> this led to a change from a noun of sentence grammar to an interactive, no longer being a semantic, syntactic, or prosodic element of a sentence—in other words, there was transfer from the sentence level to the level of discourse processing.

Depending on their particular discourse function, coopted text chunks are placed most commonly at the left or right periphery of a sentence, but they may as well appear within the sentence without being grammatically integrated in the sentence, and they can also be used as stand-alones, forming utterances of their own.

As stated in the definition, cooptation is a ‘fully productive operation,’ that is, it can involve various kinds of text chunks. Coopted text chunks need not be and frequently are not full-fledged clauses or phrases, they can be ‘reduced clauses’ like *you know* or subordinate clauses like *as it were*, both chunks having turned into English discourse markers (Section 3.3).

Two examples may illustrate the effects of cooptation, each leading to a different type of interactive. The first example concerns the English adverb *still*, cf. (2a), which was coopted as a discourse marker, cf. (2b). The shift of *still* from participant of sentence grammar to discourse marker in (2a) entailed in particular the changes listed in (3) (see Heine et al. 2021: 3–4 for more details).

- (2) English (cf. Heine et al. 2021: 3, (1))
  - a Jeff was still sick.
  - b **Still**, Jeff was sick.
- (3) Some effects of cooptation in the shift from adverb (2a) to discourse marker (2b)
  - a Meaning: From modifier or qualifier of the verb to connective device linking different parts of the discourse.
  - b Syntax: From constituent of the sentence to expression detached from the sentence syntax.
  - c Prosody: From prosodically integrated to prosodically typically set-off form.
  - d Indexical status: From non-deictic adverb to deictic form anchored to the situation of discourse.
  - e Semantic-pragmatic scope: From scope within the sentence to scope beyond the sentence.
  - f Discourse processing: The form coopted is directly accessible in discourse without being edited and processed by the machinery of sentence grammar.

In sum, with its cooptation, *still* acquired the hallmarks of interactives as proposed in Section 1.2. The effect in (3f) is of a more general nature, relating to our discussion in Section 2.1.3: *Still* in (2b) can be immediately processed without being

affected by syntactic, semantic, or prosodic constraints imposed by the clause *Jeff was sick* (cf. the direct-access hypothesis, Section 2.1.3.3).

The second example may illustrate a ‘garden-variety’ kind of cooptation as it can be observed commonly in daily interaction, leading from a proper noun of sentence grammar to a vocative expression. The form *Rumpelstiltskin* uttered by B in example (4) is a noun whose function is that of a complement of the copular verb of the sentence. In A’s response, by contrast, the very form exhibits the properties of a coopted unit: Now being a vocative expression, it is no longer a constituent of the sentence, it could be omitted without affecting the meaning content of the sentence. Being syntactically and prosodically unattached, it can form an utterance of its own (e.g., *Rumpelstiltskin!*). In other words, with its cooptation as a vocative, *Rumpelstiltskin* has been transferred to the discourse as a ‘naked’ expression, set off from the surrounding text even if relating functionally to the surrounding text.

(4) English

A: What’s your name?

B: My name is Rumpelstiltskin.

A: And where do you live, **Rumpelstiltskin**?

A common motivation for cooptation leading to interactives can be seen in the fact that certain domains of meaning provide convenient pools for designing new, expressive meanings. With regard to secondary interjections, for example, it is the ones listed in (5) that are favored cross-linguistically to be recruited as source domains for expressing emotive states of the speaker.

(5) The most common semantic source domains for secondary interjections

- a Religious terms, implying (i) avoidance, (ii) fear, (iii) moral values, and/or (iv) authority and respect
- b Terms for sexual or scatological concepts, implying (i) avoidance, (ii) disgust and aversion
- c Kin terms, implying (i) emotions (e.g., affection), (ii) moral values, and/or (iii) authority and respect

The list presented in (5) does not exhaust the range of possible source domains. Nübling and Vogel (2004) found, for example, that diseases also provide a relevant domain for swearwords, observing that in Dutch, terms for diseases are one of the main sources for swearwords—unlike German and Swedish, where diseases do not play any noticeable role. With regard to (5a), there is a distinction between celestial and diabolic expressions, well marked in Christian cultures while in Muslim societies diabolic expressions seem to be rare (Ljung 2011). Speakers of English have drawn mostly on (5a) (*God, hell, Jesus*) and (5b) (*balls, fuck, shit*), while in other languages (5c) is equally common.

Not seldom, speakers also use combinations of source domains, as in Estonian *Kuradi munn!* ('devil's dick'), Portuguese *Putá merda!* and Spanish *Qué puta mierda!* ('Shit whore'), or Hindi *Madar chod!* ('motherfucker') (Ljung 2011: 82). Via cooptation, a text piece consisting of a religious term, a sex-related term, or a kin term may be taken out of sentence grammar to create a form of respect, a swear word, or an expletive used as an utterance of its own.

A related effect of cooptation can be observed in vocatives. Once a noun (phrase) has been coopted to serve as an address form, it is now uniquely identifiable through its function of referring to the hearer in the situation of discourse, and it can be extended to express new meanings beyond its literal one. As we saw in Section 3.10.6.2, when a kin term turns into a vocative, its usage is found in a number of languages to be extended to new referents, where a semantic feature, like 'motherhood,' can give way to new features, such as 'affection,' 'intimacy,' 'tenderness,' etc.

A spectacular example of such a process in the Datooga language of north-central Tanzania was discussed in Section 3.10.6.2, where the noun *úyá* 'mother' has been coopted as a vocative and in this capacity is used to address the following kinds of persons: (a) female kin of ascending generation, (b) female non-intimate peer (equals), (c) intimates of the same generation, (d) younger sister, (e) female children and small children of both sexes, and even (f) one's own child (Mitchell 2017: 289–97). In other languages, coopted nouns for 'mother' had different effects. In the Muna language of Sulawesi in Indonesia, for example, *ina-e* ('mother!') has turned into an interjection expressing extreme pain or fright (van den Berg 1989: 189).

#### 4.1.2 Cooptation and ideophones

The effects of cooptation are quite different with some other types of interactives, such as ideophones. The way the latter can be coopted by transferring text pieces of sentence grammar to the domain of interactive grammar may be illustrated with the following English example. Verbs are a central word class of sentence grammar, having conceptual meaning and being firmly integrated in the morphosyntactic and prosodic structure of a sentence. But they can be taken out of that structure, being coopted for specific discourse functions, as can be illustrated with the English verbs *knock*, *sniff*, and *sweep*:

- (6) English (Clark 2019: 243, (24), (25), (27))
- a I knocked on the door. **Knock! Knock! Knock!** And just as I was hoping, ...
  - b And my nose went '**sniff, sniff, sniff**'
  - c I took my broom and went **sweep, sweep, sweep** up the road.

None of the three ideophones exhibits the morphosyntactic or prosodic features characteristic of verbs. While coopted units are normally syntactically independent inserts, this is seemingly not the case with (6b), where *sniff, sniff, sniff* has the appearance of a clausal complement. However, that this ideophone is introduced in the form of direct speech, where *went* combines the functions of both a predicate and a quotative marker. That *sniff, sniff, sniff* is syntactically independent would seem to be suggested, for example, by the fact that it could well be used as an utterance of its own. A conspicuous feature of the three coopted ideophones is that they are triplicated. This is a characteristic feature of ideophones, which may be duplicated, triplicated, or even quadruplicated to signal functions such as intensity, duration, or repetition (Dingemans 2015; see Sections 3.5.4, 3.12).

English is not an isolated case of a language creating new ideophones via repetition of sentence grammar material; suffice it to mention the Asmat language of Papua New Guinea. In the Flamingo Bay dialect of this language, some ideophones are coopted—or ‘transposed’ in the terminology of Voorhoeve (1965: 163)—by quadruplicating verbs:<sup>2</sup>

- |                                |   |                                                                                              |
|--------------------------------|---|----------------------------------------------------------------------------------------------|
| tep ‘to be above’ (verb)       | > | <b>tepteptepte</b> ‘the action of going lower and lower down, of sinking deeper’ (ideophone) |
| taj ‘to cease, to stop’ (verb) | > | <b>tajtajtajtaj</b> ‘the gradual ceasing of rain’ (ideophone; Voorhoeve 1965: 163)           |

Rather than repetition, speakers in some other languages draw on truncation to coopt a piece of sentence grammar to form ideophones. And more than in English, cooptation of verbs to create ideophones is highly productive in some languages, as can be shown with the East African language Swahili. What speakers of this language appear to have done and are still doing is take the first syllable or the first two or three syllables of a verb stem and coopt that piece as an ideophone for expressing a meaning related to that of the verb but having some kind of an intensifying, expressive function. Table 4.1 provides a few examples showing effects of cooptation. As is common with ideophones, the resulting form may be, and quite commonly is reduplicated.

Once cooptation has taken place, the text piece is no longer part of the sentence, acquiring the grammatical properties of interactives (see (1) of Section 1.2), such as being semantically, syntactically, and prosodically separated. While the semantic content of the Swahili ideophones relates to that of the preceding verb, the ideophones can be omitted without essentially affecting the meaning or the grammatical structure of the sentence—fully in accordance with the modifier construction proposed in Section 3.5.3.2. The usage of the ideophones can be illustrated with the following example. The ideophone *bingiri bingiri* ‘twist and turn all over’ in (7) adds a dramatic effect to the preceding verb form *akabingirika*,

<sup>2</sup> Voorhoeve (1965: 162) calls these ideophones ‘onomatopoeic interjections.’

**Table 4.1** Cooptation of some ideophones from verbs in Swahili (Bantu, Niger-Congo)

| Verbal source        | Meaning                      | Ideophone              | Expressive meaning              |
|----------------------|------------------------------|------------------------|---------------------------------|
| <i>ku-bingirik-a</i> | 'to roll (as a stone)'       | <i>bingiri bingiri</i> | 'twist and turn all over'       |
| <i>ku-bwek-a</i>     | 'to bark, bay'               | <i>bwe bwe</i>         | 'talk nonsense, talk noisily'   |
| <i>ku-chanu-a</i>    | 'to bloom, blossom'          | <i>chanu</i>           | 'blossom to perfection'         |
| <i>ku-futik-a</i>    | 'to insert sth. into sth.'   | <i>futi</i>            | 'keep a secret in one's heart'  |
| <i>ku-lew-a</i>      | 'to be drunk'                | <i>lewa lewa</i>       | 'be totally drunk'              |
| <i>ku-mwag-a</i>     | 'to pour out, spill'         | <i>mwa</i>             | 'pour water suddenly/at once'   |
| <i>ku-nyakur-a</i>   | 'to snatch'                  | <i>nyaku</i>           | 'take by force'                 |
| <i>ku-pasuk-a</i>    | 'to burst, be torn'          | <i>paa</i>             | 'burst of something very tight' |
| <i>ku-pofu-a</i>     | 'to make sb. blind'          | <i>pofu</i>            | 'blind, blinded'                |
| <i>ku-raru-a</i>     | 'to tear apart, rip'         | <i>raru raru</i>       | 'be all torn'                   |
| <i>ku-tanzuk-a</i>   | 'to be clarified'            | <i>tanzu</i>           | 'be completely untangled'       |
| <i>ku-tepete-a</i>   | 'to be weak, slack'          | <i>tepe tepe</i>       | 'be soft, loose'                |
| <i>ku-tuli-a</i>     | 'to be calm, settled, quiet' | <i>tuli tuli</i>       | 'be quiet, static, settled'     |
| <i>ku-twang-a</i>    | 'to pound'                   | <i>twa</i>             | 'hit strongly'                  |
| <i>ku-wash-a</i>     | 'to ignite, kindle, light'   | <i>waa</i>             | 'light at once'                 |

Source: Baldi and Wamitila (2014: 41–57)

Notes: *ku-* = infinitive prefix, *-a* = verbal suffix.

which acts as its T argument. (The utterance contains a second ideophone, namely *fi* 'completely dead,' coopted from the verb *ku-f-a* 'die,' which has the applicative form *ku-fi-a* 'to die at some place'.)

- (7) Swahili (Bantu, Niger-Congo; Baldi and Wamitila 2014: 52)<sup>3</sup>
- yule mkwea milima aliteleza akabingirika **bingiri bingiri**  
 (that climber mountains slipped and.rolled IDE IDE  
 'That mountain climber slipped and rolled, turning all over  
 mpaka chini, akafa **fi**.  
 up.to down and.died IDE)  
 down to the bottom and died.'

As Table 4.1 suggests, the cooptation of Swahili verbs as ideophones has given rise to the whole range of aspectual distinctions proposed in the tradition of Vendler (1967); see Section 3.5.3.1 for details.

There can be little doubt that the ideophones in Table 4.1 go back to the corresponding verbs rather than the other way round, for both synchronic and diachronic reasons. For one thing, the ideophone *lewa lewa* 'be totally drunk' must

<sup>3</sup> The example is a slightly edited version of the original. Glosses are ours, note that the authors do not use glosses for their data.

have been derived from the verb *-lew-a* ‘be drunk’ because the verb is composed of the stem *-lev-* and the passive suffix *-wa*, as is suggested by the fact that the stem also occurs in other combinations, like *-lev-y-a* ‘to make drunk’ or *m-lev-i* ‘drunkard.’ The ideophone *lewa lewa*, by contrast, is unanalyzable and cannot be related to other forms. For another, the Swahili verb *-tuli-a* ‘to be calm, settled, quiet’ is historically derived from the Proto-Bantu stem *\*túúdi-a* ‘to become quiet, to rest,’ or the verb *-twanga* ‘to pound’ from Proto-Bantu *\*-túáng-a* ‘to pound, hit,’ whilst the corresponding ideophones *tuli tuli* and *twaa* cannot, to our knowledge, be traced back to earlier historical states of the Bantu languages (Guthrie 1971).

To conclude, the diachronic process leading to the rise of the Swahili ideophones in Table 4.1 thus can be reconstructed as follows: Speakers clipped the first part of a verb off and coopted it for expressive usage as an interactive form, adding expressive features to the form and placing the piece as an ideophone after the verb phrase—in accordance with the modifier construction of Section 3.5.3.2. That Swahili does not present an isolated case can be illustrated with the following example from the related Mwera language of Kenya. As the examples in Table 4.2 suggest, the pattern of exaptation is similar to that in Swahili.

**Table 4.2** Cooptation of ideophones from verbs in Mwera (Bantu, Niger-Congo)

| Verb             | Meaning    | Verb plus ideophone          | Meaning                                         |
|------------------|------------|------------------------------|-------------------------------------------------|
| <i>ku-jend-a</i> | ‘to walk’  | <i>ku-jend-a jende jende</i> | ‘to walk far’ or ‘to keep on walking’           |
| <i>ku-tol-a</i>  | ‘to carry’ | <i>ku-tol-a tole tole</i>    | ‘to carry like anything’ or ‘to go on carrying’ |

Source: Based on Harries (1950: 125)  
Note: The prefix *ku-* is an infinitive marker.

Similarly, in the Shona language of Zimbabwe, ideophones can be turned into verbs (Section 4.2.2.5), but verbs can also productively turn into ideophones. In the latter case, the bare root is extracted, omitting prefixes and suffixes, but *e* or *eyi* is added (when the root ends in *u*, nothing is added). In this way, for example, the Shona verb *-ték-á* ‘draw water’ gave rise to the ideophone *tékeyí* ‘drawing water leasurately’ (Fortune 1962: 25).

Once again, there is historical evidence to show that in these examples the development is from verb to ideophone rather than the other way round in that the verbs but not the corresponding ideophones can be reconstructed back to Proto-Bantu, the hypothetical ancestor of all Bantu languages, such as Swahili, Mwera, and Shona. Thus, the Mwera verb stems *-jend-a* ‘walk’ and *-tol-a* ‘carry’ can be reconstructed back, respectively, to Proto-Bantu *\*-gènd-a* ‘go’ and *\*-tóód-a* ‘pick up,’ and Shona *-ték-á* ‘draw water’ has its origin in the Proto-Bantu verb *\*-ték-a* ‘draw (water)’ (Guthrie 1971).

Taking a segment of a word and coopting it for purposes of discourse processing seems to be a fairly common strategy. It was presumably also involved in the

development of the Italian interjection *beh!*, signaling ironic disapproval, clipped off from the word *bene* ‘good’ (Ameka and Wilson 2006: 15). As has been established in the relevant research, there are few limits as to the kind of material that may be picked for cooptation (see Heine et al. 2017).

#### 4.1.3 From propositional to expressive meaning

Nübling (2001: 28) suggests that the German interjection *ach* [ʔax] is historically derived from an old word for ‘pain,’ surfacing in English *to ache*, which goes back to Old English *acan* ‘to pain.’ And the German interjection *hurra* ‘hurray’ is said to be derived from the Middle High German verb *hurren* ‘to hurry,’ hence, related to Present-Day English *hurry*.

The German interjection *oje* ([ʔɔʰje:] or [ʔoʰje:]), expressing disappointment, contempt, dismay, or compassion, is historically derived from *Oh (mein) Jesus!* ‘oh (my) Jesus!’, originally expressing scare or despair (Nübling 2001). We will return to this case in Section 4.3.2.

Once coopted, an expression may acquire a meaning that has little in common with its lexical source, as can be shown with the following examples of social formulae:

- The Swahili greeting formula *shikamoo* is a polite form of greeting by an inferior to a superior, or a younger to an older person. It is a contracted form coopted from *nashika miguu yako* (1.SG.PRES.seize NC4.foot NC4-2.SG.POSS) ‘I grab your feet’ (Section 3.9.4).
- The Japanese greeting formula *dochira e* has been coopted from a sentence grammar chunk meaning ‘where to’ or ‘where are you headed/going?’ (Coulmas 1979: 256).
- The social formula *gebnguneng* ‘thank you’ of the Mayali language of Western Arnhem Land, Australia has apparently been coopted from the text piece *o-geb-ngu-neng* (I/you-nose-eat-PAST.PERF), which literally means ‘I ate your nose’ (Evans (1992: 226).<sup>4</sup>

That the meaning of the chunks coopted for deployment as interactives may have little in common with its sentence grammar meaning can also be shown with the following examples: The interjection *gbede* ‘never, under no circumstances’ of the West African language Ewe means literally ‘one day.’ Originally, it appears to have occurred as an adjunct in negative sentences, such as *nyemáva gbedé o* (1.SG.NEG.FUT.come one.day NEG) ‘I shall never (< not a day) come.’ A more common kind of meaning shift can be seen in examples like *jo-wá* (people-our)

<sup>4</sup> But see also Section 1.3.1.1.



'our people' from the Luo language of western Kenya, which was coopted as an interjection, *jo-wá!* 'Boy!' (Tucker 1994: 299).

To conclude, cooptation is a means of creating new expressive meanings by transferring text pieces of sentence grammar to the semantic space of interactive grammar.

#### 4.1.4 A note on pragmaticalization

A number of effects of cooptation have also been described in the rich work on 'pragmaticalization' or 'pragmaticization' and this raises the question of how the two terms relate to one another (e.g., Aijmer 1997; Ocampo 2006; Degand and Evers-Vermeul 2015).

Pragmaticalization has most of all been invoked in work on one specific type of interactive, namely that of discourse markers. Erman and Kotsinas (1993: 79) introduced an argument that text structuring devices at the discourse level such as discourse markers cannot reasonably be accounted for satisfactorily with reference to grammaticalization. Subsequently, it has been suggested (e.g., Aijmer 1997; Dostie 2004; Günthner and Mutz 2004: 98; Claridge and Arnovick 2010: 167) that grammaticalization describes the emergence of sentence-internal grammatical markers while pragmaticalization would be more appropriate to describe the emergence of elements serving the organization of texts (e.g., Norde 2009: 21–4). In some studies, pragmaticalization is described as a shift in meaning from the content level to the context level (Hansen 2008).

Phenomena falling within the scope of pragmaticalization have also been described by means of alternative terms such as reanalysis (Waltereit 2002: 988; 2006), discoursivization (Ocampo 2006: 307), discursization (Claridge and Arnovick 2010: 183), and proceduralization (Bolly and Degand 2013: 210; Noora and Amouzadeh 2015: 94)—signaling that the primary concern of these authors is with discourse at large rather than with sentence grammar. There are a number of definitions of the term, the following being presumably the most succinct one:

The term [pragmaticalization] refers to a process of linguistic change in which a full lexical item [...] or grammatical item [...] changes category and status and becomes a pragmatic item, that is, an item which is not fully integrated into the syntactic structure of the utterance and which has a textual or interpersonal meaning.

(Dostie 2009: 203)

At the same time, it has also been argued that both the term pragmaticalization and what it stands for are dispensable. Detges and Waltereit (2016) do not find it to be a

useful concept since it is said to obscure the differences between discourse markers and modal particles. Some authors treat pragmaticalization as a subclass of grammaticalization (Brinton 2010: 64; Prévost 2011: 391) or as falling ‘naturally under the rubric of grammaticalization’ (Brinton 2017: 35), while other authors reject the term altogether, arguing that it ‘is unnecessary as a separate type of change’ (Traugott 2007: 151).

Students of pragmaticalization have identified a number of structural features shared with the concept of cooptation discussed in this section, like syntactic detachability, distinct prosody, and ‘peculiar grammatical status.’ But they also maintain that pragmaticalization also shares various features with grammaticalization, like the fact that both are claimed to be gradual processes and involve the principles of specialization, divergence, layering, and persistence of Hopper and Traugott (2003) and a high frequency of use (see Heine et al. 2021: 20–4 for more details).

Work on pragmaticalization has drawn our attention to the relevance of pragmatics in language use. Sentence grammar also has a pragmatic component but with its immediate anchoring to the situation of discourse, interactive grammar provides a paradigm example of the pragmatic foundation of linguistic communication.<sup>5</sup>

Furthermore, that work has revealed the limits of approaches restricted to grammaticalization and has enhanced our knowledge of certain features characterizing the development of discourse markers. Nevertheless, it is not an entirely satisfactory notion to provide an explanatory account of the rise and development of discourse markers in particular and interactives in general because it does not clearly distinguish factors that are due to cooptation from those that are due to grammaticalization.

## 4.2 Grammaticalization

Once an interactive has been coopted for use on the level of discourse management, it may be extended to new contexts, being used frequently and undergoing grammaticalization. In fact, most interactives exhibit features of grammaticalization, and these features are of help for understanding why one and the same interactive can exhibit different usages depending on the contexts in which it is used. Some effects of grammaticalization have been discussed in studies based on models of generative theory with reference to the ‘syntacticization of

<sup>5</sup> An anonymous reviewer of this book is therefore wondering: ‘What would be a disadvantage of using “pragmatic grammar” instead of “interactive grammar”?’ (February 3, 2022). The answer is that pragmatics provides a necessary but not a sufficient criterion for defining interactive grammar since sentence grammar also has a pragmatic base.

speech acts' (e.g., [Munaro 2010](#); [Haegeman and Hill 2013](#); [Wiltschko 2021](#): 200), but unlike the latter studies, the present chapter is restricted to a diachronic perspective.

Examples of such effects were discussed in the course of the preceding chapter. A particularly noteworthy effect, to be observed in various types of interactives, was the subject of Section 3.9.5 with reference to social formulae. As we saw there, one and the same interactive form, in this case the English formula *please*, can exhibit two drastically different structures: On the one hand, it is a paradigm instance of an interactive in accordance with the definition in Section 1.2. But on the other hand, it also occurs as a grammaticalized variant, being integrated to quite some extent in the structure of a sentence. In short, grammaticalization theory provides a relevant parameter for explaining much of the variable behavior of interactives, and in Section 4.4 a fairly pervasive effect of grammaticalization is highlighted.

However, while grammaticalization, jointly with cooptation, is almost invariably the main mechanism involved, there may also be additional mechanisms, namely what we refer to loosely as reinforcement in Section 4.3.1 and camouflaging in Section 4.3.2.

### 4.2.1 Parameters of grammaticalization

In order to describe processes of grammaticalization, the parameters in Table 4.3 are used, also serving as diagnostics to identify instances of grammaticalization.

**Table 4.3** The parameters of grammaticalization

| Parameter                    | Kind of change                                                                                                                 |
|------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| a <i>Context extension</i>   | The rise of new meanings when linguistic expressions are extended to new contexts, leading to context-induced reinterpretation |
| b <i>Desemanticization</i>   | Loss or generalization of meaning content or functions in such contexts                                                        |
| c <i>Decategorialization</i> | Loss of morphosyntactic properties characteristic of lexical or other less grammaticalized forms                               |
| d <i>Erosion</i>             | Loss of phonetic substance, including prosodic features                                                                        |

Source: [Heine and Kuteva \(2007: 33–46\)](#)

With the exception of context extension, all parameters of Table 4.3 involve losses of grammatical substance. But in the same way as forms undergoing grammaticalization lose substance they gain new functions arising in the contexts in

which they are used. We may illustrate the mechanism with the following example of a social formula (for more examples, see Section 9.3.4). Upon hearing of the death of someone close, speakers of the Swahili language of East Africa are likely to say *Makiwa!* (cf. (8a)). This is a salutation to mourners, a social formula treated by Eastman (1992: 277) as an ‘interjection,’ which goes back historically to the subordinate clause in (8b).

(8) Swahili (Bantu, Niger-Congo; Eastman 1992: 277, (3))

a **Makiwa!**

‘My heartfelt sympathy!’

b m-ambo ya-ki-wa (ha-yo)

NC6-matter NC6-if-be DM-NC6.REF

‘things being so’ (lit. ‘if matters are (those)’)

The process leading to the rise of the interactive *Makiwa!* can be hypothesized to have been as follows. First, it involved cooptation, in that there was the conditional subordinate clause in (8b) that was coopted and transferred from the domain of sentence grammar to that of discourse processing for a specific discourse function, namely that of a formula in mourning rituals. This had the effect that the segment turned into a syntactically, semantically, and prosodically independent interactive.

Subsequently, there was grammaticalization in accordance with the parameters in Table 4.3: The segment acquired a specific function in the context of death reports (context extension), it lost its lexical meaning content in favor of the new function (desemanticization), and the three words lost their morphological status as a noun (*m-ambo*), a verb (*ya-ki-wa*), and a demonstrative (*ha-yo*) and merged into one invariable form (decategorialization). And finally, the segment lost phonetic substance (*m-ambo* > *ma-*) and morphological substance (*ya-* > zero, *ha-y-o* > zero), being reduced from five (or seven) to three syllables (erosion) in *Makiwa!*.

Grammaticalization of interactives is almost invariably the result of context extension and increased frequency of use, the emergence of new discourse function, and the gradual loss of erstwhile descriptive meanings. For example, English expletives like *motherfucker* and *cunt* were coopted for functions of social interaction. Used as ‘rude vocatives,’ they were placed in new specific contexts, and with their frequent use in these contexts they acquired functions that had little in common with their earlier meanings: Via desemanticization they lost their highly insulting content and gave way to almost neutral intensifying functions in those new contexts (Stenström 2020: 300).

Such examples are legion in the literature on interactives, and the following Section 4.2.2 will be concerned with a few more general lines of development to be observed in interactives.

## 4.2.2 Main pathways

The most general lines of development to be discussed in this section are listed in (9), and all relate to the valence structure of the interactives concerned (see Section 2.2):

- (9) Main directions of grammaticalization in interactives
  - a From speaker–hearer interaction to text organization: T gain  
(Section 4.2.2.1)
  - b From attitudes of the speaker to text organization: T gain  
(Section 4.2.2.2)
  - c From speaker–hearer interaction to attitudes of the speaker: H loss  
(Section 4.2.2.3)

The effects on valence that these pathways have are of a contrasting nature: Whereas (9a) and (9b) lead to the rise of a T (theme) argument, (9c) leads to loss of an H (hearer) argument. These pathways are captured in Table 4.4, which summarizes a number of processes of grammaticalization, namely from Stage I to Stage II and from Stage II to Stage III in both pathways. In its development, a given interactive may go through all three stages, but this need not be the case; rather, the development can be restricted to only two stages, leading only from I to II, or from II to III, or else straight from I to III.

**Table 4.4** Hypothesis on the development of some interactives towards text anchoring

| Pathway |                            | Stage I               |   | Stage II         |   | Stage III                          |
|---------|----------------------------|-----------------------|---|------------------|---|------------------------------------|
| (a)     | <i>Type of interactive</i> | Directive or vocative | > | Attention signal | > | Discourse marker                   |
|         | <i>Argument structure</i>  | S,H                   |   | S,H,T            |   | T <sub>1</sub> ,S,H,T <sub>2</sub> |
| (b)     | <i>Type of interactive</i> | Vocative              | > | Interjection     | > | Discourse marker                   |
|         | <i>Argument structure</i>  | S,H                   |   | S                |   | T <sub>1</sub> ,S,T <sub>2</sub>   |

### 4.2.2.1 From speaker–hearer interaction to text organization: T gain

This section is concerned with pathway (a) of Table 4.4. The pathway is about changes where a given interactive does not require any textual cues for its interpretation, and more specifically, where the expression has no T (theme) argument (see Section 2.2). Paradigm examples are provided by (i) directives, (ii) vocatives used as stand-alone utterances, and (iii) interjections used as negation markers. In the course of their grammaticalization, the interactives become increasingly integrated in the text and their function gradually turns from regulating the relationship between the two participants of discourse to highlighting specific information content expressed by some text piece, and that text piece may then develop into an obligatory T argument.

4.2.2.1.1 *From directive to discourse marker*

This process has been discussed in Heine et al. (2021, Section 8.4), where directives concern imperative forms:

It is especially imperatives of perception verbs for ‘look’, ‘see’, and ‘listen’ that tend to be recruited for discourse structuring purposes, such as calling for the addressee’s attention, emphasizing thematic progression and/or enabling or facilitating turn-taking.

(Heine et al. 2021: 256)

But the process is by no means restricted to perception verbs. Take the imperative form *tiens!* of the French verb *tenir* ‘to hold,’ which was presumably coopted as a directive but appears to also have developed usages as an attention signal (‘hey!’; ‘look!’) and of a discourse marker (‘well,’ ‘indeed’). While the source of *tiens!* is a second person singular imperative form, as an interactive it can be used for both singular and plural referents (cf. Wilkins 1992: 130–1; Ameka and Wilkins 2006: 15). Another fairly common source is provided by verbs of deictic motion, most of all verbs meaning ‘come’ and ‘go,’ as can be illustrated with the following example.

The German form *komm* ‘come!’ is the singular imperative form of the verb *kommen* ‘to come,’ acting as directive in (10a). It is a deictic verb, expressing movement towards the speaker, and it implies an S and an H argument. In (10b), by contrast, it is no longer clearly a deictic verb: (10b) could be uttered in a context where the hearer is next to the speaker and no movement is implied (context extension). In this context, the lexical meaning of *komm* is lost (desemanticization), and its function now is no longer that of a directive; rather, it is used to prompt the hearer to act in a spirit of mutual understanding and solidarity with the hearer (‘I am with you’)—a function that is described by Kuteva et al. (2019: 92) as a ‘solidarity imperative.’ The text piece following *komm* is an obligatory argument (T) of the interactive. Thus, in this usage, *komm* in (10b) has the argument structure (S,H,T) and is appropriately described as an attention signal.

## (10) German

- a **Komm!**  
‘Come!’
- b **Komm,** denk darüber nach!  
come think about.it after  
‘Come on, think about it!’
- c A: Ich bin schwanger.  
I am pregnant  
‘I’m pregnant.’  
B: **Komm,** das ist nicht wahr!  
come that is not true  
‘Come on, that’s not true!’

The example in (10c) is of a different kind. Rather than one T argument, *komm* now has two, namely a preceding one ( $T_1$ ), instantiated by A's turn, and a following one ( $T_2$ ), and the function of the interactive is to express a mild contradiction between  $T_1$  and  $T_2$ , that is, of establishing a discourse connection between the two text pieces: With *komm*, speaker B indicates that she is somehow shocked about what A just said and claims or pretends not to believe it. In this usage, thus, *komm* has the argument structure of a discourse marker ( $T_1, S, H, T_2$ ), and in fact, it can be classified as a discourse marker. For example, we saw in Section 3.3.2 that discourse markers are defined by Fraser (1999: 938) as elements that relate two pieces of discourse to one another, and exactly this is what *komm* does in (10c).

While we are not aware of appropriate diachronic data on the development of *komm*, the evidence that there is suggests that the history of this interactive can be hypothesized to have proceeded along the general pathways hypothesized in Table 4.4. This development can be characterized as one of 'textualization' of directives and vocatives, that is, as one leading from relative independence to increasing integration in the structure of the surrounding text.

A survey of a range of other interactives suggests, in fact, that a development of the kind sketched for German *komm* is not an isolated case. Once a directive or an interjection is taken out of its normal usage within the situation of discourse and placed within a text it is likely to grammaticalize in the direction of a discourse marker, assuming functions such as signaling connections between different pieces of discourse.

Another example is provided by the Italian discourse marker *guarda*. It is derived from the homophonous second person singular imperative *guarda!* 'look!' and can be said to be invariable, even if there are two variants in addition, namely the second person singular polite form *guardi* and the second person plural form *guardate*<sup>6</sup> (Waltereit 2002: 984; see also Heine et al. 2021: 256–60). The motivation for extending the use of the Italian directive ('imperative') *guarda!* 'look!' to serve discourse structuring functions is summarized by Waltereit thus:

The imperative 'look!' may entitle the speaker to an interruption, because it conversationally implicates that she has to say something extremely important that requires the immediate attention of the conversation partners. It is argued that speakers will tend to use this imperative even in situations where they do not have something important for the others to look at. They will do so simply because this imperative is so useful for floor-begging. At this point, *guarda!* is no longer an imperative, but a discourse marker.

(Waltereit 2002: 987)

Context extension associated with floor seeking in conversational exchanges was a crucial factor in grammaticalization, allowing the directive *guarda!* 'look!'

<sup>6</sup> Franck Floricic (p.c. of August 5, 2021) observes that there is also *gua'* in Italian, that is, a reduced variant of *guarda*.

to be used even in contexts where its lexical meaning no longer made sense (desemanticization). And it continued to shape the rise of the discourse marker by being used more frequently even in contexts where it provided the speaker with ‘an excellent justification for self-selection at turn-taking, even at non-transition-relevance places. The result of this strategical overuse was the reanalysis of the imperative as a discourse marker with the function “Listen to me, I have something important to say!”’ (Waltereit 2002: 988).

The examples provided by Waltereit (2002) suggest that the development of *guarda* did not proceed straight from directive to discourse marker but involved an intermediate stage where *guarda* was not yet a discourse marker but rather was used mainly as an attention signal, drawing attention to what follows rather than establishing a logical relation between two pieces of discourse—a function captured by Waltereit (2002: 1001) thus: ‘The topic of the ensuing discourse is highlighted as particularly important and interesting.’

Instead of the full development from directive to discourse marker, grammaticalization can be restricted to either the first part (directive > attention signal) or the second part of the development (attention signal > discourse marker).

#### 4.2.2.1.2 *From vocative to attention signal to discourse marker*

This line of grammaticalization also concerns pathway (a) of Table 4.4. Vocatives may be used as stand-alones, forming utterances of their own, e.g., *Sir!*. But quite commonly they are combined with other text pieces: ‘For example, an initial vocative can serve as an attention-getter and can also have the function of clearing space for a lengthy turn’ (Biber et al. 1999: 1113). In this way they may in fact assume the function of an attention-getter, that is, what is called here an attention signal.

For example, the form *bwana* ‘man’ or ‘husband’ of the Swahili language of East Africa is commonly used as a noun of sentence grammar, and it has also been coopted as a vocative expression meaning ‘Sir!’. But in addition it also shows uses as an attention signal (‘hey!’) and as a discourse marker, in which case it has lost its association with gender and age (desemanticization) in that it can also refer to addressees other than male or adult ones.

A special kind of socio-cultural factor appears to have contributed to the development of the form *marry* in Early Modern English. Arising as what appears to be a vocative use in the context of religion it is first attested in the second half of the fourteenth century, going back to the name of the Virgin Mary used as an oath or an invocation. In the sixteenth century *marry* freed itself from the religious context, becoming a mere interjection (desemanticization), in accordance with (b) of Table 4.4, but also acquiring functions of a discourse marker in certain contexts and no longer being perceived as a religious invocation. By the nineteenth century its use became obsolete (Jucker 2002: 227).

The Spanish noun *hombre* ‘man, person’ is used as a vocative, but has also been grammaticalized to an interjection of surprise and other functions. And it appears to have also been extended to new contexts where it acquired features



of a discourse marker: It is employed to strengthen the connection between the preceding and the following discourse as well as for speaker–hearer interaction, expressing the attitude of the speaker to the utterance, also serving as a marker of reformulation. Used in final position, *hombre* reinforces the content of the preceding utterance (Kleinknecht 2019: 299–300).

The following are a few examples involving nouns for ‘God’ as a source of grammaticalization (see Kuteva et al. 2019: 219). In a number of European languages, forms having ‘God’ as a nucleus were coopted as vocatives in invocations and developed into interjections, cf. English *God Almighty!*, *Oh God!*, German *Mein Gott!* ‘My God!’, or French *Grand Dieu!* ‘Great God!’. In Russian, the old vocative forms got lost except for forms such as *bože* ‘God!’ or *gospodi* ‘Lord!’, but these interactives are nowadays used as parts of formulae or as interjections (Sonnenhauser and Noel Aziz Hanna 2013: 12). And in the Ik language of north-eastern Uganda, the noun *nyakújí* ‘God’ has been coopted as a vocative and grammaticalized to the interjection *nyakújí* ‘God!’, expressing strong emotion.

An interesting development has been observed in the Chadic language Lele. The noun *álà* ‘God,’ borrowed from Arabic *Allah* ‘God,’ came to be coopted as a vocative and interjection expressing surprise. But it has been further grammaticalized to a discourse marker of unexpected propositions, and in certain contexts into an adversative conjunction (‘but’). Such contexts involve situations where in a sequence of two clauses the content of the second clause is interpreted as the unexpected consequence or outcome of the first clause, cf. (11).

- (11) Lele (Chadic, Afroasiatic; Frajzyngier 1996: 82; Heine and Kuteva 2007: 247)

ŋ-jè            tòb   kámyà   **álà**   né-ŋ            hómyà.  
 1.SG-HAB like milk   but make-1.SG sick  
 ‘I like milk but it makes me sick.’

While the preceding examples relate to the domain of religion, the following example concerns the domain of sexual concepts. The Spanish insulting vocatives *boludo/a* of Argentina, and *huevo/n/a* of Chile are both derived from terms for ‘testicle’ (*boludo/a* derives from the plural form *bolas* ‘testicles’ plus suffix *-udo/a* and *huevo/n/a* from the singular form *hueva* ‘testicle’ plus suffix *-ón*). The insulting effect vanishes when they are uttered in the context of friendly chats of teenagers in Buenos Aires and Santiago de Chile, respectively. In this context, *huevo/n* of Chilean Spanish, in particular, is used highly frequently to the extent ‘that it resembles a conversational “filler”, helping the speaker to get the conversation going. In other words, its role as a vocative is fading and it is developing into a discourse marker’ (Stenström 2020: 298).

Side by side with words from the domain of religion and sex there are also words for ‘man’ that have provided instances of cooptation from sentence grammar to interactive grammar, subsequently grammaticalizing from vocatives to

interjections and sometimes further to discourse markers (Kuteva et al. 2019: 278–9). The German noun *Mann* ‘man, husband’ can be used in colloquial German as a vocative *Mann!*, but also as an utterance-initial interjection expressing surprise, or as an attention signal drawing attention to what follows. And a similar development has been observed by Kleinknecht and Souza (2017) in colloquial varieties of German from (noun and subsequently to) vocative *Alter* ‘old man’ to discourse marker (Table 4.5).

Another example from Peninsular Spanish concerns terms for relatives, namely the nouns *tío* ‘uncle’ and *tía* ‘aunt.’ Already in the sixteenth century, the two nouns appear to have been used as address forms for ‘old man’ and ‘old woman’ respectively, and by the end of the nineteenth century they also denoted ‘man’ and ‘woman,’ respectively. Around 1970, the nouns *tío* ‘uncle’ and *tía* ‘aunt’ gave rise to discourse markers, often found in utterance-final position ‘with the interactional functions of leaving the floor and showing solidarity’ (Pons Bordería and Llopis Cardona 2020: 335).

A number of developments from vocative via attention signal to discourse marker concern male domestic animals, all reported from Spanish (Kleinknecht 2013; Kleinknecht and Souza 2017; Pons Bordería and Llopis Cardona 2020: 329–31). Originally a noun for ‘male animal,’ *macho* of Peninsular Spanish is documented with the meaning ‘male mule’ in the sixteenth century. At the end of the nineteenth century, *macho* acquired the meaning ‘man with a sexual urge’ and then ‘any man’ (virile or not, with or without a sexual urge). Cooptation must have taken place around 1950 when *macho* was used as a vocative with the new meaning ‘mate,’ conveying solidarity by showing proximity to the addressee. The form was used as an attention signal in utterance-final or medial position.

The following case of the Mexican Spanish form *buey* ‘ox’ also shows how a negatively valued vocative form acquires a positive social significance and gradually develops into a discourse marker (Kleinknecht 2013; Kleinknecht and Souza 2017). In Mexico of the 1970s, *buey* gave rise to the vocative form *güey* ([gwej]) or *wey*, with the transfer from presumed ox-like qualities such as stupidity, brutality, or ignorance to humans, developing from imprecation to vocative form of address among adolescents. From the 1990s on, *güey* tended to lose its pejorative connotation, and its function is described as one of a solidarity marker, identifying one as belonging to the same group. In the last third of the twentieth century, it developed features of a discourse marker covering a range of functions like attention getting, intensifying, hedging, and floor taking.

The examples in (12) illustrate the various usages of *güey* or *wey*, where (12a) shows the vocative use and (12b) that of an interjection, expressing the speaker’s emotions, while (12c) appears to be suggestive of an attention signal and (12d) of a discourse marker. Note that the different usages are not clearly separated from one another but overlap to quite some extent. For example, the usage of *güey* in (12c) can in the same way be classified as that of a discourse marker, as Kleinknecht

(2013: 253) in fact does, in that it not only relates to what follows but also reinforces what has just been said.

- (12) Mexican Spanish (Kleinknecht 2013: 246–53)
- A **Cómo** estás **güey**?  
how be.PRES.2.SG VOC  
‘How are you, *güey*?’ (9)
- B Ay **güey**, ya me quemé!  
INT INT already REFL.1.SG burn.PAST.PFV.1.SG  
‘Ouch *güey*, now I got burnt!’ (23)
- c **Wey**, no quiero ensayar con Andrea.  
DM not want.PRES.1.SG practice.INF with Andrea.F.SG  
‘*Güey*, I don’t want to practice with Andrea.’ (24)
- d Le dije **güey** que cuando  
IO.F.3.SG say.PAST.PFV.1.SG DM that when  
llegara que ya estuviera  
arrive.PAST.SUBJ.1.SG that already be.PAST.SUBJ.3.SG  
arreglada.  
dress.up.PTCP.F.SG  
‘I told her *güey* to be ready when I arrived.’ (25)

The examples in Table 4.5 show that once grammaticalized, interactives may also occur in phonetically reduced forms, as Mexican Spanish *wey* and Colloquial German *alla* do: Discourse markers are likely to be used frequently and as a result may undergo phonetic reduction (cf. the grammaticalization parameter of erosion in Table 4.3).

Table 4.5 From vocative to discourse marker

| Language           | Vocative                                | Attention signal    | Discourse marker   |
|--------------------|-----------------------------------------|---------------------|--------------------|
| Peninsular Spanish | <i>macho</i> ‘man (with a sexual urge)’ | <i>macho</i> ‘mate’ |                    |
| Peninsular Spanish | <i>hombre</i> ‘man’                     |                     | <i>hombre</i>      |
| Peninsular Spanish | <i>tío/tía</i> ‘uncle/aunt’             |                     | <i>tío/tía</i>     |
| Mexican Spanish    | <i>buey</i> ‘ox’                        | <i>güey, wey</i>    | <i>güey, wey</i>   |
| Colloquial German  | <i>Alter</i> ‘old man’                  | <i>alter</i> ‘mate’ | <i>alter, alla</i> |

Sources: Kleinknecht (2013); Kleinknecht and Souza (2017); Pons Bordería and Llopis Cardona (2020: 329–33)

4.2.2.1.3 From interjection to negation marker

Finally, attention is drawn to the fact that not uncommonly interjections in specific contexts assume the function of response signals expressing denial or

disagreement ('no'). This process is captured nicely by [Ljung \(2011\)](#) in the following way:

Many languages have developed mechanisms for expressing emphatic utterance denial by swearing that are close to the oaths. There are several such mechanisms, but the underlying idea is one and the same, viz. to deny the truth of claims put forth in an utterance by arguing that these claims have no more value than some negative and/or worthless entity. The entity in question is often of a scatological nature, as in the following English examples: *Shit it is!*, *My arse (ass) it is!*, *My (left) butt they are!* in which the scatological terms may be replaced by the euphemistic *My foot!*.

([Ljung 2011](#): 105)

This grammaticalization can be interpreted as another instance of T gain: Response signals typically express reactions to what the previous speaker has said, that is, the preceding text piece may assume the function of a T argument of the response signal (Section 3.8).

An example of this kind was discussed in Section 3.6.4.3: In an exchange between speakers of the South African Xhosa language, speaker B refuses an invitation by speaker A to smoke weed, and as a response signal he uses the laughter interjection *ha ha ha*, which in this context is equivalent to the negative response signal *hayi* 'no' ([Andrason 2021b](#): 8).

Another example of context extension of an interjection or invective assuming the function of an emphatic negation marker is reported by Franck Floricic (p.c. of July 22, 2021) from Italian: In an interview with the Italian politician Alessandra Mussolini on May 11, 2017, when asked why she did not change her family name considering its historical associations with Benito Mussolini, the late leader of Italy's National Fascist Party, she answered: *Col cazzo!* ('with the dick') 'Absolutely no!' (see also [Floricic and Milioni 2021](#)).

As shown by [Ljung \(2011: 105–6\)](#), terms for emphatic denial are taken most of all from the domain of scatological, sexual, or religious taboo expressions. In English they frequently involve a sentence grammar construction where the taboo term is placed clause-initially as a kind of complement of the following predication, as in *the hell it is*, *the heck she did*, *the devil they are*; an example from late-sixteenth-century English is the following: *The Devil she has!* (d1f.txt; [Ljung 2011: 105, \(17\)](#)).

In most cases, however, terms for emphatic denial involve interactives rather than pieces of sentence grammar. [Ljung \(2011: 106–7\)](#) provides examples from a range of European languages, such as Finnish *Hevon vittu!* 'horse cunt' or French *Mes fesses* 'my arse/buttocks.'

Another kind of example was provided in Section 3.8.3, based on [Norrick \(2009: 880\)](#): English secondary interjections such as *hell*, *fuck*, and *shit* appear

to have been weakly grammaticalized to intensifiers in the context preceding the bare response signals *yeah* and *no* in initial turn position. The following example, repeated from Section 3.8.3, illustrates this usage, where the secondary interjection *shit* apparently serves to intensify the meaning of the following response signal *no*:

- (13) English (WSC DPC238; Norrick 2009: 880)  
 DN> any, any **er** improvement on seeing client base.  
 CE > °**shit no**°.  
 DN> **no**?=  
 CE > =he's all up and down.

The way grammaticalization may proceed from interjection to negation marker can be reconstructed by means of the schema presented in a simplified format in (14). Note that there may be two alternative developments, leading either to Stage III<sub>1</sub> or Stage III<sub>2</sub>; we are restricted here to the former. In sentence grammar, this schema has been found to quite commonly apply to the rise of new demonstrative forms and, even more commonly, to negation markers (Heine et al. 2020). The present example concerns a pathway of grammaticalization hypothesized for the English form *fuck*.

- (14) A three-stage model of grammatical change via reinforcement  
 Stage I: An existing form A is used in its original state  
 Stage II: A is used to reinforce another form B  
 Stage III<sub>1</sub>: A replaces B  
 Stage III<sub>2</sub>: A and B coalesce, and AB replaces the earlier B

Stage I concerns the usage of *fuck* as an interjection, as in (15). (16) illustrates Stage II, where *fuck* (A) appears to express emotionally charged emphasis on the negative content of an adjacent text piece (B), another example was provided by the interjection *shit* in (13). In (17) finally, *fuck* (A) can be said to be used instead of a negation marker. This is a case of Stage III<sub>1</sub>, where *fuck* (A) now has turned into an emphatic negation marker.

- (15) English (Mackenzie 2019, (7i))  
 (Oh) **fuck!**
- (16) English (Mackenzie 2019, (32), (82))  
 a They didn't do **fuck**. (They didn't do nothing/They didn't do anything)  
 b He said **fucking** nothing all night.<sup>7</sup>

<sup>7</sup> Lachlan Mackenzie (p.c. of February 17, 2022) points out that (16a) and (16b) belong to different categories.

- (17) English (Mackenzie 2019, (30), (31))
- a **Fuck** else to do. (Nothing else to do)
  - b I found (sweet) **fuck** all. (I found nothing)
  - c **Fuck** if I know. (= 'I don't know')
  - d Have they **fuck**! (= 'They haven't')

The present case differs from the other cases discussed in Section 4.2.2 in that it transgresses the boundary between interactive grammar to sentence grammar:<sup>8</sup> In Stages II and III, the form *fuck* is no longer an interactive but rather has features of a participant of sentence grammar, like that of a complement in *Have they fuck!* (17d); see also Section 4.2.5.

#### 4.2.2.2 From attitudes of the speaker to text organization: T gain

This process concerns the fairly common development of interjections into discourse markers, that is, the process from Stage II to Stage III of pathway (b) in Table 4.4. What is especially interesting about this development is that it quite commonly involves *primary* interjections (see Section 3.6.4.1). In some form or other, this pathway of grammaticalization has been hinted at by a number of authors, such as Norrick (2009):

The more frequent primary interjections have generally taken on specific functions as pragmatic markers, for instance *oh* and, to a lesser degree, *ooh* and *ah*, which have assumed functions as discourse markers signaling a change in cognitive state.

(Norrick 2009: 868)

Examples involving primary interjections in other languages are not hard to come by, such as *eh* of the Maya language Q'eqchi' of Guatemala and Belize, which is often used in discourse marker function as a floor holder (Kockelman 2003: 472).

One general example can be seen in the distinction made by Ljung (2011: 83–93) between reactive interjections and pragmatic interjections. The former indicate the speaker's involuntary reaction to stimuli, as in exclamations of surprise, annoyance, or pain. Pragmatic interjections fulfill text-related functions suggestive of discourse markers, frequently serving as 'slot fillers' before a clause. It can be argued that there is a general development from reactive interjections, being paradigm instances of what is described in Section 3.6 as interjections, to pragmatic interjections, which have more in common with discourse markers than with interjections (see Section 3.6.4.2).

The process can be illustrated with the German primary interjection *oh*, expressing amazement or surprise. (18a) presents a situation where *oh* is used by

<sup>8</sup> A similar example appears to be provided by Italian *un cazzo!* ('a dick'), see Floricic and Milioni (2021).

the speaker to express surprise (even if that surprise may be feigned). In contexts like (18b), *oh* functions as an attention signal, drawing attention to the following message, and the function of surprise or amazement is backgrounded. Finally, (18c) presents a context where *oh* connects two text pieces involving different turns and speakers, and the function of *oh* can be understood to be that of a discourse connective signaling a mild contradiction between the contents of the two text pieces.

(18) German

- a [The speaker receives a large bouquet of flowers]

**Oh!**

‘Oh (I’m touched)!’

- b **Oh**, vergiss nicht, mich anzurufen!  
 oh forget not me to.give.a.ing  
 ‘Oh, don’t forget to call me!’

- c A: Das ist ja ein nettes Paar.  
 that is MP a nice couple  
 ‘That’s a nice couple.’

B: **Oh**, sie sind Schwester und Bruder!  
 oh they are sister and brother  
 ‘Oh, they are sister and brother!’

Each of the three contexts illustrated in (18) is associated with a different function of German *oh*: In (18a) it functions as an interjection expressing surprise, having only one argument (S). In contexts like (18b) it has the appearance of an attention signal taking the following text piece as a T argument, hence (S, T). In (18c), by contrast, its function is that of a discourse marker, relating two text pieces,  $T_1$  and  $T_2$ , to one another and expressing a mild adversative contrast between the two, signaling for example that the preceding text piece ( $T_1$ ) occasions a shift in B’s attention (cf. [Heritage 1998](#): 326 on English *oh*). Its argument structure thus is ( $T_1$ , S,  $T_2$ ).<sup>9</sup>

On the basis of principles of grammaticalization, the three usages of German *oh* can be hypothesized to reflect different stages of a development leading from interjection to discourse marker, with the intermediate stage of an attention signal, thus combining both pathways (a) and (b) of Table 4.4.

While we do not dispose of appropriate historical data to test this reconstruction in detail, this is different in the case of the corresponding English interjection *oh*. This form is commonly classified as a primary interjection (Section 3.6.4.1),

<sup>9</sup> A survey carried out by Tania Kuteva with her students at the University of Düsseldorf in July, 2021 suggests that all three usages of the German interactive *oh* distinguished in (13) can be observed on the medium of Twitter, expressing functions such as contrast, contradiction, surprise, amazement, disappointment, and mitigation.

portrayed as ‘basically a response cry and backward looking’ (cf. Goffman 1978). According to Biber et al. (1999: 1083) it is the most frequent English interjection (see also Heritage 1998, 2002). At the same time, it has also been described as being one of the most frequently used discourse markers of the English language, where one of its main functions is that of ‘a marker of a change of state of information’ (Jucker 2002: 217).

*Oh* (or *o*) goes back to Early Middle English but did not exist in Old English, occurring in front of a noun or name in the vocative or independently expressing a variety of emotions, such as appeal, entreaty, surprise, pain, lament, and so on (Jucker 2002: 217). In the history of *oh* from Early Modern English to Present-Day English ‘a change of focus can be observed from the exclamatory function to a function as a marker of information management,’ and Jucker (2002: 218) concludes: ‘To this extent, the Early Modern English *oh* is an interjection rather than a discourse marker, while Present Day English *oh* is a proper discourse marker.’

This interpretation is presumably not entirely correct considering that English *oh*, like German *oh*, still occurs as an interjection, as the following example shows, where it seems to express surprise and can occur as a stand-alone interjection, as demonstrated in the last line.

- (19) English (Biber et al. 1999: 21)  
 [Watching television]  
 A: Who wins?  
 B: **Pardon?**  
 A: Who wins?  
 B: I don’t know  
 A: **Oh**, you don’t know?  
 B: **No**, I wouldn’t watch it if I knew that  
 A: **Oh**. (CONV)

What is obvious is that, like German *oh*, English *oh* gradually changed from a one-argument (S) interjection to a three-argument ( $T_1$ , S,  $T_2$ ) discourse marker, in accordance with (b) of Table 4.4. And these stages of development co-exist side by side as context-dependent variants in present-day English.

**4.2.2.3 From speaker–hearer interaction to attitudes of the speaker: H loss**  
 Vocatives and directives are contrasting types of interactives, yet they share a number of features, most of all the following: Both have the argument structure (S,H), that is, they require a speaker (S) as their source and a hearer (H) as their target. In addition, there is another feature that may evolve in highly specific contexts: Their H argument may be backgrounded or disappear altogether in some uses, with the effect that an interjection-like function emerges.

The pathway looked at now is included in the grammaticalizations hypothesized in Table 4.4 but relates to a different issue. As was shown in (b) of Table 4.4,



the grammaticalization from vocatives to interjections leads from contexts involving both a speaker and a hearer to contexts no longer involving a hearer. The process can be described as one of H loss, that is, suppression of an H (hearer) argument: When vocatives change into interjections, expressing the emotive or cognitive state of the speaker rather than addressing a hearer, the role of the latter is suppressed, and the interactive concerned can now be used for example in soliloquy. [Ljung \(2011: 82\)](#) observes that this pathway is one leading from speaker–hearer interaction to attitudes of the speaker:

Thus for instance, several of the examples above contain compound words whose literal meaning is ‘son of a bitch’ or ‘son of a whore,’ expressions that are first and foremost associated with name-calling and insulting. In this their original capacity, the words in question are always used antagonistically and involve a verbal attack on an addressee. However, once they have been attracted to the category of expletive interjections they become vehicles for the speaker’s own feelings of irritation and/or surprise, as in American English *Son of a bitch!*, Hindi *Madarchod!*, Turkish *Pich uglo pich!* and Russian *Sukinsyn!*

([Ljung 2011: 82](#))

The following examples are suggestive of a grammaticalization from vocatives used in a religious speaker–hearer context to interjections no longer involving an H argument: from address form *My God* to interjection *My God!*, and from invocation *By God!* to interjection *By God!* ([Ljung \(2011: 20\)](#)). A similar example is provided by the development of the English form *Jesus* (see [Gehweiler 2008](#)). Occurring as a noun of sentence grammar in texts as early as 1125, it typically occurred sentence-initially as an invocation, a vocative directed to Jesus Christ and apparently restricted to religious contexts since 1225, if not earlier. The following example stems from the fifteenth century, suggesting that it had a two-place argument structure (S,H).

- (20) Middle English (*MED* 1470; [Gehweiler 2008: 82, \(6c\)](#))

**Jesu** defende us from dethe!

‘Jesus, defend us from death!’

But already by the fourteenth century, *Jesus* could be used in non-religious contexts without directly referring to the addressee, that is, with a backgrounded hearer (H) argument. This usage is illustrated in (21).

- (21) Middle English (*OED* 1377; [Gehweiler 2008: 83, \(7b\)](#))

**Bi Jesus**, with here ieweles, howre iustices she shendeth.

‘By Jesus, with her jewels, she disgraces our justice.’

In addition to religious concepts, another fairly widespread pathway involves concepts of kinship terminology, and most of all terms for ‘mother,’ instantiating the

pathway leading from vocative to interjection and, hence, to loss of an H argument. Italian *Mamma mia!* ('My mom!') is a case in point. Originally presumably a vocative for calling or addressing one's mother, it has turned into an interjection for a range of emotional concepts such as surprise, joy, fear, pain, and exasperation. Similar examples can be found in some other languages. In Bulgarian, for example, the noun *majčitsa*, diminutive form of the noun *majka* 'mother,' has the vocative form *majčitse* which has given rise to an interjection expressing fear (Tania Kuteva, p.c. of December 4, 2021).

The process leading to H loss is not restricted to vocatives; rather, it is a more general pathway which may also implicate a number of other types of interactives, such as directives and social formulae. The following example may illustrate the latter process. In Dutch, the social formula *Goeie morgen!* 'Good morning!' has the paradigm argument structure (S,H) of such formulae, used by the speaker (S) to address a hearer (H) (Section 3.9). However, this greeting has been extended to contexts where it expresses surprise, like interjections. In such contexts the H (hearer) argument is suppressed. Furthermore, in such contexts its association with a specific time period of the day has also been lost (desemanticization), it can be used any time of the day, and it receives 'an extreme intonation contour' (Hengeveld and Mackenzie 2008: 78).

#### 4.2.3 A network of grammaticalization

The observations made in the course of Section 4.2.2 may be of help to understand a conspicuous feature of interactives: Vocatives, interjections, discourse markers, etc. are not neatly separated from one another but are interrelated in a number of ways. On the one hand, interjections can take on the form of isolated 'response cries' as portrayed by Goffman (1978), on the other hand, they may also be deeply entrenched in the surrounding text. Accordingly, Kockelman (2003: 472) notes with reference to interjections like *ay dios* in the Mayan language Q'eqchi' of Guatemala and Belize:

Interjections are not uncommonly index signs in the *discursive* context in that they co-occur with (or serve as) a *response* to an addressee's previous utterance or a *nonresponse*. In the case of a response, the use of an interjection occurs after and makes sense only relative to the addressee's previous utterance.

(Kockelman 2003: 472)

Thus, the use of *ay dios* is much more complex than is implied by terms such as 'response cry'; it is often used, for example, to take the floor and/or initiate a new topic (Kockelman 2003: 472). It was argued in Section 4.2 that such complexity can be explained as being the result of grammaticalization processes, having

the effect that interjections lose their emotive force, and vocatives lose their ability to call or address a hearer in specific contexts in favor of text structuring functions.

The examples presented suggest that context extension as a mechanism of grammaticalization can be held responsible for much of the multifunctionality of interactives as well as for the fact that the usage of a number of them extends beyond several types of interactives: One and the same interactive, like English *listen*, can be a directive in some contexts but an attention signal in other contexts, and *oh* may function as an interjection in some of its usages but as a discourse marker in others.

Therefore the examined interactives can be described as grammaticalization chains forming an interlocking network of related, context-dependent functions, as captured in Figure 4.1. The ultimate sources of the chains of grammaticalization depicted in Table 4.4 are most commonly verbal and nominal forms of sentence grammar. With the hedge ‘most commonly,’ attention is drawn to the fact that not all directives can be traced back to imperative verbs (see Section 3.2), and not all vocatives have necessarily a nominal origin. What is common to the pathways in Figure 4.1, however, is that they both presuppose cooptation, whereby units of sentence grammar, such as verbal and nominal units, are transferred to the level of discourse processing, thereby turning into interactives like directives and vocatives.<sup>10</sup>

The presence of networks like the one in Figure 4.1 provides one of the main reasons why interactive types do not constitute neatly separated categories; rather, one interactive form can belong simultaneously to two or three different types of interactives.

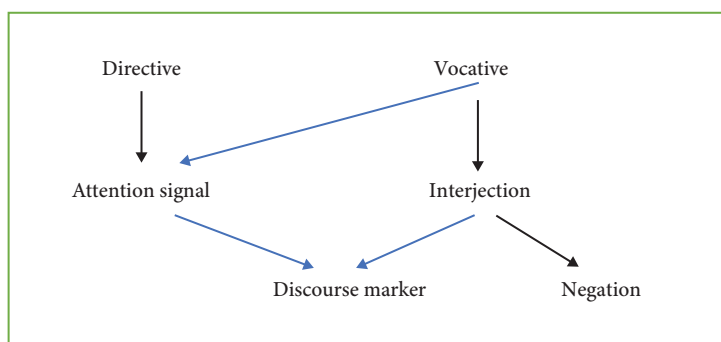


**Fig. 4.1** A network of grammaticalization chains linking types of interactives

<sup>10</sup> Obviously, no cooptation is involved in the case of interactives not derived from elements of sentence grammar, that is, primary interactives (Section 3.6.4.1).

Table 4.6 Unidirectionality in argument development

| Component of language usage  | Speaker–hearer interaction                                                                                                                                                                         | >    | Attitudes of the speaker | >    | Text organization |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|--------------------------|------|-------------------|
| Change in argument structure |                                                                                                                                                                                                    | H    |                          | T    |                   |
|                              |                                                                                                                                                                                                    | loss |                          | gain |                   |
| <i>Status of arguments</i>   | (1) S (speaker) is a stable argument, unaffected by change<br>(2) Interactives may lose an H (hearer) argument but not gain one<br>(3) Interactives may gain a T (theme) argument but not lose one |      |                          |      |                   |

Importantly, all the processes distinguished are suggestive of a general directionality, as captured in Table 4.6.

It would be tempting to relate the scenario in Table 4.6 to hypotheses that have been proposed on semantic development in grammaticalization, such as that by Traugott (1982, 1989, 1995) on the development from propositional to textual to expressive meanings. However, considering that the phenomena analyzed in that research are of a quite different nature from the ones examined here, relating for the most part to sentence grammar, such a task would be in need of a separate study.

This short survey has done hardly more than scratch the surface of the contribution made by grammaticalization in structuring the usage patterns of interactives. With the notable exception of discourse markers (see Heine et al. 2021), the evolution of interactives is still largely a *terra incognita* and more research is urgently needed.

4.2.4 Constructionalization

Grammaticalization involves, on the one hand, individual forms gradually moving from one type of interactive to another, and this was the subject of the preceding parts of Section 4.2 and some generalizations are summarized in Figure 4.1. On the other hand, it may also lead from interactive grammar to sentence grammar, as we will see in Section 4.2.5. But in addition, it also involves what, following Andrason (2021c), can be called ‘constructionalization.’ Rather than resting on the model of Construction Grammar (Goldberg 1995, 2006; Croft 2001), however, the term ‘constructionalization’ is used here in a pre-theoretical sense, referring to a process whereby interactives combine with segments of sentence grammar into new constructions, that is, new, conventionalized symbolic associations of form and meaning.<sup>11</sup>

<sup>11</sup> According to Traugott and Trousdale (2014: 257), constructionalization is the development of pairs of new forms with new meanings, irrespective of whether these are specific (i.e., substantive) signs

The following account is based on the pioneering study by [Andrason \(2021c\)](#) on interjections in Polish, and the reader is referred to that study for more detailed discussion. Polish interjections enter into construction with a wide range of text pieces apparently serving as their T arguments, classified by [Andrason \(2021c\)](#) as complements of interjections. Example (22a) illustrates such a construction, consisting of the interjection *ach* and the personal pronoun *ja* 'I.' Another kind of construction is presented in (22b), where *kurwa* (lit. 'bitch') is an expletive in collocation with the pronoun *mi*, which is a dative experiencer, constructed in the dative case, and the adverb *tu* 'here.'

- (22) Polish (Andrason 2021c, (2), (3))
- a **ach** ja ['axja]  
INT 1.SG  
'Poor me!'
- b **Kurwa** mi tu!  
INT PRON.1SG.DAT here  
'I am (here) fucking shocked!'

The complements, like *ja* in (22a) and *mi tu* in (22b), are said to instantiate a wide range of semantic roles that Polish interjections collocate with, like addressees, causers, experiencers, goals, locatives, patients, recipients, and themes. The complements consist of a number of sentence grammar constituents like personal and demonstrative pronouns, nouns, adjectives, participles, and pronominal adverbials.

Common to all combinations are the following grammatical properties: (a) The combinations have constructional meaning which may be non-compositional. (b) They are generally productive. (c) The order is typically though not invariably interjection-complement. (d) They tend to be pronounced jointly as single phrases without a pause between the two units that form the construction. (e) The complements straddle between more argument-like and more adjunct-like status. (f) The interjection-complement constructions differ in their degree of schematicity. At one end there are fairly free collocations, at the other end there are fixed, idiom-like constructions which may assume the function of new secondary interjections. (g) The interjections are analyzed by [Andrason \(2021c\)](#) as forming the head of an interjective phrase projecting two semantic core roles, namely that of experiencer and/or of causer.

Grammaticalization in this kind of constructionalization differs from the one looked at so far in Section 4.2: Rather than leading from one form to another, it leads from a collocation of text pieces to a new construction. Nevertheless,

or whether they are schemas. Our usage of the term thus is similar to that of Traugott and Trousdale (2014) but does not rely on the methodological premises that are part of the latter.

the process is in accordance with the parameters presented in Table 4.3 of Section 4.1.1:

- (a) *Context extension*: The schema head–complement is extended from sentence grammar to collocations of interjections with specific text pieces, giving rise to the new constructional schema interjection–complement.
- (b) *Desemanticization*: Each of the two units of the construction loses its specific semantic features in favor of a new constructional meaning. Furthermore, desemanticization leads to semantic generalization with the effect that the range of semantic roles expressed by the construction is gradually reduced to two roles, namely those of experiencer and causer.
- (c) *Decategorialization*: Each of the two units loses its morphosyntactic autonomy in favor of a constructional morphosyntax.
- (d) *Erosion*: The prosodic boundary (pause) between the two tends to be lost in favor of a constructional prosody (cf. *ach ja* > ['axja] in (22a)).

A final outcome of grammaticalization is that the erstwhile collocation of two separate units either fossilizes into a new secondary interjection or gives rise to a new construction of sentence grammar. With regard to the latter, attention may be drawn to the case of the Danish primary interjection *øv* (expressing vexation and/or disappointment) reported in Section 4.2.5, which exhibits some striking parallels to the case examined here.

There is reason to assume that the situation described by Andrasen (2021c) for Polish applies in some form or other also to a number of other languages, but more research is needed on this issue.

#### 4.2.5 From interactive grammar to sentence grammar

The main concern of Section 4.2 so far was with what one may loosely call internal grammaticalization, that is, with processes within interactive grammar leading from one type of interactives to another. But there is also 'external grammaticalization,' leading from interactive forms to units of sentence grammar. A spectacular example is provided by ideophones and will be looked at in more detail in Section 4.4.

Grammaticalization is a gradual process, proceeding over time in a series of small changes from less to more grammaticalized forms and constructions. But the study of interactives suggests that there is also a fairly abrupt kind of change in addition, whereby the usage of an interactive is extended to serve as a unit of sentence grammar. The following are a few casual examples illustrating the process

(see, e.g., Wilkins 1992: 131; Wharton 2003; Ameka and Wilkins 2006; Stange 2016: 36–7):

- English *Shoo!* (Middle English *Schowe!*) was originally an animal directive ('interjection') to frighten animals away. It gained a transitive verbal counterpart 'to shoe (away)', meaning 'to frighten someone or something away by saying *'shoo!'*'
  - The interjection *tsk-tsk*, *tut-tut*, /tɪsktɪsk/, or /tʌttʌt/ (a reduplicated dental click, [l]) can be extended to serve as a verb, as in *He **tutted** loudly*.
  - The interjection *ugh* is occasionally found as an adjective or noun:
- (23) British English ([KCE: 2577]; Stange 2016: 36, (22); italics in the original)  
[...]  
Andy: It's n – *it's not **ugh***. It's cider.  
Helena: *It's **ugh***. I prefer Tesco's own myself.
- The interjection *yum* can be used as an adjective, e.g., *This is **yummy*** (Ameka and Wilkins 2006).
  - The interjection *wow* can be extended to serve as a verb, for example: *Don't worry about your lecture tonight! I am sure you will **wow** them all!* (Kuteva and Heine 2021).
  - The interjection *yuck* may serve as an adjective, cf. (24).
- (24) English (Wharton 2003: 66)  
That is without doubt the **yuckiest** mouthwash I've ever tasted.
- The directive *shh* appears to have given rise to the transitive verb *shush*.
  - The interjection *hurray* can be used as a verb, *to hurray* (Foolen 1997: 25).
  - The interjection *fuck* has in certain registers of colloquial British English acquired the function of an emphatic negation marker. Thus, in the following example, *fuck* can be paraphrased as 'I definitely won't be wearing one of those' (see Section 4.2.2.1.3):
- (25) English (Kuteva and Heine 2021, (24b))  
They're all wearing kilts, but will I **fuck** be wearing one of them.
- The interjection *ouch* can be used as a noun (*It's a big ouch, an ouchie*) or as a verb, as in the following conversation.
- (26) British English ([joel25b: 2758]; Stange 2016: 36, (20); italics in the original)  
MOT: **Whoops. Uhoh.**  
INV: **Oh dear. Never mind.**  
CHI: **Ouch.**  
MOT: I think this one's just *ouched* slightly more than you.

Similar examples can be found in other languages. The German response signals *ja* ‘yes’ and *nein* ‘no’ have given rise to the verbs *be-jah-en* ‘to say yes’ and *ver-nein-en* ‘to say no,’ respectively. The interjection *øv* of Danish, expressing vexation and/or disappointment, presents the paradigm case of a primary interjection that can form utterances of its own. But it has also given rise to a number of constructions of sentence grammar, assuming functions of either a modifier or a head of the constructions concerned (see [Miltersen 2019](#) for detailed discussion). Example (27) illustrates one of these constructions, where *øv* forms the head of a ‘that’-complement clause.

(27) Danish ([Miltersen 2019](#): 54, (21))

Øv at I er syge  
 Øv that you are sick  
 ‘Sucks that you guys are sick’

In such constructions, the interjection loses salient definitional features of interactives, such as semantic, syntactic, and prosodic detachment, being a constituent and part of the meaning of the sentence. The only clear interactive feature that *øv* has retained in such constructions is that it has not lost its invariable status. A host of similar constructions, classified as interjection-complement constructions, has been identified by [Andrason \(2021c\)](#) in Polish, see Section 4.2.4.

Further examples are provided by ideophones, for these see Section 4.4.2 (Table 4.10). Finally, fairly common shifts from interactives to forms of sentence grammar can be observed in animal directives, as we saw in Section 3.11.7. One may mention, for example, that in the Pitjantjatjara language of Central Australia the directive (‘interjection’) *pai!*, directed at dogs and meaning ‘clear off!, be off!, piss off!’, has given rise to the transitive verb *paiṇi* meaning ‘to drive someone or something away; to shoo off, to hunt away’ ([Wilkins 1992](#): 131). Or in the Xhosa language of South Africa, the animal directive (‘interjection’) *hekh*, used to urge oxen to advance, is combined with the verb *ukuthi* ‘to say’ to form the fixed phrase *ukuthi hekh* (lit. ‘to.say hekh’) ‘start working with oxen’ in sentence grammar ([Andrason and Dlaki 2020](#): 199).

An interesting instance is also reported from Ayt Hadiddu, a Berber language of Central Morocco, where a whole set of animal directives has given rise to nursery terms for the animals concerned. For example, the directive *khullu!* ‘to make a chicken come,’ became the nursery term *khullu* ‘chicken,’ or the directive *rra!* ‘to make a donkey advance’ turned into the nursery noun *rrarra* ‘donkey’ ([Bynon 1976](#): 61; see Section 3.11.7).

The exact nature of the process, which tends to be discussed as a case of conversion (e.g., [Ameka and Wilkins 2006](#)), is largely unclear. Conceivably it is an instance of cooptation in reversal: Rather than leading from the plane of sentence organization to the plane of discourse organization (Section 4.1) it would have the opposite effect whereby an interactive, being anchored to the situation of



discourse, is transferred for usage as a word of sentence grammar such as a noun, a verb, or an adjective.

It would seem, however, that a more plausible account to deal with the process at the present state of research is to treat it as an instantiation of the grammaticalization parameter of context extension (Table 4.3 of Section 4.2.1), whereby expressions like *wow*, *yuck*, etc. were extended to new contexts, thereby acquiring the grammatical status characteristic of their new context and assuming a new albeit semantically related function. And having acquired a new grammatical status, the unit is now free to take on the morphological trappings characteristic of its new category, such as inflectional and derivational affixes.

### 4.3 Other mechanisms

Cooptation and grammaticalization are in fact general mechanisms shaping the form of many interactives. But they are not the only ones. Two further mechanisms are looked at in the present section, referred to as ‘expressive reinforcement’ (Section 4.3.1) and ‘camouflaging’ (Section 4.3.2). They are of more restricted scope, unlike cooptation and grammaticalization they are confined to specific types of interactives.

#### 4.3.1 Expressive reinforcement

The term ‘reinforcement’ refers to a strategy whereby the meaning of an interactive is strengthened by some additional linguistic material for specific discourse functions.

In grammaticalization theory and related frameworks, the term reinforcement is used for a mechanism whereby a linguistic form has lost part of its phonetic, semantic, and/or pragmatic substance and is ‘strengthened’ by some other element (e.g., Norde 2009: 57; Vindenes 2018; Heine et al. 2020). Accordingly, Börjars and Vincent (2011: 167–8) conclude that reinforcement is ‘a natural consequence of grammaticalization.’ In the present section, ‘expressive reinforcement’ is used in a more general sense, being restricted neither to grammaticalization nor to the final stages in the history of grammatical forms; rather, it concerns more generally the manipulation of interactives by means of additional material for expressive highlighting.

Most types of interactives can be reinforced but of all the types, expressive reinforcement is commonly associated first and foremost with ideophones, occasionally mentioned as one of the definitional features. Accordingly, Dingemanse (2011) concludes:

Ideophones may thus be freely lengthened and reduplicated to depict perceptions at various intensities, much like the shading in a drawing represents different grades of brightness.

(Dingemanse 2011: 44)

Reinforcing material can be morphophonological, morphological, or both. Morphophonological reinforcement is treated by [Dingemanse and Akita \(2017: 508\)](#) as ‘expressiveness.’ It involves most of all of the following strategies:<sup>12</sup>

- (a) Reduplication or repetition of the interactive<sup>13</sup>
- (b) Lengthening the interactive or some part of it
- (c) Adding prosodic substance relating to pitch and/or voice intensity
- (d) Phonational foregrounding by using breathy voice, creaky voice, or whisper, etc.

Example (1a) illustrates (a) and (1b) shows (b) in an ideophone of the Ghanaian language Siwu.

- (1) Siwu (Kwa, Niger-Congo; [Dingemanse 2015: 959](#))
- a wùrùfùù ‘fluffy’ vs wùrùfù~wùrùfù (~wùrùfù...) ‘fluffy here and there’
  - b wùrùfùù ‘fluffy’ vs wùrùfùùùù ‘very fluffy’

Expressive reinforcement in Japanese ideophones has been shown to involve various types of stem repetition, partial multiplication, emphatic mora augmentation, vowel lengthening, and gemination. For instance, the Japanese ideophone *don* ‘bam’ can undergo vowel lengthening (*do:n*), partial multiplication (*dododon*), or stem repetition (*don-don-don*) ([Dingemanse and Akita 2017: 510](#)).

But expressive reinforcement is by no means restricted to ideophones. Its effects on interjections can be illustrated with a few examples from South African Bantu languages. In (2), the interjection *yho* of Xhosa, expressing the speaker’s surprise, displeasure, or annoyance with one’s interlocutor, shows both (a) and (b): It is triplicated and lengthened:

- (2) Xhosa (Bantu, Niger-Congo; [Andrason and Dlaki 2020: 203, \(36b\)](#))
- Yho-o! Yho-o! Yho-o-o-o!** Asiyonyani leyo.
  - Yho-o! Yho-o! Yho-o-o-o! NEG.COP.NC9.truth DM.NC9
  - ‘Yho-o! Yho-o! Yho-o-o-o! This can’t be true.’

Another example of (b) can be seen in the interjection *tsek* of Xhosa, a swear word, which may be pronounced with the lengthening of [s] as *ts-s-sek* for emphasis, and the same applies to the ‘vocative interjection’ *wena* ‘you,’ which may be pronounced *wena*, *wenna* or *wen-na-a*. Reinforcement may also result in a slight change of meaning. Thus, the Xhosa interjection *hm* expresses contempt, surprise, and admiration, while the lengthened form *hmm* expresses contempt, disgust, repugnance, and threat ([Andrason and Dlaki 2020: 200](#)). And the German interjection *hm* expresses perplexity, but in its lengthened form

<sup>12</sup> See [Dingemanse and Akita \(2017: 508–11\)](#) for detailed discussion.

<sup>13</sup> Reduplication is understood to be a regular morphological process whereas repetition is ‘a syntactic or discourse phenomenon’ ([Dingemanse 2015: 947](#)).

*hm̩m̩* it signals agreement or confirmation (see Section 1.3.1.4). Concerning the way the English response signals *yes* and *no* can be manipulated, see Wiltschko (2021: 183).

Morphological reinforcement consists of adding a different form to the interactive. For example, in Zulu (Doke [1927] 1988: 281), the noun *inkosi* ‘chief’ has the vocative form *nkosi* ‘O chief!’. But quite commonly the social formula *e:* is prefixed to the vocative, elsewhere found as a social formula used to ask for permission to enter a house. This form of reinforcement is understood to create a respectful and polite vocative greeting, as in *e: nkosi* (SOF VOC.chief) ‘Oh chief’.

Manifestations of expressive reinforcement in discourse can be classified into the following main types:

- (i) Reinforcement is unconstrained and non-conventionalized
- (ii) Reinforcement is conventionalized to the extent that it constitutes one or a range of schematic constructions
- (iii) Reinforcement manifests itself only in frozen forms

(i) presumably constitutes the default case, being used constantly by speakers in structuring their discourses. (ii) is described by Dingemanse (2015: 958) for Siwu in terms of nine templates of expressive morphology, distinguished on the basis of differences they exhibit in the kind of reduplication and lengthening involved. For example, the Siwu ideophone *sinisini* ‘woven tightly,’ instantiating Template 1, allows for the expressive variant *sinisini~sinisinisini* signaling intensification, but not for *\*sinisini~si*.

(iii) constitutes arguably one of the most conspicuous features of ideophones: In languages across the world, many ideophones have the appearance of reduplicated forms, cf. English *bling-bling*, Swahili *lewa lewa* ‘be totally drunk,’ or Xhosa *tam-tam* ‘walk pleasantly on a soft surface.’ Reduplication may take on various shapes, such as being partial or modified, cf. English *pitter-patter* or *splish-splash*. Most likely, frozen forms are the fossilized products of either (i) or (ii).

Expressive reinforcement, as noted earlier, is especially conspicuous in ideophones, as demonstrated in particular in the seminal studies of Dingemanse (2015) and Dingemanse and Akita (2017). However, it is by no means restricted to ideophones, nor to interactives in general; rather, it is also a feature of sentence grammar. But there can be little doubt that in interactive grammar it is of utmost importance. Furthermore, expressive reinforcement appears to be sensitive to grammaticalization (Section 4.2), in that the more an interactive is grammaticalized the more it loses its ability to accept reinforcement.

The present notes could do hardly more than scratch the surface of the mechanisms and effects of reinforcement in interactives. It has been discussed in great detail by Dingemanse and associates with reference to ideophones but more research on other types of interactives is urgently needed. In concluding, mention may be made of an entirely different kind of ‘expressive reinforcement’ whereby interjections or expletives are occasionally inserted in words of sentence grammar for strengthening the expressive force of those words. The following are a few examples of insertion that have been observed in English (see [Pinker 2007](#): 361):

---

|                        |   |                                 |
|------------------------|---|---------------------------------|
| <i>absolutely</i>      | > | <i>abso-goddam-lutely</i>       |
| <i>cappuccino</i>      | > | <i>cappu-fucking-ccino</i>      |
| <i>Christ almighty</i> | > | <i>Christ al-fucking-mighty</i> |
| <i>cuckoo</i>          | > | <i>cuck-bloody-oo</i>           |
| <i>fantastic</i>       | > | <i>fan-bloody-tastic</i>        |
| <i>hot dog</i>         | > | <i>hot fucking dog</i>          |
| <i>incredible</i>      | > | <i>in-fucking-credible</i>      |
| <i>Rip van Winkel</i>  | > | <i>Rip van Fucking Winkel</i>   |

---

#### 4.3.2 Camouflaging

On December 13, 2020 the Italian football player Bryan Cristante of the club AS Roma happened to score an own goal in the match against AC Bologna, and he commented this fact by using the expletive *dio porco* (lit. ‘God pig’). As a result, he was punished for profanity by being barred from the next club match.

What this example shows is that use of some expletives or interjections can have dramatic legal implications. Swearing was a dangerous matter in the past—in the tradition of the third commandment of the bible: ‘Thou shalt not take the name of the Lord thy God in vain.’ As recounted in detail by [Ljung \(2011: 58–67\)](#), the Byzantine Emperor Justinian decreed in the sixth century that all those who ‘swore by the members of God’ should be put to death ([Ljung 2011: 60](#)), and during the twelfth century many heretics were burned at the stake for using God’s sacred name in swearing. And at the times of the French king Francis I in the sixteenth century, those caught swearing a second time were to have their lips split, third-time offenders had their tongue pierced, and fourth and definitely last-time offenders were sentenced to garroting.

But there have always been ways of avoiding such sanctions, and in the history of interactives an important way out was to draw on what one may wish to call

'camouflaging.' This mechanism has the effect that euphemisms, or 'bowdlerized alternatives' (Pinker 2007: 337) are created to replace taboo interjections relating to religion, sex, scatology, etc. that are felt to be offensive or unpleasant in certain contexts. Thus, the Italian football player mentioned could have avoided punishment by doing what some other speakers of Italian have done, namely drawing on camouflaging by replacing *dio* 'God' with *zio* 'uncle' or *due* 'two.'

There are two kinds of camouflaging strategies. Following Biber et al. (1999: 1094) they are referred to as (a) modification (e.g., *God* > *gosh*) and (b) substitution (e.g., *God* > *Goodness*). Via (a), the taboo expletive is made phonologically unrecognizable by creating 'a nonsense equivalent of a swear word' (McArthur 1992: 661; see also Gehweiler 2008: 74). By means of (b), the taboo expletive is replaced by an everyday expression of similar sound and length (e.g., *bloody* > *ruddy*, *damn* > *darn*, and *fuck* > *flip*; McArthur 1992: 661; Ljung 2011: 11). This distinction has also been described as one between moderated and euphemistic expletives, where the former involve modification and the latter substitution (see Gehweiler 2008, also for a more fine-grained analysis).

Substitution is a fairly straightforward strategy, involving the replacement of one term by another term. Modification, by contrast, is more complex, it may involve any of the following operations or their combination (cf. Gehweiler 2008: 75):

- (a) Erosion (shortening): English *Jesus!* > *gee!*
- (b) Blending: German *oh Herr Jesus!* ('oh Lord Jesus') > *ojee!*
- (c) Phoneme substitution: German *Jesus!* > *jesses!* (where *u* > *e* and *s* > *ss*)
- (d) Adding a segment: Italian *madonna!* > *madonnega!*
- (e) Rhyme change: Cockney slang *fuck!* > *Friar Tuck!*

Camouflaging, that is, using non-abusive place markers for their taboo counterparts, either via modification or substitution, has a long history in English. Referring to modifications also as 'minced oaths' (see McArthur 1992: 661), Ljung (2011) writes:

In the Middle Ages – and also later – a common English technique for creating minced oaths – especially in dealing with the English word *God* – was to replace the initial and/or the last consonant in the sensitive word by consonants with the same place of articulation: *God* became *Cock*, *Cok*, *Gog* or *Cod*, *Gad*. On occasion the vowel was changed, as in *Gad*. When more radical changes were made, the context usually supplied the information necessary for successful deciphering, for example in *Odd's blood!* and *Odd*. By means of a later and somewhat more sophisticated technique, the English oaths *By God*, *God's nails*, *God's death* and *God's wounds* were changed into *Igad!*, *'Snails!*, *'Sdeath!* and *Zounds!*, the last of which is claimed to have been the favourite oath of Queen Elizabeth I.

(Ljung 2011: 58–9)

In the nineteenth century, resistance to swearing shifted from the religious to the social domain and camouflaging became popular in the upper and middle classes: ‘interjections used in previous religious swearing like *Jesus!*, *Christ!*, and *God!* were euphemized and hidden behind linguistic peculiarities like *Jiminy Cricket!*, *Jiminy Cracket!*, *Gee Whiskers!*; *Cripes!*, *Crikey!*, *Criminy!*; *Gee!*, *Jerusalem!* *God!* became *Gosh!*, *Golly!* or *George!*’ (Ljung 2011: 66–7).

The form *gee* must have arisen fairly abruptly at the end of the nineteenth century, first attested in 1895 (cf. (3)) at a time when euphemistic form modification began to flourish as a result of the high moral standards held by the Victorian upper and middle classes (Gehweiler 2008: 85–6). There seem to have been two inter-related reasons for truncating *Jesus* to *gee*: On the one hand, it was the search for a euphemistic alternative to obscure its strong association with religion. And, on the other hand, in a time of secularization a term invoking or implying a religious context was no longer felt appropriate to express the attitudes of language users.

- (3) English (1895 OED; Gehweiler 2008: 87, (9b))

**Gee** rod! how we will thump 'em!

The cooptation of *gee* marked an abrupt change not only in form but also in meaning: Unlike *Jesus*, *gee* was not restricted to negative contexts but commonly also occurred in neutral or positive contexts, as in (4), and the taboo potential is no longer there. The meaning of *gee* has to be inferred from the context of its occurrence (Gehweiler 2008: 90).

- (4) English (1909 OED; Gehweiler 2008: 87, (9e))

‘**Gee**, what a peach of an idea!’ ‘**Oh, mother!**’

Attitudes vis-à-vis swearing have changed subsequently, but camouflaging still has not quite disappeared. Saying *sugar* instead of *shit*, *darn* instead of *damn*, or *jeezum crow* instead of *Jesus Christ!* are common usages in American English (Jay 2020: 38).

A somewhat baffling case of camouflaging via substitution, leading from one tabooed domain to another, is reported by Mackenzie (2019). The case involves a shift from the religious to the sexual domain, where the English form *God* was apparently replaced by the form *fuck*, as illustrated in the following examples (as Table 4.7 shows, *fuck* itself has also been camouflaged):

- (5) English (Mackenzie 2019: 66–7)

- |   |                                |   |                                 |
|---|--------------------------------|---|---------------------------------|
| a | <b>God</b> , that's delicious. | > | <b>Fuck</b> , that's delicious. |
| b | Thank <b>God!</b>              | > | Thank <b>fuck!</b>              |
| c | For <b>God's</b> sake          | > | For <b>fuck's</b> sake          |

Some manifestations of camouflaging as they can be found in present-day English are listed in Table 4.7. As Table 4.8 shows, the strategies used by German speakers are fairly similar; for more examples from English see Pinker (2007: 237) and from

**Table 4.7** English examples of camouflaging

| Source                   | Target                                                                                                           |
|--------------------------|------------------------------------------------------------------------------------------------------------------|
| <i>bloody</i>            | > <i>blanket, blasty, blazing, bleeding, blooming, blow, ruddy</i>                                               |
| <i>bloody hell</i>       | > <i>copulating inferno</i>                                                                                      |
| <i>bugger</i>            | > <i>bother, boy, brother</i>                                                                                    |
| <i>by God</i>            | <i>Igad</i>                                                                                                      |
| <i>by God's bones</i>    | > <i>by cokkes bones</i>                                                                                         |
| <i>Christ</i>            | > <i>Crikey</i>                                                                                                  |
| <i>Christ's wounds</i>   | > <i>Zounds</i>                                                                                                  |
| <i>damn</i>              | > <i>darn, dash, dear, drat</i>                                                                                  |
| <i>fuck</i>              | > <i>flick, foo, fudge, fug, fuzz</i>                                                                            |
| <i>for Christ's sake</i> | > <i>for crying out loud</i>                                                                                     |
| <i>God blind me</i>      | > <i>cor, blimey</i>                                                                                             |
| <i>God</i>               | > <i>cock, cod, egad, gadzooks, gawd, gog, golly, good grief, goodness gracious, Goodness, gosh, Great Scott</i> |
| <i>goddam</i>            | > <i>dadburn, dadgum, doggone, gawd, goldarn, gosh, Goodness</i>                                                 |
| <i>God's blood</i>       | > <i>S'blood</i>                                                                                                 |
| <i>God's body</i>        | > <i>S'body</i>                                                                                                  |
| <i>God's eyelid</i>      | > <i>S'lid</i>                                                                                                   |
| <i>God's nails</i>       | > <i>Snails</i>                                                                                                  |
| <i>God's life</i>        | > <i>'Od's life</i>                                                                                              |
| <i>God's truth</i>       | > <i>strewth</i>                                                                                                 |
| <i>hell</i>              | > <i>heck</i>                                                                                                    |
| <i>Jesus</i>             | > <i>gee, gee whiz, jeepers creepers, jeeze, Jiminy Cricket, jis</i>                                             |
| <i>Jesus Christ</i>      | > <i>jeezum crow</i>                                                                                             |
| <i>(Oh) Lord!</i>        | > <i>la, land's sake, lawdy, lordy</i>                                                                           |
| <i>shit</i>              | > <i>shame, sheesh, shivers, shoot, squad, sugar</i>                                                             |
| <i>son of a bitch</i>    | > <i>son of a gun</i>                                                                                            |

Sources: Allan and Burridge (1991: 36–9); Biber et al. (1999: 1083, 1094); Ameka and Wilkins (2006); O'Connell and Kowal (2008: 134); Pinker (2007: 237); Gehweiler (2008: 76); Ljung (2011: 11, 59, 78, 79); Stange 2016: 50–1; Jay (2020: 38)

both languages, see Gehweiler (2008: 81). Note that the camouflaged forms do not seem to be common outside the domain of interactive grammar.

Similar patterns of camouflaging have also been observed in other languages (see Gehweiler 2008), and Ljung (2011) observes:

The phenomenon of euphemistic taboo expletives is widespread across languages. Other examples of the same phenomenon in my data are American English *Son of a gun!* for *Son of a bitch!*, English *Copulating inferno!* for *Bloody hell!*, French *Mortbleu!* ('death of blue') for *Mort Dieu!* ('God's death'), *Sangbleu!* ('blood of blue') for *Sang (de) Dieu!* ('God's blood'), *Nom de nom!* ('name of a name') and *Nom d'un chien!* ('name of a dog') for *Nom de Dieu!* ('God's name'), German *Verflucht (nochmal)!* ('forsworn') for *Verdammt (nochmal)!* 'Damn!'),

**Table 4.8** German examples of camouflaging

| Source                                        | Translation           | Target                                                                                    |
|-----------------------------------------------|-----------------------|-------------------------------------------------------------------------------------------|
| <i>Jesus</i>                                  | ‘Jesus’               | > <i>jesses, josses</i> ‘Jesus!’                                                          |
| <i>oh, Jesu domine</i><br>(<Latin)            | ‘oh Jesus the Lord’   | > <i>ojemine</i>                                                                          |
| <i>oh mein Jesus,</i><br><i>oh Herr Jesus</i> | ‘oh my Jesus’         | > <i>ojee</i>                                                                             |
| <i>Gottes Blitz</i>                           | ‘God’s lightening’    | > <i>potzblitz</i>                                                                        |
| <i>sacré nom de dieu</i><br>(< French)        | ‘holy name of<br>God’ | > <i>sapperlot</i> ‘blimey!’                                                              |
| <i>Sakrament</i>                              | ‘sacrament’           | > <i>sapperment</i> ‘blimey!, zounds!’                                                    |
| <i>Scheisse</i>                               | ‘shit’                | > <i>Scheibenkleister</i> (‘windowpane paste’)<br>‘shit,’ or <i>schade</i> ‘what a pity!’ |
| <i>Teufel</i>                                 | ‘devil’               | > <i>taixel, teixel</i>                                                                   |

Sources: Nübling (2001: 37); Gehweiler (2008: 74, 75); Ljung (2011: 79)

Italian *Cavolo!* (‘cabbage’) for *Cazzo!* (‘dick’), Portuguese *fogo!* (‘fire’) for *fodo!* (‘fuck’),<sup>14</sup> Russian *Yaponski bog!* (‘Japanese god’) for *Yob tvojú mat!* (‘somebody fucked your mother!’), Spanish *Caramba!* for *Carajo!* (‘dick’), Swedish *Gästrikland!* (name of a province) for *Jävlar!* (‘devils’).

(Ljung 2011: 79)

Such examples do not exhaust the directions into which camouflaging, or euphemism, may move. In concluding, suffice it to mention omission as an alternative way: Rather than replacing one text segment by another, the segment suspected of profanity or blasphemy is simply dropped, as is *God* (or *Lord*) in *So help me!* or *So save us!* (Allan and Burridge 1991: 38).

#### 4.4 Grammaticalization of ideophones

Dwyer and Moshi (2003: 180–1) observe that ‘while we do consider that primary ideophones, as a type of expressive, represent a universal category, we make no such claim for grammaticalized ideophones.’ This is in fact an important observation—one that sets ideophones to some extent off from other types of interactives.

We saw in Section 4.2 that side by side with cooptation, grammaticalization is the most conspicuous mechanism shaping both the diachronic and the synchronic

<sup>14</sup> Lachlan Mackenzie (p.c. of February 17, 2022) surmises that Portuguese *fogo!* (‘fire’) is a camouflaging of *Foda-se* ‘fuck yourself’ rather than of *fodo* ‘I fuck.’



structure of interactives—so much so that the boundaries between different types of interactives are frequently blurred by pathways of grammaticalization. However, in our discussion of these pathways there was little mention of ideophones: Ideophones do not seem to participate noticeably in the grammaticalization network of interactives. There are a few exceptions, like that of the Korean ideophone *kaman* (expressing quietness), which developed into a discourse marker (Heine et al. 2021: 197), but such examples seem to be rare.

At the same time, if there is one type of interactives that commonly grammaticalizes beyond the confines of interactive grammar into sentence grammar then it is that of ideophones. In some languages they have undergone on a massive scale what Dingemanse (2017a: 373) calls ‘de-ideophonization,’ losing their extr Clausal status of syntactically, semantically, and prosodically independent forms and developing into sentence grammar categories like verbs, nouns, adjectives, or adverbs (see Dwyer and Moshi 2003). This development is summarized by Dingemanse (2017b: 198) in a nutshell thus:

The tell-tale signs of depiction that occur when ideophones are morphosyntactically independent all disappear when ideophones lose their freedom and are assimilated to become more like normal words.

(Dingemanse 2017b: 198)

Avoiding the term grammaticalization, Dingemanse and Akita (2017) use the term ‘deideophonization’ to describe salient features of this process:

So the inverse relation between expressiveness and grammatical integration provides a pathway for deideophonisation. Over time, ideophones may assimilate to ordinary vocabulary. Our account predicts that this assimilation will manifest itself as a gradual loss of expressiveness and a gradual increase in morphosyntactic integration.

(Dingemanse and Akita 2017: 527)

In more technical terms, Dwyer and Moshi (2003, (3b)) observe: ‘Some ideophones (primary) belong to the expressive dimension while others, which have been “grammaticalized” belong to the analytic dimension.’

There are, in fact, languages, such as Turkish, Korean, and Khmer, in which ideophones overlap to a significant extent with adverbials of manner (Haiman 2018: 107). And grammaticalization theory offers a way of explaining this fact: The development of ideophones is in accordance with the parameters of grammaticalization that were presented in Table 4.3 of Section 4.2.1, and part of the synchronic structure of ideophones can be accounted for with reference to this development.

## 4.4.1 A case study

The main goal of the present chapter is to provide an explanatory account for some of the diversity to be observed in the structure of interactives across languages. To this end, one specific language is looked at in more detail in this section to understand such diversity. The following example looks at the changes leading from independent utterances towards syntactically integrated words—a process dominated by increasing loss of semantic content (desemanticization), morphosyntactic integrity (decategorialization), and phonological autonomy (erosion), in accordance with the parameters in Table 4.3 of Section 4.2.1. For a detailed discussion of the process, see also Kuteva and Heine (2021).

We saw in Section 3.5.3.2 that there are three main constructions for presenting ideophones in discourse. By means of the free construction, ideophones are presented as independent utterances (e.g., *Thud!*). According to the quotative construction, ideophones are introduced as discourse complements of quotative clauses (e.g., *And the vase went thud!*). In the case of the modifier construction, by contrast, the ideophone is added to the preceding clause as a kind of afterthought or appositional parenthetical forming an independent utterance (*The vase fell to the ground – thud!*).

We suggest that the Siwu language of southeastern Ghana has made use of all three discourse constructions. Dingemanse distinguishes five morphosyntactic constructions to classify ideophones (Dingemanse 2011: 145–8; 2017a: 366–70; 2017b) and, as was shown in Section 3.5.3.2, his holophrase construction is suggestive of the free construction, while his complement construction is suggestive of the quotative construction. All the remaining constructions are manifestations of the modifier construction, and the differences among them can be explained as being the result of grammaticalization: Each of the constructions, illustrated in (6), is suggestive of a different stage of grammaticalization, as will now be demonstrated.

(6) Constructions of ideophones in Siwu (Kwa, Niger-Congo; Dingemanse 2017a: 366–70)

a The free construction<sup>15</sup>

Alɛ Kàntɔ̀ kùŋɔ̀ ðɛ                      ðbara      ũ      ara      lo.

Like Kàntɔ̀ how 3.SG-HAB 3.SG-do his things FP

↑Tsintsintsintsinsin!↑

neatly.IDE.EMPH

‘Just like Kàntɔ̀, the way he does his stuff. *Tsintsintsintsinsin!* [neatly]’

<sup>15</sup> Here and elsewhere, arrows (↑) mark the beginning and the end of prosodic foregrounding, such as markedly high pitch.

## b The adverbial construction

i-tsi            si i-fudza-ɔ                            ↑fututu~tutututu↑  
 NC1-head   if   S.NC-be.white-SG.O   IDE.pure.white~EMPH

## c The adjectival construction

bo-nà            ɔ-turi            gbógbóró-à            kere.  
 1.PL-get   ɔ-person   IDE.tough-ADJ   just  
 ‘We’ve got a sturdy person here.’

## d The predicate construction

[...]  
 ì-i-kpokporo.  
 it-NEG-IDE.hard  
 ‘It isn’t *kpokporo* [hard].’

The four constructions are now looked at in turn.

*Stage I: The free construction.* This is the stage where there is a full-fledged ideophone, as *tsintsintsintsintsin* is in (6a). Unlike the three stages to be discussed below, it corresponds to our definition of interactives. The ideophone functions as a ‘microscopic sentence’ (Diffloth 1972: 444; Dingemanse 2017a: 367–8), it is an independent unit, not part of the preceding utterance, syntactically, semantically, and prosodically complete, forming an intonation unit of its own. And Stage I ideophones are not only ‘devoid of syntactic integration’ but also ‘feature expressive morphology and prosodic foregrounding’ (Dingemanse 2017a: 373). However, the free construction in example (6a) can alternatively be interpreted also as a modifier construction where the ideophone *tsintsintsintsintsin* is added to the preceding text piece as an independent utterance as a kind of an appositional parenthetical.

*Stage II: The adverbial construction.* This can be called a stage of early grammaticalization: The use of the ideophone, like *fututu-tutututu* in (6b), has been extended to function as a modifier of the predicate (context extension), thereby losing its status as a syntactically free form. Desemanticization had the effect that semantically as well, it lost its independent status, now used to provide a depictive rendering of the scene described in the preceding predicate phrase—that is, its meaning is dependent on that of the predicate. Via decategorialization, the ideophone lost its morphosyntactically independent status by gradually turning into an adverbial modifier of the main predicate. However, grammaticalization has not been highly advanced; the ideophone is still part of what Dingemanse (2017a: 370) calls a ‘free construction.’ For one thing, the ideophone has retained a large amount of expressive and syntactic freedom, and for another, it has not proceeded to the extent that the ideophone has become an obligatory constituent of the predicate phrase, that is, its use is optional.

*Stage III: The adjectival construction.* The ideophone now shifts from what Dingemanse (2017a: 370) calls ‘free construction’ to ‘bound construction’: It may

bear inflectional morphology and undergo regular processes of tonal change. The ideophone *gbogboro* ‘tough’ has given rise to the adjective *gbogboro-à* ‘tough,’ illustrated in (6c), it functions very much like an ‘ordinary derived adjective’ directly modifying a noun. The adjectival suffix *-à*, used to derive adjectives from stative verbs, thus has been extended to the ideophone. This is suggestive of the grammaticalization parameter of context extension whereby the use of the ideophone was extended to the domain of stative verbs, thereby sharing morphological features with the latter. Furthermore, grammaticalization affected the tonal structure of the form via the parameter of erosion: The ideophone *gbogboro* ‘tough’ lost its independent tone form: It was adapted to its tonal environment, in the present example rising to extra-high tone following the ‘raising noun’ *ɔ̀-turi* ‘person’ in (6c). Finally, in accordance with the parameters of desemanticization and erosion, the ideophone also lost its ability to use expressive morphology and prosodic foregrounding, respectively.

But grammaticalization does not seem to have been concluded in that the ideophone has not been fully integrated in its new morphological class. For example, the ideophone uses the adjectival suffix *-à* often but not always, which may be taken to signal that morphological integration is not (yet) complete.

*Stage IV: The predicate construction.* Having already turned into a bound construction at Stage III, at this stage the ideophone has been grammaticalized further, turning largely into a unit of sentence grammar, which is that of a verb in the case of *kpokporo* in (6d). It has undergone decategorialization to the extent that it has lost most of the morphosyntactic independence characteristic of ideophones: Like normal verbs, it now takes verbal derivations and inflections for subject agreement, aspect, and negation, as *kpokporo* does in (6d). And via the parameter of erosion, it has lost much or all of its prosodic distinctiveness.

Importantly, a Stage IV ideophone shares with Stage III ideophones that it lacks a number of other salient properties of Stage I ideophones. In particular, it has lost the ability to foreground information prosodically and to present new information and, even more importantly, it is no longer immune to be negated and to be questioned. Note that ‘Siwu ideophones are not negated, nor are they used in question formation’ (Dingemanse 2017a: 371)—that is, Stage I ideophones of Siwu conform to the definitional properties of interactives, but not Stage IV ideophones (Section 1.2).

In short, Stage IV ideophones are fairly deeply integrated in the structure of the clause. Nevertheless, even strongly grammaticalized elements are likely to retain some features of their earlier category status, thus bearing witness to their origin.

The four-stage scenario just sketched can also be applied to other kinds of grammaticalization pathways that ideophones may undergo. According to an alternative pathway, ideophones start out as independent utterances encoded as direct speech. This pathway rests on what was described in Section 3.5.3.2 as the quotative construction, which possibly provides the cross-linguistically most

widespread strategy for encoding ideophones. Example (7) from the Shona language of Zimbabwe appears to represent a Stage-I situation, consisting of two utterances, where the first (*Nd-a-tí*) is a reporting clause instantiating the T argument, introducing the interactive *míráyí!* as an independent utterance of direct speech.

- (7) Shona (Bantu, Niger-Congo; Fortune 1962: 8; glossing and boldening ours)  
 Nd-a-tí                    **míráyí!**  
 (1.SG-PAST-say wait)  
 'I said – Wait!'

Over time, the ideophone may gradually be integrated in the preceding utterance in accordance with the parameters of grammaticalization presented in Table 4.3 of Section 4.2.1, in the end turning into an adverb-like complement, being semantically, syntactically, and prosodically integrated in the former. And the verb introducing the ideophone, most commonly 'say' (like *-tí* in (7)), 'do', 'quote', or 'think' (Childs 2003: 123), will gradually acquire features of a syntactic complementizer. On this scenario, the development can be sketched as in (8).

- (8) Hypothesized development of ideophones in the quotative constructions  
       Reporting clause - ideophone as direct speech  
     > Main clause       - adverbial complement

Note that grammaticalization is a gradual process and the four-stage scenario proposed here captures but a few structures along the chain of grammaticalization. Note further that with the rise of a new stage, the earlier stage does not disappear but rather is likely to continue to be distinguished, as witnessed in the situation of the Siwu language where all four stages distinguished co-exist side by side in present-day language use.

The observations made here confirm findings proposed earlier by Dwyer and Moshi (2003): Their distinction between 'primary' and 'grammaticalized' roughly corresponds to that between Stage I and the other stages, respectively. Since grammaticalization takes place in specific contexts, one will expect that a given ideophone is less strongly grammaticalized in some contexts but more strongly grammaticalized in other contexts. Thus, rather than with a simple distinction between a 'primary' (non-grammaticalized) usage and a grammaticalized usage of the expression concerned, we are confronted with a complex situation where a range of different, contextually defined usages are found in the language concerned, bearing witness to the various effects of one general process of grammaticalization.

And these observations also conform to the hypothesis proposed by Dwyer and Moshi (2003) according to which the development leads from an expressive to an analytic dimension of meaning: The more ideophones are grammaticalized, the more they lose their expressive force and their status as 'expressives' (Kaplan 1999).

A slightly different interpretation is proposed by [Dingemanse \(2011: 150\)](#). Rather than grammaticalization, he draws on the term ‘lexicalization’ of ideophones:<sup>16</sup>

This implies a lexicalisation scenario. Ideophones, normally free, can be pressed into service as ordinary words. The more frequently an ideophone is used, the easier it will be to recruit it as an ordinary word.

([Dingemanse 2011: 150](#))

The term ‘grammaticalization’ is also avoided by [Haiman \(2018: 104\)](#), who uses the term ‘domestication’ of ideophones instead.<sup>17</sup> A number of examples of ‘domestication’ are discussed in [Haiman \(2018: 268–81\)](#).

In accordance with our definition of interactives in Section 1.2, it is only Stage I ideophones that are considered in this book, that is, what [Dwyer and Moshi \(2003\)](#) term ‘primary ideophones.’ And here we converge with [Haiman’s \(2018\)](#) account of ideophones, which also focuses on Stage I forms.

The observations made in this section may be helpful to understanding why there are contrasting stances on how to describe the nature of ideophones. We mentioned such a case in Section 3.6.4: Unlike [Haiman \(2018: 126\)](#), [Dingemanse \(2011: 155\)](#) is hesitant to accept that deixis (indexicality) plays a role in the use of ideophones. Dingemanse’s concern is with ideophones across the various stages, whereas Haiman’s interest, like that of the present study, is first and foremost with Stage I ideophones. The evidence that there is suggests that ideophones are sensitive to deixis in their ungrammaticalized Stage I form but lose this sensitivity in the course of their grammaticalization as a result of desemanticization, but this is an issue that is in need of more research.

#### 4.4.2 A common pathway: from ideophone to lexical category

Cross-linguistically, the development of ideophones can be characterized sloppily as a ‘two-way traffic,’ involving the two processes in (8).

- (8) Common developments involving ideophones
  - a From sentence grammar to ideophone
  - b From ideophone to sentence grammar

<sup>16</sup> Concerning ‘lexicalization’ as a diachronic process, see [Brinton and Traugott \(2005\)](#); see also Section 4.4.2.

<sup>17</sup> ‘I will identify as the result of “domestication”, the process whereby ideophones become assimilated to prosaic language’ ([Haiman 2018: 104](#)), where ‘prosaic language’ roughly corresponds to the notion of ‘sentence grammar’ used here. [Haiman \(2018: 268\)](#) also characterizes the grammaticalization process as ‘a whirlwind tour of a handful of languages with ideophones in different stages of assimilation or domestication, from “rebel” to “tame” status.’

Considering that grammaticalization is defined as a unidirectional process the question is: How to account for this bi-directional traffic? The answer is that each of the traffic lines in (8) involves a different mechanism. The development in (8a) was sketched in Section 4.1.1 with reference to ideophones in Swahili. The mechanism involved is *cooptation*, whereby a text piece of sentence grammar, such as a word or part of a word, is transferred to the level of discourse management, thereby turning into an ideophone or some other interactive. For example, verbs are not uncommonly coopted as ideophones, as we saw in Section 4.1.2 with English examples like *knock*, *sniff*, and *sweep*. And as we also saw there, cooptation need not involve complete words, phrases or sentences; rather, the coopted unit can be a piece of a word. Thus, from the Swahili verb *ku-mwag-a* ‘to pour out, spill,’ the segment *mwa* was clipped to serve as the ideophone *mwa* ‘pour water suddenly/at once.’

(8b) is arguably even more widespread in the languages of the world, and it involves *grammaticalization* rather than cooptation—in accordance with what was observed in Section 4.4.1. In more general terms, (8b) can be characterized as a process whereby an ideophone is extended to new contexts reserved for lexical categories such as verbs, nouns, or adjectives and is gradually integrated in sentence grammar, in accordance with parameters of grammaticalization. This process can ultimately lead to a situation where the erstwhile ideophone ends up as a verb, a noun, an adjective, or adverb, becoming a constituent of sentence grammar and taking the morphological trappings characteristic of the word category concerned, such as derivational and inflectional affixes (see Dwyer and Moshi 2003).

Whether grammaticalization moves in the direction of new verbs, adjectives, or any other word class is contingent on the kind of strategy used by interlocutors to encode ideophones. In Section 3.5.3.2, a distinction was made between a quotative and a modifier construction. Comparative observations suggest that the former strategy is more likely to give rise to verbs or nouns, whereas the modifier construction tends to lead to the development of new adjectives or adverbs. The insightful description by Amha (2001: 254–60) shows, for example, that in the Maale language of Ethiopia the quotative construction based on the verb *ge?* ‘say’ as a quotative index is used to create derived ideophonic verbs. The modifier construction, in contrast, leads to intensifying ideophones and finally to derived ideophonic adjectives.

Table 4.9 lists a catalog of grammatical features that ideophones are likely to exhibit prior to and after grammaticalization but ignores the intermediate stages leading from initial to final stage.<sup>18</sup> The nature of intermediate stages (II and III) was illustrated in Section 4.4.1 with an example from Siwu while Stages I and IV

<sup>18</sup> Syntax and placement also concern what is referred to by Dingemanse and Akita (2017: 517) as the ‘optionality hierarchy’ in their account of Japanese ideophones, where the initial stage seems to relate to ‘syntactically optional’ and the final stage to ‘syntactically obligatory constructions.’

**Table 4.9** The initial and final stages to be expected in the grammaticalization of ideophones

|                            | Initial stage                                              | Final stage                                                                                      |
|----------------------------|------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| <i>General</i>             | Ideophones form a unit of discourse                        | Ideophones are integrated within word classes of sentence grammar                                |
| <i>Function</i>            | Expressive                                                 | Non-expressive                                                                                   |
| <i>Meaning</i>             | Does not change the sentence meaning                       | Contributes to the meaning of a sentence                                                         |
| <i>Syntax</i>              | Unattached                                                 | Constituent of a sentence                                                                        |
| <i>Morphology</i>          | Invariable                                                 | Takes inflectional and/or derivational morphology                                                |
| <i>Prosody</i>             | Prosodically distinct, forms an intonation unit of its own | Integrated in the intonation contour of the sentence                                             |
| <i>Placement</i>           | Free, can form an utterance of its own                     | Restricted to slots provided by its morphosyntactic status as a noun, verb, adjective, or adverb |
| <i>Discourse functions</i> | Resistance to being negated or interrogated                | Can freely be negated and interrogated                                                           |

discussed there relate, respectively, to the initial (Stage I) and the final stage (Stage IV) in Table 4.9.

The following example illustrates overall features of the initial and final stages of grammaticalization (see also [Dingemanse and Akita 2017](#) on Japanese). In the South African Xhosa language, the development sketched in (8a) above appears to be uncommon while (8b) is a fairly regular process ([Andrason 2021a](#): 152–3), affecting especially the rise of new verbs, as Table 4.10 shows.

Like other Bantu languages, Xhosa has a noun class system marked by prefixes, and nouns can be derived from ideophones by simply adding a noun class prefix. Thus, the prefix *um-* of noun class 1 is added to the ideophone *xhimfi* ‘hit, strike’ to form the noun *um-xhimfi* ‘boxer,’ or noun class prefix *isi-* of noun class 7 to form the noun *isi-fotho* ‘dent, dimple’ from the ideophone *fotho* ‘be dented; dent.’ In Southern Sotho, another Bantu language of southern Africa, there is a fairly productive process whereby disyllabic ideophones turn into intransitive verbs by adding the suffix *-ha*, into transitive verbs with the suffix *-la* and into causative verbs with the suffix *-tsa* ([Doke and Mofokeng \(\[1957\] 1985](#): 174–5).

Similar patterns are found in other Bantu languages. For example, in the Northern Sotho language of South Africa, names for birds are commonly derived from ideophones expressing the call made by the bird, adding a noun class prefix of class 1 (*mo-*) or class 5 (*le-*); Table 4.11 provides a few examples ([Poulos and Louwrens 1994](#): 359).

Languages differ greatly in the extent to which they have grammaticalized their ideophones. Roughly speaking, the following types of languages can be



**Table 4.10** From ideophone to verb, selected examples

| Language | Genetic affiliation            | Ideophone      |                                                  | Verb               |                                  |
|----------|--------------------------------|----------------|--------------------------------------------------|--------------------|----------------------------------|
| German   | Germanic, Indo-European        | <i>bums</i>    | ‘thud’                                           | <i>bums-en</i>     | ‘to screw sb.’                   |
| Iraqw    | Southern Cushitic, Afroasiatic | <i>plumps</i>  | ‘bump’                                           | <i>plumps-en</i>   | ‘to flop’                        |
|          |                                | <i>hiq</i>     | ‘hiccup’                                         | <i>hiq-iít</i>     | ‘to have a hiccup’               |
| Xhosa    | Bantu, Niger-Congo             | <i>tsuíf</i>   | ‘sound of wet wood in fire’                      | <i>tsuuf-iít</i>   | ‘to sizzle’                      |
|          |                                | <i>thapu</i>   | ‘appear suddenly’                                | <i>thaphu-lula</i> | ‘to unearth something hidden’    |
|          |                                | <i>xum</i>     | ‘noise of falling suddenly’                      | <i>xum-ka</i>      | ‘to stop’                        |
| Hausa    | Chadic, Afroasiatic            | <i>zùndùm</i>  | ‘sound of an object falling into a well or pond’ | <i>zundùm-ā</i>    | ‘to fall into water’             |
|          |                                | <i>tsùndùm</i> | ‘sound of an stone falling into water’           | <i>tsundùm-ā</i>   | ‘to fall into water with a plop’ |

Sources: [Andrason \(2021a: 152\)](#), [Newman \(2000: 256\)](#), [Nordbustad \(1988: 297–8\)](#)

**Table 4.11** From ideophone to noun in Southern Sotho (Bantu, Niger-Congo)

| Ideophone for call made by bird | Noun for bird  | English name   |
|---------------------------------|----------------|----------------|
| <i>bibi</i>                     | <i>le-bibi</i> | ‘Blue waxbill’ |
| <i>haha</i>                     | <i>le-haha</i> | ‘Hadedá’       |
| <i>kowê</i>                     | <i>mo-kowê</i> | ‘Grey lourie’  |
| <i>kuru</i>                     | <i>mo-kuru</i> | ‘Turtle-dove’  |

Source: [Poulos and Louwrens \(1994: 359\)](#)

distinguished based mainly on the comparative notes volunteered by [Dingemanse and Akita \(2017: 506\)](#):

- (a) *Languages with non-grammaticalized ideophones.* These are languages where ideophones are wholly ‘non-integrated,’ being in accordance with the definition of interactives in Section 1.2, exhibiting the features of the initial stage in Table 4.9. Semelai, an Austroasiatic language of Malaysia, is conceivably such a language, in that ideophones are said to never be

syntactically integrated and to usually being distinguished by an intonation break.

- (b) *Languages with partly non-grammaticalized and partly grammaticalized ideophones.* For example, in Yucatec Mayan of Mexico, ideophones are characterized by syntactic independence, expressive morphology, and marked prosody. But when ‘instantiated as verbs’ they are grammaticalized, being morphosyntactically integrated and combined with aspectual marking (Dingemanse and Akita 2017: 506).
- (c) *Languages with fully grammaticalized ideophones.* In such languages, ideophones are largely or entirely integrated in sentence grammar, in accordance with the final stage of Table 4.9. The East Cushitic language Somali of northeastern Africa appears to be such a language, where ideophones are portrayed as having become a subclass of nouns. And Haiman (2018: 267) mentions further possible type (c) languages where ‘ideophones may also be normal linguistically fully integrated nouns, verbs, adverbs, or adjectives,’ namely the Malayalam language of southwest India, Udihe (Udege) of eastern Russia, and Venda of South Africa.

Furthermore, the Sepik language Iatmul of Papua New Guinea might also qualify as a type (c) language. Ideophones are described by Jen-draschek (2012: 312–17) as a subclass of adverbs, occurring as complements of a control verb, which in most cases is the verb *wa* ‘say’. Ideophones in Iatmul, in fact, appear to be transparently derived from the quotative construction described Section 3.5.3.2, having been grammaticalized to some extent as adverbial complements.

Another language that also appears to have moved towards a grammatical structure suggestive of type (c) appears to be Hausa of Northern Nigeria: There no longer clearly is a syntactically defined class of ideophones:

[...] while one may speak of ideophonic words in Hausa as a single class from a phonological/semantic point of view, one must syntactically separate them into at least three grammatical classes: adjectival-intensifiers, verbal-intensifiers, and descriptive-adverbs.

(Newman 1968: 111; cited from Dingemanse 2018: 10)

However, Hausa ideophones are far from having been integrated into categories of sentence grammar, as the description by Newman (2000: 242–59) suggests. While they do in fact function as adverbials, adjectivals, or nominals, they clearly have retained features that set them off from categories of sentence grammar. First, they have retained specific phonological and semantic features. Second, they have not been grammaticalized to the extent that they display the whole gamut of structural features characterizing adverbs, adjectives, or nouns. And finally, they have remained prosodically distinct, being characterized by extra forcefulness or

loudness and—most of all—by extra high pitch, often also by an intonational break (Newman 2000: 256).

The situation in Japanese suggests, in fact, that grammaticalization can work differentially in ideophones in that it may affect morphosyntax distinctly more than meaning. As the description by Kita (1997: 386) shows, Japanese ideophones are syntactically but not really semantically integrated in the structure of a sentence. According to the description by Kita (1997), Japanese ideophones ('mimetics') are—at least to some extent—another example suggestive of (c), being syntactically fairly integrated:

In summary, there are two syntactically distinct classes of mimetics. Adverbial mimetics, followed by an optional complementizer, can serve only as an adverbial in a clause. Nominal mimetics are followed by either the copula to serve as the predicative element of a clause, or a postposition to form a postpositional phrase. (Kita 1997: 386)

However, the description by Dingemanse and Akita (2017) clearly shows that the situation in Japanese is more complex, exhibiting a continuum of the following kind: At one end there is the least grammaticalized quotative construction, characterized by minimal decategorialization and erosion, presenting instances of ideophones close to the prototype. At the other end there are the predicative constructions, showing the highest amount of grammaticalization and integration in the structure of the predicate.

To conclude, there is justification to maintain a distinction between primary and grammaticalized ideophones as Dwyer and Moshi (2003) do, but the boundary between the two is gradient. These authors are also right in observing that the former represent a universal category while such a claim cannot be made for grammaticalized ideophones. Considering that ideophones are a universal category and that grammaticalization has also been shown to be a universally valid mechanism this might seem to be surprising. As a matter of fact, however, it is not, for the following reason: The way grammaticalization affects ideophones differs greatly from one language to another. Differences concern, first, the relative degree of grammaticalization, second, the kind of ideophones affected, third, the construction selected for presenting an ideophone, and fourth, the contexts in which ideophones grammaticalize. Accordingly, it is fairly safe to assume that there are no two languages in the world showing exactly the same kind of grammaticalized structures in ideophones.

The reason for dealing with ideophones in this chapter in more detail is that grammaticalization theory provides a tool for explaining why ideophones behave typologically the way they do. For example, Dingemanse (2017b: 197) and Dingemanse and Akita (2017) draw attention to the fact that ideophones show 'an inverse relation between prosodic foregrounding and morphosyntactic

integration: Ideophones that are more deeply integrated in the structure of the clause lose their prosodic foregrounding.’ This observation is far from surprising; rather, it is a predictable product of grammaticalization, as we saw in the course of this section. Thus, in accordance with the grammaticalization parameters presented in Section 4.2.1, an ideophone undergoing grammaticalization can be expected, on the one hand, to gradually lose its ‘syntactic’ independence and to be increasingly integrated morphosyntactically within the structure of a sentence (decategorialization). On the other hand, it will also lose its prosodic distinctiveness, such as ‘prosodic foregrounding’ (erosion).

And in addition to these two predictions, grammaticalization theory also allows formulating a third prediction, based on the parameter of desemanticization: The more an ideophone is grammaticalized the more it will lose its expressive function, becoming semantically integrated in the meaning of a sentence. In sum, ideophones may undergo the kind of development sketched in Table 4.9, leading gradually from the initial to the final stage, even if few may make it all the way to the final stage.

Ameka (2001) characterizes ideophones cross-linguistically thus:

[Ideophones] are first and foremost a type of words – a lexical class of words – which need not belong to the same grammatical word class in a particular language nor across languages.

(Ameka 2001: 26)

One main topic of the present section was to explain why ideophones ‘need not belong to the same grammatical word class’: The more ideophones undergo grammaticalization the more they are likely to be integrated in the existing word classes of sentence grammar. With reference to the typology proposed, the following generalization can be proposed: While ideophones in languages of type (c) are likely to be part of the word classes of sentence grammar, ideophones of type (a) are full-fledged interactives and therefore do not belong to any word class of sentence grammar.

For good reasons, the development of ideophones into words of sentence grammar has also been argued to involve lexicalization (Dingemanse 2011: 150). However, since this development is in accordance with parameters of grammaticalization, as we saw in this chapter, we follow Dwyer and Moshi (2003) in proposing a grammaticalization account here, but more research is needed on this issue.

## 4.5 Conclusions

Presenting a survey of sentence grammar usages of the English interjections *ouch*, *ugh*, *whoops*, *wow*, and *yuck*, Stange (2016: 37) concludes: ‘This range of

examples provides compelling counterevidence to the claim that interjections are immune to morphological processes.' According to the view proffered here, this is by no means counterevidence to the definition proposed in Section 1.2: Morphological processes as described by Stange do not concern the status of interjections, or of other types of interactives for that matter. Rather, they demonstrate that interactives can cease being interactives by developing into something else, such as nouns, verbs, adjectives, or adverbs of sentence grammar, and the concern of the present chapter was to identify the mechanisms underlying such developments.

The goal of Chapter 4 thus was to provide a diachronic perspective of interactives. The findings made can be summarized in the following way: The two outstanding mechanisms characterizing the grammatical structure of interactives are cooptation (Section 4.1) and grammaticalization (Section 4.2). Via cooptation, interactives are 'exapted' from the rich pool of text material provided by sentence grammar, being cut off syntactically and prosodically from their etymological sources, now being anchored to the situation of discourse rather than the structure of a sentence. Grammaticalization has two main effects. On the one hand, it may lead to a change in argument structure: Vocatives lose their H (hearer) argument when turning into interjections, while interjections gain T (theme) arguments when grammaticalized into discourse markers, etc. (Section 4.2). On the other hand, grammaticalization is responsible for shifts from one type of interactive to another, and these shifts have dramatic effects on the synchronic structure of the forms concerned: A given interactive may qualify in some of its usages as an interjection but as an attention signal in other contexts, and in still other contexts it may well be described as a discourse marker.

Compared to cooptation and grammaticalization, other mechanisms shaping the development of interactives are of a more local nature. This applies especially to what was called 'camouflaging' in Section 4.3.2, a strategy restricted essentially to interjections.

As the chapter may have shown, the diachronic processes discussed provide some explanations on the synchronic behavior of interactives. Cooptation accounts for the set of properties that define them as a category of their own. Grammaticalization, in contrast, is responsible for the fact that interactives form a network of related, context-dependent functions. This network provides one of the main reasons why interactive types do not constitute neatly separated categories; rather, one and the same interactive form can function, e.g., as a directive in some contexts, as an attention signal in other contexts, and as discourse markers in still other contexts. The result is that a given interactive form can simultaneously belong to two or three different interactive types.

Diachrony thus provides a partial, if not the main explanation for the interlocking nature of the network of interactives. But it also captures the bi-directional development that was sketched in (8) for ideophones: On the one hand,

ideophones can be coopted from sentence grammar, as we saw most clearly in the Swahili examples in Section 4.1. On the other hand, it was shown in Section 4.4 that ideophones have moved freely from interactive grammar to sentence grammar in a number of languages.

Such bi-directional developments, while common in ideophones, is rare in other types of interactives—nevertheless is not unheard of. Take the case of the English item *fuck*, first attested as a word of sentence grammar at least since the sixteenth century, at some later time coopted as an expletive and interjection, but later on grammaticalized to usages as a negation marker (Section 4.2.2.1.3).

In concluding, it should be emphasized that many of the hypotheses proposed in this chapter must be taken with care, being based on the diachronic interpretation of patterns of present-day language use rather than on systematic historical reconstruction—the history of interactives, be that in English or in any other language, is still poorly documented.

## Related approaches

So far we were concerned with describing interactives, be that from a synchronic perspective (Chapters 1 to 3) or a diachronic perspective (Chapter 4). This chapter aims at searching for a wider perspective by relating the observations made to previous research. The idea that there is a class of linguistic forms that differs from the rest of grammar and, hence, should be treated in its own right in fact surfaces in some form or other in most traditions of linguistic analysis, and the present chapter provides an overview of work that has been done in this direction. To this end, the main questions asked in this chapter are as follows:

- (a) What is the place of interactives in linguistic monographs and grammars?
- (b) How does interactive grammar relate to other frameworks on dual-processing?
- (c) Is there extra-linguistic evidence in support of interactives as a category of their own?

Question (a) is the subject of Section 5.1, while Section 5.2 is devoted to question (b), dealing with a more theoretical issue, namely with the claim that language structure exhibits a dualistic organization. Question (c) is dealt with in Sections 5.3 and 5.4, being concerned, on the one hand, with correlations between linguistic and neurolinguistic categorization (Section 5.3) and, on the other hand, with a related dualistic contrast that has been identified in social psychology (Sections 5.4, 5.5).

### 5.1 Notions of interactive grammar

The present section is devoted to finding out how interactives have been treated in traditional accounts of grammatical analysis and description.

#### 5.1.1 Grammar traditions

Common to most monographs of linguistics and grammars of English and other European languages is, first, that interactives are treated as a marginal phenomenon that may not even be considered worthy of having a chapter devoted

to them or of having a place in grammar in the first place. Second and relatedly, even if they are not ignored, their description is generally poor, there is hardly any information on the grammatical or functional behavior of interactives. And third, as far as they are distinguished at all, they tend to be lumped together under some specific label, which is most commonly that of ‘interjections.’

The following overview is restricted to a few studies for illustrative purposes, and no claim is made that these studies are representative in any way. However, studies that fail to deal with interactives or deal with them only in passing are ignored. To begin with, some studies devoted to the grammatical description of the English language are examined (Section 5.1.1.1) before looking at grammars of languages in other parts of the world (Section 5.1.1.2).

5.1.1.1 Some studies on interactives in English

In his classic monograph of linguistics, Bloomfield ([1933] 1962: 176) proposes a category of interjections, which includes English forms like *ouch*, *oh*, *shh* (*sh*), *gosh*, *hello*, *sir*, *ma’am*, and *yes*, as well as ‘secondary interjections,’ many said to be ‘often of peculiar construction,’ such as *dear me*, *goodness gracious*, *by golly*, *you angel*, *please*, *thank you*, *goodbye*. As Table 5.1 shows, there are five types of interactives included by this author within his category of ‘interjections.’

Table 5.1 ‘Interjections’ distinguished by Bloomfield ([1933] 1962)

|                 |                                                                                                     |
|-----------------|-----------------------------------------------------------------------------------------------------|
| Directive       | <i>sh</i>                                                                                           |
| Interjections   | <i>by golly</i> , <i>dear me</i> , <i>goodness gracious</i> , <i>gosh</i> , <i>oh</i> , <i>ouch</i> |
| Response signal | <i>yes</i>                                                                                          |
| Social formulae | <i>goodbye</i> , <i>hello</i> , <i>please</i> , <i>thank you</i>                                    |
| Vocatives       | <i>ma’am</i> , <i>sir</i>                                                                           |

Source: Bloomfield ([1933] 1962: 176)

A similar treatment is found in some other earlier grammars of English, like the *English Grammar for Today* of Leech et al. (1982), who write:

Interjections are rather peripheral to language: words like *ugh*, *phew*, *oh*, *ah* and *ouch* are linguistically somewhat primitive expressions of feeling, only loosely integrated into the linguistic system. We can include here too swear words (*damn* etc.), greetings (*hello*) and other signalling words like *goodbye*, *yes*, *no*, *okay* etc.  
(Leech et al. 1982: 53)

Table 5.2 shows that in this grammar four types of interactives are distinguished.

A particularly rich perspective of interactives is presented by Zwicky (1985) in his discussion of clitics and particles. He proposes a category of forms that he calls ‘discourse markers,’ containing a range of interactives, as Table 5.3 shows. Zwicky



**Table 5.2** ‘Interjections’ distinguished by [Leech et al. \(1982\)](#)

|                  |                                      |
|------------------|--------------------------------------|
| Discourse marker | <i>okay</i>                          |
| Interjections    | <i>ah, damn, oh, ouch, phew, ugh</i> |
| Response signals | <i>no, yes</i>                       |
| Social formulae  | <i>goodbye, hello</i>                |

Source: [Leech et al. \(1982: 53\)](#)

adds that ‘[o]n distributional grounds, the traditional class of exclamatory “interjections” in English – items like *ouch, boy, gosh, holy cow, wow, my goodness, me,* and *hell* – should also be grouped with these markers’ ([Zwicky 1985: 303](#)).

**Table 5.3** ‘Discourse markers’ distinguished by [Zwicky \(1985\)](#)

|                   |                                                        |
|-------------------|--------------------------------------------------------|
| Attention signal  | <i>hey</i>                                             |
| Directives        | <i>listen, look</i>                                    |
| Discourse markers | <i>like, oh, okay, now, say, uh, well, why, y’know</i> |
| Interjections     | <i>gosh, holy cow, my goodness, oh, wow</i>            |
| Response signals  | <i>no, yes</i>                                         |
| Social formula    | <i>please</i>                                          |

Source: [Zwicky \(1985: 303–4\)](#)

Zwicky notes that ‘discourse markers and their kin are syntactically insulated from the rest of the sentences in which they occur. Beyond the fact that they are located within a constituent, they form no sort of unit with adjacent words,’ and he adds that they are also ‘prosodically independent’ ([Zwicky 1985: 303–4](#)). Importantly, Zwicky shows that his ‘discourse markers’ can occupy the same paradigmatic slot, giving the following examples:

- (1) English ([Zwicky 1985: 303–4](#))
  - a. Kim will want, **well/oh/like/uh/say/why**, a golden penguin.
  - b. **Well/Hey/Okay/Yes/Y’know/Look/Listen**, let’s go to Pismo Beach.

A concept that in some ways resembles interactives is that of pragmatic markers of [Fraser \(1990\)](#). Using what may be called a primarily semantic-pragmatic approach, Fraser distinguishes three types of pragmatic markers, namely (a) basic, (b) commentary, and (c) parallel pragmatic markers. (a)-markers signal with what (illocutionary) force the speaker intends the sentence content to be taken, they include expressions like *please, I claim, I promise*, and *I admit*. (b)-markers encode an entire message constituting a comment on the basic message, such markers being *frankly, incredibly, amazingly, if I may be personal*, and *well*. (c)-markers encode entire messages which are separate from and are added to the basic and/or

commentary messages. (c)-markers are *damn* (as in *Take your damn shoes off the table*) and vocatives (*Sir, My friend*), but they also include discourse markers, such as *and, so, anyway, and well*.

Furthermore, Fraser (1990: 386, fn. 3) distinguishes between structural and lexical pragmatic markers. The examples above relate only to the latter while the former include, for example, the sentence moods declarative, interrogative, and imperative and need not concern us here.

The similarity between Fraser's (1990) and our approach can be seen in the fact that his catalog of pragmatic markers includes instances of some types of interactives (see Table 5.4). Beyond this fact, there is little that the two approaches have in common, most of all for the following reason: Whereas interactives are defined here strictly in terms of grammatical criteria, such criteria play essentially no role in Fraser's framework. For example, for Fraser (1990: 388) coordinating conjunctions like *and, or, and but* are discourse markers (see also Schiffrin 1987), while our category of discourse markers excludes these conjunctions since in most of their uses they do not conform to our definition of neither discourse markers nor of interactives in general (Section 1.2).

**Table 5.4** 'Pragmatic markers' distinguished by Fraser (1990; 1999) qualifying as interactives

|                   |                              |
|-------------------|------------------------------|
| Discourse markers | <i>I admit, anyway, well</i> |
| Interjections     | <i>damn, huh, phoo</i>       |
| Social formula    | <i>please</i>                |
| Vocatives         | <i>Sir, My friend</i>        |

Source: Fraser (1990; 1999)

In short, Fraser's (1990) notion of pragmatic marker characterizes a group of linguistic expressions having related pragmatic functions as they can also be observed in interactives, but it does not define this group as a grammatical category of the kind proposed here.

Also approaching the analysis of interactives from the perspective of 'interjections,' Norrick (2009: 868) also includes discourse markers in this category. He describes the relationship between interjections and discourse markers, arguing that the more frequent primary interjections have generally taken on specific functions as pragmatic markers, assuming functions as discourse markers (Norrick 2009: 868). While the topic of Norrick's insightful analysis is interjections, the majority of units listed in Table 5.5 are discourse markers. Note that, in the tradition of Schiffrin (1987), the forms *and, but, and (be)cause* are also listed by Norrick (2009: 868) but are not included here because in most of their uses these are connectives of sentence grammar and do not correspond to our definition of interactives in Section 1.2.

**Table 5.5** ‘Interjections’ distinguished by [Norrick \(2009\)](#)

|                   |                                                                            |
|-------------------|----------------------------------------------------------------------------|
| Attention signal  | <i>hey</i>                                                                 |
| Discourse markers | <i>anyway, I mean, like, mm, okay, really, right, so, uh, well, y’know</i> |
| Interjections     | <i>ah, huh, oh, ooh, wow</i>                                               |
| Response signals  | <i>hm, uh-huh, mhm, right, yeah</i>                                        |

Source: [Norrick \(2009: 868\)](#)

While the studies so far summarized are restricted to English interactives, [Ameka and Wilkins \(2006\)](#) adopt a worldwide perspective in proposing the six pragmatic classes of ‘interjections’ depicted in [Table 5.6](#). This classification takes care of seven types of interactives, the only types missing being evaluatives, ideophones, and vocatives.

**Table 5.6** Classes of ‘interjections’ distinguished by [Ameka and Wilkins \(2006\)](#)

| Class | Functions                                                                         | English examples                         | Corresponding types of interactives |
|-------|-----------------------------------------------------------------------------------|------------------------------------------|-------------------------------------|
| (i)   | To express a speaker’s mental states and actions                                  | <i>agh, ow, ugh, wow</i>                 | Interjections                       |
| (ii)  | They are directed at an interlocutor                                              | <i>hey, yoohoo, shh</i>                  | Attention signals, directives       |
| (iii) | To monitor the conversation or discourse                                          | <i>huh?, hmmm, mmmm, uh-huh</i>          | Reponse elicitors, response signals |
| (iv)  | To respond to something a participant says                                        | <i>mhm, no, ok, yeah, yes</i>            | Response signals                    |
| (v)   | They belong to a ritual pair or set from which speaker and addressee might select | <i>bye-bye, hi-hi, sorry, ta, thanks</i> | Social formulae                     |
| (vi)  | To give a meta-commentary, e.g., to signal problems in formulation or production  | <i>uh, um</i>                            | Discourse markers                   |

Finally, building on earlier work, especially on [Ameka \(1992a\)](#), [Stange \(2009: 54; 2016: 19, Figure 2.3\)](#) proposes a classification that captures some salient features of the category of interactives. With reference to the way her classification relates to this category, her proposal can be summarized in the following way:

- (a) Her class label ‘interjection’ includes the following types: Interjections, directives, and response signals.

- (b) Her type of ‘interjections *proper*’ is similar to that of interjections as proposed here in that it includes the primary interjections *ugh*, *whoops*, and *ow* and the secondary interjections *oh dear* and *shit*. There are two exceptions, however. First, ‘Interjections *proper*’ also include the response signals *yes* and *no*. And second, her class of ‘cognitive interjections’ (*aha!*, *ah!*, and *I see!*) is not part of the category of ‘interjections *proper*’.
- (c) Her type of ‘phatic interjections’ (*uh-huh*, *mhm*, *right*, and *yeah*) corresponds to our type of response signals, with the exceptions mentioned in (b).
- (d) Her type of ‘conative interjections’ (*shh*, *psst*, and *shut it*) echoes our type of directives.

To conclude, [Stange’s \(2016\)](#) concept of ‘interjection’ is more restrictive than that of [Ameka and Wilkins \(2006\)](#) in that it excludes types such as attention signals, discourse markers, response elicitors, and social formulae—for obvious reasons since the concern of her study is with emotive interjections in British English.

#### 5.1.1.2 Studies on interactives in other parts of the world

In a similar way as in the work on English grammar just mentioned, interjections or some related notions are treated as a highly inclusive ‘part of speech’ or class in many grammars of languages across the world, largely, though not entirely corresponding to the category of interactives. The following discussion is restricted to a few languages.

For Korean, [Sohn \(1999: 399, 399–400\)](#) proposes to distinguish a category of ‘discourse adverbials’. They usually occur either alone or preceding a sentence, and they ‘occur syntactically outside the boundary of a sentence,’ being ‘conditioned by discourse contexts or speech situations’ ([Sohn 1999: 399](#)). They include instances of four types of interactives, as [Table 5.7](#) shows.

**Table 5.7** Korean ‘discourse adverbials’ distinguished by [Sohn \(1999\)](#)

|                  |                                                                                                                                                  |
|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| Interjections    | <i>acha</i> ‘Heavens!, By Jove!’, <i>aiko</i> (or <i>aikwu</i> ) ‘Oh My!, Dear Me!, Good heavens!, Alas!, Gosh’, <i>a.yas</i> ‘Shucks!’, ‘Ouch!’ |
| Response signals | <i>yey/ney</i> ‘yes’, <i>ani-(y)o</i> ‘no’, etc.                                                                                                 |
| Social formulae  | <i>ipwa</i> , <i>ipwa.yo</i> , <i>yepo</i> , <i>yepokey</i> ‘Hello!’                                                                             |
| Vocatives        | <i>sensayng-nim</i> ‘Sir, Professor!’, <i>Kim-kwun</i> ‘Mr. Kim!’                                                                                |

Source: [Sohn \(1999: 399–400\)](#)

Moving to another continent, in his grammar of Wampis, a language of the Peruvian Amazon, [Peña \(2015: 412–20\)](#) includes a category of ‘interjections’ which

contains not only genuine interjections but also a social formula, a response signal, commands (imperative-like forms like *hastá* 'Wait!'), discourse markers, and ideophones, cf. Table 5.8.

**Table 5.8** Wampis 'interjections' distinguished by Peña (2015)

|                   |                                                                                  |
|-------------------|----------------------------------------------------------------------------------|
| Directive         | <i>hastá</i> 'Wait!'                                                             |
| Discourse markers | <i>kami</i> 'so' (repair), <i>tuki</i> 'well, OK, then,' <i>naa</i> (hesitation) |
| Ideophones        | <i>kusuú</i> (spitting), <i>tapít</i> 'grabbing forcefully or suddenly'          |
| Interjections     | <i>mah</i> (speaker surprise or admiration), <i>sí</i> 'what a shame, pity'      |
| Response signal   | <i>haʔá</i> 'yes'                                                                |
| Social formula    | <i>sí</i> 'thanks'                                                               |

Source: Peña (2015: 412–20)

Similar observations have been made in African languages. In his grammar of the Atlantic language Kisi of Guinea, Sierra Leone, and Liberia, Childs (1995) stands out in including most types of interactives, but he does not treat interactives as a uniform entity. Like many other authors of reference grammars he proposes a category of 'interjections,' which includes the seven types in Table 5.9.

**Table 5.9** Kisi 'interjections' distinguished by Childs (1995)

|                   |                                                                                                                                                                                                                                                                                                                                  |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Attention signal  | <i>s-s-s</i> (attract someone's attention)                                                                                                                                                                                                                                                                                       |
| Directive         | <i>m</i> 'here, take this (the object)'                                                                                                                                                                                                                                                                                          |
| Ideophone         | <i>koŋ-koŋ</i> (imitating the knock on the door)                                                                                                                                                                                                                                                                                 |
| Interjections     | <i>b̂</i> (amazement and elation at remembering something), <i>èèè</i> (disappointment), <i>kó</i> (doubt or uncertainty), <i>kpá</i> , <i>kpá</i> (irritation, doubt, lack of surety, hesitation), <i>kpóù</i> (surprise or astonishment)                                                                                       |
| Response elicitor | <i>nèé</i> 'right?'                                                                                                                                                                                                                                                                                                              |
| Response signals  | <i>ěě</i> 'yes,' (?) <i>ě?ě?</i> 'no'                                                                                                                                                                                                                                                                                            |
| Social formulae   | <i>báléká</i> 'thank you,' <i>hàkátó</i> 'excuse me,' <i>idíyó</i> , <i>idíyálè</i> 'goodbye,' <i>ísè</i> , <i>làwósè</i> 'hello' (greetings, SG and PL), <i>mbàà</i> (response to salutation), <i>núwálí</i> , <i>níwálí</i> 'thank you,' <i>ŋ' cìikáŋ</i> 'goodbye,' <i>wóníwálí</i> 'congratulations,' <i>yandíí</i> 'please' |

Source: Childs (1995: 145–7)

Being aware that interactives form a more comprehensive category than what is traditionally known as 'interjections,' Doke ([1927] 1988: 279) proposes a class of 'interjectives,' 'which conveys a complete concept without the implication of any subject' for Zulu and other Nguni Bantu languages of South Africa. As Table 5.10 shows, Doke's category of 'interjectives' includes six of the ten interactive types.

Finally, mention may be made of the Salishan language Thompson (Nlaka'pamuctsin) of British Columbia, for which a category of 'expletives'

**Table 5.10** Zulu ‘interjectives’ distinguished by [Doke \(\[1927\] 1988\)](#)

|                  |                                                                                                                                                                                                                               |
|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Attention signal | <i>we</i> : ‘hey!’                                                                                                                                                                                                            |
| Directives       | <i>he</i> :: ‘look out!’, <i>we</i> :: ‘take care!’, <i>Woza lapa!</i> ‘Come here!’                                                                                                                                           |
| Interjections    | <i>halala</i> ‘hurray!’, <i>hawu</i> ‘is it so?’ (of wonder, surprise), <i>hayi</i> ‘is that so?’ (of amazement), <i>ngcingci</i> (of delight, pleasure), <i>wewu, wou</i> ‘ah!’ (of wonder), <i>x</i> [lateral click] ‘ach!’ |
| Response signals | <i>ehé, yebo</i> ‘yes!’, <i>hayi, cha, qha</i> ‘no!’                                                                                                                                                                          |
| Social formulae  | <i>e</i> : (exclamation used in place of knocking at the door), ( <i>h</i> ) <i>a</i> , ( <i>h</i> ) <i>abé</i> ‘why did you do this?’, <i>nxephepha</i> ‘excuse me!’, I beg your pardon!’                                    |
| Vocatives        | <i>baba</i> ‘father!’, <i>bafana</i> ‘boys!’, <i>nkosi</i> ‘O chief!’                                                                                                                                                         |

Source: [Doke \(\[1927\] 1988: 279–83\)](#)

is proposed, which are characterized by their independent status. While they may appear as parts of longer sentences, ‘expletives’ are not integrated in the structure of a sentence and ‘they are regularly set off by intonational marking’ ([Thompson and Thompson 1992: 184](#)). The eight types of interactives classified as ‘expletives’ are illustrated in [Table 5.11](#).

**Table 5.11** Thompson ‘expletives’ distinguished by [Thompson and Thompson \(1992\)](#)

|                  |                                                                                                                                        |
|------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Attention signal | <i>héyè</i> ‘hey!’, ‘say!’                                                                                                             |
| Directives       | <i>mé</i> ‘give me (something)!’ (baby word), <i>né(h)</i> ‘here!, take this!, have some!’                                             |
| Discourse marker | <i>y’éý</i> (response of audience encouraging narrator to continue story)                                                              |
| Ideophones       | <i>ʔéʔ</i> ‘caw’ (crow’s call), <i>klák</i> (sound of grasshopper spitting)                                                            |
| Interjections    | <i>ʔené</i> ‘alas! oh, dear!’, <i>ʔú</i> ‘oh!, aha!’                                                                                   |
| Response signals | <i>ʔý, ʔéý</i> ‘yes’, <i>téʔ</i> ‘no’                                                                                                  |
| Social formulae  | <i>ní</i> ‘please, well, alright?’ (soliciting hearer’s good feeling, making courteous request), <i>húm’et</i> ‘goodbye, alright, yes’ |
| Vocatives        | <i>kík!</i> ‘mother!’, <i>sín!</i> ‘younger brother!’                                                                                  |

Source: [Thompson and Thompson \(1992: 184–6\)](#)

The question underlying the small survey just presented was: Are the authors concerned somehow aware that in the language studied by them there is some notion of an interactive category? It would seem that the answer is essentially in the affirmative, for the following reasons. Since they include between four and seven of the ten types in their category, most commonly referred to summarily as ‘interjections,’ there appears to be an awareness of some kind of category that is called here interactives. And there also is some convergence on which types are most central within the category, namely interjections, response signals, and social formulae. Perhaps interestingly, evaluatives are distinguished by none of the authors.

## 5.1.2 Inserts

Among all the studies on English grammar surveyed, that by Biber et al. (1999: 225) stands out to coming closest to the one presented here and, in fact, we have benefitted greatly from these authors, adopting part of their terminology and drawing on the rich data base presented by them. This study is therefore discussed here in greater detail.

In order to segment dialogues in spoken English, the authors propose a category of 'C-units,' that is, maximal (or independent) units of conversational syntax, 'in the sense that they cannot be *syntactically* integrated with the elements which precede or follow them.' These units are divided into 'clausal units' and 'non-clausal units.' The former are structures 'consisting of an independent clause together with any dependent clauses embedded within it,' that is, they correspond to what is called here 'sentence grammar.' Non-clausal units by contrast are segments consisting entirely or partly of non-clausal material (Biber et al. 1999: 1069–70).

Furthermore, the authors classify non-clausal units into two categories, namely syntactic non-clausal units and inserts. The former are 'segments which are neither clausal units nor part of clausal units, while inserts 'are stand-alone words which are characterized in general by their inability to enter into syntactic relations with other structures... They comprise a class of words that is peripheral, both in the grammar and in the lexicon of the language.'

The following piece of conversation taken from Biber et al. (1999: 1069–70) may illustrate the distinctions made. The text segments in (2) consist of two clausal units (marked in plain print), two syntactic non-clausal units (*my father's*, *your father's mother?*), and three inserts (*no*, *yeah*, *uh huh*).

- (2) English conversation (excerpt from Biber et al. 1999: 1070; A and B are women aged 30 and 85, respectively. '||' marks the boundary of a maximal unit, non-clausal units are in bold)<sup>1</sup>
- B: || So this was your mother's? ||
- A: || No, || **my father's**. ||
- B: || **Your father's mother?** ||
- A: || **Yeah**. || Her name was Martha <name> ||
- B: || **Uh huh**. ||

The concept 'interactive' proposed here significantly overlaps with that of 'inserts' of Biber and associates, who characterize inserts as behaving 'like atomic wholes' (Biber et al. 1999: 1083). Biber et al. (1999: 225) classify inserts as one type of 'non-clausal units' (p. 1069) or 'non-clausal material,' expressing a variety of

<sup>1</sup> Biber et al. (1999: 1069) say that '||' marks boundaries of clausal units, but they use '||' more broadly also for non-clausal units, that is, for what they call maximal units or C-units.

speech functions (statement, question, command, exclamation, etc.). One important group consists of inserts such as response words, interjections, greetings, apologies, and thanks. In addition, conversation frequently uses ordinary words and phrases which are not integrated in clauses (Biber et al. 1999: 1069). Inserts are characterized thus:

[They] are **stand-alone words** which are characterized in general by their inability to enter into syntactic relations with other structures... They comprise a class of words that is peripheral, both in the grammar and in the lexicon of the language.

(Biber et al. 1999: 1082; boldening in the original)

There are six defining features that apply to its more central members (Biber et al. 1999: 1082):

- (3) Features defining inserts (Biber et al. 1999: 1082)
  - a They may appear on their own, i.e., not as part of a larger grammatical structure.
  - b They may appear unattached.
  - c They rarely occur *medially* in a syntactic structure.
  - d They are morphologically simple.
  - e They are not homonyms of words in other word classes.
  - f They have no denotative meaning, their use is defined rather by their pragmatic function.

Biber et al. (1999: 1083–99) propose nine ‘major functional types’ of inserts, adding that these types have a role ‘both as separate non-clausal units and as peripheral parts of other units’ (p. 1083). These types are listed in Table 5.12 with examples given by the authors. As the overview in Table 5.12 suggests, both the classes and the labels used for them are similar to the ones proposed here (compare the first and the second columns). Differences are minor, except for the fact that we introduce four types that were not distinguished by these authors, namely directives, evaluatives, ideophones, and vocatives (see also Chapter 3, Table 3.2). While vocatives are not treated as inserts, they are said to be related to them. For example, like discourse markers, vocatives have a ‘discourse management’ function and ‘are generally loosely attached to the beginning or end of a clausal or non-clausal unit,’ and they ‘have not only an identifying but also an attitudinal function’ (Biber et al. 1999: 1047).

The minor differences that there are concern especially the following: First, we follow most other authors in using a wider concept of ‘interjections,’ which includes both primary and secondary interjections. Biber et al.’s (1999) type of interjection is restricted to primary interjections, whereas secondary interjections are treated as expletives. Second, we assume that greetings and farewells are also



‘polite speech-act formulae’ or, as we say here, social formulae. And third, we classify hesitators as belonging to discourse markers for reasons given in [Heine et al. \(2021: 260–3\)](#); see Section 3.3.4.

**Table 5.12** English inserts distinguished by [Biber et al. \(1999\)](#)

| Type proposed by <a href="#">Biber et al. (1999)</a> | Roughly corresponding type proposed here | Examples of inserts                                                                                                                                                                                   |
|------------------------------------------------------|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Attention signals                                    | Attention signals                        | <i>hey, say, yo</i>                                                                                                                                                                                   |
| Discourse markers                                    | Discourse markers                        | <i>I mean, now, right, oh, uh, um, well, you know, you see</i>                                                                                                                                        |
| Expletives                                           | Secondary interjections                  | <i>bloody hell, fuck, gee, geez, (my) God, fuck, goddammit, golly, good grief, Good Lord, gosh, heavens, heck, Jesus, my God, my gosh, shit</i>                                                       |
| Greetings and farewells                              | Social formulae                          | <i>(bye) bye, (good) morning, good night, goodbye, hello, hi, see you (later)</i>                                                                                                                     |
| Hesitators                                           | Discourse markers                        | <i>er er, erm, uh, um</i>                                                                                                                                                                             |
| Interjections                                        | Primary interjections                    | <i>aargh, ah, ha, oh, ooh, oops, ouch, tt, ugh, urgh, whoopee, whoops, wow, wowee, yippee, yuk</i>                                                                                                    |
| Polite speech-act formulae                           | Social formulae                          | <i>beg your pardon, bye, excuse me, congratulations, good afternoon, good evening, good luck, no problem, pardon me, please, see you later, sorry, thank you, thanks, that's okay, you're welcome</i> |
| Response elicitors                                   | Response elicitors                       | <i>eh?, huh?, isn't it?, okay?, right?, see?</i>                                                                                                                                                      |
| Response forms                                       | Response signals                         | <i>huh uh, I see, mm, no, nope, okay, really?, right, sure, uh huh, thank you, yeah, yep, yes</i>                                                                                                     |

Source: [Biber et al. \(1999: 1082–99\)](#)

One of the important observations made by [Biber et al. \(1999: 1083\)](#) is that the major functional types belonging to the category of inserts ‘can be versatile in taking on different conversational roles. For example, *oh*, as by far the most common of the **interjections** (A), shares also in the functions of **discourse markers** (C) and **response forms** (F). *Okay* can act as a discourse marker, response elicitor, and response form’ ([Biber et al. 1999: 1083](#); bold print in the original); these kinds of ‘sharing’ were discussed in Section 4.2 in terms of grammaticalization.

To conclude, there are substantial overlaps between the framework of [Biber et al. \(1999\)](#) and that proposed here, and this also applies to the overall classification of inserts and interactives, as Figure 5.1 shows (the categories of [Biber et al. 1999](#) are printed in bold, the corresponding categories of discourse grammar are in plain print; concerning ‘discourse grammar’ and ‘thetical grammar,’ see Section 5.2.3).

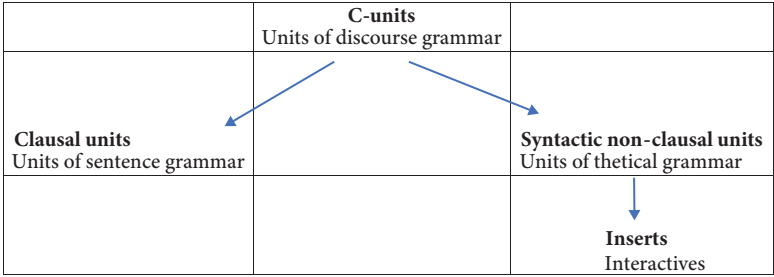


Fig. 5.1 Classification of the C-units and inserts of Biber et al. (1999)

5.1.3 Expressives

Some convergence with the present notion of an interactive category is also found in philosophical and formal-semantic research on expressives, as we saw in Section 2.1.2 (Kaplan 1999; see also Gutzmann and Henderson 2019).<sup>2</sup> Observing that the meaning of certain linguistic expressions cannot reasonably be described in terms of truth-conditional logic, they propose the concept of ‘use’ as an additional parameter of defining meaning. Thus, Kaplan (1999: 1) notes that ‘[t]here are also words that don’t seem to have a meaning, words like *goodbye* and the first person pronoun *I*... Still, they have a use...’ And he adds: ‘I don’t ask, ‘What does *goodbye* mean?’ Instead I ask, what are the conditions under which the expression would be correctly or accurately used?’ With the phrase ‘correctly or accurately used’ Kaplan highlights the central role of the situation where an expression is used, that is, what is called here the situation of discourse:

Now, I take it that the expressive (in this case, the word *goodbye*) has as its expressive content something like ‘You and I are now parting from one another for a significant period of time’...*Goodbye* seems perfectly acceptable to be used in that situation. So that’s the expressive content.

(Kaplan 1999: 16)

The expression *goodbye* is a social formula, but Kaplan extends his analysis also to interjections, like *ouch* and *oops*:

The semantic information in the word *ouch* is represented by the set of those contexts at which the word *ouch* is expressively correct, namely, the set of those contexts at which the agent is in pain. That set of contexts represents the semantic information contained in the word *ouch*.

(Kaplan 1999: 10)

<sup>2</sup> Kaplan (1999: 5) defines ‘expressives,’ somewhat vaguely, thus: ‘Let us call an expression an **expressive** if it expresses or displays something which either is or is not the case’ (boldening in the original).

Distinguishing between ‘a semantics of Meanings and a semantics of Use,’ Kaplan says that ‘an expression is descriptively correct if what it describes is the case,’ while it ‘is expressively correct if what it expresses or displays is the case’ (Kaplan 1999: 5). He pleads for ‘a correct Semantic Theory’ for certain expressions of natural language, and that theory ‘would state *rules of use* rather than something like a concept expressed’ (Kaplan 1999: 6).

To be sure, dependence on their use is by no means a feature that is restricted to the meaning of ‘expressives,’ that is, of interactives like interjections and social formulae; rather, it can also be observed in forms of sentence grammar. However, there are differences. For one thing, ‘rules of use’ are distinctly more pronounced in interactives than in forms of sentence grammar. For another, the latter are constrained by the syntactic roles they play as parts of phrases, clauses, and sentences, as was already pointed out by Kaplan (1999: 14).<sup>3</sup> Such constraints do not apply to interactives since they are not syntactic constituents of any kind, and rather than ‘describing’ meaning, they are said to ‘express,’ ‘display,’ or ‘depict’ meaning, as has been pointed out by some authors (Kaplan 1999: 5; Gutzmann and Henderson 2019: 120; see also Dingemanse 2011: 17; Haiman 2018: 57 on ideophones).

With this perspective, Kaplan proposes a sketch of the semantic space characterizing the grammar of interactives, namely that of ‘expressives’ like *ouch*, *oops*, and *goodbye*, which ‘display something about a state or attitude of the speaker’ and have use-conditional content. However, the range of expressives proposed does not coincide with that of interactives; rather, it is more extensive, also covering deictic forms (referred to by him as indexicals) like the first person pronoun *I*. The class of ‘expressives’ thus extends beyond that of interactives, and it also includes some forms of sentence grammar such as the evaluative adjective *damn* in (4).

- (4) English (Kaplan 1999: 5)  
That *damn* Kaplan was promoted.

The distinction between expressives, conveying use-conditional meaning, and linguistic forms conveying truth conditional meaning and between the two kinds of meaning content involved was taken up and developed further by Gutzmann and associates. They propose a framework called hybrid semantics which employs both use-conditions and truth-conditions, and where sentences do not only receive a truth value (true or false) but also a use value (felicitous or infelicitous) (Gutzmann 2015; Davis and Gutzmann 2015). For example, Gutzmann and Henderson (2019: 119) write that in examples like (4) above ‘*damn* is present

<sup>3</sup> Kaplan (1999: 14) notes, for example, that *ouch* and *I am in pain* are informationally equivalent but not logically equivalent, because one cannot derive *ouch* since it is not a sentence and syntax therefore blocks derivation.

only at the use-conditional layer and that the truth-conditional meaning is unaffected by it.’ Gutzmann and Henderson (2019: 107, 130) distinguish between truth-conditional meaning (‘X is true’) and use-conditional meaning (‘X is felicitously used’), and each of these meanings is associated with its own layer, level, or dimension. More specifically, use-conditional meaning content is viewed ‘as a set of the contexts in which the sentence is uttered felicitously’ (Gutzmann and Henderson 2019: 119).

While the interest of Kaplan (1999) is mainly with interactives, which are the topic of this book, that of Gutzmann and associates (see also Gutzmann 2015, 2019) is not centrally with interactives but rather with the meaning and syntax of sentence grammar structures. What all this research on the significance of use conditionality suggests, however, is that interactives are expressives in the sense of Kaplan (1999).

The present account converges with many of the points made in the tradition of Kaplan (1999), but there are also differences. First, interactives are ‘expressives’ to the extent that they conform to our definition in Section 1.2. This means, for example, that the English interjection *damn* is an interactive, in accordance with that definition. But this does not apply to the form *damn* in (4) above since it is a modifier of sentence grammar and hence is at variance with the definition. And second, we agree that interactives are expressives but argue that in addition to their expressive content they also have descriptive content, and that the latter content can be elucidated by means of paraphrase, as we saw in Section 2.1.

A somewhat different understanding of expressives can be found in Wharton (2003, 2009, 2016). Based on the framework of Relevance Theory, Wharton (2016: 21) proposes a category of expressives having the following properties:

- (a) They contribute to speaker meaning in a manner that is somehow independent from the utterance(s) in which they appear.
- (b) The kind of meaning they convey is ‘descriptively ineffable.’
- (c) Their meaning has ‘a flavour of the non-linguistic’ about it, it is often conveyed by non-linguistic means.

Wharton’s approach differs from the one proposed here, on the one hand, in that it is strictly semantic and, on the other hand, in that it contains a number of grammatically disparate units. First, most of the linguistic units discussed by him are theticals,<sup>4</sup> like the reporting clause *they say* or interjections, like *damn* or *ouch*. Second, there are also text pieces of sentence grammar (*I am cross*), noun phrases (*that bloody so-and-so*), or connectives (*but*). And third, they include also non-segmental units such as emotional intonational contour patterns.

Wharton proposes the interesting hypothesis that interjections ‘are partly natural and partly coded. They fall at various points along a continuum between

<sup>4</sup> Concerning the term ‘thetical,’ see Section 5.2.3.

display and language proper,' and they 'might move along the continuum from 'non-linguistic' to 'partly linguistic' to 'linguistic' without radically altering the type of information it conveys' (Wharton 2003, 2016: 29). More information would be needed on how such a continuum is structured, and in particular how exactly the non-linguistic pole and the intermediate stages of the continuum are to be defined. But as we saw in Section 2.1.3, this idea of a continuum, or a scale, is compatible with the model of representing information in discourse processing as sketched in Table 2.3, where three modes of representation were distinguished, namely physical depiction, vocal depiction, and description.

To conclude, research on expressives has come up with a wealth of relevant observations and, as was observed in Section 2.1, this book adopts Kaplan's distinction between descriptive and expressive meaning but argues that interactives are the combined product of both meanings. What distinguishes interactives from most other linguistic forms is their expressive meaning: Interactives are said to both express and depict meaning (Kaplan 1999: 5; Clark 2016). But more importantly, interactives, like gestures, differ from forms of sentence grammar in being immediately anchored to the situation of discourse rather than being edited and processed by the machinery of sentence grammar (see the direct-access hypothesis, Section 2.1.3.3).

#### 5.1.4 Discussion

In the preceding discussion, a number of approaches were presented that in some form or other highlight features of interactive grammar. The overview provided was restricted to a collection of such approaches.

In concluding, mention should also be made of the notion of 'functional performance structures' proposed by Noel Aziz Hanna and Sonnenhauser (2013, 2019), that is, 'syntagmatic categories with indexical function.' The focus of the authors is not only on vocatives as a type of interactives but also on other thetical categories in the sense of Heine et al. (2013) such as 'parentheticals, citations and delocutive particles.'

### 5.2 Dual-process frameworks

The main theme of the preceding chapters was to argue that sentence grammar and interactive grammar provide interlocutors with two contrasting resources for processing their discourses. This is not a new view, and in the present section, the distinction is related to previous discussions on this issue. To this end, Section 5.2.1 gives an overview of relevant research, while Sections 5.2.2 and 5.2.3 discuss two

approaches that are consistent with the analysis proposed here but provide a wider perspective of discourse analysis.

### 5.2.1 An overview

Dual-process frameworks have been discussed in detail in [Heine \(2019\)](#) and [Heine, Kuteva, and Long \(2020\)](#) and the reader is referred to those studies for more information. The following overview is restricted to a concise characterization of approaches that are immediately relevant to a better understanding of the nature of interactives. This means that a number of other dualistic distinctions that have been proposed are not considered here. Such distinctions include that between conceptual and procedural meaning ([Wilson and Sperber 1993](#); [Blakemore 2002](#); [Ifantidou-Trouki 1993](#); [Furkó 2014](#): 290–1), between clausal and extra-clausal constituents ([Dik 1997](#)), between a primary and a collateral track of performance ([Clark and Fox Tree 2002](#)), or between a representational and an interpersonal level of language structure ([Hengeveld and Mackenzie 2008](#)).

The order in which the frameworks are discussed below reflects, as far as this is feasible, the chronological order in which, to our knowledge, they first appeared in print.

An early study on dualistic processing is that by [Van Dijk \(1980\)](#), who distinguishes between a *microstructure* and a *macrostructure* of discourse. The former is about the structure of sentences and sequences, it concerns local information, such as the meanings of words, phrases, clauses, and simple actions: ‘microstructures are the actually and directly ‘expressed’ structures of discourse’ ([Van Dijk 1980](#): 29). Macrostructures are higher-level semantic and conceptual structures that organize the ‘local’ microstructures of discourse, interaction, and their cognitive processing. The underlying role played by macrostructures is ‘to establish global meanings and global coherence in a discourse.’ For [Van Dijk \(1980](#): 29), some means for expressing macrostructures are summaries, short paraphrases, and conclusions.

The relationship between microstructures and macrostructures is conceived by Van Dijk as one of derivation and inferencing via macrorules, which ‘derive macrostructures from microstructures’ ([Van Dijk 1980](#): 13, 46). While discourse markers and some other types of interactives provide paradigm tools for macrostructural tasks, they are not identified as such by [Van Dijk \(1980\)](#).

That interactives constitute a higher level of discourse processing than units of sentence grammar is argued for in some form or other in a number of studies. [Evans \(1992](#): 235) suggests, for example, with reference to interjections: ‘One might argue that we are dealing here with some sort of higher-level organization, parallelling at the level of discourse the sequential organization of words within the sentence.’ Similarly, based on a syntactic model

of language competence, Wiltchko (2021: 3) argues that interactional language is hierarchically structured: ‘interactional structure dominates propositional structure.’

A related observation can be found in the work by Nespoulous (1980; see also Nespoulous et al. 1998) on the distinction between *referential* vs. *modalizing speech*. Referential discourse conveys the expected content of a discourse task, while modalizing discourse is defined as discourse about discourse or a subject’s comments about their performance during a discourse task. Once again, this distinction captures features relating to that between sentence grammar and interactive grammar, but the particular role played by interactives in modalizing discourse is not addressed in that work.

Of more immediate interest to the present study is the distinction made by Maschler (1994, 2009) between *linguaging* and *metalinguaging*. Proposing what she calls the theory of metalanguage, Maschler (1994: 327) distinguishes the following two levels: (a) ‘linguaging,’ i.e., the process of using language, and (b) ‘metalinguaging,’ that is, using language in order to communicate about the process of using language. Maschler’s concepts are semantic-pragmatic, and her interest is with language as an ongoing process rather than an accomplished product (Maschler 1994: 326; 2009: 1). Every act of linguaging involves both levels: ‘at the same time that the speaker is linguaging about the world, he/she is also communicating information about how the utterance is intended; that is, the speaker is metalinguaging – using language to communicate information about linguaging’ (Maschler 1994: 327). But some utterances are particularly high in their metalingual dimension, and in bilingual situations, linguaging and metalinguaging can be accomplished each in a different language (Maschler 2009: 6).

An important element of metalanguaging is provided by discourse markers, which she views as ‘utterances occurring at verbal-activity boundaries which are metalingual at the level of discourse’ (Maschler 1994: 333).<sup>5</sup> She distinguishes between a frame, which is metalingual (or metacommunicative), and a message, which is framed by the frame. Accordingly, the ‘frame is the metamessage.’ It is the function of utterances in conveying information about how linguaging is organized at the discourse level which qualifies them as discourse markers.

While Maschler’s concern is with discourse markers, this term is used more generally for metalingual devices, which also include other types of interactives like interjections, directives, and response signals.

Rather than with discourse markers, interjections, directives, and response signals, Diffloth’s (1994) interest is with another type of interjectives, namely with ideophones, called ‘expressives’ by him. Observing that words such as adjectives

<sup>5</sup> Maschler (2009: 117) distinguishes between textual, interpersonal, and cognitive discourse markers.

and adverbs function in the ‘conventional semiotic domain,’ where having truth values and being negatable are fundamental characteristics, he concludes:

Expressives do not operate in that domain; instead, they represent an attempt at fully exploiting the semiotics of iconicity, in order to convey various sensations in as direct a manner as speech makes possible. Because of this, it is not sufficient to say that expressives form a distinct basic part of speech: they actually constitute a parallel sublanguage grafted on, and parasitic on, the conventional one.

(Diffloth 1994: 108)

A related distinction between two levels of discourse processing is also proposed in research on metadiscourse (e.g., Hyland 1998, 2005, 2017; Ädel 2006; Mauranen 2010; Zhang 2016), even if this research is largely confined to written discourse. For example, Hyland (1998) proposes a distinction between *propositional discourse* and *metadiscourse*. Whereas the former involves the presentation of texts as it is achieved typically in terms of sentence organization, the latter deals with ‘discourse about discourse,’ being defined by Hyland thus:

Metadiscourse refers to aspects of a text which explicitly organise the discourse, engage the audience and signal the writer’s attitude. [...] Its role in establishing and maintaining contact between the writer and the reader and between the writer and the message also makes it a central pragmatic concept.

(Hyland 1998: 437)

The significance of metadiscourse lies in its role of explicating a context for interpretation, ‘making linguistic choices which that audience will conventionally recognise as persuasive’ (Hyland 1998: 438). Metadiscourse involves a wide range of linguistic expressions, consisting of particles, words, phrases, clauses, as well as reduced phrases and clauses. Paradigm examples of metadiscourse units serving the organization of texts are what is called ‘phoric markers’ (e.g., *above, following, first*), ‘discourse labels’ (e.g., *aim, answer, speak*), and ‘code glosses’ (e.g., *namely, i.e.*) (see Ädel 2006: 98; Zhang 2016: 2005–6).

The concept of ‘metadiscourse’ is compatible with most of the other approaches on the macrostructure of discourse, relating the text to the situation of discourse, but its concern is largely restricted to text organization. Accordingly, other than discourse markers, functions and forms of interactive grammar are not of central interest in this research.

In the discourse analysis of Johnstone (2002: 64), the distinction between two structures is described as one between the sentence-level grammar and discourse-level grammar of a language. The former includes the set of conventions speakers employ for forming phrases, clauses, and sentences, whereas the latter concerns ‘utterances that are longer than a sentence – conversations, paragraphs, stories,



and arguments' (Johnstone 2002: 64). The two levels both influence and are influenced by the way in which the respective conventions are deployed in discourse. Like Van Dijk (1980) and Nespoulous (1980), however, the author is not centrally concerned with how her findings relate to interactives.

In some lines of psycholinguistic research, what one may wish to call a comprehension framework is proposed, highlighting the mental representations that readers create to understand a text. Two kinds of representation are distinguished (Kintsch 1988; Gernsbacher 1990; Graesser, Singer, and Trabasso 1994; Greene, McKoon, and Ratcliff 1992; McKoon and Ratcliff 1980, 1992, 1998; Prat, Long, and Baynes 2007), namely (a) a *propositional representation* (or textbase) and (b) a *discourse model* (or situation model).

The propositional representation is organized as a network of the explicit ideas or propositions derived from the text as well as the relations among them (Kintsch 1974). It is 'locally coherent,' in that incoming propositions are integrated with those propositions that are active in working memory. The relations among propositions are primarily referential (Kintsch 1974; McKoon and Ratcliff 1980). Propositional representation serves as a foundation for the discourse model (Graesser, Singer, and Trabasso 1994; Greene, McKoon, and Ratcliff 1992; McKoon and Ratcliff 1980, 1992, 1998).

The discourse model is a model of representation of what a text is about. Explicit text information is integrated with relevant world knowledge to reflect the important features of the real or the imaginary situation depicted in the text. To construct a discourse model, readers must engage in active inferential processing to interpret and organize text information in light of their prior understanding of the knowledge domain. The discourse model is 'globally coherent,' that is, ideas in the text are connected by means of some overarching theme (Prat, Long, and Baynes 2007: 283–4).

With the exception of Maschler (1994, 2009), the theoretical models sketched so far are not highly specific with regard to how the distinctions proposed precisely reflect discontinuities of linguistic categorization. This is different in the frameworks looked at now. One of these frameworks is proposed by Dwyer and Moshi (2003), who distinguish two semiotic dimensions, namely an analytic and an expressive dimension. The interactives studied by them (especially ideophones) belong to the expressive dimension, whereas the analytic dimension is the concern of sentence grammar.

A dual-process framework that is particularly relevant to the distinction between the two grammars proposed here is fuzzy-trace theory of memory and reasoning in psychology (e.g., Brainerd and Reyna 1992, 2002; Reyna and Brainerd 1995). Central to the theory, which is rooted in Gestalt psychology, is the distinction between two types of information or mental representation, referred to as *verbatim* and *gist* information. On this theory, verbatim and gist are encoded in two separate traces. *Verbatim* information concerns exact details of memory and

reasoning, such as distinguishing the individual parts of a pattern. Information encoded as gist is conceptual and concerns the pattern as a whole, as proposed in Gestalt psychology. Taking the printed English form *banana* as an example, verbatim features of this form can be said to be that it consists of six letters that are printed all in lowercase. A gist feature, in contrast, would be that the form refers to a type of fruit.

While fuzzy-trace theory has so far not been applied to grammatical or discourse material such as that examined in this book, salient elements of it can immediately be extended to the latter: Sentence grammar constitutes a mode of encoding verbatim information by means of freely combinable parts, whereas interactive grammar encodes gist information by means of conventionalized unanalyzable expressions, that is, by means of interactive forms. Take the English interjection *Wow!* that was discussed in Section 2.1.1: Like other interactives, it encodes gist information in a holistic format. Its sentence grammar paraphrase *I'm delighted!*, by contrast, represents one of the ways in which the conceptual content of the interjection can be rendered in an analytic format: It is composed of a number of individual morphemes in accordance with rules prescribed by the grammar of English.

Another framework focusing on linguistic phenomena is that of [Pawley \(2009\)](#). Distinguishing between *analytic* and *holistic processing*, Pawley summarizes his dualistic framework thus:

It appears that competent speakers of a language know many linguistic entities in two ways: holistically and analytically, and can move between the two. People are good at generalising, at perceiving patterns, and the generalising capacity is essential to the learning of general rules. On the other hand, people have severely limited rapid processing capacity but they have an enormous memory, which allows them to store and retrieve, or recognize familiar complex form-meaning pairings. Thus, a realistic account of the cognitive processes that underpin nativelike command of a language should accommodate this kind of dual knowledge.

([Pawley 2009](#): 21)

It is above all what [Pawley \(2009\)](#) describes as formulaic language and, more specifically, as speech act formulas (SAFs) that reflect holistic knowledge. In his model of linguistic competence there is 'a continuum of more or less lexicalized constructions, with their own 'grammar,' lying between the extremes of abstract constructions and unanalysable lexical units.' Speech act formulas exhibit significant correlates with interactives, namely the following:

- (a) Like interactives, SAFs are defined as social institutions, they serve discourse functions over and above those of referring and predicating.

- (b) Like interactives, each of the ‘productive (speech) formulas’ proposed by Pawley (2009: 20) ‘has its own mini-grammar’.
- (c) Both concern ‘prototypical discourse functions such as greeting, welcoming, farewelling, introducing, complimenting, insulting, apologising, and so on.’ Thus, SAFs involve in particular social formulae like *Hello*, *Sorry*, *Congratulations*, *Excuse me*, and *Yours sincerely*, but they also include many other formulaic expressions (Pawley 2009: 6).
- (d) SAFs are by definition formulaic, and so are interactives.
- (e) Like SAFs, interactives tend to be associated with a certain intonation and prosodic pattern (Pawley 2009: 7).

SAFs and interactives overlap significantly; nevertheless, they need to be distinguished, for the following reasons. First, SAFs are mostly but not always interactives; they also include a range of other expressions, such as *You can say that again*, *I wouldn’t dream of it*, or *That’s another story*. Such formulas do not conform to our definition in Section 1.2. For example, being syntactically unattached, interactives cannot be embedded—the social formula *Hello!* in (5a) is both an interactive and an SAF in the sense of Pawley (2009). But other kinds of SAFs, like *That’s another story*, are not interactives: They exhibit the propositional structure of sentence grammar and, hence, can be embedded. Thus, whereas *That’s another story* can easily be subordinated as a complement clause in (5a), this is not possible with interactives, like *Hello!* in (5b).

- (5) English (cf. Pawley 2009)
  - a John believes that that’s another story.
  - b #John believes that **Hello!**

To conclude, SAFs consist of a wide range of pre-fabricated expressions; Pawley (2009: 8) estimates that ‘the ordinary mature native speaker of English’ knows thousands of them. And they also include interactives, but the latter constitute only a fraction of them, namely those that conform to the definition in Section 1.2.

SAFs are classified essentially on the basis of whether or not they exhibit formulaic behavior, or idiomaticity constraints, even if there are also other features in addition (see Pawley 2009: 6–8). Note that Pawley’s (2009) term ‘speech act formula’ is not exactly the same as the ‘speech formula’ of Bannard and Lieven (2009: 299), defined as ‘a multiword piece of a language that occurs a lot.’ But what was just observed on SAFs also applies to the ‘speech formulas’ of Bannard and Lieven (2009: 299): The latter include a number of interactives but are not restricted to them, not being compatible with our definition of interactives.

Pawley (2009) was not the first to propose the term ‘speech act formula.’ For example, the term was used earlier by Biber et al. (1999: 1093–4) in the form ‘polite speech act formula.’ But unlike Pawley’s (2009) use of the term, that by Biber et al.

(1999) is exclusively reserved for interactives, being restricted more specifically to social formulae for thanking, apologizing, requesting, and congratulating (see Section 5.1.2).

An account resembling that of Pawley (2009) can be found in the work of Van Lancker Sidtis and associates (e.g., Van Lancker Sidtis 2004, 2009). Using a neurolinguistic perspective, she proposes what she calls a dual-process model of speech. Central to the model is the distinction between *novel speech* and *formulaic speech* or automatic speech, also referred to as a distinction between propositional and non-propositional speech.<sup>6</sup> With reference to the processing of information, Van Lancker Sidtis describes the difference as one between two modes, where each mode places different processing demands on speech production and comprehension. Whereas propositional expressions require lexical retrieval and arrangement according to grammatical rules, non-propositional production involves activating and retrieving prepackaged units or schemata<sup>7</sup> (Van Lancker Sidtis 2004: 29).

This dual-process model features an analytic mode for the generation of novel speech and a holistic mode for processing formulaic speech. The distinction is one between newly created or propositional language and fixed or non-propositional language—a distinction that somehow echoes the one made here between sentence grammar and interactive grammar. Unlike the former, the latter is not newly created from the operation of grammatical rules on lexical items (Van Lancker Sidtis 2009: 445).

Novel speech is concerned with fully productive language use—that is, it corresponds largely to the concept of sentence grammar proposed here. Formulaic speech has the characteristics listed in (6), easily identified instances of it being (a) interjections, (b) expletives, (c) pause fillers (*uh, um*), (d) discourse elements (*well, so*), (e) nonliteral lexical meanings of idioms (*He was at the end of his rope*), and (f) proverbs. Except for (e) and (f), all these groups consist of interactives.

- (6) Characteristics of instances of formulaic speech (Van Lancker Sidtis 2004, 2009)
  - (a) They have a stereotyped form, they are fixed and unitary
  - (b) They have a set intonation contour
  - (c) They have conventionalized meaning which is complex and usually nonliteral, rife with nuance and connotations, and depends in special ways on social context
  - (d) People know them intuitively

<sup>6</sup> A detailed account of this model is found in various publications of Van Lancker Sidtis and associates (Van Lancker 1988, 1990, 1997; Van Lancker Sidtis 2004, 2009; Van Lancker Sidtis and Postman 2006; Sidtis, Canterucci, and Katsnelson 2009).

<sup>7</sup> By 'schemata,' Van Lancker Sidtis (2004: 28) understands fixed forms with at least one open lexical slot.

There is substantial neurological evidence to suggest that novel speech is represented in the left hemisphere of the human brain whereas formulaic speech is facilitated by a subcortical right hemisphere circuit. For example, right hemisphere and/or subcortical damage leads to a selective impairment of formulaic language, while left hemisphere damage leads to a selective impairment of novel language with relative preservation of formulaic language (Van Lancker Sidtis 2009: 460). We will return to this issue in Section 5.3.

### 5.2.2 Microgrammar vs. macrogrammar

Proposing a wider conception of grammar, Haselow (2013, 2016a, 2016b, 2017) argues that linguistic processing and linguistic activity exhibit a dualistic organization, which is reflected in linguistic discourse. He distinguishes two domains of grammar or levels of grammatical relations, called *microgrammar* and *macrogrammar*. The former deals with internal linearization and hierarchization and the latter with the organization of language based on cognitive, discourse-structural, and interactive principles of language use outside hierarchically organized structures (Haselow 2017: 30). While microgrammar refers to sentence-internal structural patterns, macrogrammar refers to grammatical patterns across sentences and larger text units and serves to establish information structure and textual coherence (Haselow 2013: 386).

The two kinds of grammar use different serialization principles (Haselow 2016a: 81; 2016b: 386): In microgrammar, each constituent is morphosyntactically related to at least one other constituent in hierarchical configurations (Haselow 2016a: 82). Macrogrammar, by contrast, is based on serialization principles that rest upon (i) speech planning, (ii) processibility, (iii) textual coherence and (iv) contextual embeddedness, and macrogrammatical elements ‘do not change propositional content’ (Haselow 2016a: 82).

The English final particles *actually*, *anyway*, *but*, *even*, *so*, *then*, and *though* are classified by Haselow as elements of the macrogrammatical level. What distinguishes these particles from those of microgrammar is that (a) they may create a link between a discourse unit and an implied proposition, (b) they express ‘subjective meaning’ and may modify illocutionary force, (c) rather than categorical values such as person or case, they express conversation-actional values such as agreement, disagreement, concession, correction, surprise, etc., (d) they link two independent subsequent units of spoken discourse, they are syntactically peripheral, and (e) they have low key intonation (Haselow 2013: 378, 382, 412). The final particles discussed by Haselow (2016a, 2016b) are discourse markers, that is, they belong to one of the types of interactives, discussed in Section 3.3. But in Haselow (2017) more types are included.

Haselow's characterization of the two kinds of grammar is in fact compatible with the distinction made here, in that microgrammar relates to sentence grammar while interactive grammar is included in macrogrammar, as shown in (7).

- (7) Macrogrammar vs. microgrammar and sentence grammar vs. interactive grammar
  - a Both domains of grammar are constantly available and provide interlocutors with alternative options of discourse processing.
  - b Microgrammar corresponds to the domain of sentence grammar as proposed here, in that microgrammatical forms are characterized by syntactic integration and hierarchical relationship. Interactives, such as discourse markers, interjections, response elicitors (tags), and vocatives, in contrast, belong to macrogrammar.
  - c The distinction between the two domains exhibits significant correlations with lateralized brain activity in that microgrammar is strongly associated with the left hemisphere, whereas macrogrammar relates more strongly to right-hemisphere activity (see Section 5.3).

Haselow's (2017) approach is based on findings from grammatical analysis, conversation analysis, cognitive psychology, and neurolinguistics. It has a broader perspective than the one used here since social communication as it surfaces in interactive grammar concerns only one part of discourse processing in spontaneous spoken English. And much the same applies to the approach sketched in Section 5.2.3, where interactive grammar is also included within a wider domain of discourse processing.

### 5.2.3 Sentence grammar vs. thetical grammar

The concept of interactives is based on the framework of discourse grammar (Kaltenböck, Heine, and Kuteva 2011; Heine et al. 2013). Central to the framework is the distinction between sentence grammar and thetical grammar. Sentence grammar roughly corresponds to what is referred to in Haiman's (2018) study of ideophones as 'prosaic language' or 'propositional language.' It is organized in terms of propositional concepts and clauses and their combination—in the sense of Biber et al. (1999; 1069), it consists essentially of 'clausal units,' that is, clauses together with any dependent clauses embedded within them (see Section 5.1.2, Figure 5.1). It has been the only, or the main subject of mainstream theories of linguistics. The concern of thetical grammar is with theticals, that is, with linguistic discourse units which have metatextual functions and are immediately anchored to the situation of discourse.

While separate in principle, sentence grammar and thetical grammar interact in multiple ways in shaping linguistic discourse. The main way of interaction is via

cooptation, an operation whereby chunks of sentence grammar such as clauses, phrases, words, or any other units are deployed for use in thetical grammar (Heine et al. 2017; see Section 4.1).

Interactives are not an isolated class of metatextual expressions; rather, they belong to the large class of theticals, or what in other research traditions is called parentheticals (e.g., Reinhart 1983; Dehé and Kavalova 2006, 2007; Blakemore 2007; Kaltenböck 2007; Schneider 2007a, 2007b; Dehé and Wichmann 2010; Dehé 2014). Interactives thus are theticals, where the latter are defined as follows:

Theticals are coopted units whose functions are determined by the situation of discourse, serving (a) to overcome constraints imposed by linearization in structuring texts, (b) to package text segments, (c) to provide the source of information, (d) to place a text in a wider perspective, e.g., by elaborating, providing an explanation, a comment or supplementary information, (e) to describe the attitudes of the speaker, and/or (f) to interact with the hearer.

(Heine et al. 2021: 72)

Theticals are characterized by a set of grammatical properties, namely the ones in (8) (Heine et al. 2020).<sup>8</sup>

- (8) Grammatical properties characterizing theticals
  - a Meaning: Their meaning is not part of the meaning of the sentence they are associated with
  - b Function: Their function is metatextual
  - c Syntax: They are not a syntactic constituent of the sentence in which they occur
  - d Prosody: They are likely to be set off prosodically from the rest of the sentence
  - e Semantic-pragmatic scope: They have scope beyond the sentence
  - f Placement: Depending on their discourse function they are flexible in their placement behavior

Theticals have been subclassified in a number of ways (Kaltenböck, Heine, and Kuteva 2011; Heine et al. 2013). For the purposes of the present study, it is most of all a classification based on the constructional format of theticals that is of interest. With regard to their constructional format, three main types of theticals are distinguished (Heine, Kaltenböck, and Kuteva 2016: 188; Heine et al. 2017: 819–21), namely instantaneous, constructional, and formulaic theticals. *Instantaneous theticals* are fully compositional and can be designed spontaneously any time and anywhere, and quite a few are nonce formations being uttered only once and never

<sup>8</sup> Most of the properties listed also figure in discussions of parentheticals. But there is a wide range of notions associated with the term ‘parenthetical,’ each highlighting specific manifestations of theticals.

again. Thus, the following example of an instantaneous thetical, printed in bold, has the appearance of a freely composed clause—one that may never have been used before in this form and, conceivably, will never be produced again:

- (9) English (DCPSE: DL-E02, #114; [Kavalova 2007](#): 148)

Mr Heath's government **and I'm not complaining because I'd advocated this at a previous time** introduced the threshold system ...

*Constructional theticals* are recurrent patterns or constructions of theticals having a schematic format and function. Accounts of a number of constructional theticals of spoken English are found in [Kaltenböck \(2007\)](#) and [Dehé \(2014: 3–6\)](#);<sup>9</sup> for another catalog see [Kaltenböck, Heine, and Kuteva \(2011: 852–3, Table 1\)](#). In the terminology of Construction Grammar ([Goldberg 1995, 2006; Croft 2001](#)), they include both fully schematic and partially substantive constructions. An example of fully schematic theticals is provided, for example, by comment clauses, like *I believe* in (10) (see [Brinton 2008](#)).

- (10) English (ICE-GB: s1b-035-28)

And the Labour Party **I believe** want sanctions to work

*Formulaic theticals* are frozen chunks used in stereotyped communication as conversational routines. They are invariable units, and they correspond essentially to the category of interactives (see [Kaltenböck, Heine, and Kuteva 2011: 870–2; Heine et al. 2013; Heine, Kuteva, and Long 2020: 68](#)). For example, [Heine et al. \(2013\)](#) provide the list in Table 5.13 exemplifying formulaic theticals.

In short, all interactives are theticals but not all theticals are interactives—it is only formulaic theticals that are.

**Table 5.13** English examples of formulaic theticals

| Formulaic theticals         | English examples                                       | Corresponding type of interactive |
|-----------------------------|--------------------------------------------------------|-----------------------------------|
| Conceptual theticals        | <i>as it were, for example, if at all, if you will</i> | Discourse markers                 |
| Formulae of social exchange | <i>good morning, hello, please, thank you</i>          | Social formulae                   |
| Vocatives                   | <i>Sir!, Waiter!</i>                                   | Vocatives                         |
| Imperatives                 | <i>Come on!, Piss off!</i>                             | Directives                        |
| Interjections               | <i>boy, damn, fuck, hell</i>                           | Interjections                     |

Source: [Heine et al. \(2013: 178\)](#)

<sup>9</sup> The types of constructional theticals treated by [Dehé \(2014: 44–6\)](#) in detail are non-restrictive (appositive) relative clauses, nominal appositions, comment clauses, reporting clauses, question tags, *and*-parentheticals, and *as*-parentheticals. The thetical in (9), printed in bold, is an instance of an *and*-parenthetical.



A notion related to discourse grammar is proposed by [Noel and Sommerhauser \(2019\)](#), making a distinction between two layers of syntax, namely a ‘primary’ or ‘basic’ layer and a ‘secondary,’ ‘supplementary,’ or ‘additional’ layer of syntax. The former layer concerns what is called here sentence grammar while the latter concerns the syntax of pragmatics and includes parenthetical categories and vocatives, that is, structures belonging to thetical grammar in the framework of discourse grammar ([Heine et al. 2013](#)). The framework of [Noel and Sommerhauser \(2019\)](#) differs from the one proposed here in that dualism is framed in terms of two kinds of syntax while it is based here on the distinction between syntax in the case of sentence grammar and principles of discourse organization in the case of interactive grammar.

### 5.2.4 Propositional structure vs. interactional structure

Although from a different perspective, dual-process frameworks have also figured in work based on generative grammar, especially in works dealing with the interface between conversational pragmatics and syntax and the ‘syntacticization of speech acts’ ([Munaro and Poletto 2004](#); [Munaro 2010](#); [Haegeman and Hill 2013](#); [Wiltschko 2021](#): 200).<sup>10</sup> The most pronounced and elaborated study in that tradition is [Wiltschko \(2021\)](#), which is couched in a formal syntactic theory and contains a detailed review and analysis of other work presented in formal linguistics and other research traditions. Wiltschko’s concern is also with interactives, her goal is to develop a syntactic analysis for a specific subset of interactives, namely confirmationals such as utterance-final *eh?*, *huh?* eliciting a reacting move, that is, a subset of response elicitors, and response markers such as *no*, *yeah*, and *yup*, called here response signals.

Ultimately building on the performative hypothesis of [Ross \(1970\)](#) and [Sadock \(1969\)](#),<sup>11</sup> [Wiltschko \(2021\)](#) distinguishes two main syntactic levels, called the interactional and the propositional level, where the former encodes illocutionary force and contextual variables and consists of a responding and a grounding layer or function. Responding is located above grounding; hence the organization can be depicted as in (11), referred to by the author as the Interactional Spine Hypothesis.

- (11) The Interactional Spine Hypothesis of [Wiltschko \(2021](#): 3, (3))
- |          |                         |            |
|----------|-------------------------|------------|
| Level 1a | Interactional structure | Responding |
| Level 1b |                         | Grounding  |
| Level 2  | Propositional structure |            |

<sup>10</sup> We are grateful to an anonymous reviewer of Oxford University Press for having drawn our attention to this work.

<sup>11</sup> For example, [Ross \(1970\)](#) argues that every clause is dominated by a speech act structure, in that even when we say things, we do things: we tell others about the world.

On this hypothesis, interactional language is hierarchically structured: ‘interactional structure dominates propositional structure’ (Wiltschko 2021: 3). The highest level consists of the layer of responding, whose core function it is ‘to manage the moves that serve to synchronize the interlocutors’ knowledge states,’ enabling the speaker to configure the propositional content of the utterance so that the addressees can update their knowledge state to include that content (Wiltschko 2021: 72). The level or layer of grounding serves to manage the common ground between the interlocutors, which plays an important role in the synchronization of minds. Adding the interactional layer above the propositional layer is motivated by the fact that utterances may contain linguistic objects that refer to the proposition expressed in the utterance.

The present section is restricted to examining how the notion ‘interactional language’ proposed by Wiltschko (2021) relates to dual-processing as conceived here. Her model in fact qualifies as a dual-process framework to the extent that it also distinguishes between kinds of structure. The following remark echoes what we argue in the present study:

There are thus two aspects of language: the role it plays in the thoughts about the world and the role it plays in the communication of those thoughts. This dual function sets up a dichotomy that has pervaded the study of language; it has defined different research agendas and methodologies. There are those traditions that take as their object of study the form of language used in the expression of thought, and there are those that take it to be its communicative function.

(Wiltschko 2021: 1)

In fact, the two structures distinguished by her exhibit significant parallels to the distinction between sentence grammar and interactive grammar made here. ‘Interactional language’ falls squarely within what is called here ‘interactive grammar’—her conclusion that ‘just as propositional structure is required to talk about the world, so is interactional structure required to talk to people’ (Wiltschko 2021: 199) is in full accordance with the corresponding distinction proposed here.

There are further similarities between Wiltschko’s and our framework. For example, she assumes ‘that we can use the same model for both propositional and interactional language,’ and that ‘the language of interaction makes use of similar building blocks as the propositional grammar it embeds’ (Wiltschko 2021: 8, 75). As we saw in Section 1.4, our framework is also based on the assumption that in spite of all the differences separating sentence grammar from interactive grammar, there is a common base shared by the two grammars, namely what was described in Section 5.2.3 as discourse grammar. In Wiltschko’s model, by contrast, the common base is provided by syntax, in ‘that the same module, namely syntax, is responsible for the construction of sentences and the construction of linguistic interaction’ (Wiltschko 2021: 71). And we also agree that backchannels can

only be understood outside the sentential domain (Wiltschko 2021: 46) but argue that this applies not only to backchannels but to interactives in general.

Differences to the perspective adopted here relate, on the one hand, to the fact that her interest is mainly with language competence whereas we are concerned with how interlocutors structure their linguistic discourses on the basis of the linguistic resources available to them. Furthermore, Wiltschko's (2021: 73) concern is with turn-sequencing characterized by adjacency pairs, consisting minimally of an initiating move and a reacting move. Turn sequencing and adjacency pairs are also important in the present study (see, e.g., Section 7.3.2) but are not immediately relevant to all types of interactives. For example, interactives like ideophones and interjections do not typically presuppose any turn sequencing or adjacency pairs.

And finally, differences relate, in particular to the following points. Both response elicitors and response signals belong to the types of interactives that exhibit the highest amount of interference with sentence grammar, for the following reason: They take a piece of text as their discourse argument (T), both expressing roughly that the speaker (S) addresses the hearer (H) with reference to some text piece (T). Hence, both are amenable to an analysis in terms of relations obtaining between pieces of text (Sections 3.7, 3.8). This is in fact a major issue in Wiltschko's (2021) study, and she is able to establish that there are noteworthy constraints on combining, that is, on which kinds of T arguments response elicitors ('confirmational', e.g., *eh?*, *huh?*) and response signals (e.g., *yes*, *no*) may or may not take.

Problems that we have with an analysis in terms of syntax are most of all the following. First, how would a syntactic analysis cope with types of interactives not taking a T argument, that is, interjections (e.g., *ouch*, *wow*) or many social formulae (e.g., *hi*, *good morning*), that is, expressions which frequently occur on their own without involving any material of sentence grammar? It would seem that such interactives typically do not involve any propositions or propositional content. In some studies of formal linguistics (especially Munaro 2010) though, it is argued that rather than by linguistic material, information relating to propositional content can also be provided by the extra-linguistic context.

Second, and relatedly, T arguments—as conceived here—are not syntactic but discourse arguments, hence, their functions may extend beyond the variables identified by Wiltschko (2021) in her Interactional Spine Hypothesis, that is, responding, grounding, and propositional structure: How can interactives, such as response elicitors and response signals, be accounted for with reference to the overall structure of discourse far beyond the confines of propositional structures?

Third, another issue associated with a formal framework like that of Wiltschko (2021) relates to explanation. As we saw in Chapter 4, many 'peculiar' usages of interactives can be accounted for with reference to mechanisms such as cooption and grammaticalization. For example, response elicitors ('confirmational')

may in some of their usages function as discourse markers, or interjections as response signals (Section 4.2). There do not seem to be appropriate explanatory parameters available for such usages on the basis of the Interactional Spine Hypothesis.

Wiltschko (2021: 103) notes that there are units of language ‘that are attitudinal only and do not involve a response layer,’ suggesting that not all three levels of the Interactional Spine Hypothesis are compulsory ones. The question then is: Why do response elicitors in some of their uses not involve a response layer? For example, we observed in Sections 3.7.4 that there are uses of response elicitors, like English *right?*, *eh?* or question tags, where the speaker does not expect, nor even welcome a response, that is, where apparently no response layer is involved. A reason for this behavior was proffered in Section 3.7.4 and, more generally, in Section 4.2: Like other types of interactives, response elicitors exhibit the structure of grammaticalization scales or chains, where at one end of the scale there is a prototypical function of response elicitor, namely that of eliciting a response. At the other end of the scale there are highly grammaticalized instances of the same form where a response elicitor no longer is a ‘response elicitor,’ that is, where its primary function no longer relates to speaker–hearer interaction but rather to the organization of texts, that is, where no response is expected. In other words, in order to describe and understand the functions expressed by interactives, such as response elicitors, knowledge of both the whole range of its uses and the relative degree of its grammaticalization would seem to be required.

And finally, it was argued in Section 2.1 that interactives differ from units of sentence grammar in being accessed by interlocutors directly from the situation of discourse (cf. the direct-access hypothesis, Section 2.1.3.3). This accounts for their definitional properties presented in Section 1.2, such as the fact that they are indexicals that are syntactically detached, prosodically set off, etc. Empirical evidence in support of these definitional properties from languages across the globe was provided in Section 1.3. The question then is how such properties can be explained on the basis of the Interactional Spine Hypothesis.

To be sure, Wiltschko’s analysis pushes our understanding of the constraints underlying the usage of specific interactives far beyond what could be achieved in the present book. Her seminal analysis of grounding, in particular, is relevant for a comprehensive understanding not only of response elicitors and response signals but of interactives in general and thus provides an account that complements the one proffered here.

At the same time we suspect that extending the concept ‘syntax’ to principles underlying the organization of discourse at large might run the risk of turning that concept into a somewhat vacuous notion. Interlocutors are constantly confronted with deciding on when, where, and how to draw on the potential provided by interactive grammar as opposed to that of sentence grammar in order to design their discourses in the best way possible. This affects the preferred and dispreferred

options on usage, placement, and combining, and in a few cases the choices made are conventionalized to the extent that they acquire the character of constraints. But reducing most or all of that behavior to the template of syntax may not be very helpful for understanding how linguistic texts are built.

### 5.3 Neurolinguistic correlates

Of all types of interactives, it is especially interjections that have been associated with distinct neural activity. This applies most of all to expletive interjections like *Fuck!*, *Shit!*, *Damn!*, and *Bugger!*, which are commonly viewed as paradigm instances of swearing (Carter and McCarthy 2006: 226; Ljung 2011, chapter 4; see also the contributions in Nassenstein and Storch 2020). That swearing is strongly associated with right-hemisphere activity has by now become almost a truism among some linguists, as voiced in the following remark by Pinker (2007):

It's not that the right hemisphere contains a profanity module, but that its linguistic abilities are confined to memorized formulas rather than rule-governed combinations.

(Pinker 2007: 334)

Furthermore, swearing has been demonstrated to produce a hypoalgesic effect, affecting pain tolerance, pain perception, and heart rate, and nullifying the link between fear of pain and pain perception (Stephens, Atkins, and Kingston 2009).

But it is not only interjections that are of neurolinguistic interest. There is by now substantial evidence to show that the distinction between sentence grammar and thetical grammar is sensitive to brain activity, and more specifically to brain lateralization (e.g., Haselow 2021; Heine 1919; Heine et al. 2019; Heine, Kuteva, and Kaltenböck 2014; Heine, Kuteva, and Long 2020). And in Heine 2019, the following generalization is proposed, phrased in terms of the distinction microstructure vs. macrostructure, where the former concerns sentence grammar and the latter includes interactive grammar:

- (12) Whereas the microstructure implicates mainly the left hemisphere, building a macrostructure of discourse is a task that cannot be achieved without participation of the right hemisphere.

Neurolinguistic research establishing that the two halves of the human cortex each performs different functions can be traced back to the nineteenth century, and this research also concerns speech, or language processing. Both hemispheres are needed in normal communicative settings. At the same time, it is fairly uncontroversial that the left hemisphere (LH) is 'language-dominant,' playing an essential role in processing speech phenomena, and not only spoken or written but also

signed sentence input (Sakai et al. 2005). The advantage that the left hemisphere offers for most language tasks is one of the best-established facts about the human brain (Jung-Beeman 2005), and ‘speech functions’ are lateralized in the left hemisphere in most adults regardless of hand preference (Branch, Milner, and Rasmussen 1964).

Research in fields such as neuropsychology, neuroimaging, and neuroanatomy has brought about a more differentiated perspective of brain lateralization and its contribution to language-related behavior, especially with reference to the contribution of the right hemisphere to specific linguistic phenomena. Under certain conditions, language participation of the right hemisphere (RH) increases significantly: Lesion studies of stroke patients, work on dichotic listening, and functional imaging studies of healthy persons show that some language functions are mediated, or are mediated more by the right rather than the left hemisphere.

As demonstrated by Van Lancker Sidtis (e.g., 2004, 2009), her dual-process model of speech, summarized in Section 5.2, is supported by neurolinguistic findings on brain lateralization. Since aphasia following right hemisphere damage is extremely rare, aphasia is almost exclusively associated with left hemisphere damage and, in fact, precursors of the dual-process model and brain lateralization can already be found in research on aphasia in the nineteenth century (see Van Lancker Sidtis 2004, 2009: 460 for detailed discussion).

Building on the work of Van Lancker Sidtis (e.g., 2004, 2009) and associates, Heine, Kuteva, and Kaltenböck (2014) and Heine et al. (2015) argue that activity in the right hemisphere of the human brain involves more strongly thetical grammar than sentence grammar, and support for this finding has been found in Haselow’s (2021) research on discourse markers. Comparing the speech production of unilaterally brain-damaged persons with that of neurologically unimpaired speakers, Haselow concludes with reference to the use of discourse markers (DMs) and the way speech production is affected in speakers with left-hemisphere damage (LHD) and right-hemisphere damage (RHD):

The results show significant differences in the use of DMs between speakers with LHD, [who] do not differ significantly from unimpaired control speakers, and those with RHD. The empirical observation that macrostructuration – as far as the use of DMs is concerned – is the domain in which linguistic skills appear to be largely preserved with LHD speakers, whereas a higher degree of impairment can be observed with RHD speakers, is congruent with the large body of evidence for hemispheric differences in language processing. The findings point to a division of labor between sentence-structuration, or *microgrammar*, which is associated with the LH, and discourse-structuration, or *macrogrammar*, which strongly involves RH activity.

(Haselow 2021: 185; italics in the original)

And, more generally, there is also evidence to show that right hemisphere activity is also strongly associated with types of interactives other than discourse markers, as the following example may illustrate. The example is about the speech behavior of persons suffering left hemisphere damage, and in particular of aphasic persons. While aphasia may in rare cases be associated with right hemisphere damage, it is almost exclusively a phenomenon of left hemisphere dysfunction in the distribution of the middle cerebral artery. Extending over most of each hemisphere, this artery excludes a narrow strip on the anterior frontal lobe and another narrow area on the posterior parietal lobe (Van Lancker Sidtis 2009). Accordingly, the term 'aphasic speech' as used here refers exclusively to cases of aphasia caused by left hemisphere damage.

The text piece in (13) is taken from spontaneous speech by an aphasic subject who had recovered some speech from early non-fluency after five treatment sessions. The text piece has already been analyzed by Van Lancker Sidtis (2009) with reference to her distinction between novel speech and formulaic speech (Section 5.2) and by Heine, Kuteva, and Kaltenböck (2014) with reference to the distinction sentence grammar vs. thetical grammar. The goal of the present analysis is restricted to finding out what the contribution of interactives is in structuring aphasic discourse like that in the text piece of (13). In this text piece, interactives are written in bold italics.<sup>12</sup>

- (13) English aphasic speech sample (Van Lancker Sidtis, 2009: 455, Table 1)
- Uh.. uh good morning.. uh.. um .. me uh I want a.. big big ter //uh television, alright? Um, big. Alright ? And uh.. money? Yes. Fine..um.. big and . uh ... small um.. TV . yes.. uh small um.. Uh.. sky and cricket and .. uh soccer and movies and news and.. alright? Um.. right. Uh. Where? Ah! Alright! Boah! nice! Wow! Big! And small! Ho-ho, Jesus! Uh.. price? What? two thousand.. oh Jesus! hm.. wait. um.. hm hm hm. yes. alright.. maybe uh. two thousand? Oh, Jesus. Alright. Uh phone and wait, alright? Uh.. oh, Jesus! Hi! Jane um.. phew.. uh what is the matter? Money? Oh, Jesus.. alright.. alright! thank you! see you! Uh salesman .. uh.. money, yes.. fine..*

Table 5.14 shows that in the speech of the left-hemisphere damaged aphasic patient concerned, there are altogether 116 information units in the text piece of (13), of which 39 are classified as belonging to sentence grammar and 77 to

<sup>12</sup> For the purposes of the present study, we are ignoring in this text sample the classification proposed by Van Lancker Sidtis (2009: 455), who marks formulaic speech in bold italics and novel expressions underscored.

interactive grammar. Thus, 66.4% of the units, that is roughly two thirds, are interactives as against 33.6% units that can be classified as being suggestive of sentence grammar.<sup>13</sup>

With the exception of attention signals and ideophones, all types of interactives are represented in the text piece of (13). The lack of ideophones is hardly surprising since ideophones are generally of extremely low text frequency throughout English discourse of all kinds. Absence of attention signals might be due to the fact that this type of interactive requires some amount of discourse planning, which does not appear to be pronounced in persons suffering right hemisphere damage. The clearly predominant type is that of discourse markers. This observation is seemingly surprising, as a matter of fact, however it is not: The majority of discourse markers consists of the hesitation fillers *uh* and *um* (see Section 3.3.4), which is not unexpected considering that aphasics are notoriously faced with problems of discourse processing. The second largest group of interactives consists of interjections, like *boah!*, *Jesus*, *phew*, and *wow!*, which also does not seem to be surprising since the right hemisphere has been shown to be strongly associated with emotional behavior (e.g., [Ley 1980](#); [Moscovitch 1983](#); [Blonder, Bowers, and Heilman 1991](#); [Borod et al. 1992](#); [Borod, Bloom, and Haywood 1998](#); [Sherratt 2007](#); see [Heine, Kuteva, and Long 2014](#): 166–7 for further references).

These findings confirm what has been observed in other research on hemisphere-related brain activity, which suggests, on the one hand, that sentence grammar is of paramount importance for processing meaning contents organized in the propositional format of clauses and sentences, that is, ‘novel,’ compositional speech in the wording of Van Lancker Sidtis. On the other hand, it also suggests that text pieces of what is classified variously as formulaic, holistic or thetical speech, and most centrally as interactives, appear to draw more heavily on right hemisphere activity. Accordingly, persons suffering left hemisphere damage are most likely to structure their discourses with interactives and, conversely, persons suffering right hemisphere damage tend to have a poor input of interactives (e.g., [Van Lancker Sidtis 2004, 2009](#); [Heine, Kuteva, and Kaltenböck 2014](#); [Heine et al. 2015](#)).

To conclude, the right hemisphere and the left hemisphere have been found to jointly process language at all levels and to support one another in discourse processing (e.g., [Hagoort, Brown, and Swaab 1996](#); [Beeman and Chiarello 1998](#): 6). At the same time, there are also some positive correlations between the functions subserved by the two grammars distinguished here in that interactive grammar implicates right hemisphere activity more pronouncedly than sentence grammar.

<sup>13</sup> In [Heine, Kuteva, and Kaltenböck \(2014: 157\)](#), units such as *big*, *big and*, *me*, *money*, *small*, *TV*, etc. were left unclassified while in the present study all units that are not clearly interactives are treated as belonging in some way to sentence grammar.



**Table 5.14** Sentence grammar vs. interactive grammar in a speech sample of an English-speaking aphasic patient

| Sentence grammar |         | Interactive grammar |                  | Total   |                |
|------------------|---------|---------------------|------------------|---------|----------------|
| Number of tokens | Percent | Type of interactive | Number of tokens | Percent | Number Percent |
|                  |         | Directives          | 2                | 1.7     |                |
|                  |         | Discourse markers   | 30               | 25.9    |                |
|                  |         | Evaluatives         | 4                | 3.5     |                |
|                  |         | Social formulae     | 4                | 3.5     |                |
|                  |         | Interjections       | 16               | 13.8    |                |
|                  |         | Response elicitors  | 5                | 4.3     |                |
|                  |         | Response signals    | 15               | 12.8    |                |
|                  |         | Vocatives           | 1                | 0.9     |                |
| Total            | 39      |                     | 77               | 66.4    | 116 100        |

Source: Data taken from [Van Lancker Sidtis \(2009: 455, Table 1, Test 5\)](#)

5.4 Dualism in social psychology

In addition to linguistic and neurolinguistic studies there are also studies in other academic fields proposing a related dualistic perspective, especially in psychological studies. One distinction reflecting that between sentence grammar and interactive grammar is made in fuzzy-trace theory, where verbatim encoding relates to the former and gist encoding to the latter grammar ([Brainerd and Reyna 1992, 2002](#); [Reyna and Brainerd 1995](#); Section 5.2.1). But most notably, this applies to social and cognitive psychology, where it has been argued that there are two more or less distinct processing modes, or reasoning systems, or cognitive systems available for many cognitive tasks. On this ‘system view,’ human cognition is composed of two distinct systems, with different structures, functions, and evolutionary histories ([Frankish 2010: 919](#)). Rather than systems, these are modes of reasoning in the framework of [Henrich, Heine, and Norenzayan \(2010: 24–7\)](#), namely a holistic and an analytic mode of reasoning.

For example, in the dual-process model of [Kahneman \(2012\)](#) it is argued that we apprehend the world in two radically opposed ways, employing two different modes of thought or clashing decision-making processes. These modes are referred to by him as System 1 and System 2. System 1 is fast and largely unconscious and cannot be switched off. It is intuitive, associative, metaphorical,

automatic, and impressionistic, and it is highly sensitive to subtle environmental cues. Its operations involve no sense of intentional control, resting on basic emotions, memory, and hard-wired rules of thumb. System 2, by contrast, is mostly detached from emotions, and it is credited with the continuous monitoring of one's behavior. Being in charge of self-control, it is slow, deliberate, effortful, laboriously checking the facts and doing the math, and it usually accepts what System 1 tells it (Kahneman 2012: 24–6).

Instead of System 1 and System 2, the distinction is described by some researchers as one between Type 1 and Type 2 (e.g., Evans 2008; Frankish 2010; see Evans and Stanovich 2013 for discussion).<sup>14</sup> Both terminologies refer essentially to the same phenomena, and we will use here the less committal terms Type 1 and Type 2 for both, and rather than 'dual system theories,' the approaches are here called summarily dual-process frameworks.<sup>15</sup>

The two types of processing exhibit clear correlations with the two grammars proposed here: Type 1 reflects the main features of interactive grammar and Type 2 those of sentence grammar (see Section 5.5). The two types are argued to operate in concert but to suit different kinds of knowledge, and they serve complementary functions (Sloman 1996: 18). Nevertheless, the two also share some features. Inferences and decisions can reflect processes of either type.

Altogether more than two dozen features have been adduced by various authors to describe the distinction between Type 1 and Type 2 (see Frankish 2010: 922; Evans and Stanovich 2013: 225), and the nature of the features differs greatly from one author to another. For example, Epstein's (1994) experiential and rational processing systems correlate with the Type 1 vs. Type 2 distinctions non-conscious vs. conscious, holistic vs. analytic, associative vs. logical, concrete vs. abstract, and rapid vs. slow processing. For Sloman (1996), the distinction is one between a reflexive associative system (Type 1), which draws inferences from statistical regularities in the environment, and a deliberate rule-based system (Type 2), which operates on symbolic structures and aims to describe the underlying logical and causal structure.

Evans and Over (1996) again propose a dual system model of reasoning and judgment, building on dual-process theories of reasoning and work on implicit learning. Their model posits implicit and explicit cognitive systems, where the implicit system (Type 1) is non-conscious or pre-conscious, rapid, parallel, low effort, high capacity, and shaped by biologically constrained, domain-specific learning. The explicit system (Type 2), by contrast, is conscious, slow, serial,

<sup>14</sup> For a more detailed discussion of the difference between Type 1 and Type 2, see Heine, Kuteva, and Long (2020).

<sup>15</sup> With the term 'framework,' we wish to avoid the problem of whether, or to what extent, dual-process frameworks qualify as 'theories,' 'models,' or something else (see Gawrowski and Creighton 2013: 307 for discussion of this problem).

high effort, limited capacity, and responsive to verbal instruction. Inferences and decisions can reflect processes in either system.

In the default-interventionist model of [Evans and Stanovich \(2013: 236–7\)](#), the defining characteristic of Type 1 processes is their autonomy, they make minimal demands on working memory resources. Type 2, on the other hand, involves higher order reasoning processes. Defining features of processing are working memory capacity and cognitive decoupling ([Evans and Stanovich 2013: 236](#)). The differential role played by working memory is also addressed by [Frankish \(2010: 922\)](#) and some other authors. [De Neys \(2006\)](#) found that System 1 (Type 1) works independent of working memory, and when System 2 (Type 2) was impeded due to lack of working memory space, System 1 took over. Furthermore, [Frankish \(2010: 922\)](#) draws attention to the differential role played by context, arguing that Type 1 is contextualized whereas Type 2 is decontextualized.

Table 5.15 provides a list of frequently mentioned features found to characterize distinctions in reasoning and decision-making processes associated with the two types.

**Table 5.15** Some distinguishing features associated with processes of reasoning and judgment

| Type 1 (System 1) | Type 2 (System 2) |
|-------------------|-------------------|
| Holistic          | Analytic          |
| Automatic         | Controlled        |
| Non-conscious     | Conscious         |
| Intuitive         | Rule-based        |
| Experiential      | Rational          |
| Contextualized    | Decontextualized  |
| Associative       | Logical           |
| Heuristic         | Reflective        |
| Reflexive         | Deliberate        |
| Rapid processing  | Slow processing   |
| Implicit system   | Explicit system   |

*Main sources:* [Frankish \(2010: 922\)](#); [Kahneman \(2012\)](#); [Evans and Stanovich \(2013: 225\)](#)

Most of the research findings surveyed converge on arguing that both types of processing are constantly available, they interact and, importantly, that Type 1 is likely to apply before Type 2. Furthermore, some domains of behavior have been identified where both Type 1 and Type 2 are needed. The following in particular are argued to require the two types to work in concert: Judgment, evaluation, complex planning, complex decision making, complex problem solving, critical thinking, strategic thinking, hypothetical thinking, systematic thinking, innovation, and technical design ([Evans 2003](#)). In the dual system model of

reasoning and judgment of [Evans and Over \(1996\)](#), decisions can reflect processes in either system, but there is also interaction between the systems.

In the dual-process model of [Kahneman \(2012\)](#); see also [Kahneman and Tversky 1972, 1973](#); [Tversky and Kahneman 1974](#); [Kahneman and Frederick 2002](#)), System 2 somehow presupposes the presence of System 1. Type 1 is likely to apply before Type 2, and there may be a hierarchy of precedence between the two, in that [Kahneman \(2012: 24, 26\)](#) argues that his System 2 usually accepts what System 1 tells it, it is mobilized when a question arises for which System 1 does not offer an answer.

Authors proposing dual-process frameworks draw on a large body of theoretical premises and methodologies, including experimental, psychometric, and neuroscientific techniques. While their interest is generally with human behavior, neurophysiological experiments have also been carried out, e.g., to provide supporting evidence that the two processes in dual-process theory have separate and distinct neuronal substrates; see [Evans and Stanovich \(2013: 233–4\)](#) for an overview of this work.

A common thread in a number of dual-process frameworks can be seen in the claim that Type 1 is suggestive of an evolutionarily old form of cognition, one that includes innate skills that are shared with non-human animals. In contrast, Type 2 is claimed to be more recent, being unique to humans (e.g., [Reber 1993](#); [Epstein 1994](#); [Evans and Over 1996](#); [Stanovich 1999, 2004](#); [Kahneman 2012: 21](#); [Evans and Stanovich 2013: 225](#)). In the dual system framework of [Stanovich \(1999, 2004\)](#), System 1 was designed for the promotion of narrowly genetic goals, such as reproductive success, while the more flexible System 2 serves the goals of the individual person and allows us to rebel against genetic imperatives. As intuitively appealing such views may be, more data are needed to assess their empirical significance.

Research on psychological dual-process frameworks has not gone unchallenged. The following problems in particular are pointed out: First, not all dual-process frameworks are essentially the same; rather, there is a wide range of alternative approaches, even if most of them subscribe to a common core. Second, the boundary between the two types is not clearly delineated and there is also some vagueness in the way the two systems are characterized. Third, some of the features proposed to characterize Types 1 and 2 have been claimed to be problematic, and it has been argued that certain kinds of reasoning phenomena have been misclassified by dual-process theorists.

Fourth, it has also been argued that there is no need for a dualist distinction (see especially [Osman 2004](#); [Kruglanski et al. 2006](#)), or that the range of phenomena covered by dual-process frameworks is better taken care of by a continuum model extending from implicit to explicit processes of learning and reasoning. For students of continuum models, the distinction in the performance of reasoning tasks relates to differences in degree rather than in the kind of reasoning system used.

For example, [Cleeremans and Jiménez \(2002\)](#) and others propose a single-system learning framework incorporating a continuum between implicit and explicit processes; see also [Osman \(2004: 993\)](#). And finally, there remain some unresolved questions. For example, whether we store, retrieve, encode, and forget information separately in the two types or completely independently of each other remains to some extent controversial.

## 5.5 Discussion

The frameworks examined in the preceding sections do not exhaust the range of notions that have been proposed to account for a dualistic organization of mental processing. Mention should be made especially of a more general distinction between two kinds of reasoning processes influencing language structure, namely *successive* vs. *associative thinking* ([Havers 1931](#); see [Schneider 2007b: 26, 39](#)), which strongly correlates both with that between Type 2 and Type 1 in Section 5.4 and between sentence grammar and interactive grammar, respectively. Among the dual-process frameworks discussed in Section 5.2, it is most of all the ones listed in Table 5.16 that are immediately relevant for an understanding of the distinction between two kinds of grammar as it is made in this book. And of these, (a) and (d) are fully compatible with the distinction while (b) and (c) are largely though not entirely so: Both holistic speech of [Pawley \(2009\)](#) and formulaic speech of [Van Lancker Sidtis \(2009\)](#) also include some phenomena that do not qualify as interactives.

**Table 5.16** Dual-process frameworks relating most strongly to interactive grammar

|   | Kind of framework              | Category relating to             |                                | Sources                                                                      |
|---|--------------------------------|----------------------------------|--------------------------------|------------------------------------------------------------------------------|
|   |                                | Sentence grammar                 | Interactive grammar            |                                                                              |
| a | Semiotics                      | Analytic dimension               | Expressive dimension           | <a href="#">Dwyer and Moshi (2003)</a>                                       |
| b | Model of linguistic competence | Analytic processing              | Holistic processing            | <a href="#">Pawley (2009)</a>                                                |
| c | Social psychology              | Analytic, controlled, rule-based | Holistic, automatic, intuitive | <a href="#">Frankish (2010)</a> , <a href="#">Kahneman (2012)</a> and others |
| d | Neurolinguistics               | Novel speech                     | Formulaic speech               | <a href="#">Van Lancker Sidtis (2009)</a>                                    |
| e | Discourse grammar              | Sentence grammar                 | Thetical grammar               | <a href="#">Heine et al. (2013)</a>                                          |

In spite of the problems mentioned in Section 5.4, the dual-process frameworks proposed in social psychology have provided a seminal perspective of mental

activity, having been built on a substantial body of experimental, psychometric, and neuroscientific evidence (cf. [Evans and Stanovich 2013](#): 224). What may have become evident in particular is that the distinction made in these frameworks is immediately relevant to the one that is the central concern of this book in that there are a number of similarities between Type 1 and interactive grammar, on the one hand, and Type 2 and sentence grammar, on the other: Interactives have many of the features that have been adduced by students of social psychology to describe Type 1, whereas Type 2 exhibits features that can be observed in text pieces of sentence grammar (see also Section 7.1).

These similarities can be illustrated with the following text example of a conversation between four female speakers. There are four interactives in this example, namely an interjection (*oh*), an attention signal (*hey*), a discourse marker (*um*), and a response signal (*yeah*), and all are placed at the beginning of the turn in which they occur.

(14) English (LSWEC-AC (122001); [Norrick 2009](#): 881)

Ann: **oh** are you going to drink that out of the bottle?

Catherine: I always drink them out of the bottle.

Anna: **hey** Jenn there's a diet coke out if you want one.

Jennifer: **um, yeah** I'll have a diet coke.

The similarities can be described as follows. First, we observed that both Type 1 and Type 2 are constantly available and interact with one another. As example (14) shows, this also applies to interactives and sentence grammar, which frequently co-occur and complement one another (Section 7.3). There appears to be a kind of framing structure involved in (14) where the interactives *oh*, *hey*, *um*, and *yeah* set the frame, to be completed by the following unit of sentence grammar. Not uncommonly, the interactive provides a stimulus followed by a more elaborate comment added by a segment of sentence grammar. While interactives are syntactically independent and may be used as utterances of their own, in most of their usages they co-occur with text pieces of sentence grammar and precede the latter in this way.

Related observations have been made in the psychological frameworks surveyed in that Type 1, corresponding to interactive grammar, is likely to apply before Type 2, corresponding to sentence grammar—that is, the order is Type 1 – Type 2. For [Evans and Stanovich \(2013: 228\)](#), Type 1 processing generates intuitive default responses on which subsequent reflective Type 2 processing may or may not intervene, and Kahneman observes:

System 1 continuously generates suggestions for System 2: impressions, intuitions, intentions, and feelings. If endorsed by System 2, impressions and intuitions turn into beliefs, and impulses turn into voluntary actions. When all goes smoothly, which is most of the time, System 2 adopts the suggestions of

System 1 with little or no modification ... When System 1 runs into difficulty, it calls on System 2 to support more detailed and specific processing that may solve the problem of the moment.

(Kahneman 2012: 24)

While the placement of interactives is on the whole flexible, their main position is with few exceptions, such as response elicitors and ideophones, at the left periphery of an utterance.

Second, Type 1 (System 1) is described as being holistic and Type 2 (System 2) as analytic and rule-based. Thus, Type 2 can follow rules or compare objects on several attributes and it ‘constructs thoughts in an orderly series of steps’ (Kahneman 2012: 24, 36). This concerns in fact one of the main features distinguishing interactives from sentence grammar. As the example in (14) shows, interactives such as *oh*, *hey*, *um*, and *yeah* are holistic, invariable forms whereas the units of sentence grammar following them are linearized strings composed analytically by means of morphosyntactic rules. And this difference also takes care of other features characterizing the Type 1 vs. Type 2 distinction: Interactives, such as the ones just mentioned, are produced fairly spontaneously, whereas units of sentence grammar are somehow the result of controlled, conscious, and rule-based grammatical planning.

Third, the distinction contextualized vs. decontextualized is discussed in some form or other by a number of social psychologists. For Sloman (1996), the reflexive associative system of Type 1 draws inferences from statistical regularities in the environment, that is, in the context, and according to Kahneman (2012), System 1 (Type 1) is highly sensitive to subtle environmental cues. This feature relates to a core property of interactive grammar: Interactives are indexicals, they are immediately anchored to the situation of discourse (the ‘context’), that is, their function is determined by the context in which they are used (see Section 1.2). The function of units of sentence grammar, by contrast, is defined with reference to their syntactic and semantic role within the sentence, and in this sense, they may be said to be ‘decontextualized.’

And fourth, another feature separating Type 1 and Type 2 concerns the distinction rapid vs. slow processing. Obviously, an interactive like *hey* in example (12) is a short, monosyllabic form which is distinctly more likely to be processed rapidly than the following unit of sentence grammar *Jenn there’s a diet coke out if you want one*. The latter unit is made up of a series of eleven forms structured in terms of syntactic relations and it is safe to assume that this unit requires distinctly more processing effort and time than the preceding interactive *hey*.

These are but a few examples showing that, on the one hand, the phenomena captured by psychological dual-process frameworks are not the same as those figuring in the linguistic frameworks sketched in Section 5.2. But on the other hand, there are striking correlations between the two kinds of frameworks—note that

these correlations are not restricted to single features but to pairs of features. For a more detailed account of these correlations, see [Heine, Kuteva, and Long \(2020\)](#).

What these observations suggest is that underlying the distinction there is a more basic strategy of human behavior to process phenomena in two contrasting ways, employing two different modes of thought and decision making: These phenomena are treated either as holistic depictions for rapid processing. Alternatively, they are treated by monitoring one's behavior and viewing the phenomena as conceptually complex entities that need to be processed laboriously and analyzed in terms of the grammatical architecture and the building blocks of which they are composed (see e.g. [Pawley 2009](#): 21; [Kahneman 2012](#): 24–6).

## 5.6 Conclusions

This chapter was concerned with previous research on interactives. As we saw in Section 5.1, there are a number of grammars and grammatical analyses that have come up with notions of or related to interactive grammar, and in Section 5.2 we saw that the idea of a dualistic organization of grammar is far from new. Of all the approaches surveyed, two were singled out as being directly in line with the analysis proposed here, namely the one adopted in the *Longman Grammar of Spoken and Written English* ([Biber et al. 1999](#)) and that of discourse grammar ([Kaltenböck, Heine and Kuteva 2011](#); [Heine et al. 2013](#)).

Furthermore, we saw that interactive grammar can be reconciled with findings made in disciplines other than linguistics. Section 5.3 showed that the distinction between two kinds of grammar exhibits significant correlations with activity in the two hemispheres of the human brain in that interactive grammar is more strongly associated with right hemisphere activity than sentence grammar. And remarkable correlations were also observed in Section 5.5, suggesting that interactive grammar relates to Type 1 and sentence grammar to Type 2 distinguished in research by social psychologists on reasoning and judgment.



# 6

## The status of interactives

So far we have been dealing with interactives in grammar and discourse. The present chapter is concerned with two kinds of subject matter. On the one hand, it takes up issues that were mentioned in preceding chapters but could not be treated there in any detail. This concerns two questions that have been raised in much of previous research, namely what the place of interactives is in the overall space of grammar and discourse (Section 6.1), and whether interactives qualify to be classified as ‘words’ (Section 6.2).

On the other hand, there are a few more general questions concerning the ontological status of the category of interactives, and specifically the question whether this category forms a ‘system’ on its own (Section 6.3). Section 6.4 takes up an issue raised in Section 6.1, namely whether interactives are ‘marginalia’ of language structure. The issue of Section 6.5 is the behavior of interactives in language contact, while Section 6.6 addresses the question of how interactive grammar relates to first language acquisition. In Section 6.7, finally, a list of general features of interactives is proposed.

### 6.1 Where are interactives located in grammar and discourse?

Interactive grammar serves, in the wording of Wiltchko (2021: 2–3), to manage the common ground between interlocutors, playing an important role in the synchronization of minds. But where exactly are interactives located in the space of language structure and discourse processing?

There is a wide array of views that surface in the relevant research on what the place of interactives is vis-à-vis language, language structure, and discourse. Many, though not all views that have been voiced on this issue can be reduced to one or more of the hypotheses listed in (1).

- (1) Five views on the status of interactives
  - (a) They are part of language structure
  - (b) They are located at the periphery of language
  - (c) They are intermediate between language and something else
  - (d) They are not part of language
  - (e) They form a domain of grammar separate from sentence grammar

These views are now discussed in turn. Note that some of the authors to be cited do not subscribe to one particular view but may favor one view in some contexts and another view in other contexts.

### 6.1.1 They are part of language structure

Proponents of the hypothesis in (1a) may admit that interactives exhibit some peculiar features but that this also applies to some extent to other parts of language structure as well. Hence, interactives can somehow be described in terms of standard methods of linguistic analysis.

A position to this effect can be found in many grammatical descriptions of languages, be that grammars of European or other languages. But it surfaces also in some studies devoted to interactives. For example, in his seminal study of ‘interjections,’ Wilkins (1992: 153) argues that while ‘interjections’ tend to have quite a complex semantic structure, that structure is nevertheless ‘amenable to semantic decomposition.’

This view is shared in some way or other by a number of other students of interjections, in that interjections are classified as part of language structure, being semantically rich and having a conceptual structure that can be explicated, even if they also differ from other language material (e.g., Norrick 2009). In the ‘conceptualist’ literature drawing on the model of natural semantic metalanguage of Wierzbicka and associates, many peculiarities of interactives are pointed out, yet the meaning of interjections and other interactives is explicated in much the same format as other meanings (Ameka 1992a, 1992b; Wierzbicka 1992; Wilkins 1992).

### 6.1.2 They are located at the periphery of language structure

According to the actually prevailing hypothesis, interactives are claimed to be peripheral or marginal to the language system, being located in some fringe area of the linguistic domain (cf. Wharton 2003). This position surfaces in some form or other in a substantial body of linguistic work up to the present. Already about a century ago, Sapir (1921: 6) relegated interjections to being at best a ‘decorative edging to the ample, complex fabric’ of language structure.

For Cuenca (2000), interjections form a context-sensitive ‘peripheral class’ of the category ‘sentence,’ one that typically encodes pragmatic meanings. Goffman (1978: 99) views interjections such as *oops* as ‘response cries’ which are not full-fledged words, while others describe them as conventionalized cries, ‘typically indicating the speaker’s emotional response’ (Dixon 2010b: 27), or as ‘somewhat

primitive expressions of feeling, only loosely integrated into the linguistic domain' (Leech et al. 1982: 53).

Defining a class of linguistic elements that to quite some extent coincides with that of interactives, Ameka (2020) proposes to locate that class at the periphery of language structure, referring to it as 'marginalia'; we will deal with that position in more detail in Section 6.4.

A procedure widely adopted in grammar writing, be that in textbooks or linguistic analysis, grammars of European languages or languages in other parts of the world, is to reduce interactives to some kind of appendage placed towards the end of the grammar. And rather than describing their grammatical structure, functions, and usage, many authors are satisfied with listing a few interactives, sometimes providing a few examples showing how they are used. The general impression conveyed is somehow that these forms are not worth analyzing but should at least be mentioned in a treatment of language structure.

### 6.1.3 They are intermediate between language and something else

An alternative position, surfacing especially in linguistic work since the 1980s, holds that interactives are not clearly part of the structure of language but rather are intermediate between language and some other domain of human behavior. Interactives have been said to be located at the boundary between verbal and non-verbal, or between linguistic and paralinguistic communication (Goffman 1978; Ameka 1992a: 112; 2006: 745; Nübling 2004; Stange and Nübling 2014: 1986). For some authors, interactives belong to the phenomena classified as paralinguistic. Hayashi and Yoon (2010: 43), for example, call elements such as hesitators or fillers, that is, a sub-class of discourse markers (Section 3.3.4), paralinguistic signals.

A century ago, Sapir (1921: 6–7) observed that 'interjections' are 'the nearest of all language sounds to instinctive utterance.' Interjections and other interactives have also been viewed as 'asystematic' elements, classified as non-words, semi-words, quasi-words, response cries (Goffman 1978; Ameka 2006: 744; Ameka and Wilkins 2006: 5; cf. Andrason and Dali 2020: 167), or 'ritualized acts' (Goffman 1981: 100). Interjections, Wharton (2003, 2016: 29) concludes, are partly natural and partly coded, they fall at various points along a continuum between language proper and display (see also Section 2.1.3).

The bewilderment experienced by some authors when it comes to dealing with interactives can be illustrated with the following example. In his grammar of

the Paamese language of Vanuatu, Crowley notes with reference to ‘interjections,’ which also include some other types of interactives:

The boundaries between what is linguistic and what is not linguistic become difficult to judge when dealing with interjections as many interjections do not fit into the grammatical structure of an utterance, as well as being phonologically unusual in many cases [...] they may have some abnormal intonation pattern.

(Crowley 1982: 76)

#### 6.1.4 They are not part of language

Other authors again are satisfied with simply noting that interactives, and in particular interjections, do not belong to grammar, nor to language. Amazement about the peculiar nature of interjections in fact induced earlier authors to deny them a place within language. Already in the mid nineteenth century, Müller (1861: 346, 352) declared that ‘interjections’ ‘are playthings, not the tools of language’ and that ‘language begins where interjections end.’ Jakobson (1960: 354) considered them exemplars of the ‘purely emotive stratum of language,’ and in his history of linguistics, Benfey (1869: 295) went so far as claiming that interjection is ‘the negation of language.’

A pronounced modern representative of this view is Goffman (1978, 1981), for whom ‘response cries,’ that is, the interjections studied by him, are a variety of ‘non-words’ and as such, ‘can’t quite be called part of a language,’ and the same conclusion is drawn by Wharton (2003: 84): ‘The answer, then, is no, interjections are not part of language.’ And discussing the linguistic abilities of certain grave aphasics, he draws the following important conclusion: ‘If one can have interjections, but not language, it is hard to see how the former can be viewed as part of the latter’ (Wharton 2003: 84–5).

The view that interactives cannot really be classified as belonging to language surfaces in some form or other in studies portraying them as ‘extra-linguistic,’ or extra-grammatical elements. In his analysis of English holophrases, Mackenzie (1998) also discusses several types of interactives, such as ‘fixed expressions,’ which include social formulae like *Cheers!*, *How do you do?*, and *Thank you*, and indications of assent or dissent, that is, response signals like *yes*, *no*, *absolutely*, and *hardly*. Mackenzie (1998: 283) concludes that these holophrases ‘can be safely left outside the grammar.’

Eastman (1992: 273) describes ‘interjections’ in the East African language Swahili ‘as elements of culturally specific extra-linguistic systems of communication in context,’ where her term ‘interjection’ covers a number of different types of interactives. And arguing for the non-linguistic character of ‘interjections,’ Wharton (2003: 84–5) observes that there is aphasiological evidence for

a dissociation between interjections and ‘language proper’: Interjections remain within the repertoire of certain grave aphasics. He draws attention to neurology, arguing that ‘use of interjections is associated with phylogenetically ancient limbic sub-cortical circuitry linked with emotion, as opposed to the more recent cortical structures implicated in the production of language proper’—a position being close to the one argued for in Section 5.3.

### 6.1.5 They form a domain of grammar separate from sentence grammar

This is the position adopted in many of the dual-process frameworks discussed in Section 5.2. For example, [Dwyer and Moshi \(2003: 174–5, 179\)](#) propose a distinction between an analytic dimension and an expressive dimension of semiotic categories, where the former corresponds essentially to sentence grammar and the latter to interactive grammar. A similar view can be found already in the work of earlier linguists. For example, [Karcevski](#) wrote in [1941](#): ‘As for interjections, ... they move on the *non-conceptual* plane of language, which is opposed to all other semiotic planes.’<sup>1</sup>

Also adopting a semiotic perspective, [Diffloth \(1994: 108\)](#) maintains that ideophones, called ‘expressives’ by him, do not function in what he refers to as the ‘conventional semiotic domain.’ Rather, ‘they actually constitute a parallel sub-language grafted on, and parasitic on, the conventional one.’ Also dealing with ideophones, [Zwicky and Pullum \(1987: 338\)](#) observe that they were best analyzed as falling ‘in a domain orthogonal to the grammar,’ with their own rules and regularities of expressive morphology (see [Dingemanse 2018: 14](#)). And [Dingemanse \(2017b\)](#) goes on to propose a two-systems hypothesis which he sketches thus:

What we see here is the encounter of two distinct and partly incommensurable methods of communication: the discrete, arbitrary, descriptive system represented by ordinary words, and the gradient, iconic, depictive system represented by ideophones. These two systems place different requirements on the material use of speech, yet both are part of one linearly unfolding speech stream.

([Dingemanse 2017b: 197](#))

Similarly, in work on Lexical Anthropology, interjections are conceived as belonging to a semantic domain of their own. For example, ‘saying *wow!* is not a realisation of the semantics of *astonishment*, but a realisation of the semantics of *wow!*’ ([Levisen 2019: 115](#)):

<sup>1</sup> ‘Quant aux interjections, ... [e]lles se meuvent sur le plan *non-conceptuel* du langage lequel s’oppose par là à tous les autres plans sémiologiques’ ([Karcevski 1941: 61](#); italics in the original). We owe this reference to Franck Floricic (p.c. of August 7, 2021).

In Lexical Anthropology, interjections are not viewed as subordinates in a superordinate system. Instead, they are viewed as independent semantic contributions to discourse, and consequently they deserve to be studied in their own right, and not as an appendix to something else.

(Levisen 2019: 114)

A related dualistic perspective is also proposed in some lines of generative linguistics. In particular, we saw in Section 5.2.4 that in the work of Wiltschko (2021) two main syntactic levels or structures are distinguished, called the interactional and the propositional level or structure, where the former includes two types of interactives, namely response elicitors and response signals, and the latter roughly corresponds to that of sentence grammar.

### 6.1.6 Conclusion

The classification of hypotheses proposed in Section 6.1 is not meant to represent an exhaustive account of all the positions that surface in work on interactives, some of which are hard to classify in terms of these hypotheses. Each of the positions discussed draws attention to specific features characterizing the nature of interactives. Nevertheless, we argue that it is the position in Section 6.1.5 that accounts best for analyzing and describing interactives—reflecting in some way or other features that have been identified in work on dual-process frameworks discussed in Section 5.2. That interactives form a domain separate from sentence grammar is most of all in line with the duality hypothesis proposed in the model of discourse grammar (Section 5.2.3).

On the one hand, the grammar of interactives and sentence grammar differ significantly from one another in their structure, as captured by the properties discussed in Section 1.3. On the other hand, we saw in Section 1.4 that the two also have a set of properties in common, suggesting that the two have one and the same foundation: Both are part of language, and both are part of one encompassing grammar, referred to here as discourse grammar (Section 5.2.3).

In sum, while sentence grammar and interactive grammar share a common base, they are sufficiently different to justify treating them as distinct tools of discourse processing on the basis of both their functions and grammatical properties. We will deal with this position in more detail in Section 6.4.

## 6.2 Are interactives words?

Throughout this study we avoided calling interactives ‘words,’ referring to them with the less committal terms ‘expressions,’ or ‘forms.’ This, however, is not

the convention adopted by the many students of interactives who converge on classifying interactives as ‘words,’ and some of them have in fact made an effort to justify this convention. Since this is an issue that relates to the ontological status of the category, it is looked at in more detail in the present section.

### 6.2.1 On terminology

Grammarians, both of English and of other languages, are hard-pressed when it comes to searching for appropriate labels to account for the grammatical status of interactives: Are English forms like *ouch*, *thud*, or *um* words and, if not, what are they? The following example, restricted to interjections in one particular language, may illustrate the problem surrounding the grammatical taxonomy of interactives.

In their reference grammar of English, Quirk et al. (1985: 853) call interjections ‘purely emotive words.’ For example, the interjections *oh*, *ah*, *ugh*, and *phew* are listed as one of their word classes, even if a ‘marginal and anomalous class’ (Quirk et al. 1985: 67). Other interjections again, like *ouch* and *hush*, are treated as both interjections and ‘quasi-linguistic noises’ (Quirk et al. 1985: 88, note). Other forms again, commonly treated as interjections by other authors, including expletives like *damn (it)*, *my gosh*, *shit*, etc., are classified neither as words nor as a marginal or anomalous class but rather as ‘formulae’ (Quirk et al. 1985: 852). To add to this indeterminate situation, Quirk et al. (1985: 88, note) conclude that ‘we must acknowledge that statements and questions, too, can occur as interjections.’

And to further add to this complexity, interjections and other kinds of interactives are also called ‘semi-words,’ ‘quasi-words,’ ‘response cries,’ ‘non-word vocalizations,’ or simply ‘non-words’ (Goffman 1978: 811, 814–15; 1981: 90), ‘vocal gestures’ or ‘linguistic gestures’ (Ameka 1992a: 106), ‘particles’ (e.g., Nübling 2004: 11), or ‘ritualized acts’ (Goffman 1981: 100) (see Andrason and Dlaki 2020: 167). Wharton (2016: 21) concludes that interjections are ‘partly natural, partly conventional semi-words,’ and Haiman (2018: 135) concurs with Goffman’s usage of ‘response cries’ as ‘semi-words.’

Interjections have also been called ‘lexemes’ (e.g., Wilkins 1992: 153)—Mackenzie (2019: 76) classifies English interjections as a lexeme class. Most commonly however, they are called ‘words’ (e.g., Zwicky 1985: 283; Wilkins 1992; Evans 1992: 226; Ameka 1992a: 102; Kockelman 2003: 469; Norrick 2009: 867). Even though they may consist simply of a dental click ([|]), rendered in English writing as *tsk-tsk!*, *tut-tut!*, /tsktisk/, or /tattat/, Wilkins (1992: 148, fn. 21) emphasizes that interjections are words. This view is challenged though by Goddard (2014: 54), who suggests that interjections can be classified into non-word, ‘noise-like’ ones, ‘word-like’ ones, and secondary interjections, the last ones behaving like words (see also Section 3.6.4.1).

With regard to ideophones, Ameka (2001: 26) asserts that they ‘are first and foremost a type of words – a lexical class of words – which need not belong to the same grammatical word class in a particular language nor across languages.’ Güldemann (2008: 280) defines ideophones as ‘expressive and symbolic lexemes,’ and for Dingemanse (2019: 13, 15) they are an ‘open lexical class.’ He adds that one of the ‘four key properties of ideophones that are recurrent across languages’ is that ‘they are words, i.e., conventionalized lexical items that can be listed and defined.’ Accordingly, Dingemanse, Torreira, and Enfield (2013) call, for example, the other-initiated repair marker *huh?* of English or other languages ‘a lexical word’ (Section 3.7.5). Similarly, English fillers like *uh* and *um* (Section 3.3.4) have been argued to be legitimate English words (Clark and Fox Tree 2002).

Use of the term ‘word’ for interactives has a long tradition: Already a century ago, Jespersen (1922: 415) assumed that interjections, classified by him as particles, ‘are learnt like any other words.’ While being said to be ‘rather peripheral to language,’ Leech et al. (1982: 53) maintain that interjections like *ugh*, *phew*, *oh*, *ah*, and *ouch*, swear words (*damn* etc.), and social formulae like *hello*, *goodbye*, or response signals like *yes*, *no*, *okay* are nevertheless words. In a similar way, Biber et al. (1999: 225) classify their ‘inserts,’ which include response signals, interjections, greetings, apologies, etc. as ‘stand-alone words.’

There are other authors, however, who have reservations calling interjections ‘words.’ As early as 1765, the French writer De Brosse observed that interjections ‘are not words but something more’ (De Brosse 1765: 223).<sup>2</sup> For later generations of scholars, however, interjections were not ‘something more’ but more likely ‘something less’ than words. About a century after De Brosse, Benfey (1869) is satisfied to characterize interjections in the following way in his history of linguistics:

Between interjection and word there is a chasm wide enough to allow us to say that the interjection is the negation of language, for interjections are employed when one either cannot or will not speak.

(Benfey 1869: 295)

And some present-day linguists are not happy either with raising them to the status of words.

According to Goffman (1981: 99), ‘exclamatory interjections’ are not ‘full-fledged words.’ And the headline of sub-chapter 2.3 in Haiman’s (2018: 104) monograph is ‘Ideophones are not quite words,’ adding that they are also not ‘a part of speech’ (see also Meinard 2015: 160–2):

<sup>2</sup> ‘... les interjections ... ne sont pas des mots mais quelque chose de plus’ (De Brosse 1765: 223).



Finally, although both ideophones and response cries are words, they can be called 'semi-words' insofar as they share to some extent the properties of lacking some of the design features of language. They are not entirely conventionalized, they exhibit analog patterning, and they are less standardized than conventional words.

(Haiman 2018: 135)

And there is also the position that some interactives are words while others are not. Thus, Biber et al. (1999: 1082) conclude with reference to English inserts: 'In fact it may be questioned whether some inserts are words at all: this applies to interjections (such as *ugh*, *ooh*), response forms (such as *uh huh*, *mhm*) and hesitators (such as *mm*, *uh*).'

But this is not the only view expressed. Güldemann (2008: 280) characterizes ideophones metaphorically as 'a performance or a gesture in the disguise of a word.' More commonly, interactives are conceived as utterances or sentences rather than as words. For Jakobson (1960: 354), interjections 'are not components but equivalents of sentences.' In their grammar of English, Carter and McCarthy (2006) define English interjections, like *bother*, *crikey*, *damn*, *heavens*, *hooray*, *jeez*, *ouch*, *ugh*, *wow*, and *yuk* as a category of utterances rather than a word class. Adding that interjections normally refer to 'exclamative utterances consisting of single words that do not fit easily into the major word classes,' they highlight the relation between swearing and interjections:

Swearing often takes the form of interjections. This can involve single words, or short phrases or clauses that are used to express a variety of strong feelings, in particular annoyance, frustration and anger.

(Carter and McCarthy 2006: 226)

Diffloth (1972: 444) portrays ideophones as 'microscopic sentences' and according to Ljung (2011: 82–3), evidence that interjections are utterances rather than words can be seen, first, in the fact that some of them are or are historically derived from sentences:

In Swedish and German, for instance, both sudden surprise and a negative attitude to a following clause may be expressed by *Ta mig fan!* and *Hol mich der Teufel!*, both meaning 'The devil take me!'

(Ljung 2011: 83)

Second, Ljung (2011) found that English taboo words are used predominantly as word combinations rather than as single-word items. For example, he found the taboo word *God* in over 80%, and *Hell* even in over 99% of their occurrences only in combinations. And third, when discussing the modification of the English

expletives *God!* and *Jesus!* as *gosh!* and *gee!*, respectively, [Ljung \(2011: 11\)](#) emphasizes that it ‘is not the words *God* and *Jesus*, but interjectional utterances containing these words which are used in swearing.’ A compromise would be to define interactives as utterances consisting of one or only a few words. Still, this would imply that they are words or combinations of words.

There are alternative proposals. Interjections have been called ‘non-words,’ ‘semi-words,’ or ‘quasi-words’ ([Ameka 2006: 744](#)). [Goffman \(1978: 811\)](#) described them as ‘free-standing ‘natural expressions’ or signs meant to be taken to index directly the state of the transmitter,’ adding that they qualify as ‘semi-words’ or ‘response cries.’ In a similar fashion, [Wharton \(2016: 21\)](#) concludes that interjections are ‘partly natural, partly conventional semi-words.’

According to another alternative, interjections and discourse markers have been classified as ‘particles,’ that is, as an indeclinable word class like adverbs, conjunctions, prepositions, etc. or, alternatively, as invariant words which are not adverbs, conjunctions, or prepositions (cf. [Bussmann 1996: 867](#)).

### 6.2.2 What is a word?

There may be reasons to call interactives ‘words’—but what is a word? The problem is, first, that there are too many definitions of words, second, that phonological definitions are frequently incompatible with grammatical ones and, third, that none of the definitions that there are is widely accepted. Nevertheless, there are a few features that may be helpful to decide on this issue in a given case. In addition to the properties that were mentioned in [Section 1.4](#), interactives share with words most of all the following features:

- (a) A word, in the sense of [Bloomfield \(1962 \[1933\]: 178\)](#), is a ‘minimum free form,’ and depending on how one defines a ‘free form,’ this also applies to interactives.
- (b) A word is a conventionalized combination of form and meaning that can be separated in some way from surrounding material. This applies also to interactives, which are form–meaning combinations that are generally separated from surrounding material.
- (c) Words have to be learnt and are stored as such in memory, and the same appears to apply generally to interactives.

But there are also criteria for words that do not readily apply to interactives:

- (d) Words are commonly understood to conform to a specific syllable structure. There are some interactives that do not clearly conform to this criterion. This applies for example to primary interjections consisting only

- of consonants (e.g., *mm*, *mhm*), or a click consonant. An English example is provided by the dental click ([l]) (Section 6.2.1).
- (e) According to Dixon (2010b: 7), a phonological word is a phonological unit larger than the syllable. This does not readily apply to some English discourse markers like *uh* or *um*, to interjections (or ideophones) like *brrr* ('I feel cold,' 'It's cold'), a bilabial trill, nor to primary interjections like *hm*, attention signals like *pst*, or to directives like *shh* (Biber et al. 1999: 1094–5; Haiman 2018: 114). With regard to ideophones, Childs (2003: 122) concludes that they 'are immune to or unaffected by phonological rules.'
  - (f) 'Words frequently are morphologically complex, in the sense that they are to be analysed as being composed of two or more morphemes' (Zwicky 1985: 288). More specifically, (grammatical) words are defined by Dixon (2010b: 5) as 'inflected forms of a lexeme.' Clearly, this does not apply to interactives: They are by definition uninflected and unanalyzable (Section 1.2); hence, they cannot be or contain 'inflected forms of a lexeme' (Dixon 2010b: 5).
  - (g) While words *describe* meaning, interactives have been shown to *express* and *depict* meaning (Section 2.1).<sup>3</sup>
  - (h) In some accounts, words are defined or characterized with reference to the sentence. For example, for Sapir (1921: 34) words are 'one of the smallest, completely satisfying bits of isolated 'meaning' into which the sentence resolves itself.' This criterion does not apply to interactives: Since they are not part of a sentence, they cannot be part of the material into which sentences may resolve themselves.

These observations do not exhaust the terminological problems surrounding interactives. In many language descriptions, for example, interjections have been classified as one of the 'parts of speech' of grammar or of a language on the same level with parts of speech like nouns, pronouns, verbs, conjunctions, prepositions, etc. (cf. Ameka 1992a: 102). And in fact, since the time of early Latin grammar, interactives such as interjections have been listed as one of the parts of speech of grammar (e.g., Matthews [1997] 2007: 289). Parts of speech again have also been called 'lexical categories' (Trask 1993: 155), or word classes (e.g., Matthews [1997] 2007: 289), where the latter stand for 'groups of words which share the same syntactic and morphological properties' (Crystal 1994: 420).

There are, however, also views that question the claim that interactives are parts of speech. For example, Haiman (2018: 104) concludes that ideophones are not quite words, nor are they a part of speech (see also Meinard 2015: 160–2).

<sup>3</sup> Franck Floricic (p.c. of August 5, 2021) notes that words can also be used expressively and, as was argued in Section 2.1, interactives also have a descriptive meaning component.

### 6.2.3 Discussion

Problems with extending the notion ‘word’ to interactives have a long history. For the philosopher [Cassirer \(1944\)](#), interjections do not qualify as words since only words form part of a sentence with a logical and syntactic structure.

On account of the observations made in the preceding [Section 6.2.2](#) one may wonder whether the term ‘word’ is really appropriate when referring to interactives, and the same applies to terms such as ‘part of speech’ and related notions like ‘lexeme,’ ‘word class,’ etc. (see [Meinard 2015](#): 160–2 on ‘part of speech’). And much the same applies also to terms like ‘semi-word’ or ‘response cry’ (cf. [Goffman 1978](#): 811, 814; [Haiman 2018](#): 135), which are also not entirely felicitous to describe the nature of interactives. Not only do they have a bit of the ring of calling donkeys ‘semi-horses’ or pears ‘semi-apples,’ it also remains usually unclear what exactly the component ‘semi-’ stands for. ‘Response cry,’ while capturing certain features of a group of interjections, is also not an entirely satisfactory label, for the following reasons: First, it does not apply to a larger segment of the ten types of interactives, let alone to interactives in general. Second, it does not seem to do justice to the fact that interactives are full-fledged linguistic units rather than ‘cries,’ even if they are of a different kind than words of sentence grammar ([Section 1.4](#)). And finally, it has so far remained largely unclear what exactly the grammatical properties of ‘response cries’ or ‘semi-words’ are.

Arguably a more suitable term is ‘particle.’ Note that like interactives, particles are uninflected, invariant forms. Nevertheless, there are reasons not to extend this term to interactives. For one thing, the status of ‘particle’ as a relevant term of linguistic taxonomy has been shown to be debatable ([Zwicky 1985](#)). And for another, the term is typically related in some way or other to the notion ‘word,’ be that as forming a subgroup of words or with reference to its cooccurrence with words, as in English phrasal verbs (*look at, make up, put down, take off*). And the term ‘particle’ has also been used by a number of authors for word classes like adverbs, prepositions and conjunctions—it is widely assumed to require other words or phrases to impart meaning ([Trask 1993](#): 202; [Busmann 1996](#): 867; [Matthews \[1997\] 2007](#): 289).

None of these factors applies to paradigm instances of interactives, that is, to interjections like *ouch*, social formulae like *goodbye*, ideophones like *boing*, vocatives like *darling*, or directives like *shh*. Thus, extending the term ‘particle’ to interactives is likely to blur the particular nature of the grammar of interactives. But most of all, interactives typically can form utterances of their own—a feature that is at variance with most definitions of particles.

Considering all the problems surrounding the classification of interactives by means of labels taken from or associated with the tradition of sentence grammar, terms like ‘word,’ ‘semi-word,’ ‘particle,’ ‘lexical class,’ ‘part of speech,’ etc., are avoided here. Rather, interactives are called in general ‘expressions,’ or ‘forms.’ The

latter term is generally employed here for entities sharing the properties that were listed in [Section 1.4](#), including taxa of sentence grammar such as words. And on account of their expressive meaning ([Section 2.1.2](#)), interactives are also referred to as ‘expressions.’

In addition to ‘(linguistic) form’ there is a second term that was used in the preceding chapters for units of both sentence grammar and interactive grammar, namely the term ‘utterance.’ An ‘utterance’ is perhaps most commonly understood to be a bounded piece of speech produced by a particular individual on a particular occasion. Such a bounded piece can be, on the one hand, a sentence. But, on the other hand, it may be an interactive like *Wow!*, *Hi!*, or *Out!*. And we follow many other authors in extending the term ‘utterance’ beyond speech to written texts. Utterances have been said to contrast with sentences, e.g., in that the former are taken to relate to ‘language-behavior’ while the latter relate to the ‘language system’ (cf. [Lyons 1977](#): 239). On the perspective adopted in this book, utterances are understood to relate to the structure of linguistic discourse—be that spoken or written discourse, and interactives in some of their usages do qualify as utterances.

### 6.3 Do interactives form a system?

In a paper entitled ‘The Pragmatics of Interjections: The Case of Polish,’ published in the renowned *Journal of Pragmatics*, the author [Kryk \(1992\)](#) discusses only one linguistic element in detail, but that element is not an interjection but rather the Polish discourse marker *no*, similar in its functions to the English discourse marker *well*. Linguists are well aware that interjections, such as English *ouch* and *wow*, and discourse markers, such as *after all* and *indeed*, are quite different types of interactives, and so is Kryk herself, as she demonstrates in the course of her paper. In many studies, the two types, as well as other types, are on the one hand distinguished but on the other hand treated as the same, or as being closely interrelated: discourse markers have been called interjections, and interjections have been called discourse markers, etc. ([Ameka 1992a](#): 102).

What these observations suggest is that there seems to be some awareness of the underlying base shared by the different types of interactives. The question to be looked into in this section is whether such an underlying base, reflected, for example, in the set of properties that were identified in [Section 1.2](#), justifies calling interactives a system.<sup>4</sup> Note that, like ‘word,’ the term ‘system’ was avoided in the preceding chapters.<sup>5</sup>

<sup>4</sup> For example, [Eastman \(1992: 273\)](#), describes ‘interjections’ in the East African language Swahili ‘as elements of culturally specific extra-linguistic systems of communication in context.’

<sup>5</sup> Franck Floricic (p.c. of July 26, 2021) suggests treating interactive grammar and sentence grammar as ‘coexisting subsystems within the diasystem of a language.’

Systems have been conceived of in a number of ways. Fairly commonly they tend to be described generally as consisting of groups of interrelated entities serving some specific purpose. In order to establish that some phenomenon in fact qualifies as a system, it is in particular the questions in (2) that have been asked.

- (2) What is a system?
- a Is there a set of entities making up the system?
  - b Do the entities interact with one another?
  - c If there is a system, which purpose does it serve?
  - d How is the system set off from other systems or phenomena?

These questions are now looked at in turn.

### 6.3.1 Is there a set of entities making up the system?

The entities making up the category of interactives include, on the one hand, all the ten types distinguished in this study, as discussed in detail in Chapter 3, and, on the other hand, all the individual interactives of which the types are composed. While there may be additional interactives not discussed here, it is safe to assume that there is a specific set of entities of which the category is composed, as defined in [Section 1.2](#).

### 6.3.2 Do the entities interact with one another?

The various entities interact with one another in a number of ways. First, there are no clear boundaries separating the types or individual interactives from one another. Thus, we noted in the introduction to [Section 6.3](#) that in taxonomic treatments one and the same interactive has been allocated to two or more different types of interactives by different authors, or even by the same author. For example,

- response signals expressing confirmation ('yes') or disconfirmation ('no') are treated by some authors as discourse markers, but other authors, like [Fraser \(1990: 391\)](#), list them as interjections,
- the English forms *oh* and *okay* are treated by [Fraser \(1990: 391\)](#) as interjections but by others as discourse markers,
- the form *ahem* is treated by [Fraser \(1990: 391–2\)](#) as both an interjection and as a vocative,
- the form *sorry* is treated by some as a discourse marker (e.g., [Wichmann 2011](#)), by others as an interjection, but is classified here as a social formula.

Second, interactives share the same principles of placement. These principles are of two kinds: On the one hand, they concern the ability of essentially all types to form utterances of their own in specific usages, syntactically and semantically separated from the surrounding text. And on the other hand, there is a clear preference for most interactives to be placed at the left or right periphery of sentences.<sup>6</sup> Third, they can occur in the same paradigm. Thus, in Table 6.1 (reprinted from Table 1.2) there are nine types of interactives occupying the same placement slot at the left periphery of a sentence and being prosodically separated from the following text segment.

**Table 6.1** Examples of English interactives in the left periphery slot of an utterance

| Type of interactive | Example                                             |
|---------------------|-----------------------------------------------------|
| Attention signal    | <i>Hey</i> , there is no alien in the fridge.       |
| Directive           | <i>Listen</i> , there is no alien in the fridge.    |
| Discourse marker    | <i>After all</i> , there is no alien in the fridge. |
| Evaluative          | <i>Great</i> , there is no alien in the fridge.     |
| Interjection        | <i>Oh</i> , there is no alien in the fridge.        |
| Response elicitor   | <i>Right?</i> There is no alien in the fridge?      |
| Response signal     | <i>No</i> , there is no alien in the fridge.        |
| Social formula      | <i>Sorry</i> , there is no alien in the fridge.     |
| Vocative            | <i>Darling</i> , there is no alien in the fridge.   |

And fourth, interactives are interconnected by a network of grammaticalization. Rather than constituting neatly separated categories, we saw in Section 4.2.3 (Figure 4.1) that interactives, like attention signals, directives, discourse markers, interjections, and vocatives, can be described in terms of chains of grammaticalization, forming an interlocking network of related, contextually defined units. For example, in a number of languages, imperative forms for ‘look!’ have acquired a range of different interactive usages, occurring in some contexts as directives, in other contexts as attention signals, and in still other contexts as discourse markers. Such forms include but are by no means restricted to English *look*, Italian *guarda*, Spanish *mira*, French *regarde*, and Portuguese *olha* (Waltereit 2002: 997; Fagard 2010).

That there is an interlocking structure relating interactives to one another has been pointed out in some form or other by many authors. For example, Kryk (1992: 194) notes that Polish *no*, resembling the English discourse marker *well*, ‘cuts across the alleged border between discourse particles and interjections,

<sup>6</sup> Ideophones and response elicitors are notable exceptions (see Sections 3.5, 3.7).

indicating that we are dealing not with discrete categories, but with a cline (scale or squish), a phenomenon observed with other word classes as well.

### 6.3.3 If there is a system, which purpose does it serve?

While interactives serve a range of different discourse functions, there is an overarching purpose common to them, namely to express essential functions relevant to social communication in a holistic format. What [Pawley and Syder \(1983\)](#) write with reference to formulaic speech applies most of all to the functions of interactives: They ‘offer processing efficiency because single memorized units, even if made up of a sequence of words, are processed more quickly and easily than the same sequences of words which are generated creatively’ ([Conklin and Schmitt 2008](#): 75).

### 6.3.4 How is the system set off from other systems or phenomena?

The category of interactives is clearly set off from other linguistic phenomena, more specifically from sentence grammar: First, their functions are anchored directly to the situation of discourse rather than to the structure of sentences. Second, they are distinguished grammatically from sentences and other linguistic material by the definitional properties in [Section 1.2](#), being grammatically separated from the sentences they may be associated with. Third, whereas sentence grammar involves description, interactive grammar involves depiction ([Section 2.1.3](#)).

Fourth, interactives tend to exhibit phonetic and phonotactic features not found in sentence grammar ([Section 1.3.2.2](#)). And fifth, the category also differs from sentence grammar in showing a distinction between primary and secondary forms, as shown in Table 3.14 of [Section 3.6.4.1](#). There is no clear equivalent to this distinction in sentence grammar.

In sum, irrespective of whether one may wish to define sentence grammar as a system, it would seem that there are reasons for describing interactive grammar as a system on its own as distinct from sentence grammar. This is a position that has in fact been voiced in some form or other in studies of interactives. For example, [Eastman \(1992: 273\)](#) describes interjections and other interactives of the Swahili language ‘as elements of culturally specific extra-linguistic systems of communication in context.’ And [Dwyer and Moshi \(2003: 174–5, 179\)](#) propose with reference to ideophones a distinction between two ‘autonomous semiological systems,’ or between an analytic dimension and an expressive dimension of semiotic categories, where the former corresponds essentially to sentence grammar and the latter to interactive grammar.



## 6.4 Are interactives linguistic marginalia?

Both the term and an interest in linguistic marginalia can be traced back for the most part to [Joseph \(1997\)](#). Joseph argues that marginal, or peripheral phenomena, such as onomatopoeias, borrowings, particular phonetic features, or specific varieties of language use, are often hard to integrate into the broader, more far-reaching analyses, yet they have led to at least as many useful insights as the study of the ‘core’ alone has yielded. For a better understanding of the nature of language in general, Joseph concludes, phenomena located at the periphery of language structure and language usage deserve a more prominent place in linguistic debate and analysis.

Building on this study, [Ameka \(2020\)](#) argues that marginalia can contribute to linguistics most of all in two ways: On the one hand, they are needed to formulate generalizations and, on the other, they are of help in explaining certain aspects of language structure. Note that marginalia need to be distinguished from ‘rara,’ that is, typologically exceptional phenomena. Marginalia, by contrast, are not typologically exceptional but ‘many linguists think [they] can be ignored without harm to linguistic inquiry’ ([Dingemanse 2017b](#): 195–6). Ameka proposes to reduce the vast range of phenomena classified by [Joseph \(1997\)](#) as marginalia to a more restricted set of linguistic elements, namely the ones in (1).

Note, however, that Felix Ameka does not conceive of marginalia as having an ontological status of their own. Rather, he views them as a heterogeneous set of elements that linguists have placed on the margins of language structure—in other words, they are not meant to form a category comparable to that of interactive grammar (Felix Ameka, p.c. of March 6, 2022).

- (1) Linguistic marginalia proposed by [Ameka \(2020\)](#)
  - a Idioms
  - b Particles
  - c Interjections and vocalizations
  - d Swear words, taboo words
  - e Conversational routines
  - f Ideophones
  - g Manual gestures

As can be seen in (1), the elements classified by [Ameka \(2020\)](#) as marginalia exhibit significant overlap with interactive grammar. While this does not clearly apply to (1a) and (1b), all other kinds of elements relate to interactives. Interjections and ideophones are paradigm types of interactives (Sections 3.5, 3.6), and so are vocalizations in (1c), which, according to [Ameka \(1992a](#): 104), include the response signals *mhm*, *uh-huh*, and *yeah* (Section 3.8). Swear words and taboo words in (1d) also concern at least to some extent interjections, and the conversational routines in (1e) include social formulae (Section 3.9). (1g) is a feature of both sentence

grammar and interactive grammar but, as was pointed out in [Section 1.3.2.3](#), gesture is distinctly more prominent and presumably also to some extent of a different quality than in sentence grammar.

And much the same applies to the properties in (2) characterizing or defining the elements in (1). The properties do not fully match the definitional properties that were proposed in [Section 1.2](#) but none seems to be at variance with what can be observed in interactives, even if some properties also apply to elements of sentence grammar.<sup>7</sup>

- (2) The ‘value’ of peripheral elements ([Ameka 2020](#))
- a They are marked prosodically, phonologically, morpho-syntactically, and are semantically vague
  - b They are expressive, hence they show ‘distinctive phonological, morphological, and syntactic behavior’
  - c They are interactional tools as they help language users solve coordination problems in communication, as collateral signs, turn-taking management, and stance taking
  - d They provide clues for the understanding of the nature of (semiotic) signs as they tend to be composite utterances and be multimodally packaged
  - e They provide a window on mechanisms of language change and of language spread in contact zones

In sum then, there is massive overlap between interactives and the marginalia proposed by [Ameka \(2020\)](#). The question then is: Are interactives linguistic marginalia relative to the ‘core,’ or the ‘center’ of language structure? In other words, is interactive grammar located at or forms the periphery of sentence grammar—in accordance with the position sketched in [Section 6.2.1](#)? On the basis of the perspective adopted in this book, the answer proffered is in the negative, for the following interrelated reasons.

The first, and the most important reason is provided by the definition proposed in [Section 1.2](#). On this definition, interactives contrast with the units of sentence grammar in their grammatical structure: They are neither a semantic nor a syntactic part of sentences they may be associated with, and they are typically also set off prosodically from that sentence.

Second, in the same way as sentences, interactives can form utterances of their own as well as independent turns in conversational exchanges. Thus, the fact that they are self-contained pieces of text that consist either of a sentence or of an interactive would seem to be suggestive of a fair amount of mutual independence of the two.

<sup>7</sup> This concerns in particular (2e): While there are so far no appropriate data to allow for an overall comparison, it seems that language change and language diffusion are in much the same way accessible via sentence grammar as they are via interactive grammar (see [Section 6.5](#)).

Third, rather than being processed by the morphosyntactic machinery of sentence grammar, interactives are immediately available in the situation of discourse (cf. the direct-access hypothesis, [Section 2.1.3.3](#)). Accordingly, as has been pointed out independently in a number of studies, interactives do not describe meaning; rather, they enact, display, index or, as we say here, depict meaning.

Fourth, sentence grammar and interactive grammar provide interlocutors with two in some way equivalent alternatives for designing their discourse contributions. For example, a speaker may respond to a comment made by the previous speaker by saying either *You are absolutely right* or *Yes*, either *I feel disgusted* or *Yuck!*, either *There just was a gunshot* or *Bang!*, or either *I haven't clearly heard what you just said* or *huh?*, or either *I appreciate what you have done for me* or *Thanks*.

Fifth, compared to sentence grammar, interactive grammar provides a distinctly more economical tool for designing pieces of discourse, saving time and processing energy. For example, how to describe the experience expressed by an interjection like *phew*, an ideophone like *boing*, or a discourse marker like *well* in a concise way by means of the resources available in sentence grammar?

Sixth, sentence grammar constitutes an optimal device for conveying conceptual information. Interactive grammar, by contrast, constitutes a close to indispensable tool for structuring the space of social communication—a space that is widely believed to concern a central function of language usage. Without relying on the resources available in interactive grammar, people would be hard pressed to express salient functions of social interaction, like finding appropriate ways to initiate and maintain social contacts, being polite, greet, part, thank, apologize, bless, curse, congratulate, wish good health or recovery from sickness, etc.

Seventh, we saw in [Section 2.4](#) that in combinations of interactives with pieces of sentence grammar, the former tend to present foregrounded information while the latter tend to present backgrounded information.

Eighth, the two hemispheres of the human brain jointly process language and support one another in discourse processing. Sentence grammar crucially involves left hemisphere activity whereas interactive grammar implicates right hemisphere activity distinctly more than sentence grammar ([Van Lancker Sidtis 2004, 2009](#); [Heine, Kuteva, and Kaltenböck 2014](#); [Heine et al. 2015](#); [Section 5.3](#)).

And finally, research findings made in social psychology also appear to be in support of the distinction between the two grammars. As was shown in [Section 5.4](#), two types of reasoning and judgment are proposed in that research (e.g., [Kahneman 2012](#); see [Heine 2019](#) for an overview). Whereas Type 2 involves, for example, analytic, controlled, slow, and rule-based processes, Type 1 processes are holistic, automatic, rapid, and intuitive. This distinction shows significant parallels to that between the two kinds of grammar, in that Type 2 reflects features commonly associated with sentence grammar whereas Type 1 processes relate to features to be expected in interactive grammar.

To conclude, there are reasons not to treat interactives as marginalia of any kind; these reasons suggest, rather, that interactive grammar be best analyzed as a system within discourse grammar that is distinct from sentence grammar—in the wording of Franck Floricic (p.c. of July 26, 2021), interactive grammar and sentence grammar can be treated as ‘coexisting subsystems within the diasystem of a language.’

## 6.5 Interactives in language contact

Are interactives sensitive to contact-induced transfer of linguistic material from one language to another, be that transfer via borrowing or replication?<sup>8</sup> Felix Ameka (p.c. of July 25, 2021) suggests that the answer is in the affirmative. He observes, for example, that what we described in [Section 1.3.2.3](#) as the attention signal *kss...!* of the West African Ewe language, used to call someone’s attention over a distance, is not restricted to Ewe but is found across the languages in Ghana, and that social formulae are also prone to spread via cultural contact.

Research on discourse markers is in fact in full support of Ameka’s suggestion. In particular, it has demonstrated in detail in [Heine et al. \(2021, Chapter 7; see Section 3.3.6\)](#) that discourse markers can be borrowed on a massive scale. As is shown there, for example, the Spanish discourse marker *entonces* ‘then, therefore, thus’ and the Italian discourse marker *allora* ‘then, at that time’ have both been borrowed in a range of languages, and much the same applies to the Arabic discourse marker *yaa’ni* ([jæʕni] or [jaʕni]), to be found across three continents, as *ya’ni* ‘that is’ in Persian, *yaani* ‘that is, that is to say, I mean’ in Swahili, and *yani* ‘that is, in other words’ in Turkish. Furthermore, this research suggests that contact-induced transfer is by no means restricted to discourse markers but extends to other types of interactives.

This hypothesis is in accordance with observations that have been made in the relevant literature on other types of interactives. For example, in his discussion of three diachronic characteristics of ‘politeness formulas,’ that is, social formulae in the sense of the term as used here, [Ferguson \(1981\)](#) writes:

The third feature of diachrony to be noted is the strong tendency for the structure and use of politeness formulas to diffuse with other elements of culture across language boundaries. Thus, for example, a striking number of Arabic greetings and thank you formulas have spread along with Islam to speech communities which have not shifted to Arabic. In South Asia, where patterns of politeness formulas

<sup>8</sup> Borrowing means that form–meaning units, such as words, are taken from one language and inserted in another language. Replication, by contrast, involves the transfer of meaning and/or structure but not of form ([Heine and Kuteva 2005](#)).

are quite different from those in the Arab world, Muslim populations use Perso-Arabic formulas quite at variance with the usage of Hindu speakers of the same language.

(Ferguson 1981: 32)

In fact, what was observed in Section 3.3.6 on discourse markers applies presumably to interactives at large: They are invariable expressions, syntactically unattached, typically separated prosodically from surrounding text material, not being a semantic part of sentence structure, at the same time having a strong emblematic value. They are thus technically much easier to handle and to integrate in another language than elements of sentence grammar. For the learner of another language, therefore, interactives are most readily identifiable to be deployed for discourse processing.

To conclude, there can be little doubt that interactives are highly sensitive to language contact. However, beyond what was just observed on discourse markers, our knowledge about contact-induced transfer of interactives is still scant—there are, to our knowledge, hardly any substantial cross-linguistic contributions on this issue. Whether interactive grammar differs significantly from sentence grammar in its magnitude of contact sensitivity must therefore remain an open question at the present state of research.

## 6.6 Interactives and language acquisition

Another task for future research concerns the question of how interactive grammar relates to first language acquisition. During the period roughly between 9 and 17 months of age, children, or toddlers, develop the ability to design form–meaning units expressing complete, meaningful thoughts and having the appearance of self-contained utterances. Such units tend to be called ‘holophrases’ and the relevant period is referred to as the holophrastic stage, or the ‘one-word stage’ of language acquisition. The question then is whether such holophrases qualify as interactives.

Such units produced during the holophrastic stage are in fact in accordance with our definition of interactives in [Section 1.2](#). All evidence that we were able to access suggests, first, that the units are invariable and, second, that they are presumably deictic forms, being restricted in their functions and scope to the here-and-now of the situation of discourse. Third, the units are frequently described as utterances of their own and therefore are by default set off semantically, syntactically, and prosodically from any surrounding material. And finally, we are not aware of any clues to suggest that a child at the holophrastic stage is able to negate the holophrase produced nor to turn it into a question.

In addition to these grammatical features there are also semantic cues that may be taken to lend support to the hypothesis that units produced by children during

the holophrastic stage can be classified as interactives. For example, the data presented by Clark (2003: 82, Table 4.2) show that among the forms ('early words') uttered by English-speaking children there are some forms which are suggestive of ideophones (*vroom*), interjections (*ouch*, *yum yum*), response signals (*no*), social formulae (*bye*, *hi*), and vocatives (*mommy*, *daddy*). Furthermore, a number of units, like *banana*, *cookie*, *juice*, etc. would seem to have, at least in some of their uses, the function of directives relating to the children's requests.

These observations indicate that child language provides a promising field of study for understanding the nature of interactives. But such a study is beyond the scope of the present book, especially since it would be faced with a number of problems. For one thing, not all of the holophrases produced by children are immediately suggestive of functions commonly expressed by interactives. For example, in addition to the forms mentioned there is a range of other forms figuring in holophrases of young English-speaking children, many of which are hard to reconcile with functions associated with the category of interactives. Such forms capture what is going on around the children, they include terms for animals (e.g., *dog*, *kitty*), body parts (e.g., *eye*, *nose*), clothing (e.g., *hat*, *shoe*), household objects (e.g., *bottle*, *keys*), toys (e.g., *ball*, *boat*), vehicles (e.g., *car*, *truck*), etc. These forms are quite commonly polyfunctional, expressing a number of different meanings some of which may have little in common with the meaning of their 'lexical sources.'

And for another, in spite of all the work that has been done on first language acquisition it remains at times unclear what exactly the communicative intents of children are when producing some specific holophrase. For example, the form *door* was found to be used consistently for opening boxes, taking lids off jars, removing nuts from bolts, and taking clothes off dolls (Clark 2003: 87). But what exactly is the communicative goal of the child when using this form in a given context: Does the child want the object concerned to be manipulated, does it want to get access to the object, or does it simply wish to indicate the existence of the object or the relevant activity?

To conclude, in spite of the potential that first language acquisition offers, there remain problems to be looked into by the student of interactives before tackling this subject matter.

## 6.7 Are interactives fossils of language evolution?

Interactives, especially interjections and ideophones, have figured in various studies on the origin of human language or languages, such studies having a long tradition going back to Jean-Jacques Rousseau and beyond (see D'Atri 1995). Both their grammatical form and their functions set interactives off from sentence grammar and have inspired researchers to build hypotheses on earlier states of language evolution—for Rousseau (1990: 71), 'protolanguage' was 'entirely interjectional.'

A paradigm example is provided by Jackendoff (1999, 2002), who argues that certain design features of modern languages are suggestive of features of early layers of language evolution. Such features include primary interjections like *ouch* and *oops*, the secondary interjections *dammit* and *oboy*, the directive *shh* (*ssh*), the attention signals *hey* and *pst* (*psst*), the response signals *yes* and *no*, and the social formulae *goodbye*, *hello*, and *thanks*. These items are described by him thus (see also Jackendoff 2002: 131–2):

- (a) They have phonology and semantics but no syntax
- (b) They can be used on their own as meaningful utterances
- (c) They cannot be combined with other words into sentences.

Being ‘single-word utterances,’ these items are claimed to be ‘fossils’ of the one-word stage of language evolution (Jackendoff 2002: 240). A more recent and detailed hypothesis that can be reconciled with that of Jackendoff is proposed by Haiman (2018: 216, 263), whose main concern is with ideophones and the passage from ‘dramatic showing’ to ‘narrative telling’ and from gestures to words. Dealing with a wider range of interactives, Kuteva and Heine (2021) argue that language evolution must have proceeded gradually from a holistic mode, consisting of expressions such as interjections (*wow*, *ouch*), ideophones (*bang*, *splish-splash*), social formulae (*hello*, *sorry*, *goodbye*), etc., to an analytic mode as it can be observed in sentence grammar.

The findings presented in the course of the preceding chapters would, in fact, seem to lend support to hypotheses of this kind, and in this section the main observations to this effect made in the book are summarized, namely the ones in (3).

- (3) Features that may be suggestive of an archaic structure as suggested by interactives
  - (a) Interactives are universal to the extent that they can be expected to be found in some form or other in any language
  - (b) They contain a set of primary interactives, whose origin is opaque
  - (c) They form a system separate from sentence grammar
  - (d) They have a holistic rather than an analytic structure
  - (e) They are indexicals
  - (f) They lack conceptual displacement
  - (g) They cannot normally be negated
  - (h) The meanings expressed by them concern most of all the attitudes of the speaker and speaker–hearer interaction
  - (i) Rather than describing meaning, they depict meaning
  - (j) They contain a set of conventional expressions dedicated for coding stancetaking
  - (k) They may consist of clicks
  - (l) They may exhibit sound symbolic features

These features can be taken to form generalizations on the behavior of interactives, even if some of the features are in need of more detailed empirical evidence.

(a) Interactives are universal to the extent that they can be expected to be found in some form or other in any language

There are languages that lack categories of sentence grammar, such as adjectives or articles. This is different with interactive grammar: While so far only a portion of the world’s languages have been documented, we are not aware of any well described language that would lack interactives such as interjections, discourse markers, social formulae, or ideophones.

(b) They contain a set of primary interactives, whose origin is opaque  
As we saw in Section 3.6.4.1, primary interactives are expressions which, unlike secondary interactives, are not transparently related to any form of sentence grammar. For example, English *shh* is a primary interactive, whereas *damn* is a secondary interactive since it is transparently related to the verb *damn*. Table 6.2, based on Table 3.14 of Section 3.6.4.1, contains a list of English primary interactives.

Table 6.2 Examples of English primary interactives

| Type               | Primary forms                  |
|--------------------|--------------------------------|
| Attention signals  | <i>hey, oh, oi, pst, yo(o)</i> |
| Directive          | <i>shh</i>                     |
| Discourse markers  | <i>uh, um</i>                  |
| Ideophones         | <i>bang, plop, thud</i>        |
| Interjections      | <i>oh, ouch, wow, yuck</i>     |
| Response elicitors | <i>eh?, huh?</i>               |
| Response signals   | <i>mhm, uh huh, unh unh</i>    |

The history of primary interactives is as a rule beyond the scope of established methodology of historical linguistics, that is, it cannot be reconstructed back to any earlier diachronic states.

(c) They form a system separate from sentence grammar  
Throughout this book the term ‘system’ was avoided. As we saw in Section 6.3, however, there are reasons to describe interactives as a system of its own in which primary interactives form a sub-system.

(d) They have a holistic rather than an analytic structure  
In accordance with the definition in Section 1.2, interactives are holistic in the sense that they are invariable forms, having no internal analytic structure and being separated from the surrounding text (see also Kuteva and Heine 2021).



## (e) They are indexicals

This feature is captured by the definitional property in [Section 1.2](#): The usage of interactives is restricted to the here-and-now of the situation of discourse.

## (f) They lack conceptual displacement

As was observed in [Section 1.3.2.6](#), this feature is derivative of (e), that is, of the indexical status of interactives. Displacement is an inherent property of sentence grammar, enabling interlocutors to move from the here-and-now of the situation of discourse to refer to persons, places, times, and worlds that are outside of that situation.

Displacement thus is a central feature distinguishing sentence grammar from interactive grammar. Whereas the former disposes of a wealth of lexical, grammatical, and syntactic means to refer to the world outside the situation of discourse, displacement is essentially absent in interactive grammar.

This is an observation that is of immediate interest for research on language evolution. Since [Hockett \(1960\)](#), displacement has been argued to distinguish human languages from other kinds of communication systems, even if forms of displacement have also been observed in some animal species like bees and ants. Its absence in interactive grammar may therefore be suggestive of a stage of evolution before early humans acquired the ability to take communication out of the here and now—that is, a stage prior to the emergence of sentence grammar (see, e.g., [Bickerton 2009](#): 217).

## (g) They cannot normally be negated

This is another definitional property of interactives, discussed in [Section 1.3.1.6](#).

## (h) The meanings expressed by them concern most of all the attitudes of the speaker and speaker–hearer interaction

Most central to the functions of interactives is to draw attention to what is conceived to be important or relevant to the hearer, expressing the speaker's emotive or cognitive state, and depicting phenomena of the environment, and to interact with other people. What is blatantly absent is the ability to analytically describe conceptual discontinuities.

## (i) Rather than describing meaning, they depict meaning

This feature was discussed in [Section 2.1.3](#). While they also have a descriptive component in addition to their expressive meaning, interactives were interpreted as vocal depictions. Thus, they resemble performative physical depictions ([Clark 2016](#)), the difference is that the former draw on interactives while the latter rely on gesture and other kinds of physical activity rather than on linguistic material.

(j) They contain a set of conventional expressions dedicated to coding stancetaking

Stancetaking, as described by Du Bois (2007) invokes a shared framework for co-action with others (Du Bois 2007: 171). It includes three main concepts, namely alignment, evaluation, and positioning. Interactives present fully conventional expressions for all three kinds of concepts of stancetaking. For example, there are response signals for ‘yes’ and ‘no’ expressing alignment, evaluatives like ‘excellent’ or ‘horrible’ expressing evaluation, and interjections like ‘wow’ and ‘yuck’ expressing functions of positioning, where positioning may be either affective (‘phew’) or epistemic (‘aha’).

(k) They may consist of clicks

This feature was mentioned in Section 1.3.2.8. As was noted there, clicks serving as conventionalized form–meaning units are argued to be amongst the earliest linguistic achievements in the evolution of human languages (Poulos 2021). Click consonants are massively found in the phoneme systems of about thirty languages of southern and eastern Africa. But the clicks occurring in interactives are of a different kind: For one thing, click-only form–meaning units appear to be essentially absent in the African click languages just mentioned, but in interactive grammar they constitute meaningful forms rather than being parts of meaningful forms. And for another, those click languages have phonemic click consonants which are usually complex, being composites of click features, their accompaniments plus vowels and/or consonants. ‘Naked clicks’ in contrast, that is, clicks lacking distinctive phonetic accompaniments and other phonetic material but forming meaningful units in and of themselves, have so far been found only in interactives.

Click-only interactives are fairly common. They occur in particular in expressions for the response signals for ‘yes’ and/or ‘no’ in a substantial number of languages across the world (Section 3.8.4; Gil 2013) as well as in interjections (Section 3.6). Furthermore, they are widespread as form–meaning units in animal directives, that is, in calls used in the interaction with animals (Section 3.11.5).

(l) They may exhibit sound symbolic features

While (l) does not apply generally to interactives, we saw in Chapter 3 that, across languages, some types show significant correlations between their function and the phonetic substance employed for their expression. Sound symbolic features were observed in particular in attention signals (Section 3.1.5), in a subtype of response elicitors, called other-initiated repair markers (Section 3.7.5), and in animal dispersal calls, that is, a class of animal directives (Section 3.11.6). In addition, we found sound symbolic features also in response signals (Section 3.8.4).

The generalizations proposed in this section suggest, first, that interactives seem to qualify as universals of linguistic activity and, second, that they might belong to the earliest achievements in human language evolution, as has been claimed

independently in a number of studies. But pursuing this issue further here would be beyond the scope of the present book.

## 6.8 Conclusions

The chapter was concerned more generally with the status of interactives, and most of all with questions that have plagued previous research. Such questions relate to the place of interactives in the overall space of grammar and discourse, to whether interactives qualify as ‘words,’ as a ‘system,’ or as ‘marginalia.’ Furthermore, we were briefly concerned with how interactives behave in language contact and with their relevance in first language acquisition. Finally, [Section 6.7](#) summarized some general properties of interactives, suggesting that these properties might reflect an archaic design feature of human language evolution, but no attempt was made to relate these observations to existing hypotheses on language genesis.

## Deciding between two grammars

The preceding chapter had a focus on the ontological status of interactives as a category, or a system. The main interest of the present chapter is with exploring what the presence of two competing grammars means to a speech community with reference to the way discourses are built. In more general terms, the two grammars can be related to the distinction between two modes of processing information made by [Van Lancker Sidtis \(2004\)](#), where one mode involves novel or propositional speech while the other involves formulaic, automatic, or non-propositional speech:

Each mode, propositional and non-propositional, places different processing demands on speech production and comprehension. Production of propositional expressions requires lexical retrieval and arrangement according to grammatical rules; non-propositional production involves activating and retrieving prepackaged units or schemata.

([Van Lancker Sidtis 2004](#): 29)

It is the following questions that are the main subject of this chapter:

- (a) How is interactive grammar to be distinguished from sentence grammar?
- (b) What is the contribution of each of the two grammars in structuring discourses?

Question (a) is the subject of Section 7.1, while question (b) is addressed in Section 7.3. The questions relate to issues that have been discussed widely within the tradition of research on competing motivations (e.g., [Haiman 1983](#); [Du Bois 1985](#); [MacWhinney, Malchukov, and Moravcsik 2014](#) and the contributions therein). A wide spectrum of possible motivations has been identified in that research. That the two grammars provide interlocutors with contrasting tools for text organization has been argued in the framework of discourse grammar ([Kaltenböck and Heine 2014](#)). Another question central to the present chapter therefore is whether interactive grammar and sentence grammar can in fact be said to compete with one another in structuring discourse.

7.1 Distinguishing features

In previous research a number of features were identified on how interactive forms differ from words, phrases, and clauses of sentence grammar. As we saw in Sections 5.4 and 5.5, these features show striking similarities to those identified by students of social psychology in distinguishing two kinds of decision-making processes (cf. Kahneman 2012: 28), and the distinctions listed in Table 7.1 have been informed by findings made in that research paradigm.

Table 7.1 Features prototypically distinguishing sentence grammar from interactive grammar

|   | Parameter                                        | Sentence grammar                 | Interactive grammar          |
|---|--------------------------------------------------|----------------------------------|------------------------------|
| a | <i>Conceptual space</i>                          | Ideational                       | Social                       |
| b | <i>Meaning</i>                                   | Descriptive                      | Expressive and depictive     |
| c | <i>Mental representation</i>                     | Verbatim                         | Gist                         |
| d | <i>Form</i>                                      | Analytic                         | Holistic                     |
| e | <i>Speed of processing</i>                       | Slow                             | Fast                         |
| f | <i>Sequence of activation</i>                    | Typically activated subsequently | Typically activated first    |
| g | <i>The locus of semantic-pragmatic anchoring</i> | Sentence                         | Situation of discourse       |
| h | <i>The role of context</i>                       | Highly decontextualized          | Highly context-sensitive     |
| i | <i>Constraints on usage</i>                      | Follows morphosyntactic rules    | Follows pragmatic principles |
| j | <i>Use of gesture</i>                            | Moderate                         | Pronounced                   |

The features listed in Table 7.1 are now looked at in turn.

7.1.1 Conceptual space

Both grammars presuppose a speaker and a hearer, but the communicative objectives are not the same, and in accordance with their respective objectives, each of the two grammars has its own way of classifying the universe of experience. In sentence grammar, meaning serves to talk about how we experience the world—how it is construed and classified in accordance with conceptual discontinuities to be observed in the world, including phenomena not immediately accessible to the senses.

In interactive grammar, by contrast, meaning is interactive, the universe is perceived and conceptualized with reference to experiences located in the situation of

discourse, and more specifically in speaker–hearer interaction, the attitudes of the speaker, and the way social interaction shapes linguistic discourse. Accordingly, interactives are tailored to the needs arising in situations where people meet, greet, wish to be polite, apologize, bless, congratulate, pray, curse, expose their feelings, impact on others, etc.

### 7.1.2 Meaning

As observed in Section 2.1, authors concerned with semantic differences between sentence grammar and interactive grammar have come up with four generalizations: (a) Forms of sentence grammar are claimed to be sensitive to truth conditions while interactives are not. (b) The former are said to describe, whereas the latter depict meaning (Section 2.1.2). (c) Whereas meaning in sentence grammar is determined by the structure of sentences, that of interactive grammar is anchored to and determined by the situation of discourse. And (d) and relatedly, representation relies on description in sentence grammar but on depiction in interactive grammar (Section 2.1.3).

The claim in (a) is not entirely unproblematic though, considering for example that there is a body of literature showing that not all forms of sentence grammar are amenable to truth conditions. (b) and (c), by contrast, capture important generalizations, as shown in Section 2.1. A sentence like *I am in pain* can well be said to be a description of the bodily state I am in. An expression like *Ouch!*, by contrast, expresses an emotive state (Section 2.1.2), and it is said to represent a depiction rather than a description of that state (Section 2.1.3). Rather than drawing on the structure of a sentence, expressions like *Ouch!* are activated without being processed by the machinery of sentence grammar (cf. the direct-access hypothesis, Section 2.1.3.3).

Interactives are therefore conceived as expressive depictions. But, as was argued in Section 2.1, they have a second kind of meaning content in addition, which in the case of *Ouch!* can be paraphrased roughly as ‘I am in pain’ and is descriptive in nature. In this sense, thus, the meaning of interactives is more complex than that of sentence grammar units, combining two semantic components.

### 7.1.3 Mental representation

In the fuzzy-trace theory of memory and reasoning of Brainerd and Reyna (1992, 2002) and Reyna and Brainerd (1995), an elementary distinction is made between two types of information or mental representation, namely between *verbatim* and *gist* information; see Section 5.2.1 for a more detailed account. Verbatim information contains exact details of memory and reasoning, such as

distinguishing the individual parts of a pattern. Information encoded as gist, by contrast, deals with patterns as wholes.

This distinction echoes that between the two kinds of grammar proposed here. The concern of sentence grammar is with a verbatim representation of information in the form of phrases, clauses, and sentences, analyzing propositional content as being composed of interrelated parts such as morphological, syntactic, and semantic elements. Interactive grammar, in contrast, represents conceptual information as *gestalts* in the form of unanalyzable wholes. The English interjection *Ouch!*, for example, can be argued to constitute a gist representation, forming an unanalyzable whole, whereas its paraphrase *I am in pain* can be interpreted as a verbatim representation of roughly the same content in the form of a series of morphological, syntactic, and semantic components, having the propositional format of a sentence.

#### 7.1.4 Form

Information encoded by means of sentence grammar is analytically structured, being composed by means of morphosyntactic principles, processed in a linear format, and determined by deliberate semantic and syntactic planning. Interactives, by contrast, express at times complex contents of meaning in the form of holophrases—that is, ready-made compact, unanalyzable wholes. These wholes have been characterized as more or less ‘automatic’ entities (see, e.g., [Wierzbicka 1992](#): 188 on interjections, or [Ferguson 1976](#): 24 on social formulae).

This distinction might have a parallel in processes of decision making as proposed in the work of some social psychologists (Sections 5.4 and 5.5), where Type 2 (System 2), showing parallels to sentence grammar, is said to be analytically structured, being slow, deliberate, effortful, laboriously checking the facts ([Kahneman 2012](#): 28), operating on symbolic structures and aiming to describe the underlying logical and causal structure. Type 1 (or System 1), by contrast, showing features of interactive grammar, is described as holistic, being fast, intuitive, associative, automatic, and impressionistic ([Evans and Over 1996](#)). The distinction mirrors that made in fuzzy-trace theory between verbatim and gist encoding (respectively [Brainerd and Reyna 1992, 2002](#); [Reyna and Brainerd 1995](#)); see Section 5.2.1.

#### 7.1.5 Speed of processing

Compared to text pieces of sentence grammar, interactives are easy to process and save time in encoding and decoding. Interactives are included in what [Pawley](#)

(1992) calls ‘speech formulas,’ and what he says about the latter also applies to interactives:

In the production of extended discourse, formulas are essential building blocks: ready-made units which free speakers and hearers from the task of attending consciously to each word. Thus freed, they are able to focus on the larger structure and sense of the discourse, or on nuances of wording or sound.

(Pawley 1992: 23)

Being formulaic expressions, interactives can be assumed to be processed more quickly than corresponding nonformulaic phrases of sentence grammar, thus offering a clear advantage in processing efficiency. Especially, in spoken discourse that occurs under heavy time constraints, formulaic language in general and interactives in particular appear to be preferred over freely combined sequences of sentence grammar (Conklin and Schmitt 2008: 72–3); see also Section 2.1.3.3.

### 7.1.6 Sequence of activation

Being syntactically independent from one another, sentence grammar can form utterances without involving any interactives, and vice versa. But when the two co-occur, as they commonly do, interactives are most likely to precede (e.g., *Okay, let’s go back to work*).<sup>1</sup> This arrangement can be called the interactive-first collocation pattern, whereby the interactive is placed at the left periphery of an utterance to set the frame or to present the focal point to be filled by the following utterance or larger text piece presenting a comment on the use of the interactive.

Similar observations were made in social psychology with reference to two types of decision-making processes, as we saw in Sections 5.4 and 5.5: There is a division of labor between the ‘fast’ Type 1 (or System 1), corresponding to interactive grammar, and the ‘slow’ Type 2 (System 2), corresponding to sentence grammar, in that the former usually precedes, if not triggers Type 2, or even is a prerequisite for Type 2 to be activated.

At the same time there is also a less common opposite pattern of activation which can be referred to as the interactive-last collocation pattern (e.g., *He was your first husband, right?*), where the interactive modifies or elaborates on a preceding piece of sentence grammar or places it in a wider context or is intended to involve the hearer.

<sup>1</sup> Response elicitors and ideophones are notable exceptions to this preference (see Sections 3.5, 3.7).



### 7.1.7 The locus of semantic-pragmatic anchoring

In sentence grammar, the main locus of semantic-pragmatic anchoring is typically provided by the sentence. Interactives, by contrast, are located in the situation of discourse, and compared to forms of sentence grammar, they are therefore highly context-dependent. As [Kahneman \(2012: 24–\)](#) writes with reference to decision-making processes, Type 1 (System 1) ‘is highly sensitive to subtle environmental cues,’ and the same applies to the usage of interactive grammar. Interactives do not allow for context-free semantic description: They are inherently entrenched in the pragmatic environment in which they operate.

#### 7.1.8 The role of context

The differential role played by context is highlighted in some work of social psychologists on reasoning and decision-making processes (e.g., [Frankish 2010: 922](#)), where it is suggested that Type 2, which is suggestive of sentence grammar, is ‘decontextualized’ while Type 1, which relates to interactive grammar, is contextualized. Similar observations have been made in some linguistic studies. [Kita \(1997: 379\)](#), in particular, identifies two kinds of semantic dimensions in his analysis of Japanese ideophones (‘mimetics’). Distinguishing between an analytic and an affecto-imagistic dimension, he argues that in the former dimension, namely that of sentence grammar, meaning is represented as a hierarchical structure of decontextualized semantic primitives. In the affecto-imagistic dimension of ideophones, by contrast, meaning is contextualized in terms of affect and various kinds of imagery, and it relates to what is called here the expressive function of language (see Section 2.1.2).

#### 7.1.9 Constraints on usage

For a speaker of English, using words like *tree* or *eat* requires a knowledge of which word class they belong to, which word classes they may combine with, what their syntactic function is, where they can be placed, and what kind of affixes they may take—their usage being constrained by morphosyntactic rules and norms. Not being processed by the machinery of sentence grammar, interactives are largely free from such constraints (cf. the direct-access hypothesis, Section 2.1.3.3). And being expressives, they have a remarkably unconstrained potential for creativity:

These observations are consistent with the semantics of expressives as [they] call for individual creativity and usage of expressives, which stands in direct

opposition to the semantics of analytic language which calls for the fixed relationship between the signified and signifier in a closed system.

(Dwyer and Moshi 2003: 179)

Rather than following morphosyntactic rules, the usage of interactives is guided by how a discourse is globally designed, and how they best contribute to structuring that discourse, being preferably placed at the periphery of an utterance, and they may form utterances of their own.

### 7.1.10 Use of gesture

An observation made in many studies of interjections, ideophones and some other interactives is that gesture, that is, movement of the body, hands, arms, head, and/or eyes, is a central component in the usage of interactives, the two frequently forming depictive composites (Section 2.1.3). In Section 1.3.2.3, a voice-gesture model was proposed according to which interlocutors have the choice between usage of an interactive, a gesture, or a combination of both in many situations.

To be sure, text pieces of sentence grammar are also frequently accompanied by gestures, and this applies especially to deictic adverbs and pronouns. But as the relevant literature on this subject suggests, interactives rely distinctly more heavily on gesture than text pieces of sentence grammar. For example, Haiman (2018) notes:

Recall for now our English examples – the ‘interjection’ *yuk* or any of its synonyms is typically, and easily, accompanied by the speaker’s favorite facially pantomimed expression of disgust: the ‘interjection’ *brrrrr* by a pantomimed shudder and/or wrapping one’s arms about oneself as if to shelter oneself from the cold. The same kind of observation can be made about the ideophonic performance of negation in *unhunh*.

(Haiman 2018: 93)

To conclude, each of the two kinds of grammar provides interlocutors with contrasting means of voicing their communicative intents, allowing them either to choose between the two or, more commonly, to combine both for structuring their discourses.

## 7.2 Types of discourse

In the present section, a few salient types of conversational discourse are discussed. We are restricted to a rough overview of usages as they surface in texts, and exemplification is restricted to text pieces suggestive of alternative ways in which the

two grammars combine in structuring discourse. With regard to the relative contribution made by the two types of grammar, the three prototypical situations in (1) can be singled out.

- (1) Three types of combining interactive grammar with discourse grammar
  - a Usage of interactive grammar is predominant
  - b Both grammars are equally important
  - c Use of sentence grammar is predominant

### 7.2.1 Usage of interactive grammar is predominant

Parts of conversations consisting mostly or entirely of interactives are not hard to come by. What characterizes such parts is that the amount of conceptual information contained is generally poor; rather, the main concern of interlocutors in this kind of texts appears to be to establish, maintain, or improve social relations.

A common example of this kind of situation is provided by short encounters of people, for example, in telephone exchanges or in meetings on the road, where communication focuses on exchanging greetings and related formulae. The following exchange in Italian illustrates such a situation: With the possible exception of the piece *ci si sente* 'we will speak again,' the exchange consists of interactives only, namely the interjection *ahah*, the discourse markers *insomma* and *va bene*, and the social formula *ciao*.

- (2) Italian, conversation (Frank-Job 2006: 371, (19))

B: **ahah** **Giovanna insomma** *ci si sente*  
 'ahah Giovanna, so, we will speak again'

A: **va bene**  
 'okay'

B: **va bene**  
 'okay'

A: **okay**  
 'okay'

B: **ciao ciao**  
 'bye, bye'

A: **ciao**  
 'bye'

The text piece in (2) appears to be taken from a telephone conversation in Italian. An example of the opening section of a paradigm telephone conversation in English (already mentioned in Section 1.1) is presented in Table 7.2. It consists entirely of three interactives, namely the social formulae *hello* and *hi* and the interjection *oh*. For detailed analysis of the greeting ritual in Table 7.2 based on the

**Table 7.2** Example of an opening section in English telephone conversations

| Turn | Interlocutor | Text                                         | Function                                              |
|------|--------------|----------------------------------------------|-------------------------------------------------------|
|      | C:           | (C causes telephone to ring at R's location) | Summons                                               |
| 1    | R:           | <i>Hello</i>                                 | Answer + display for recognition                      |
| 2    | C:           | <i>Hi</i>                                    | Greetings, first part: Claim that C has recognized R  |
| 3    | R:           | <i>Oh hi::</i>                               | Greetings, second part: Claim that R has recognized C |

Source: Levinson (1983: 311–12)

framework of conversation analysis, see Levinson (1983: 311–12); salient features of the ritual are summarized by him thus:

- (a) telephone (and other related) conversations begin with summons–answer pairs; (b) reciprocal greetings are relevant at the very beginning of calls; (c) also at the very beginning of calls, recognition (or identification) is a prime concern.  
(Levinson 1983: 312)

Similarly detailed rituals can be observed in the closing section of telephone conversations, which in English tend to be dominated by pre-closing formulae like *okay*, *alright*, *so::*, etc., followed by terminal formulae like *bye*, *righteo*, *cheers*, etc. (Levinson 1983: 317).

The text piece in Figure 7.2 may give an impression of the rich communicational content conveyed by the sequential arrangement of three interactives in a conventional situation.

A number of other kinds of interactive-only settings were mentioned in the preceding text of this book. Suffice it to cite the following excerpt of a conversation in the Mayali language of Western Arnhem Land, Australia, that we saw already in Section 1.1. It consists of four instances of interactives ('interjections'), namely an attention signal or social formula (*boi!*), a directive (*njudj!*), and two instances of the interjection *njonj-njonj!*:

- (3) Mayali (Gunwinyguan, Arnhem; Evans 1992: 227, (3); repeated from Section 1.1, (2))

A: **Boi!**

Hey, come here!

[Old woman to youth:] 'Hey, come here!' [Old woman holds up baby]

- B: **Njonj-njonj!**                      **Njudj!**                      **Njonj-njonj!**  
 What a little sweetheart!    Blow your nose!    What a little sweetheart!  
 ‘Isn’t she a little sweetheart! Blow your nose! Isn’t she cute!’  
 [...]

Another kind of setting concerns specific social functions expressed by interactives, such as initiating social contacts, congratulating, well-wishing, or apologizing. But there also is a range of other settings in addition. In example (4) from the Siwu language of southeastern Ghana (Dingemanse 2011: 156), the setting is one in which two people are working together squeezing palm fruit fiber in a net to produce palm oil. At one point, a bit of oil squirts out onto B’s arm, and a brief exchange ensues. The following extract reproduces the first part of the exchange.

- (4) Siwu (Kwa, Niger-Congo; Dingemanse 2011: 156, Extract 6.1)  
 [Some oil squirts onto B’s arm]  
 A: **oh, ñdzo ló**  
     INT    pardon    FP.ADV  
     ‘Oh, sorry!’  
 B: **oh!**  
     INT  
     ‘Oh! (No worries)’

This exchange once again consists entirely of interactives: *oh* is an interjection, *ñdzo* is a social formula expressing apology, and *ló* is a discourse marker.<sup>2</sup>

## 7.2.2 Both grammars are equally important

Compared to interactive-only text pieces, a distinctly more common kind of conversational exchange involves combinations of sentence grammar and interactive grammar. An English example of this type is the following: There are four interactives (*o:h*, *ye:s*, *oh*, *boy*) and four words of sentence grammar (*remember*, *was* *I soaked*).

- (5) English (SBCSAE 0011; Norrick 2009: 879)  
 Doris: re[member]?  
 Sam: [o:h] ye:[s].  
 Doris: [oh boy] was I soaked.

A patterning like in (5), where a series of interactives is framed by pieces of sentence grammar is not typically expected at the beginning or the end of a

<sup>2</sup> Dingemanse (2011: 156) glosses *ló* ‘FP.ADV’ (advise (lo)). The data contained in his study strongly suggest that *ló* is an utterance-final discourse marker (see Heine et al. 2021: 6, (7) for a definition of discourse markers), translated by Dingemanse (2011) by ‘mind you!’ or ‘I advise you.’

conversation, unless in highly specific contexts. In bounded encounters of interlocutors, the expected patterning is more likely to be the one in (6).

- (6) A common three-stage structure of short conversational exchanges
  - a Opening: interactives, most of all social formulae and vocatives
  - b Exchange of conceptual information: sentence grammar text
  - c Closure: interactives, most commonly social formulae

(6a) and (6c) can jointly be referred to as a *framing structure*, signaling the opening and the closure of exchanges. The shortened telephone exchange in (7) is an example of such a conversation: It is framed by the social formula *pronto?* as an opener and the discourse marker *va bene* and the social formula *grazie* in the concluding contribution. After the opening there is an exchange of conceptual information. At the core of the exchange is B's request for information (*c'è Paolo?*) and A's complying with the request (*Paolo è uscito ha detto che tornava verso le sei*). Both are text pieces of sentence grammar, but the second piece is introduced by an interjection (*eh*) and a response signal (*no*).

- (7) Italian, telephone conversation (Frank-Job 2006: 367, Table 1, 1–4)

A: **pronto?**

Hello?

B: <?> c'è Paolo?

'<?> Is Paolo there?'

A: **eh no** Paolo è uscito ha detto che tornava verso le sei

'Ah, no, Paolo's gone. He said that he would be back around six'

B: **va bene grazie**

'Okay, thank you'

A paradigm case of a division of labor between the two grammars can be seen in cases where one of the two interlocutors involved is restricted to interactives. The mechanism most commonly employed is *backchannel* (Yngve 1970), where the role of speaker A is largely restricted to providing conceptual information while speaker B reacts to that by means of response signals or discourse markers used as backchannels, typically signaling that B has understood and agrees with A, and/or encourages A to continue. Thus, in the following brief exchange in British English between the three interlocutors A, B, and C, it is only A who provides conceptual information, whereas B and C are restricted to the backchannel forms *mm* and *yeah*:

- (8) British English (Biber et al. 1999: 1091)

A: and their lawn is high up, up the road, [**you know**]

B: **Mm**

C: **Yeah**

A: so he dug up the lawn and put a little garage in.

A similar example can be seen in the following exchange in Italian, where speaker A presents the main conceptual information, merely interspersed with an interjection (*eh*) and a few discourse markers (*adesso*, *comunque*, and *insomma*), while the role of B is restricted to the response signals *mh* and *sì* as backchannels:

- (9) Italian, conversation (Frank-Job 2006: 363, (4))

A: [...] **comunque** se vuole far la terapia magari  
 ‘[...] well, if you want to do the therapy perhaps’

B: **mh**  
 ‘mh’

A: **eh** chiama\_ **insomma adesso** ci pensa un attimo e vede un pochino  
 ‘eh call\_ okay, now, think about it for a moment and look it over  
 a little bit’

B: **sì**  
 ‘yes’

Backchannels, also called continuers, are by no means restricted to response signals, they may involve a number of other types of interactives. Take the following conversation, in which Brianne provides the conceptual information while Addie’s is restricted to backchanneling, using a primary interjection (three instances of *oh*), two secondary interjections (*jeez*, *god*), and only one response signal (*no*).

- (10) English (SCoSE, Addie and Brianne; Norrick 2009: 886)

Brianne: it took me an hour to get out to O’Hare.

Addie: **oh [jeez]**,

Brianne: [and then] I was standing on the aisle for an hour.  
 ‘cause it’s packed.

Addie: **oh no.**

Brianne: and-

Addie: **oh [god.]**

Brianne: [then-] **okay**, we left on time.

### 7.2.3 Use of sentence grammar is predominant

There are many areas of discourse where interactives play hardly any or no role. This applies most of all to written descriptive texts. Academic discourse constitutes one of the areas in which interactives tend to be poorly represented. But this applies to written discourse, not necessarily to spoken discourse. Take the following case

of a public lecture by the well-known philosopher David Kaplan—a seminal contribution to semantic theory which has also informed the present study: In his Howison Lecture in Philosophy that he delivered in 1999 at the University of California at Berkeley, he said for example:

**So** here I kind of wave my hands at developing a sort of model theory for this stuff; that's a rather technical mathematical kind of development, as to how a formal language would be developed. What we would need to know about any expressive in the language is what must be the case in order for the expressive to be expressively correct. **Right, so** for it to be expressively correct is for its expressive content to hold, **right, so** this is where the rules of use come in, **right?**

(Kaplan 1999: 10; boldening ours)

In this short piece of 89 words there are six tokens of interactives (printed in bold), namely the discourse markers *so* (three times) and *right* (twice) and the response elicitor *right?*, that is, almost 7% of the forms used are interactives. Still, it is hard to find instances of interactives such as interjections or ideophones in academic discourse, be that in written or spoken discourse.

The classification proposed in this section provides a first general overview of the relative contribution made by the two grammars to structuring discourses. More research, based on the analysis of text corpora is urgently needed.

## 7.3 How do the two grammars contribute to discourse processing?

### 7.3.1 General observations

Comparing interactive grammar with sentence grammar was one of the main topics of the preceding chapters. What the findings may suggest is in particular the generalizations proposed in (11).

- (11) Sentence grammar and interactive grammar compared
  - a Both grammars are constantly available for designing discourses
  - b Translating from one grammar into another is a problem
  - c The two grammars have complementary functions
  - d Interactives can be inserted in pieces of sentence grammar but not the other way round

The points made in (11) are now briefly discussed in turn.



(a) Both grammars are constantly available for designing discourses. Speakers are able to draw on both grammars any time: The two are always active and can be called on, either one after the other or by inserting one in another. There are, however, clear preferences on how and where the two are employed. To this end, the generalizations in (12) can be proposed.

- (12) The relative contribution of interactives in discourse
- a Use of interactives is common in spoken discourse but rare in written discourse
  - b Use of interactives in discourse devoted to social exchange is essentially mandatory
  - c Use of interactives constitutes a close to indispensable component of conversations
  - d There is substantial clustering of interactives at the beginning and the end of social exchanges

(b) Translating from one grammar into another is a problem

The meaning of interactives has been said to be ‘descriptively ineffable’ (Potts 2007a; Blakemore 2011) or ‘insusceptible to formalization’ (Kaplan 1999) and, in fact, the two grammars are not readily intertranslatable. While their descriptive meaning content can be made explicit by means of paraphrase (Section 2.1.1), interjections, ideophones, discourse markers, and other interactives cannot readily be translated into notions of sentence grammar, and Haiman (2018) observes:

When we say, for example, that *thud* is the ‘impression of a heavy solid object falling to the ground,’ that *lu:ng* is (in Cambodian) the ‘impression of a dog vomiting’ or more generally that X is an ‘impression of’ some vivid scene, it is because we are forced to give linear expression to a number of concepts expressed by verbs and nouns and adjectives that cannot but occur in order in propositional language, but are all expressed at once in the onomatopoe or, more generally, the ideophone.

(Haiman 2018: 104)

As aptly observed by Haiman (2018: 138), ideophones and interjections ‘cannot be translated into propositional language,’ and Goffman (1978: 126) notes with reference to exclamatory interjections (‘response cries’) that they are ‘communications specifically designed for use outside states of talk.’

And while the meaning of interactives is portrayed as being ineffable or insusceptible to formalization, much the same applies the other way round: It is equally hard, if not impossible, to express concepts of sentence grammar in terms of forms available in the grammar of interactives—in other words, the two grammars appear to be largely mutually incompatible in how they encode meaning (see Section 2.1). At the same time, the availability of two contrasting modes

endows speakers with the ability to simultaneously draw on two different meaning resources to structure their discourses.

(c) They have complementary functions

Sentence grammar has no clearly dedicated constructions for the expression of functions central to interactive grammar. What [Pawley \(2009: 21\)](#) writes about analytic and holistic processing also applies to the relationship between the two grammars: Speakers can and constantly do move between the two. They use sentence grammar for presenting meaning contents about the world in a propositional format and interactive grammar for social communication and the structuring of discourse.

(d) Interactives can be inserted in pieces of sentence grammar but not the other way round

When pieces of sentence grammar and interactives are combined, the preferred position of the latter is at the left periphery of the former. Exceptions are provided by response elicitors and ideophones, which both tend to appear at the right periphery (Sections 3.5.3.2, 3.7.3). And in addition, most interactives can also be inserted within pieces of sentence grammar. Insertion is normally restricted to word boundaries, that is, interactives are unlikely to appear within a word. Nevertheless, there are few exceptions. English speakers have been found infixing interjections like *Goddam*, *fucking*, or *bloody* in words of sentence grammar, e.g., *abso-goddam-lutely*, *cappu-fucking-ccino*, *fan-bloody-tastic* (see Section 4.3.1; [Pinker 2007: 361](#)).

Conversely, however, pieces of sentence grammar cannot be inserted in interactives. There is an obvious reason for this: Since interactives are unanalyzable forms, they do not normally allow any text material to be placed within them.

### 7.3.2 Interactives in adjacency pairs

Many usages of interactives involve what, following [Schegloff and Sacks \(1973\)](#), are called ‘adjacency pairs.’ Such pairs are schemas of conversational routines consisting of a sequence of two parts, where the first part, produced by speaker A, entails that speaker B is expected to respond in the second part in a conventionalized way.

Adjacency pairs can be found in a range of different schemas of social interaction. With reference to the usage of interactives, it is especially the schemas presented below that tend to trigger the usage of function-specific responses. Each of the schemas is illustrated with an English example.

(a) Greeting rituals:

(13) American English (Biber et al. 1999: 1109, (17))

A: **Morning Diane.**

B: **Hi, Joyce, how are you?**

(b) Leave-taking rituals:

(14) A: **See you!**

B: **Yeah, see you later!**

(c) Apologizing rituals:

(15) American English (Biber et al. 1999: 1093)

A: **Sorry**, I didn't mean to scare you.

B: **That's okay.**

(d) Doing someone a favor:

(16) A: Here's your horse.

B: **Thanks.**

(e) Calling somebody:

(17) A: **Waiter!**

B: **Yes, sir.**

In adjacency pairs like (17), the response signal tends to be supplemented by a vocative adjunct (*sir*) referring to speaker A of the preceding turn.

(f) Polar questions: They are likely to prompt responses by means of response signals:

(18) British English (Biber et al. 1999: 1093)

A: Would you like another drink **Adam?**

B: **Yes please - .**

For a detailed analysis of adjacency pairs involving response signals ('response markers'), see Wiltschko (2021, Chapter 6).

(g) Response requests expressed by response elicitors (English *eh?*, *huh?*, *right?*, etc.) followed by a reaction by the hearer, cf. (19). This adjacency pair has been the subject of a detailed analysis by Wiltschko (2021, Chapter 5) on what she calls 'confirmational', that is, a subset of response elicitors.

(19) English (Wiltschko 2021: 118, (34))

A: You have a new dog, **eh?**

B: **Yeah**, I just got him last week.

Adjacency pairs are understood to be composed of two parts of text, uttered, respectively, by speaker A and speaker B. However, since interactives are anchored

immediately to the situation of discourse, the contribution by speaker A may consist of an action performed rather than of a text piece. For example, (20) provides a perfect alternative to (16).

- (20) A: [Speaker A hands a horse over to B]  
B: **Thanks.**

Furthermore, instead of an interactive, the response by B may, and not seldom does consist of a gesture, such as a nod in English for confirmation instead of the response signal *yes*, or a raised-hand movement instead of a social formula such as the response to a greeting (*Hi!*) or expressing one's gratitude (*Thank you!*).

The examples just provided by no means exhaust the range of conventional adjacency pairs, suffice it to mention another common type to be observed in court proceedings. The following case of courtroom interaction, taken from transcripts of the O.J. Simpson trial in the 1990s, may illustrate this type. The examination is introduced by the discourse marker *now* and exhibits a fixed format: It consists of an assertion by the lawyer (speaker A) and ends with a response elicitor, namely either a question tag (*hadn't you?*, *wasn't it?*, or *didn't she?*) or *true?*, and Simpson (speaker B) invariably answers with the response signal *no*.

- (21) American English (Mortensen 2019: 155)  
Lawyer: **Now**, in 1989, when you had this altercation with Nicole,  
you had beaten Nicole in the past, **hadn't you?**  
Simpson: **No.**  
Lawyer: And in this 1989 incident, it was the last straw for Nicole,  
**wasn't it?**  
Simpson: **No.**  
Lawyer: And she told you that, **didn't she?**  
Simpson: **No.**  
Lawyer: You had beaten her in the past, and on one occasion, you and  
she went to a doctor and lied about what happened, **true?**  
Simpson: **No.**  
Lawyer: And told the doctor that she fell off a bicycle. **True?**  
[...]

In all adjacency pairs examined, speaker B has two basic options to respond, namely drawing either on interactive grammar or on sentence grammar. Take the example in (22) (doing a favor): Speaker B may respond either by using the social formula *thanks*, cf. (i), or by drawing on an utterance of sentence grammar, cf. (ii).

- (22) A: Here's your horse.  
B: (i) **Thanks.**  
(ii) I appreciate what you have done for me.

The question then is: What determines or influences the choice made by interlocutors in deciding on a response in terms of either interactive grammar or sentence grammar? There is at least a general answer of the following kind: Utterances in terms of interactive grammar are short whereas corresponding utterances of sentence grammar are not only distinctly longer but also internally more complex. Take the responses of B in (22): Whereas *Thanks* in (i) consists of a single monosyllabic and unanalyzable meaningful unit, the corresponding utterance *I appreciate what you've done for me* in (ii) is not only much longer, consisting of eight meaningful units, but also morphosyntactically more complex, being composed of a main clause and a subordinate complement clause. On the assumption that the two responses in (22) convey roughly the same information content, interactive grammar thus provides a more economical and a more efficient tool of discourse processing.

With reference to the cooperative principle of Grice (1975), interactive grammar clearly provides the option to be preferred: On Grice's Maxim of Quantity ('Say no less than the conversation requires. Say no more than the conversation requires'), response (ii) in (22) can—on the assumption mentioned—be argued to violate this maxim and, hence, to constitute a dispreferred discourse option.

Nevertheless, the question raised is in need of much further investigation. But at least for (f), relating to response signals, there are some empirical data. Take the following citation from the New Testament: 'Let what you say be simply "Yes" or "No"; anything more than this comes from evil.' (Matthew 5:37). With reference to the framework proposed here, one way of interpreting this citation is to argue that response signals like 'yes' and 'no' are as a rule the most unmarked, or the default discourse options dedicated to responding to polar questions—note that polar questions are also commonly called 'yes-no-questions.'

That this is in fact correct is suggested by Enfield et al. (2019) in their cross-linguistic study of question-answer sequences. The interest of the authors was with the speaker's choice between what they refer to as 'interjection-type responses' and 'repetition-type responses' in confirming answers to polar questions. In the first type, an 'interjection' or 'particle' is used, that is, a response signal like 'yes' or its variants (*yeah, yep, uh-huh, mm-hm, certainly*, or a nod). The second type, also called 'the echo strategy,' involved answers that fully or partially repeated elements of the question, without qualitative semantic alteration (for example, *Is John coming?—John's coming, John is, He is, He is coming, or He's coming*). Note that the authors were restricted to confirming answers, that is, disconfirming answers (e.g., 'no') or non-answers ('I don't know,' 'I can't remember,' or laughter) were excluded.<sup>3</sup>

<sup>3</sup> Also excluded was what they call 'the transformative strategy,' involving answers that added or replaced elements of the question, that is, responses such as the following: *Is John coming?—It's Tom who's coming, John's coming later, or John's in London* (Enfield et al. 2019: 288).

Analyzing fourteen languages of world-wide distribution, the authors observed a global predominance of the interjection-type answer strategy, that is, of the use of response signals. Out of the fourteen languages, speakers of eleven languages were found to rely on response signals in over 80% of their answers to polar questions. Exceptions were observed in the remaining three languages, namely Brazilian Portuguese, ǀAkhoe Haillom of Namibia, and Tzeltal of Mexico, which all showed percentages lower than 80%, but [Enfield et al. \(2019\)](#) suggest that there are specific cultural or linguistic factors to explain these lower percentages.<sup>4</sup>

In a similar way, [Mackenzie \(1998: 279\)](#) found in a count of yes–no-questions in English fictional writing 71% of the answers to be ‘holophrases,’ that is, response signals like *yes*, *no*, *absolutely*, and *hardly*, thus offering ‘the most natural answer to questions.’ Such findings lend support to the hypothesis that response signals provide the default strategy for answering polar questions of the kind investigated by the authors. These signals are said to be ‘designed not to introduce any pragmatic turbulence, thus allowing the dialogue to get quickly back on track,’ and to represent a solution to the problem of how to answer a polar question and do nothing more than that:

We argue that this correlates with their being pragmatically unmarked. We argue that the semantic-pragmatic design of interjections [...] makes them intrinsically better suited to serving as an unmarked answer type, regardless of cultural or linguistic context. Because the interjection answer type semantically codes the function of confirming, it can be said to be a dedicated answer form.

([Enfield et al. 2019: 289](#))

In contrast, repetition-type answers, that is, answers based on drawing on material of sentence grammar, are said to independently assert the proposition to claim a degree of thematic agency, or independent interest over that proposition ([Enfield et al. 2019: 292](#)), or as [Enfield, Stivers, and Levinson \(2010: 2616\)](#) put it, they ‘assert relatively more agency or authority over the proposition than interjection responses.’

In sum, there is reason to assume that in adjacency pairs of the kind examined here, interactives constitute the default solution for forming responses, being dedicated to this purpose and providing the most ‘pragmatically unmarked’ option. But in order to verify this hypothesis, more studies of the kind carried out by [Enfield et al. \(2019\)](#); see also [Stivers and Enfield 2010](#)) are needed. Such studies should be concerned in particular with the question of what precisely the motivations of language users are for deciding either in favor of interactive grammar or of sentence grammar.

<sup>4</sup> Interestingly, ‘combination answers’ of the type *Yes, he is* were found to be of low frequency ([Enfield et al. 2019: 289](#)).

Claiming thematic agency or authority over the discourse are conceivably strong motivations to opt for the means available in sentence grammar, but without any further information this must remain an unresolved issue, especially for the following reasons. For one thing, the authors were, as was just mentioned, restricted in their research design to confirming answers—that is, disconfirming answers (e.g., ‘no’) were excluded. The generalization might have been different if disconfirming answers had been included, considering that ‘no’-answers can have the opposite effect from ‘yes’-answers. For example, unlike ‘yes,’ a disconfirming ‘no’ on what has just been said can in specific contexts be understood as a strong signal of the speaker aiming to take the floor and claiming more agency or authority over the discourse.

And for another, there is reason to assume that choices between responses in terms of material from either sentence grammar or interactive grammar are very likely to be shaped by factors like the ones proposed in Table 7.1 of Section 7.1. Factors such as the conceptual space involved (interactive vs. ideational), the nature of meaning (expressive vs. descriptive), the nature of forms (holistic vs. analytic), or the speed of processing intended (fast vs. slow) are equally likely to guide interlocutors in deciding on either sentence grammar or interactive grammar.

## 7.4 Conclusion

Chapter 7 was concerned with how the two grammars perform in discourse processing relative to one another. We saw that each has its own preferred space of operation, but there are also areas where interlocutors are free to choose, and in such areas the two can be said to be in competition.

The conclusion to be drawn from the observations made is, in a nutshell, that both grammars are needed jointly for successful communication.

## 8

# Conclusions

At the beginning of Chapter 1, it was stated that the goal of the present study was to take issue with a view widely held in monographs of linguistics according to which expressions like interjections, response signals, social formulae, discourse markers, etc. are a mixed-bag that is highly language specific and does not form any unified grammatical category. It is hoped that the preceding chapters have shown that this view is in need of revision. Based on cross-linguistic observations it was argued rather that these expressions, called interactives, do form a category of their own and have a grammar distinct from sentence grammar, referred to as interactive grammar. Interactive grammar crucially involves stancetaking, invoking ‘a shared framework for co-action with others’ (Du Bois 2007: 171), and rather than being analytic like sentence grammar, it is organized holistically.

According to another widely held view surfacing in some form or other in many linguistic studies, interactives occupy a marginal place at the periphery of language structure, being referred to as ‘marginalia’ (Joseph 1997; Ameka 2020). As we saw in Section 6.4, this view as well is at variance with findings presented in this book according to which interactive grammar is best analyzed as a distinct system within the grammar of discourse.

The reasons for describing interactive grammar as a system of its own were presented in Sections 6.3 and 6.4, they can be summarized as below:<sup>1</sup>

- (a) Interactive grammar differs from sentence grammar in subscribing to the properties mentioned in definition (1) of Section 1.2.
- (b) Like sentences, most interactives can form utterances of their own as well as independent turns in conversational exchanges.
- (c) Rather than being edited and processed by the morphosyntactic machinery of sentence grammar, interactives are immediately accessible (cf. the direct-access hypothesis, Section 2.1.3.3).
- (d) Sentence grammar and interactive grammar may provide interlocutors with two in some way equivalent alternatives for designing their discourse contributions.
- (e) Compared to sentence grammar, interactive grammar provides a distinctly more economical tool for designing discourses, saving time, and processing energy.

<sup>1</sup> For a discussion of the term ‘system,’ see Section 6.3.



- (f) Sentence grammar constitutes an optimal device for conveying conceptual information. Interactive grammar, by contrast, constitutes a close to indispensable tool for structuring the space of social communication.
- (g) In combinations of interactives with pieces of sentence grammar, the former tend to present foregrounded information while the latter tend to present backgrounded information.
- (h) There is neurolinguistic evidence in support of the distinction in that interactive grammar differs from sentence grammar in drawing more heavily on brain activity in the right hemisphere (Section 5.3).
- (i) There is also support from research in social psychology, where two types of reasoning and judgment are distinguished, and this distinction shows strong parallels with the distinction between the two grammars (Section 5.5).

The way the distinction between the two kinds of grammar correlates with findings made in neurolinguistic and psychological work was summarized in Section 5.5 thus: Underlying the distinction there is a more basic strategy of human behavior to process phenomena in two contrasting ways, employing two different modes of thought and decision making: These phenomena are treated either as holistic depictions for rapid processing. Alternatively, they are treated by monitoring one's behavior and viewing the phenomena as conceptually complex entities that need to be processed laboriously and analyzed in terms of the grammatical architecture and the building blocks of which that architecture is composed.

To be sure, the two grammars are not always neatly separated and the boundary between the two is at times blurred. But such cases can as a rule be accounted for in terms of either principles of discourse processing or of diachronic mechanisms. On the one hand, the units of the two grammars constantly co-occur side by side in discourse and it therefore comes as no surprise that there are interferences between the two. For example, a prosodic boundary separating an interactive from a piece of sentence grammar may disappear in rapid speech or when the two co-occur regularly.

On the other hand, there are diachronic connections between the two: As we saw in Chapter 4, units of sentence grammar can give rise to interactives, and interactives may be grammaticalized into units of sentence grammar, e.g., when the use of an interjection like *wow* or an ideophone like *bang* were extended to serve as verbs, that is, *to wow*, *to bang* (Sections 4.2.2.5 and 4.4).

The concern of the book was with interactives as a category of discourse processing different from sentence grammar. What Dingemans (2021b) says about interjections applies to interactives in general: 'They offer insights that can help linguistics extend its reach from isolated sentences to interactional sequences.'

We saw that in previous research, this category was perspectivized and classified in a number of ways, either as located at the periphery of language, as

being intermediate between language and some other phenomena, or as forming a domain separate from sentence grammar (Section 6.1). Interjections were called discourse markers and discourse markers were called interjections. And interjections have also been dubbed ‘response cries,’ but they have also been called ‘words’ (Section 6.2). It is hoped that the book has shown that such seemingly conflicting classifications are by no means the result of questionable analyses by the authors concerned; rather, they are indicative of the complex interlocking pattern of relations characterizing the space of interactive grammar.

In the presentation of the text, we proceeded from the specific to the general, from synchronic to diachronic observations, from linguistic observations to findings made in other disciplines, and from description to explanation. Chapter 1 described the scope of the book, providing a definition of interactives in Section 1.2, discussing issues that had plagued their analysis, and preparing the reader for what to expect in the chapters to follow.

The meaning of interactives was discussed in Section 2.1, where two important features were distinguished. It was argued that interactives, while having a component of descriptive meaning, are expressives, and they are depictions—in short, they were classified as expressive depictions. But the main claim made in Section 2.2 of Chapter 2 was that one of the characteristics of interactives is that they have a valence structure, that is, they differ from one another in which and how many arguments they take. Rather than being determined by the syntax of a clause or sentence, or the valence of a verb, arguments are an inherent part of the functions of interactives, and were referred to as ‘discourse arguments’ to distinguish them from syntactic arguments in sentence grammar.

In Chapter 3, the ten types of interactives were looked at in more detail. Of all these types, ideophones turned out to constitute the most peripheral one. For one thing, they have been claimed not to clearly subscribe to the property of indexicality proposed in Section 1.2. And for another, their grammaticalization behavior is also different. On the one hand, they grammaticalize more readily in the direction of sentence grammar than other types of interactives (Section 4.4). And, on the other hand, they do not participate in the network of grammaticalization characterizing interactives (Section 4.2).

The goal of Chapter 4 was to look for diachronic regularities in the development of interactives with a view to understanding their synchronic behavior. And we saw that there are in fact some diachronic explanations for why some interactives are structured the way they are, and for why the boundaries between different types of interactives are frequently fluid and overlapping.

While the first four chapters were devoted to description, the ensuing chapters aimed at adopting a wider perspective on the subject matter of the book. To this end, Chapter 5 was concerned, on the one hand, with previous research on interactives. As we saw in Sections 5.1 and 5.2, there are a number of approaches that have come up with notions of or related to interactive grammar. At the same time, there

are only two approaches that are fully compatible with the analysis proposed here, namely the one adopted in the *Longman Grammar of Spoken and Written English* (Biber et al. 1999) and that of discourse grammar (Kaltenböck et al. 2011; Heine et al. 2013).

On the other hand, an attempt was made in Chapter 5 to examine how interactive grammar can be reconciled with findings made in disciplines other than linguistics. In Section 5.3, it was shown that the distinction between two kinds of grammar exhibits significant correlations with activity in the two hemispheres of the human brain, and noteworthy parallels were also found in the distinction between two types or systems of reasoning and judgment in social psychology.

Taking up issues that had surfaced in earlier research, Chapter 6 looked at interactives by identifying their salient properties and arguing that interactive grammar has the features of a system distinct from sentence grammar. Finally, the question asked in Chapter 7 was how the two grammars perform both relative to one another and to discourse processing. The conclusion reached was that both grammars are needed for successful communication.

The analysis proposed here is in accordance with that of Evans (1992) in his discussion of interactives ('interjections') in the Australian language Mayali:

One might argue that we are dealing here with some sort of higher-level organization, parallelling at the level of discourse the sequential organization of words within the sentence. This would be problematic for the view that interjections are only loosely integrated into the grammar.

(Evans 1992: 235)

From the perspective of sentence grammar, interactives have, in fact, commonly been said to be loosely integrated, if integrated at all. According to the perspective proffered here, both sentences and interactives form building blocks within the overall structure of discourse processing but the latter are said to represent a 'higher-level organization' in the sense of Evans (1992). To be sure, sentences form important building blocks in this organization. But, as was shown in the preceding chapters (especially Section 7.2), its potential has limits. For one thing, interactives provide a more efficient tool for certain tasks of social communication. For another, sentences play only a minor role in certain parts of discourse (Section 7.3). And finally, since their activation bypasses the machinery of sentence grammar, interactives are depictions enabling interlocutors to establish an immediate link with the situation of discourse.

We saw that, depending on the communicative intents of interlocutors, we find segments of text that contain no forms of interactive grammar, but there are also text segments that are devoid of any material of sentence grammar (e.g., Section 7.2). What this suggests is that the two kinds of grammar are to quite some extent independent from one another.

The main emphasis in this book was on describing the grammar of interactives. But in Chapter 4, we were also looking for explanations, where the latter were exclusively of a diachronic nature. It was observed, for example, that linguistic units sometimes exhibit two contrasting patterns of usage, where one shows the grammatical features of a word of sentence grammar, whereas the other is suggestive of an interactive as defined in Section 1.2. This situation was accounted for, first, in terms of cooptation in Section 4.1. Cooptation was argued to be of help for understanding why an English item like *boy* occurs in two contrasting usages: On the one hand, it forms a paradigm instance of the lexical category of nouns. On the other hand, it occurs as an extra-clausal unit commonly described as an interjection, and the mechanism to be held responsible for the change from the former to the latter was described as one involving cooptation.

And second, this situation was also described and explained in Sections 4.2 and 4.4 by means of grammaticalization theory. One kind of example was the subject of Section 3.9.5, where two contrasting structures of the English social formula *please* were discussed, and more examples were presented in Section 4.2. Networks like the one in Figure 4.1 of Section 4.2.3 provide one of the main reasons why the types of interactives do not each constitute neatly separated categories; rather, one and the same interactive form can function, e.g., as a directive in some contexts, as an attention signal in other contexts, and as a discourse marker in still other contexts. Accordingly, a given interactive form can simultaneously be part of two or even three different interactive types, reflecting different degrees of grammaticalization.

While these processes appear to have taken place within interactive grammar, we saw that there is also a second kind of example showing that grammaticalization may also lead from interactive grammar to sentence grammar. This process is most conspicuous in ideophones. In a number of languages, ideophones straddle between prototypical interactive usages and usages that are more or less integrated in the structure of a sentence. As was argued in Section 4.4, this situation can also be explained as being due to differential effects of grammaticalization.

The book is based on a comparative-typological perspective, drawing on data from well over a hundred languages, and the generalizations proposed are assumed to be cross-linguistically significant. Nevertheless, no claim can be made to the effect that these generalizations are representative of the world's languages at large. For one thing, for thousands of languages we have no descriptions or only inadequate ones. And for another, even the descriptions that exist are for the most part silent on the structure of the grammar of interactives. Accordingly, the findings presented must be taken with care, subject to verification by future research.

There remain a number of issues where the approach used in this book offers no satisfactory account, or no account at all. One issue concerns the nature of 'expressive meaning,' which is part of the discourse functions of interactives. As we saw in Section 2.1, quite some efforts have been made to understand what

exactly this meaning is about. What seems obvious is that, rather than in terms of truth conditions, this issue should be approached from the use conditions and, more specifically, from ‘discourse placedness conditions’ underlying the structure of interactives. But, as far as we see, there is so far no convenient methodology that would allow coming to grips with the notion of use conditions in analyzing expressive meaning.

In the course of the preceding chapters, numerous examples were presented showing that interactives quite frequently combine with one another. The question, however, what the principles governing such collocations are had to be ignored except for a few casual observations made at various points in Chapter 3. And much the same applies to the question of what constraints there are on combining interactives with pieces of sentence grammar (but see also [Wiltschko 2021](#)). These are central questions that are in need of answers for a more comprehensive understanding of the grammar of interactives. But answering them would require a study of its own, building on larger corpus-based sets of text data and quantitative analyses.

Another issue relates to the following question: What determines the choice of speakers or writers when drawing either on sentence grammar or on interactive grammar? This was a central topic in Chapter 7 but, as we saw there, so far hardly any empirically sound data are available to answer this question. To be sure, a few general observations do exist. For example, talking or writing about the world is almost entirely reserved for sentence grammar, while regulating social communication is hard to achieve without a participation of interactive grammar. But what guides interlocutors in situations where both kinds of grammar provide acceptable options for social exchange, such as responding to greetings or apologies, or answering polar questions? As we saw in [Section 7.3.2](#), there are a few first insights, but such insights have achieved hardly more than scratching the surface.

Another issue, one that did not quite receive the attention it deserves, concerns the relation between linguistic and paralinguistic forms of expression. For one thing, spoken discourse usually implicates gesture of various kinds, and this applies to both types of grammar. But, as has been pointed out independently in many studies, in interactive grammar gesture plays an outstanding role ([Section 1.3.2.3](#)). For example, a nod may provide a perfect alternative to the response signal *yes* in many contexts, and so does a function-specific hand gesture as an alternative to a social formula like *Hello!*, *Goodbye!*, or *Sorry!*, or to directives like *Shh!*, *Come!*, or *Out!*. It would seem, in fact, that interlocutors dispose of a kind of expression model as it was sketched in [Section 1.3.2.3](#) (see also [Haiman 2018: 298–9](#)), reprinted for convenience in (1), that they can activate any time to decide on whether to draw on a vocal form (A), a gesture (C), or a combination of the two (B).

## (1) A voice-gesture model for the expression of interactive concepts

(from [Section 1.3.2.3](#))

| A    | B              | C       |
|------|----------------|---------|
| Form | Form + Gesture | Gesture |

In addition to the questions just mentioned there remains a more general one: Do the findings presented have a bearing on grammar writing? While this question raises fundamental issues of linguistic theory far beyond the concerns of this book, the answer is essentially in the affirmative. It would seem that most conventions of linguistic analysis as they surface in reference grammars or textbooks are hard-pressed to appropriately reflect how language is used in social interaction. And considering that social interaction is arguably one of the main functions subserved by language, reducing interactives to a kind of appendage of grammatical treatment does not seem to do justice to the role they play in our daily life. The findings made in this study suggest, rather, that interactive grammar be best viewed as forming a domain separate from sentence grammar—after all, like sentences, interactives can form utterances and other self-contained segments of discourse, as we saw in the preceding chapters. To put it bluntly, grammars of English or any other language, as we conceive of them, need to be longer in order to appropriately account for social communication.

Furthermore, there remains a range of other topics that could not be discussed in the book, because they would need to be studied separately. One topic concerns the argument structure of interactives, examined in more detail in [Section 2.2](#), in which three categories of arguments were distinguished, namely S (speaker), H (hearer), and T (theme). As was observed in [Section 2.2.2](#), it would seem that the S argument can further be subcategorized into a number of ‘semantic roles’ such as experiencer and agent (cf. [Andrason 2021c](#)). But once again, this is an issue that would need a study of its own.

Another topic concerns language contact, which was at least mentioned in [Section 6.5](#). Interactives are easily transferred from one language to another when people speaking different languages interact. But the data available do not allow to determine, on the one hand, whether there are regularities characterizing such transfers and, on the other hand, whether interactive grammar differs significantly from sentence grammar in the structure and the magnitude of transfer.

A further topic—one that is crucial for a better understanding of the nature of interactive grammar—concerns language acquisition ([Section 6.6](#)). Do children in their early language development rely more heavily on interactive notions, as one might suspect? And how do the two kinds of grammar evolve relative to one another? These, as well as a number of other questions associated with both first and second language acquisition, however, cannot be answered without comprehensive investigation.

Finally, there is the question of whether work on interactives can shed light on earlier stages in the evolution of human languages (Section 6.7). The reconstruction of languages back to their beginnings was for a long time a playground for conjecture, but over the last decades it has become a field of sound scholarship and interdisciplinary cooperation. It is especially the nature of primary interactives, such as interjections (*wow*, *ouch*, *yuck*), directives (*shh*), attention signals (*hey*), and response elicitors (*huh?*), that has inspired researchers since Jean-Jacques Rousseau and earlier to hypothesize that they reflect archaic features of languages: Primary interactives are not only morphologically unanalyzable and syntactically independent, their usage relates to elementary functions of human behavior, and they are etymologically opaque, that is, their history remains as a rule unknown.

However, the question of how such observations relate to all the work that has been done on language genesis is beyond the scope of this book and requires a separate analysis, and such an analysis is well advised to take the catalog of general features into account that were identified in Section 6.6.

The goal of this study was to provide a first attempt at understanding the behavior of interactives and to identify elements of their overall structure. But as we saw in the preceding chapters, the range of questions asked outnumbered that of the answers that could be provided. It is hoped that the book will encourage research to find answers to the remaining questions and to come up with a more comprehensive framework of interactive grammar.

# References

- Acharya, Jayaraj 1991. *A Descriptive Grammar of Nepali and an Analyzed Corpus*. Washington, D.C.: Georgetown University Press.
- Ädel, Annelie 2006. *Metadiscourse in L1 and L2 English*. Amsterdam, Philadelphia: Benjamins.
- Aijmer, Karin 1997. 'I think' – an English modal particle. In T. Swan and O. Jansen Westvik (eds.), *Modality in Germanic Languages. Historical and Comparative Perspectives*. Berlin, New York: Mouton de Gruyter, pp. 1–47.
- Aijmer, Karin 2002. *English Discourse Particles: Evidence from a Corpus*. Amsterdam, Philadelphia: Benjamins.
- Aijmer, Karin and Anne-Marie Simon-Vandenberg 2003. The discourse particle 'well' and its equivalents in Swedish and Dutch. *Linguistics* 41, 1: 1123–61.
- Aikhenvald, Alexandra Y. 2008. *The Manambu Language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2010. *Imperatives and Commands*. (Oxford Studies in Linguistic Theory.) Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2020. 'Damn your eyes!' (Not really): Imperative imprecatives, and curses as commands. In Nico Nassenstein and Anne Storch (eds.), *Swearing and Cursing: Contexts and Practices in a Critical Linguistic Perspective*. (Language and Social Life, 22.) Berlin: De Gruyter Mouton, pp. 53–74.
- Allan, Keith and Kate Burridge 1991. *Euphemism and Dysphemism: Language Used as Shield and Weapon*. New York, Oxford: Oxford University Press.
- Alpher, Barry 2001. Ideophones in interaction with intonation and the expression of new information in some indigenous languages of Australia. In Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*. Amsterdam: John Benjamins, pp. 9–24.
- Ameka, Felix K. 1992a. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18: 101–18.
- Ameka, Felix K. 1992b. The meaning of phatic and conative interjections. *Journal of Pragmatics* 18: 245–71.
- Ameka, Felix K. 1994. Interjections. In R. E. Asher (ed.), *The Encyclopedia of Language and Linguistics*. Volume 2. Oxford: Pergamon Press, pp. 1712–15.
- Ameka, Felix K. 2001. Ideophones and the nature of the adjective word class in Ewe. In Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*. Amsterdam: John Benjamins, pp. 25–48.
- Ameka, Felix K. 2006. Interjections. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, pp. 743–46.
- Ameka, Felix K. 2020. Linguistic marginalia and 'grammars of use'. Abralín Ao Vivo Linguists. Brazilian Linguistics Association, July 7, 2020.
- Ameka, Felix K. and David Wilkins 2006. Interjections. In Jan-Ola Östman and Jef Verschueren (eds.), *Handbook of Pragmatics*. Amsterdam, Philadelphia: Benjamins, pp. 1–19.
- Amha, Azeb 2001. *The Maale Language*. (CNWS Publications, 99.) Leiden: Research School of Asian, African, and Amerindian Studies, Leiden University.



- Amha, Azeb 2013. Directives to humans and to domestic animals: The imperative and some interjections in Zargulla. In Marie-Claude Simeone-Senelle and Martine Vanhove (eds.), *Proceedings of the 5th International Conference on Cushitic and Omotic Languages, Paris, April 16–18, 2008*. Cologne: Köppe, pp. 211–29.
- Andrason, Alexander 2020. Ideophones as linguistic ‘rebels’ – The extra-systematicity of ideophones in Xhosa. Part 1. *Asian and African Studies* 29, 2: 119–65.
- Andrason, Alexander 2021a. Ideophones as linguistic ‘rebels’ – The extra-systematicity of ideophones in Xhosa. Part 2. *Asian and African Studies* 30, 1 (in press).
- Andrason, Alexander 2021b. Laughter interjections in Xhosa. *Studies in African Languages and Cultures* 55: 31–71.
- Andrason, Alexander 2021c. Argument structure of interjections: Evidence from Polish. To appear in the journal *Lingua*.
- Andrason, Alexander 2022. The non-arbitrariness of some conative calls used to chase animals. Typescript.
- Andrason, Alexander and Mawande Dlali 2020. The (crucial yet neglected) category of interjections in Xhosa. *STUFF* 73, 2: 159–217.
- Andrason, Alexander, Anne-Maria Fehn, and Admire Phiri 2020. Interjections in Tjwao. *Bulletin of the School of Oriental and African Studies (SOAS)* 2020: 1–27. (To appear).
- Andrason, Alexander and Michael Karani 2021a. Emotive interjections in Arusa Maasai. Typescript.
- Andrason, Alexander and Michael Karani 2021b. Conative calls to animals: From Arusa Maasai to a cross-linguistic prototype. *Lodz Papers in Pragmatics* 17, 1: 3–40. <https://doi.org/10.1515/lpp-2021-0002>.
- Andrason, Alexander and Haile Matutu 2019. The syntax of interjections in isiXhosa—a corpus-driven study. *Stellenbosch Papers in Linguistics* 58: 1–16. <https://doi.org/10.5842/58-0-833>.
- Andrason, Alexander, Admire Phiri, and Anne Maria Fehn 2021. Onomatopoeias in Tjwao. Typescript.
- Anstatt, Tanja 2011. Der slavische Vokativ im europäischen Kontext. In Ljudmila Geist and Grit Melhorn (eds.), *Linguistische Beiträge zur Slavistik: XIV: Jungslavist Innen-Treffen*. Munich: Sagner, pp. 9–26.
- Armstrong, David, William Stokoe, and Sherman Wilcox 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.
- Austin, Peter 1981. *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- Baldi, Sergio and Kyallo Wadi Wamitila 2014. Ideophones in Swahili: a preliminary survey. In Iwona Kraska-Szlenk and Beata Wójtowicz (eds.), *Current Research in African Studies: Papers in Honour of Mwalimu Dr. Eugeniusz Rzewuski*. Warsaw: Dom Wydawniczy Elipsa, pp. 39–60.
- Bally, C. 1909. *Traité de stylistique française*. Paris: Klincksieck.
- Bannard, Colin and Elena Lieven 2009. Repetition and reuse in child language learning. In Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen M. Wheatley (eds.), *Formulaic Language. Volume 2: Acquisition, Loss, Psychological Reality, and Functional Explanations*. (Typological Studies in Language, 83.) Amsterdam, Philadelphia: Benjamins, pp. 297–321.
- Barth-Weingarten, Dagmar and Elizabeth Couper-Kuhlen 2002. On the development of final *though*: A case of grammaticalization? In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*. Amsterdam, Philadelphia: Benjamins, pp. 345–61.

- Beeching, Kate 2016. *Pragmatic Markers in British English*. Cambridge: Cambridge University Press.
- Beeman, Mark and Christine Chiarello 1998. Complementary right- and left-hemisphere language comprehension. *Current Directions in Psychological Science* 7, 1: 1–8.
- Benfey, Theodor 1869. *Geschichte der Sprachwissenschaft*. Munich: J.G. Cotta.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Bickerton, Derek 2009. *Adam's Tongue: How Humans Made Language, How Language Made Humans*. New York: Hill and Wang.
- Blakemore, Diane 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. (Cambridge Studies in Linguistics, 99.) Cambridge: Cambridge University Press.
- Blakemore, Diane 2007. 'Or'-parentheticals, 'that is'-parentheticals and the pragmatics of reformulation. *Journal of Linguistics* 43: 311–39.
- Blakemore, Diane 2011. On the descriptive ineffability of expressive meaning. *Journal of Pragmatics* 43: 3537–50.
- Blecke, Thomas 1996. *Lexikalische Kategorien und grammatische Strukturen im Tigemaxo (Bozo, Mande)*. (Mande Languages and Linguistics.) Cologne: Köppe.
- Blonder, Lee Xenakis, Dawn Bowers, and Kenneth M. Heilman 1991. The role of the right hemisphere in emotional communication. *Brain* 114, 3: 1115–27.
- Bloomfield, Leonard [1933] 1962. *Language*. London: George Allen & Unwin.
- Bolinger, Dwight 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. London: Edward Arnold.
- Bolly, Catherine and Liesbeth Degand 2013. Have you seen what I mean? From verbal constructions to discourse markers. *Journal of Historical Pragmatics* 14, 2: 210–35.
- Borchmann, Simon 2019. Non-spontaneous and communicative emotive interjections – ecological pragmatic observations. *Scandinavian Studies in Language* 10, 1: 7–40.
- Börjars, Kersti and Nigel Vincent 2011. Grammaticalization and unidirectionality. In Heiko Narrog and Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press, pp. 163–76.
- Borod, J., R. Bloom, and C. Santschi Haywood 1998. Verbal aspects of emotional communication in the right cerebral hemisphere. In M. Beeman and C. Chiarello (eds.), *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 285–307.
- Borod, J. C., F. Andelman, L. K. Obler, J. R. Tweedy, and J. Welkowitz 1992. Right hemisphere specialization for the identification of emotional words and sentences: Evidence from stroke patients. *Neuropsychologia* 30, 9: 827–44.
- Brainerd, Charles J. and Valerie F. Reyna 1992. Explaining 'memory free' reasoning. *Psychological Science* 3, 6: 332–9. doi:10.1111/j.1467-9280.1992.tb00042.x. S2CID 145481025.
- Brainerd, Charles J. and Valerie F. Reyna 2002. Fuzzy-trace theory and false memory. *Current Directions in Psychological Science* 11, 5: 164–9.
- Branch, C., B. Milner, and T. Rasmussen 1964. Intracarotid sodium amytal for the lateralization of cerebral speech dominance. *Journal of Neurosurgery* 21: 399–405.
- Brandão, Ana Paula Barros 2014. *A Reference Grammar of Paresi-Haliti (Arawak)*. Ph.D. dissertation, The University of Texas at Austin.
- Brauner, Siegmund 1993. *Einführung ins Schona*. Cologne: Köppe.
- Brenzinger, Matthias and Sheena Shah 2021. *Writing Clicks*. London: EL Publishing.
- Brenzinger, Matthias and Sheena Shah 2022. Click phonemes versus naked clicks. Type-script.

- Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. (Studies in English Language.) Cambridge: Cambridge University Press.
- Brinton, Laurel J. 2010. The development of *I mean*: Implications for the study of historical pragmatics. In Susan M. Fitzmaurice and Irma Taavitsainen (eds.), *Methods in Historical Pragmatics*. Berlin: Mouton de Gruyter, pp. 37–80.
- Brinton, Laurel. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: Cambridge University Press.
- Brinton, Laurel J. and Elizabeth Closs Traugott 2005. *Lexicalization and Language Change*. (Research Surveys in Linguistics.) Cambridge: Cambridge University Press.
- Brown, Penelope and Stephen C. Levinson 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Buechel, Eugene 1939. *A Grammar of Lakota: The Language of the Teton Sioux Indians*. Saint Francis, P.O.: Saint Francis Mission.
- Bühler, Karl [1934] 1965. *Sprachtheorie*. Stuttgart: Fischer. Reprint of Jena: Fischer, 1934.
- Burridge, Kate and Margret Florey 2002. 'Yeah-no, he's a good kid': a discourse analysis of Yeah-no in Australian English. *Australian Journal of Linguistics* 22, 2: 149–71.
- Bussmann, Hadumod 1996. *Routledge Dictionary of Language and Linguistics*. London and New York: Routledge.
- Bynon, James 1976. Domestic animal calling in a Berber tribe. In William McCormack and Stephen Wurm (eds.), *Language and Man: Anthropological Issues*. The Hague: Mouton, pp. 39–66.
- Caffi, Claudia and Richard W. Janney 1994. Toward a pragmatics of emotive communication. *Journal of Pragmatics* 22, 3–4: 325–73. doi:10.1016/0378-2166(94)90115-5
- Carter, Ronald and Michael McCarthy 2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
- Cassirer, Ernst 1944. *An Essay on Man*. New Haven, London: Yale University Press.
- Cecil, Matthew Jones 2010. *Cross-linguistic variation in turn taking practices: A computational study of the Callhome Corpus*. MA thesis, University of Colorado.
- Childs, G. Tucker 1994. African ideophones. In Leanne Hinton, Johanna Nichols, and John J. Ohala (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 178–204.
- Childs, G. Tucker 1995. *A Grammar of Kisi, a Southern Atlantic Language*. (Mouton Grammar Library, 16.) Berlin, New York: Mouton de Gruyter.
- Childs, G. Tucker 2003. *An Introduction to African Languages*. Amsterdam, Philadelphia: Benjamins.
- Clancy, Patricia, Sandra A. Thompson, Ryoko Suzuki, and Hongyin Tao 1996. The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics* 26: 355–87.
- Claridge, Claudia 2013. The evolution of three pragmatic markers: *As it were*, *so to speak/say* and *if you like*. *Journal of Historical Pragmatics* 14, 2: 161–84.
- Claridge, Claudia and Leslie Arnovick 2010. Pragmaticalisation and discursification. In H. Jucker and Irma Taavitsainen (eds.), *Historical Pragmatics*. (Handbook of Pragmatics, 8.) Berlin: de Gruyter Mouton, pp. 165–92.
- Clark, Eve V. 2003. *First Language Acquisition*. Cambridge: Cambridge University Press.
- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123, 3: 324–47.
- Clark, Herbert H. 2019. Depicting in communication. In Peter Hagoort (ed.), *Human Language: From Genes to Behavior*. Cambridge, MA: The MIT Press, pp. 235–47.
- Clark, Herbert H. and Jean E. Fox Tree 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84: 73–111.

- Clark, Herbert H. and Richard J. Gerrig 1990. Quotations as demonstrations. *Language* 66, 4: 764–805.
- Cleeremans, A., and L. Jiménez 2002. Implicit learning and consciousness: A graded, dynamical perspective. In R. M. French and A. Cleeremans (eds.), *Implicit Learning and Consciousness: An Empirical, Philosophical and Computational Consensus in the Making*. Hove, U.K.: Psychology Press, pp. 1–40.
- Cole, Desmond [1955] 1987. *An Introduction to Tswana Grammar*. 5th impression. Cape Town: Longman Penguin Southern Africa.
- Conklin, Kathy and Norbert Schmitt 2008. Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics* 29, 1: 72–89.
- Cornips, Leonie 2022. The animal turn in postcolonial linguistics: The interspecies greeting of the dairy cow. *Journal of Postcolonial Linguistics* 6: 209–31.
- Coulmas, Florian 1979. On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics* 3: 239–66.
- Coulmas, Florian (ed.) 1981. *Explorations in Standardized Communication Situations and Prepatterned Speech*. Volume 2: *Conversational Routine* (Janua Linguarum. Series Maior, 96.) Berlin: De Gruyter Mouton.
- Creider, Jane Tapsubei and Chet A. Creider 2001. *A Dictionary of the Nandi Language*. (Nilo-Saharan, 16.) Cologne: Köppe.
- Croft, William 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Crowley, Terry 1982. *The Paamese Language of Vanuatu*. (Pacific Linguistics Series B, 87.) Canberra: Australian National University.
- Crystal, David 1994. *An Encyclopedic Dictionary of Language and Languages*. London: Penguin Books.
- Cuenca, Maria Josep 2000. Defining the indefinable? Interjections. *Syntaxis* 3: 29–44.
- Cuenca, Maria Josep 2008. Pragmatic markers in contrast: The case of *well*. *Journal of Pragmatics* 40: 1373–91.
- D'Atri, Annabella 1995. The theory of interjections in Vico and Rousseau. In L. Formigari and D. Gambarara (eds.), *Historical Roots of Linguistic Theories*. Amsterdam, Philadelphia: Benjamins, pp. 115–27.
- Davies, Eirlyse 1987. A contrastive approach to the analysis of politeness formulas. *Applied Linguistics* 8, 1: 75–88.
- Davis, Christopher and Daniel Gutzmann 2015. Use-conditional meaning and the semantics of pragmaticalization. *Proceedings of Sinn und Bedeutung* 19: 197–213.
- De Brosses, Charles 1765. *Traité de la formation mécanique des langues, et principes physiques de l'étymologie*. Two volumes. Paris: Chez Saillant/Vincent/Desaint.
- De Neys, W. 2006. Dual-processing in reasoning: two systems but one reasoner. *Psychological Science* 17, 5: 428–33.
- Degand, Liesbeth and Jacqueline Evers-Vermeul 2015. Grammaticalization or pragmaticalization of discourse markers?: More than a terminological issue. *Journal of Historical Pragmatics* 16, 1: 59–85.
- Dehé, Nicole 2014. *Parentheticals in Spoken English: The Syntax-Prosody Relation*. (Studies in English Language.) Cambridge: Cambridge University Press.
- Dehé, Nicole and Bettina Braun 2013. The prosody of question tags. *English Language and Linguistics* 17, 1: 129–57.
- Dehé, Nicole and Yordanka Kavalova 2006. The syntax, pragmatics, and prosody of parenthetical *what*. *English Language and Linguistics* 10: 289–320.

- Dehé, Nicole and Yordanka Kavalova 2007. *Parentheticals*. (Linguistics Today, 106.) Amsterdam, Philadelphia: Benjamins.
- Dehé, Nicole and Anne Wichmann 2010. Sentence-initial *I think (that)* and *I believe (that)*: Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in Language* 34, 1: 36–74.
- Dér, Csilla Ilona 2010. On the status of discourse markers. *Acta Linguistica Hungarica* 57, 1: 3–28.
- Dér, Csilla Ilona and Alexandra Markó 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53, 2: 135–80.
- Detges, Ulrich and Richard WALTERIT 2016. Grammaticalization and pragmaticalization. In S. Fischer and C. Gabriel (eds.), *Manual of Grammatical Interfaces in Romance*. (Manuals of Romance Linguistics.) Berlin: de Gruyter Mouton, pp. 635–57.
- Diffloth, Gérard 1972. Notes on expressive meaning. *Papers of the Chicago Linguistic Society* 8: 440–47.
- Diffloth, Gérard 1994. ‘i: big, a: small’. In Leanne Hinton, Johanna Nichols, and John J. Ohala (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 107–14.
- Dik, Simon C. 1997. *The Theory of Functional Grammar, Part 2: Complex and Derived Constructions*. (Functional Grammar Series, 21.) Berlin, New York: Mouton de Gruyter, pp. 379–409.
- Dimmendaal, Gerrit Jan 1983. *The Turkana Language*. Dordrecht: Foris Publications.
- Dingemanse, Mark 2011. *The Meaning and Use of Ideophones in Siwu*. Ph.D. dissertation, Nijmegen: Radboud University.
- Dingemanse, Mark 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6, 10: 654–72.
- Dingemanse, Mark 2013. Ideophones and gesture in everyday speech. *Gesture* 13, 2: 143–65.
- Dingemanse, Mark 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language* 39, 4: 946–70. (doi:10.1075/sl.39.4.05din)
- Dingemanse, Mark 2017a. Expressiveness and system integration: On the typology of ideophones, with special reference to Siwu. *STUF – Language Typology and Universals* 70, 2: 363–84.
- Dingemanse, Mark 2017b. On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In Nicholas J. Enfield (ed.), *Dependencies in Language*. Berlin: Language Science Press, pp. 195–202.
- Dingemanse, Mark 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics* 3, 1: 1–30.
- Dingemanse, Mark 2019. ‘Ideophone’ as a comparative concept. In Kimi Akita and P. Pardeshi (eds.), *Ideophones, Mimetics, and Expressives*. Amsterdam: Benjamins, pp. 13–33.
- Dingemanse, Mark 2020. Between sound and speech: Liminal signs in interaction. *Research on Language and Social Interaction* 53, 1: 188–96.
- Dingemanse, Mark 2021a. Ideophones. In Eva van Lier (ed.), *The Oxford Handbook of Word Classes*. Oxford: Oxford University Press (in press). (doi:10.31234/osf.io/u96zt)
- Dingemanse, Mark 2021b. Interjections. In Eva van Lier (ed.), *The Oxford Handbook of Word Classes*. Oxford: Oxford University Press (in press). (doi:10.31234/osf.io/ngcrs)
- Dingemanse, Mark and Kimi Akita 2017. An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics* 53, 3: 501–32.

- Dingemanse, Mark, Francisco Torreira, and Nicholas J. Enfield 2013. Is 'Huh?' a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PLoS ONE* 8, 11: e78273.
- Dixon, R. M. W. 2004. *The Jarawara Language of Southern Amazonia*. Oxford: Oxford University Press.
- Dixon, R. M. W. 2010a. *Basic Linguistic Theory. Volume 1: Methodology*. Oxford: Oxford University Press.
- Dixon, R. M. W. 2010b. *Basic Linguistic Theory. Volume 2: Grammatical Topics*. Oxford: Oxford University Press.
- Doke, Clement M. 1935. *Bantu Linguistic Terminology*. London: Longman.
- Doke, Clement M. [1927] 1988. *Textbook of Zulu Grammar*. Sixth edition. Cape Town: Maskew Miller Longman.
- Doke, Clement M. and S. M. Mofokeng [1957] 1985. *Textbook of Southern Sotho Grammar*. Cape Town: Maskew Miller Longman.
- Dorvlo, Kofi 2008. *A Grammar of Logba (Ikpana)*. Ph.D. dissertation. Utrecht: LOT (Netherlands Graduate School of Linguistics).
- Dostie, Gaétane 2004. *Pragmaticalisation et marqueurs discursifs: Analyse sémantique et traitement lexicographique*. Brussels: De Boeck & Larcier.
- Dostie, Gaétane 2009. Discourse markers and regional variation in French: A lexico-semantic approach. In Kate Beeching, Nigel Armstrong, and Françoise Gadet (eds.), *Sociolinguistic Variation in Contemporary French*. Amsterdam, Philadelphia: Benjamins, pp. 201–14.
- Du Bois, John 1985. Competing motivations. In John Haiman (ed.), *Iconicity in Syntax*. (Typological Studies in Language, 6.) Amsterdam: Benjamins, pp. 343–66.
- Du Bois, John W. 2007. The stance triangle. In Robert Engelbretson (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam, Philadelphia: Benjamins, pp. 139–82.
- Dwyer, David, and Lioba Moshi 2003. Primary and grammaticalized ideophones. In John M. Mugane (ed.), *Linguistic Typology and Representation of African Languages*. Trenton: Africa World Press, pp. 173–85.
- Eastman, Carol M. 1992. Swahili interjections: Blurring language-use/gesture-use boundaries. *Journal of Pragmatics* 18: 273–87.
- Eckardt, Regine 2014. *Dear Ede!* Semantics and pragmatics of vocatives. In Gutzmann, Daniel, Jan Kopping, and Cecile Meier (eds.), *Approaches to Meaning: Compositions, Values, Interpretation*. Leiden: Brill, pp. 223–49. DOI: 10.1163/9789004279377\_011.
- Ehlich, Konrad 1986. *Interjektionen*. Tübingen: Niemeyer.
- Ekman, Paul, Wallace V. Friesen, and Phoebe Ellsworth 1972. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon.
- Enfield, N. J. 2007. *A Grammar of Lao*. Berlin: Mouton de Gruyter.
- Enfield, N. J. 2017. *How We Talk: The Inner Workings of Conversation*. New York: Basic Books.
- Enfield, N. J., Tanya Stivers, Penelope Brown, Christina Englert, Katariina Harjunpää, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, Tiina Keisanen, Mirka Rauniomaa, Chase Wesley Raymond, Federico Rossano, Kyung-Eun Yoon, Inge Zwitterlood, and Stephen C. Levinson 2019. Polar answers. *Journal of Linguistics* 55: 277–304.
- Enfield, N. J., Tanya Stivers, and Stephen C. Levinson (eds.) 2010. Question–response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics* 42, 10: 2615–19.

- Epstein, Seymour 1994. Integration of the cognitive and the psychodynamic unconscious. *The American Psychologist* 49, 8: 709–24.
- Erman, Britt and Ulla-Britt Kotsinas 1993. Pragmaticalization: The case of *ba* and *you know*. *Studier i modern sprakvetenskap* 10: 76–92.
- Erman, Britt and Beatrice Warren 2000. The idiom principle and the open-choice principle. *Text* 20: 29–62.
- Espinal, Teresa 2013. On the structure of vocatives. In Barbara Sonnenhauser and Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing Between System and Performance*. (Trends in Linguistics. Studies and Monographs, 261.) Berlin: De Gruyter, pp. 109–32.
- Evans, Jonathan St. B. T. 2003. In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences* 7, 10: 454–9.
- Evans, Jonathan St. B. T. 2008. Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology* 59: 255–78.
- Evans, Jonathan St. B. T. and David E. Over 1996. *Rationality and Reasoning*. Hove: Psychology Press.
- Evans, Jonathan St. B. T. and K. E. Stanovich 2013. Dual-process theories of higher cognition: advancing the debate. *Perspectives on Psychological Science* 8, 3: 223–41.
- Evans, Nicholas 1992. 'Wanjh! Bonj! Nja!': Sequential organization and social deixis in Mayali interjections. *Journal of Pragmatics* 18: 225–44.
- Evans, Nicholas 2003. *Bininj Gun-Wok: A Pan-Dialectal Grammar of Mayali, Kunwinjku and Kune*. Two volumes. Canberra: Australian National University.
- Evans, Nicholas, Henrik Bergqvist, and Lila San Roque 2018a. The grammar of engagement I: framework and initial exemplification. *Language and Cognition* 10, 1: 110–40. doi:10.1017/langcog.2017.21
- Evans, Nicholas, Henrik Bergqvist, and Lila San Roque 2018b. The grammar of engagement II: typology and diachrony. *Language and Cognition* 10, 1: 141–70. doi:10.1017/langcog.2017.22
- Fagard, Benjamin 2010. *Évida, olha...: Imperatives as discourse markers and grammaticalization paths in Romance: A diachronic corpus study*. *Languages in Contrast* 10, 2: 245–67.
- Feldman, Harry 1986. *A Grammar of Awtuw*. (Pacific Linguistics Series B, 94.) Canberra: Australian National University.
- Ferguson, Charles A. 1976. The structure and use of politeness formulas. *Language in Society* 5: 137–51.
- Ferguson, Charles A. 1981. The structure and use of politeness formulas. (Reprint of Ferguson 1976.) In Florian Coulmas (ed.), *Explorations in Standardized Communication Situations and Prepatterned Speech. Volume 2: Conversational Routine*. (Janua Linguarum. Series Maior, 96.) Berlin: De Gruyter Mouton, pp. 21–36.
- Fivaz, Derek 1986. *A Grammar of Oshindonga (Wambo)*. (African Studies of the Academy, 1.) Windhoek: The Academy.
- Florici, Franck 2011. Le vocatif et la périphérie du système des cas: entre archaïsmes et innovations. *Mémoires de la Société Linguistique de Paris (Nouvelle Série)* 19: 103–34.
- Florici, Franck 2015. Reflections on Holophrastic Negation. Typescript, Université de Paris 3 – Sorbonne Nouvelle.
- Florici, Franck and Georgia Milioni 2021. *Un cazzo, un cavolo, un cacchio, un corno, un tubo*, ecc. come termini di polarità negativa? Typescript, Paris.
- Florici, Franck and Lucia Molinu 2012. Romance monosyllabic imperatives and markedness. In Thomas Stolz, Nicole Nau, and Cornelia Stroh (eds.), *Monosyllables: From Phonology to Typology*. (STUF Studia Typologica.) Berlin: Akademie Verlag, pp. 149–72.

- Floricić, Franck and Lucia Molinu 2018. Are Sardinian vocatives perfectly regular? In M. Grimaldi, R. Lai, L. Franco, and B. Baldi (eds.), *Structuring Variation in Romance Linguistics and Beyond: In Honour of Leonardo M. Savoia*. Amsterdam, Philadelphia: Benjamins, pp. 271–82.
- Foolen, Ad 1997. The expressive function of language: towards a cognitive semantic approach. In Susanne Niemeier and René Dirven (eds.), *The Language of Emotions*. Amsterdam, Philadelphia: Benjamins, pp. 15–31.
- Foolen, Ad 2012. The relevance of emotion for language and linguistics. In A. Foolen, U. M. Lüdtke, T. P. Racine and J. Zlatev (eds.), *Moving Ourselves, Moving Others. Motion and Emotion in Intersubjectivity, Consciousness and Language*. Amsterdam: Benjamins, pp. 349–68.
- Foolen, Ad 2016. Expressives. In Nick Riemer (ed.), *The Routledge Handbook of Semantics*. London, New York: Routledge, pp. 473–90.
- Foolen, Ad 2018. Modalpartikeln in Grammatiktheorien. In Elisabeth Leiss and Sonja Zeman (eds.), *Die Zukunft von Grammatik – Die Grammatik der Zukunft. Festschrift für Werner Abraham anlässlich seines 80. Geburtstags*. (Studien zur deutschen Grammatik, 92.) Tübingen: Stauffenburg, pp. 133–52.
- Fortune, G. 1962. *Ideophones in Shona*. An Inaugural Lecture given in the University College of Rhodesia and Nyasaland on April 28, 1962. London: Oxford University Press.
- Frajzyngier, Zygmunt 1996. *Grammaticalization of the Complex Sentence: A Case Study in Chadic*. Amsterdam, Philadelphia: Benjamins.
- Frankish, Keith 2010. Dual-process and dual system theories of reasoning. *Philosophy Compass* 5, 10: 914–26.
- Frank-Job, Barbara 2006. A dynamic-interactional approach to discourse markers. In Kerstin Fischer (ed.), *Approaches to Discourse Particles*. Amsterdam: Elsevier, pp. 359–74.
- Fraser, Bruce 1988. Types of English discourse markers. *Acta Linguistica Hungarica* 38: 19–33.
- Fraser, Bruce 1990. An approach to discourse markers. *Journal of Pragmatics* 14, 3: 383–98.
- Fraser, Bruce 1996. Pragmatic markers. *Pragmatics* 6, 2: 167–90.
- Fraser, Bruce 1999. What are discourse markers. *Journal of Pragmatics* 31: 931–52.
- Furkó, Bálint Péter 2005. *The Pragmatic Marker—Discourse Marker Dichotomy Reconsidered: The Case of well and of course*. Ph.D. dissertation, Debreceni Egyetem, Bölcsészettudományi Kar, Debrecen.
- Furkó, Bálint Péter 2014. Cooption over grammaticalization. *Argumentum* 10: 289–300.
- Gawrowski, Betram and Laura A. Creighton 2013. Dual-process theories. In D. E. Carlston (ed.), *The Oxford Handbook of Social Cognition*. New York, NY: Oxford University Press, pp. 282–312.
- Gehweiler, Elke 2008. From proper name to primary interjection: The case of *gee!* *Journal of Historical Pragmatics* 9, 1: 71–93.
- Gernsbacher, M. 1990. *Language Comprehension as Structure Building*. Hillsdale, NJ: Erlbaum.
- Ghaderi, Soleiman 2021. Outside the clause: Functions of Persian *na* ‘no’. Typescript.
- Ghaderi, Soleiman and Mohammad Amouzadeh 2020. Aspects of *bale* (‘yes’) in Persian discourse: The case of its functions, positions, and evolution. Typescript.
- Gil, David 2013. Para-linguistic usages of clicks. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Chapter 142. <http://wals.info/chapter/142>.
- Givón, T. 1984. *Syntax: A Functional-Typological Introduction*. Volume 1. Amsterdam, Philadelphia: Benjamins.



- Givón, T. 2001. *Syntax: An Introduction*. Volume 1. Amsterdam, Philadelphia: Benjamins.
- Givón, T. 2011. *Seadock, or the Last Battle of Berkeley*. (The Boz Trilogy, 1.) Durango: White Cloud Publishing.
- Goddard, Cliff 2014. Interjections and emotion (with special reference to 'surprise' and 'disgust'). *Emotion Review* 6, 1: 53–63.
- Goffman, Erving 1978. Response cries. *Language* 54, 4: 787–815.
- Goffman, Erving 1981. *Forms of Talk*. Oxford: Blackwell.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Graesser, A. C., M. Singer, and T. Trabasso 1994. Constructing inferences during narrative text comprehension. *Psychological Review* 101: 371–95.
- Greenberg, Joseph H. 1963. *The Languages of Africa*. The Hague: Mouton.
- Greene, S. B., G. McKoon, and R. Ratcliff 1992. Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning Memory and Cognition* 18: 266–83.
- Grenoble, Leonore A., Rebekah Baglini, and Martina Martinović 2015. Verbal gestures in Wolof. In Ruth Kramer, Elizabeth C. Zsiga, and One Tlale Boyer (eds.), *Selected Proceedings of the 44th Annual Conference on African Linguistics*. Somerville, MA: Cascadia Proceedings Project, pp. 110–21.
- Grice, H. Paul 1975. Logic and conversation. In Peter Cole and Jerry L. Morgan (eds.), *Speech Acts: Syntax and Semantics III*. New York: Academic Press, pp. 41–58.
- Güldemann, Tom 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey* (Empirical Approaches to Language Typology, 34). Berlin: Mouton de Gruyter.
- Guma, S. M. 1971. *An Outline Structure of Southern Sotho*. Pietermaritzburg: Shuter & Shooter.
- Günthner, Susanne and Katrin Mutz 2004. Grammaticalization vs. pragmaticalization? The development of pragmatic markers in German and Italian. In Walter Bisang, Nikolas Himmelmann, and Björn Wiemer (eds.), *What Makes Grammaticalization? A Look from its Fringes and its Components*. Berlin: Mouton de Gruyter, pp. 77–107.
- Guthrie, Malcolm 1971. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*. Volume 2. Farnborough: Gregg.
- Gutzmann, Daniel 2015. *Use-Conditional Meaning. Studies in Multidimensional Semantics*. Oxford: Oxford University Press.
- Gutzmann, Daniel 2019. *The Grammar of Expressivity*. Oxford: Oxford University Press.
- Gutzmann, Daniel and Robert Henderson 2019. Expressive updates, much? *Language* 95, 1: 107–35.
- Haegeman, Liliane and Virginia Hill 2013. The syntacticization of discourse. In Raffaella Folli, Christina Sevdali, and Robert Truswell (eds.), *Syntax and Its Limits*. Oxford: Oxford University Press, pp. 370–90.
- Hagoort, Peter, Colin M. Brown, and Tamara Y. Swaab 1996. Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain* 119: 627–49.
- Haiman, John 1980. *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. (Studies in Language Companion Series, 5.) Amsterdam: Benjamins.
- Haiman, John 1983. Iconic and economic motivation. *Language* 59: 781–819.
- Haiman, John 2018. *Ideophones and the Evolution of Language*. Cambridge: Cambridge University Press.

- Halliday, Michael A. K. and Ruqaiya Hasan 1976. *Cohesion in English*. London: Longman.
- Hansen, Maj-Britt Mosegaard 2008. *Particles at the Semantics/Pragmatics Interface: Synchronic and Diachronic Issues*. Oxford: Elsevier.
- Harjula, Lotta 2004. *The Ha Language of Tanzania: Grammar, Texts and Vocabulary*. (East African Languages and Dialects, 13.) Cologne: Köppe.
- Harries, Lyndon 1950. *A Grammar of Mwera*. Johannesburg: Witwatersrand University Press.
- Haselow, Alexander 2011. Discourse marker and modal particle: The functions of utterance-final *then* in spoken English. *Journal of Pragmatics* 43: 3603–23.
- Haselow, Alexander 2013. Arguing for a wide conception of grammar: The case of final particles in spoken discourse. *Folia Linguistica* 47, 2: 375–424.
- Haselow, Alexander 2016a. A processual view on grammar: Macrogrammar and the final field in spoken syntax. *Language Sciences* 54: 77–101.
- Haselow, Alexander 2016b. Intensifying adverbs ‘outside the clause’. In Gunther Kaltenböck, Evelien Keizer, and Arne Lohmann (eds.), *Outside the Clause*. (Studies in Language Companion Series, 178.) Amsterdam, Philadelphia: Benjamins, pp. 379–415.
- Haselow, Alexander 2017. *Spontaneous Spoken English: An Integrated Approach to the Emergent Grammar of Speech*. Cambridge: Cambridge University Press.
- Haselow, Alexander 2021. Discourse markers and brain lateralization: Evidence for dual language processing from neurological disorders. In Sylvie Hancil and Alexander Haselow (eds.), *Studies at the Grammar–Discourse Interface*. Amsterdam, Philadelphia: Benjamins, pp. 157–93.
- Havers, Wilhelm 1931. *Handbuch der erklärenden Syntax. Ein Versuch zur Erforschung der Bedingungen und Triebkräfte in Syntax und Stylistik*. Heidelberg: Carl Winter.
- Hayashi, Makoto and Kyung-Eun Yoon 2006. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formulation trouble. *Studies in Language* 30: 485–540.
- Hayashi, Makoto and Kyung-Eun Yoon 2010. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formulation trouble. In Nino Amiridze, Boyd H. Davis, and Marboyegaret Maclagan (eds.), *Fillers, Pauses and Placeholders*. (Typological Studies in Language, 93.) Amsterdam, Philadelphia: Benjamins, pp. 33–65.
- Heine, Bernd 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. New York, Oxford: Oxford University Press.
- Heine, Bernd 1997. *Cognitive Foundations of Grammar*. Oxford, New York: Oxford University Press.
- Heine, Bernd 2002. On the role of context in grammaticalization. In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*. (Typological Studies in Language, 49.) Amsterdam, Philadelphia: Benjamins, pp. 83–101.
- Heine, Bernd 2016a. On non-finiteness and canonical imperatives. In Claudine Chamoreau and Zarina Estrada-Fernández (eds.), *Finiteness and Nominalization*. (Typological Studies in Language, 113.) Amsterdam, Philadelphia: Benjamins, pp. 245–70.
- Heine, Bernd 2016b. Language contact and extra-clausal constituents: The case of discourse markers. In Gunther Kaltenböck, Evelien Keizer, and Arne Lohmann 2016. *Outside the Clause*. (Studies in Language Companion Series, 178.) Amsterdam, Philadelphia: Benjamins, pp. 243–72.
- Heine, Bernd 2019. Some observations on the dualistic nature of discourse processing. *Folia Linguistica* 53, 2: 411–42.

- Heine, Bernd, Gunther Kaltenböck, and Tania Kuteva 2013. On the origin of grammar. In Claire Lefebvre, Bernard Comrie, and Henri Cohen (eds.), *New Perspectives on the Origins of Language*. Amsterdam, Philadelphia: Benjamins, pp. 379–405.
- Heine, Bernd, Gunther Kaltenböck, and Tania Kuteva 2015. Some observations on the evolution of utterance-final particles. In Sylvie Hancil, Alexander Haselow, and Margje Post (eds.), *Final Particles*. (Trends in Linguistics, Studies and Monographs, 284.) Berlin, Boston: Walter de Gruyter, pp. 111–40.
- Heine, Bernd, Gunther Kaltenböck, and Tania Kuteva 2016. On insubordination and cooptation. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*. (Typological Studies in Language). Amsterdam, Philadelphia: Benjamins, pp. 39–63.
- Heine, Bernd, Gunther Kaltenböck, and Tania Kuteva 2020. On the status of *wh*-exclamatives in English. *Functions of Language* 27, 2: 207–33.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long 2013. An outline of discourse grammar. In Shannon Bischoff and Carmen Jany (eds.), *Functional Approaches to Language*. Berlin: Mouton de Gruyter, pp. 175–233.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long 2015. On Some Correlations between Grammar and Brain Lateralization. *Oxford Handbooks Online in Linguistics*. New York: Oxford University Press.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long. 2017. Cooptation as a discourse strategy. *Linguistics* 55: 1–43.
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press.
- Heine, Bernd, Christa König, and Karsten Legère 2015. On institutional frames in Akie: A discourse grammar approach. In Osamu Hieda (ed.), *Information Structure and Nilotic Languages*. (Studies in Nilotic Linguistics, 10.) Tokyo: Tokyo University of Foreign Studies, Research Institute for Languages and Cultures of Asia and Africa, pp. 141–57.
- Heine, Bernd and Tania Kuteva 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva 2005. *Language Contact and Grammatical Change*. (Cambridge Approaches to Language Contact, 3.) Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva 2007. *The Genesis of Grammar: A Reconstruction*. (Studies in the Evolution of Language, 9.) Oxford: Oxford University Press.
- Heine, Bernd, Tania Kuteva, and Gunther Kaltenböck 2014. Discourse Grammar, the dual-process model, and brain lateralization: Some correlations. *Language and Cognition* 6, 1: 146–80.
- Heine, Bernd, Tania Kuteva, and Haiping Long 2020. Dual-process frameworks on reasoning and linguistic discourse: a comparison. In Alexander Haselow and Gunther Kaltenböck (eds.), *Grammar and Cognition: Dualistic Models of Language Structure and Language Processing*. Amsterdam, Philadelphia: Benjamins, pp. 59–89.
- Hengeveld, Kees and J. Lachlan Mackenzie 2008. *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford: Oxford University Press.
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33, 2–3: 61–83.
- Heritage, John 1998. *Oh*-prefaced responses to inquiry. *Language in Society* 27, 3: 291–334.
- Heritage, John 2002. *Oh*-prefaced responses to assessments: A method of modifying agreement/disagreement. In Cecilia Ford, Barbara Fox, and Sandra Thompson (eds.), *The Language of Turn and Sequence*. New York: Oxford University Press, pp. 196–224.
- Hewitt, B. G. 1995. *Georgian: A Structural Reference Grammar*. Amsterdam, Philadelphia: Benjamins.

- Hinds, John 1986. *Japanese*. London, New York: Routledge.
- Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203: 88–111.
- Hopper, Paul J. and Elizabeth C. Traugott 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huddleston, Rodney and Geoffrey K. Pullum 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hyland, Ken 1998. Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics* 30: 437–55.
- Hyland, Ken 2005. *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.
- Hyland, Ken 2017. Metadiscourse: What is it and where is it going? *Journal of Pragmatics* 113: 16–29.
- Ifantidou-Trouki, Elly 1993. Sentential adverbs and relevance. *Lingua* 90, 1–2: 69–90.
- Jackendoff, Ray 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Science* 3, 7: 272–9.
- Jackendoff, Ray 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jacobsen, William H., Jr. 1994. Nootkan vocative vocalism and its implications. In Leanne Hinton, Johanna Nichols, and John J. Ohala (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 23–39.
- Jakobson, Roman 1956. Metalanguage as a linguistic problem. In Roman Jakobson (ed.), *The Framework of Language*. Ann Arbor: Horace H. Rackham School of Graduate Studies, pp. 81–92.
- Jakobson, Roman 1960. Linguistics and poetics. In Thomas A. Sebeok (ed.), *Style in Language*. Cambridge, Ma.: M.I.T. Press, pp. 350–77.
- Jakobson, Roman 1971. Zur Struktur des russischen Verbums. In Roman Jakobson, *Selected Writings. Vol. II*. The Hague: Mouton, pp. 3–15.
- Janney, Richard W. 1999. Words as gestures. *Journal of Pragmatics* 31: 953–72.
- Jay, Timothy 2020. Ten issues facing taboo word scholars. In Nico Nassenstein and Anne Storch (eds.), *Swearing and Cursing: Contexts and Practices in a Critical Linguistic Perspective*. (Language and Social Life, 22.) Berlin: De Gruyter Mouton, pp. 37–51.
- Jendraschek, Gerd 2001. Semantic and structural properties of Turkish ideophones. *Turkic Languages* 5: 88–103.
- Jendraschek, Gerd 2012. *A Grammar of Iatmul*. Habilitationsschrift, Universität Regensburg.
- Jespersen, Otto 1922. *Language: Its Nature, Development and Origin*. London: Allen and Unwin.
- Jespersen, Otto 1924. *The Philosophy of Grammar*. London: Allen and Unwin.
- Johnstone, Barbara 2002. *Discourse Analysis*. Malden, Oxford: Blackwell.
- Joseph, Brian D. 1997. On the linguistics of marginality: The centrality of the periphery. *Chicago Linguistic Society* 33: 197–213.
- Jucker, Andreas H. 1993. The discourse marker *well*: a relevance-theoretical account. *Journal of Pragmatics* 19, 5: 435–52.
- Jucker, Andreas H. 1997. The discourse marker *well* in the history of English. *English Language and Linguistics* 1, 1: 91–110.
- Jucker, Andreas H. 2002. Discourse markers in Early Modern English. In Richard Watts and Peter Trudgill (eds.), *Alternative Histories of English*. London, New York: Routledge, pp. 210–30.
- Jung-Beeman, Mark 2005. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences* 9: 512–18.

- Kahneman, Daniel 2012. *Thinking, Fast and Slow*. [First published 2011 by Farrar, Straus and Giroux, USA]. London: Penguin Books.
- Kahneman, Daniel and S. Frederick 2002. Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, and D. Kahneman (eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge, M.A.: Cambridge University Press, pp. 49–81.
- Kahneman, Daniel and A. Tversky 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3: 430–54.
- Kahneman, Daniel and A. Tversky 1973. On the psychology of prediction. *Psychological Review* 80: 237–51.
- Kaltenböck, Gunther 2007. Spoken parenthetical clauses in English. In Nicole Dehé and Yordanka Kavalova (eds.), *Parentheticals*. (Linguistics Today, 106.) Amsterdam, Philadelphia: Benjamins, pp. 25–52.
- Kaltenböck, Gunther and Bernd Heine 2014. Sentence grammar vs. thetical grammar: Two competing domains. In Brian MacWhinney, Andrej Malchukov, and Edith Moravcsik (eds.), *Competing Motivations in Grammar and Usage*. Oxford: Oxford University Press, pp. 348–63.
- Kaltenböck, Gunther, Bernd Heine, and Tania Kuteva 2011. On thetical grammar. *Studies in Language* 35, 4: 848–93.
- Kaplan, David 1999. The meaning of *ouch* and *oops*: Explorations in the theory of meaning as use, transcribed by Elizabeth Coppock. Howison Lecture in Philosophy, delivered at the University of California at Berkeley.
- Karani, Michael and Alexander Andrason 2021. Ideophones in Arusa Maasai—syntax, morphology and phonetics. Typescript.
- Karcevski, Serge 1941. Introduction à l'étude de l'interjection. *Cahiers Ferdinand de Saussure* 1: 57–75.
- Kavalova, Yordanka 2007. And-parenthetical clauses. In Nicole Dehé and Yordanka Kavalova (eds.), *Parentheticals*. (Linguistics Today, 106.) Amsterdam, Philadelphia: Benjamins, pp. 145–72.
- Kilian-Hatz, Christa 2001. Universality and diversity: Ideophones from Baka and Khoe. In Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*. (Typological Studies in Language, 44.) Amsterdam: Benjamins, pp. 155–63.
- Kilian-Hatz, Christa 2006. Ideophones. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, pp. 508–12.
- Kim, Stephanie Hyeri and Sung-Ock Sohn 2015. Grammar as an emergent response to interactional needs: A study of final *kuntay* 'but' in Korean conversation. *Journal of Pragmatics* 83: 73–90.
- Kintsch, W. 1974. *The Representation of Meaning in Memory*. Hillsdale, NJ: Erlbaum.
- Kintsch, W. 1988. The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review* 95: 163–82.
- Kita, Sotaro 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35: 379–415.
- Kleinknecht, Friederike 2013. Mexican *güey*—from vocative to discourse marker: A case of grammaticalisation? In Barbara Sonnenhauser and Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing between System and Performance*. Berlin: De Gruyter Mouton, pp. 235–68. <https://doi.org/10.1515/9783110304176.235>
- Kleinknecht, Friederike 2019. *Der Vokativ und seine Verwendung im gesprochenen Spanish*. Ph.D. dissertation, University of Munich.

- Kleinknecht, Friederike and Miguel Souza 2017. Vocatives as a source category for pragmatic markers: From deixis to discourse marking via affectivity. In Chiara Fedriani and Andrea Sansò (eds.), *Pragmatic Markers, Discourse Markers and Modal Particles*. (Studies in Language Companion Series, 186.) Amsterdam, Philadelphia: Benjamins, pp. 257–87.
- Knowles, John 1980. The tag as a parenthetical. *Studies in Language* 4, 3: 379–409.
- Kockelman, Paul 2003. The meanings of interjections in Q'eqchi' Maya: From emotive reaction to social and discursive action. *Current Anthropology* 44: 467–90.
- König, Christa, Bernd Heine, and Karsten Legère 2015a. *The Akie Language of Tanzania: A Sketch of Discourse Grammar*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (Tokyo University of Foreign Studies).
- König, Christa, Bernd Heine, and Karsten Legère 2015b. Discourse Markers in Akie, a Southern Nilotic Language of Tanzania. In Osamu Hieda (ed.), *Information Structure and Nilotic Languages*. (Studies in Nilotic Linguistics, 10.) Tokyo: Tokyo University of Foreign Studies, Research Institute for Languages and Cultures of Asia and Africa, pp. 117–39.
- Koo, Hyun Jung and Seongha Rhee 2018. Ideophones and attenuatives in Korean. Paper presented at the 51st Annual Meeting, Societas Linguistica Europaea, Tallinn University, Estonia, August 29–September 1, 2018.
- Koroma, Regine 1994. *Die Morphosyntax des Gola*. (Afrikanistische Monographien, 4.) Cologne: Institut für Afrikanistik, Universität zu Köln.
- Kruglanski, A. W., H. P. Erb, A. Pierro, L. Mannetti and W. Y. Chun 2006. On parametric continuities in the world of binary either ors. *Psychological Inquiry* 17: 153–65.
- Kryk, Barbara 1992. The pragmatics of interjections: The case of Polish. *Journal of Pragmatics* 18: 193–207.
- Kunene, Daniel 1978. *The Ideophone in Southern Sesotho*. Berlin: Reimer.
- Kuryłowicz, Jerzy [1949] 1973. Le problème du classement des cas. In Jerzy Kuryłowicz (ed.), *Esquisses Linguistiques I*. Munich: Wilhelm Fink, pp. 131–50.
- Kuteva, Tania and Bernd Heine 2021. On the structure of early language. Analytic vs holistic language processing and grammaticalization. In Nathalie Gontier, Andy Lock, and Chris Sinha (eds.), *The Oxford Handbook of Human Symbolic Evolution*. Oxford: Oxford University Press, pp. 1–45.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, and Seongha Rhee 2019. *World Lexicon of Grammaticalization*. Second extensively revised and updated edition. Cambridge: Cambridge University Press.
- Lakoff, George 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lambrecht, Knud 1996. On the formal and functional relationship between topics and vocatives: evidence from French. In Adele E. Goldberg (ed.), *Conceptual Structure, Discourse and Language*. (CSLI Publications.) Stanford: Center for the Study of Language and Information, pp. 267–88.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford, New York: Oxford University Press.
- Lébikaza, Kézié Koyenzi 1999. *Grammaire kabiyè: une analyse systématique: phonologie, tonologie et morphosyntaxe*. (Grammatische Analysen afrikanischer Sprachen, 11.) Cologne: Köppe.
- Leech, Geoffrey, Margaret Deuchar, and Robert Hoogenraad 1982. *English Grammar for Today*. London: Longman.
- Lee-Goldman, Russell 2011. No as a discourse marker. *Journal of Pragmatics* 43: 2627–49.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

- Levisen, C. 2019. Laughter interjections: Contributions to a lexical anthropology of humour (with special reference to Danish). *Scandinavian Studies in Language* 10, 1: 110–30.
- Lewis, Diana M. 2000. *Some Emergent Discourse Connectives in English: Grammaticalization via Rhetorical Patterns*. Ph.D. dissertation, University of Oxford, Faculty of English Language.
- Ley, R. G. 1980. *Emotion and the Right Hemisphere*. Doctoral Dissertation, University of Waterloo.
- Li, Charles N. and Sandra A. Thompson 1981. *Mandarin Chinese. A Functional Reference Grammar*. Berkeley, Los Angeles: University of California Press.
- Liljegen, Henrik 2016. *A grammar of Palula*. (Studies in Diversity Linguistics, 8.) Berlin: Language Science Press.
- Lionnet, Florian 2020. Paralinguistic use of clicks in Chad. In Bonny Sands (ed.), *Click Consonants*. Leiden: Brill, pp. 422–37. Doi:10.1163/9789004424357\_015
- Ljung, Magnus 2011. *Swearing: A Cross-Cultural Linguistic Study*. Basingstoke: Palgrave Macmillan.
- Lusekelo, Amani 2021. Linguistic aspects of the forms of address in Nyakyusa. *Language in Africa* 2, 1: 62–90.
- Lydall, Jean 2000. Having fun with ideophones: A sociolinguistic look at ideophones in Hamar. In Baye Yimam, Richard Pankhurst, David Chapple, Yonas Admassu, Alula Pankhurst, and Birhanu Teferra (eds.), *Ethiopian Studies at the End of the Second Millennium*. Volume 2. Addis Ababa: Addis Ababa University, pp. 886–911.
- Lyons, John 1977. *Semantics*. 2 volumes. Cambridge: Cambridge University Press.
- Mackenzie, J. Lachlan 1998. The basis of syntax in the holophrase. In Mike Hannay and A. Machtelt Bolkestein (eds.), *Functional Grammar and Verbal Interaction*. Amsterdam, Philadelphia: Benjamins, pp. 267–95.
- Mackenzie, J. Lachlan 2019. The syntax of an emotional expletive in English. In J. Lachlan Mackenzie and Laura Alba-Juez (eds.), *Emotion in Discourse*. Amsterdam: Benjamins, pp. 55–86.
- MacWhinney, Brian, Andrej Malchukov, and Edith Moravcsik (eds.) 2014. *Competing Motivations in Grammar and Usage*. Oxford: Oxford University Press.
- Martin, Samuel 1975. *A Reference Grammar of Japanese*. New Haven: Yale University Press.
- Maschler, Yael 1994. Metalanguaging and discourse markers in bilingual conversation. *Language in Society* 23: 325–66.
- Maschler, Yael 2009. *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam, Philadelphia: Benjamins.
- Matisoff, James A. [1979] 2000. *Blessings, Curses, Hopes, and Fears: Psycho-Ostensive Expressions in Yiddish*. Second Edition. Stanford, CA: Stanford University Press.
- Matras, Yaron 2009. *Language Contact*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.
- Matthews, Peter H. [1997] 2007. *The Concise Oxford Dictionary of Linguistics*. Second edition. Oxford: Oxford University Press.
- Mauranen, A. 2010. Discourse reflexivity—a discourse universal? The case of ELF. *Nordic Journal of English Studies* 9, 2: 13–40.
- Mauri, Caterina and Andrea Sansò 2011. How directive constructions emerge: Grammaticalization, constructionalization, cooptation. *Journal of Pragmatics* 43, 14: 3489–3521.
- McArthur, Tom 1992. *The Oxford Companion to the English Language*. Oxford: Oxford University Press.
- McCready, Elin 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3, 8: 1–57. doi: 10.3765/sp.3.8.

- McGregor, William 1990. *A Functional Grammar of Gooniyandi*. (Studies in Language Companion Series, 22.) Amsterdam, Philadelphia: Benjamins.
- McKoon, G. and R. Ratcliff 1980. Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior* 19: 369–86.
- McKoon, G. and R. Ratcliff 1992. Inference during reading. *Psychological Review* 99: 440–66.
- McKoon, G. and R. Ratcliff 1998. Memory based language processing: Psycholinguistic research in the 1990's. *Annual Review of Psychology* 49: 25–42.
- Meinard, Maruszka Eve Marie 2015. Distinguishing onomatopoeias from interjections. *Journal of Pragmatics* 76: 150–68.
- Méndez-Naya, Belén 2006. Adjunct, modifier, discourse marker: On the various functions of *right* in the history of English. *Folia Linguistica Historica* 27, 1–2: 141–95.
- Michaelis, Laura A. 2001. Exclamative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language Typology and Language Universals*. Berlin, New York: de Gruyter, pp. 436–56.
- Miltersen, Ehm Hjorth 2019. Øv – a problem for categorization. *Scandinavian Studies in Language* 10, 1: 41–62.
- Mitchell, Alice 2017. The pragmatics of a kinship term: The meaning and use of 'mother' in Datooga (Nilotic). In Raija Kramer and Roland Kießling (eds.), *Mechthildian Approaches to Afrikanistik: Advances in Language Based Research on Africa, Festschrift für Mechthild Reh*. Cologne: Köppe, pp. 287–301.
- Mitchell, Alice 2020. 'Oh, bald father!': Kinship and swearing among Datooga of Tanzania. In Nico Nassenstein and Anne Storch (eds.), *Swearing and Cursing: Contexts and Practices in a Critical Linguistic Perspective*. (Language and Social Life, 22.) Berlin: De Gruyter Mouton, pp. 79–101.
- Möhlhlig, Wilhelm J. G. 1967. *Die Sprache der Dciriku: Phonologie, Prosodologie und Morphologie*. Ph.D. dissertation, University of Cologne.
- Moore, Erin 2007. *Yeah-No: A Discourse Marker in Australian English*. Honours thesis, University of Melbourne, Melbourne.
- Morris, H. F. and B. E. R. Kirwan 1957. *A Runyankore Grammar*. Nairobi, Kampala and Dar-es-Salaam: Eagle Press.
- Mortensen, Sune Sønderberg 2019. Interjections in American and Danish courtroom interaction: A linguistic and legal cultural comparison. *Scandinavian Studies in Language* 10, 1: 152–73.
- Moscovitch, M. 1983. The linguistic and emotional functions of the normal right hemisphere. In E. Perecman (ed.), *Cognitive Processing in the Right Hemisphere*. New York: Academic Press, pp. 57–82.
- Müller, Friedrich Max 1861. *Lectures on the Science of Language: Delivered at the Royal Institution of Great Britain in April, May, & June 1861*. London: Longman.
- Munaro, Nicola 2010. On the edge-feature of particles, interjections, and short answers. In Anna Maria Di Sciullo and Virginia Hill (eds.), *Edges, Heads, and Projections*. Amsterdam, Philadelphia: Benjamins, pp. 67–86. Doi: 10.1075/la.156.07mun.
- Munaro, Nicola and Cecilia Poletto 2004. Sentential particles and clausal typing in the Veneto dialects. In Benjamin Shaer, Werner Frey, and Claudia Maienborn (eds.), *ZAS Papers in Linguistics* 35, 2: 375–97.
- Mwinlaaru, Isaac Nuokya-Ire 2017. *A Systemic Functional Description of the Grammar of Dagaare*. Ph.D. dissertation, The Hong Kong Polytechnic University.



- Narrog, Heiko and Bernd Heine 2021. *Grammaticalization*. Oxford: Oxford University Press.
- Nassenstein, Nico and Anne Storch (eds.) 2020. *Swearing and Cursing: Contexts and Practices in a Critical Linguistic Perspective*. (Language and Social Life, 22.) Berlin: De Gruyter Mouton.
- Nespoulous, J. L. 1980. De deux comportements verbaux de base: référentiel et modalisateur. De leur dissociation dans le discours aphasique. *Cahiers de Psychologie* 23: 195–210.
- Nespoulous, J. L., C. Code, J. Virbel, and A. R. Lecours 1998. Hypotheses on the dissociation between ‘referential’ and ‘modalizing’ verbal behaviour in aphasia. *Applied Psycholinguistics* 19: 311–31.
- Nevalainen, Terttu and Helena Raumolin-Brunberg 1995. Constraints on politeness: The pragmatics of address formulae in Early English correspondence. In Andreas H. Jucker (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English*. Amsterdam, Philadelphia: Benjamins, pp. 541–601.
- Newman, Paul 1968. Ideophones from a syntactic point of view. *Journal of West African Languages* 5: 107–17.
- Newman, Paul 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven, London: Yale University Press.
- Noel Aziz Hanna, Patrizia and Barbara Sonnenhauser 2013. Vocatives as functional performance structures. In Barbara Sonnenhauser and Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing between System and Performance*. (Trends in Linguistics. Studies and Monographs, 261). Berlin: De Gruyter, pp. 283–305.
- Noel Aziz Hanna, Patrizia and Barbara Sonnenhauser 2019. Introduction: Approaching a syntax of pragmatics. *Sprachwissenschaft* 44, 2: 1–11.
- Noora, Azam and Mohammad Amouzadeh 2015. Pragmaticalization of Yæ’ni in Persian. *International Journal of Language Studies* 9, 1: 91–122.
- Nordbustad, Frøydis 1988. *Iraqw Grammar: An Analytical Study of the Iraqw Language*. (Language and Dialect Studies in East Africa.) Berlin: Reimer.
- Norde, Muriel 2009. *Degrammaticalization*. Oxford: Oxford University Press.
- Norricks, Neal R. 2009. Interjections as pragmatic markers. *Journal of Pragmatics* 41: 866–91.
- Nübling, Damaris 2001. Von *oh mein Jesus!* zu *oje!* Der Interjektionalisierungspfad von der sekundären zur primären Interjektion. *Deutsche Sprache* 29, 1: 20–45.
- Nübling, Damaris 2004. Die prototypische Interjektion: ein Definitionsvorschlag. *Zeitschrift für Semiotik* 26: 11–46.
- Nübling, Damaris and Marianne Vogel 2004. Fluchen und Schimpfen kontrastiv: zur sexuellen, krankheitsbasierten, skatologischen und religiösen Fluch- und Schimpfwortprototypik im Niederländischen, Deutschen und Schwedischen. *Germanistische Mitteilungen* 59: 19–33.
- Nuckolls, Janis B. 1996. *Sounds like Life: Sound-Symbolic Grammar, Performance, and Cognition in Pastaza Quechua*. New York: Oxford University Press.
- Nuyts, Jan 2005 Modality: Overview and linguistic issues. In William Frawley (ed.), *The Expression of Modality*. (The Expression of Cognitive Categories 1.) Berlin, New York: Mouton de Gruyter, pp. 1–26.
- O’Connell, Daniel C. and Sabine Kowal 2004. The history of research on the filled pause as evidence of *The written language bias in linguistics* (Linell, 1982). *Journal of Psycholinguistic Research* 33: 459–74.

- O'Connell, Daniel C. and Sabine Kowal 2005. *Uh* and *um* revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research* 34: 555–76.
- O'Connell, Daniel C. and Sabine Kowal 2008. *Cognition and Language: A Series in Psycholinguistics. Communicating with One Another: Toward a Psychology of Spontaneous Spoken Discourse*. Berlin: Springer Science and Business Media.
- O'Connell, Daniel C., Sabine Kowal, and C. Ageneau 2005. Interjections in interviews. *Journal of Psycholinguistic Research* 34: 153–71.
- Ocampo, Francisco 2006. Movement towards discourse is not grammaticalization: the evolution of /claro/ from adjective to discourse particle in spoken Spanish. In Nuria Sagarra and Almeida Jacqueline Toribio (eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Somerville, M.A.: Cascadilla Proceedings Project, pp. 308–19.
- Ogi, Naomi 2017. *Involvement and Attitude in Japanese Discourse: Interactive Markers*. Amsterdam, Philadelphia: Benjamins.
- Onodera, Noriko Okada 2011. The grammaticalization of discourse markers. In Heiko Narrog and Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press, pp. 611–20.
- Orozco, Inti Aedo 2016. Evaluatives in Grammar. M.A. dissertation, University of Mainz.
- Osman, Magda 2004. An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review* 11: 988–1010.
- Padilla Cruz, Manuel 2009. Might interjections encode concepts? More questions than answers. *Lodz Papers in Pragmatics* 5, 2: 241–70.
- Parker, Steve 2021. A cross-linguistic corpus of forms meaning 'yes'. Typescript.
- Paton, W. F. 1971. *Ambrym (Lonwolwol) Grammar*. (Pacific Linguistics Series B, 19.) Canberra: The Australian National University.
- Pawley, Andrew 1992. Formulaic speech. In William H. Bright (ed.) 1992. *Oxford International Encyclopedia of Linguistics*. Volume 2. New York, Oxford: Oxford University Press, pp. 22–25.
- Pawley, Andrew 2009. Grammarians' languages versus humanists' languages and the place of speech act formulas in models of linguistic competence. In Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen M. Wheatley (eds.), *Formulaic Language. Volume 1: Distribution and Historical Change*. (Typological Studies in Language, 82.) Amsterdam, Philadelphia: Benjamins, pp. 3–26.
- Pawley, Andrew and F. Syder 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards and R. Schmidt (eds.), *Language and Communication*. London: Longman, pp. 191–225.
- Peirce, Charles Sanders 1932. The icon, index, and symbol. In C. Hartshorne and P. Weiss (eds.), *Collected Papers of Charles Sanders Peirce*. Volume 2. Cambridge, M.A.: Harvard University Press, pp. 156–73.
- Peña, Jaime Germán 2015. *A Grammar of Wampis*. Ph.D. dissertation, University of Oregon.
- Pennington, Ryan 2016. *A Grammar of Ma Manda: A Papuan Language of Morobe Province, Papua New Guinea*. Ph.D. dissertation, James Cook University, Cairns.
- Pinker, Steven 2007. *The Stuff of Thought: Language as a Window into Human Nature*. New York: Viking.
- Pitkin, Harvey 1984. *Wintu Grammar*. (University of California Publications in Linguistics, 94.) Berkeley and Los Angeles: University of California Press.
- Pons Bordería, Salvador and Ana Belén Llopis Cardona 2020. Some reflections on semantic–pragmatic cycles. *Journal of Historical Pragmatics* 21, 2: 317–47.
- Poplack, Shana 1980. Sometimes I'll start a sentence in English Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics* 18, 7-8: 581–618.

- Potts, Christopher 2007a. The expressive dimension. *Theoretical Linguistics* 33: 165–98.
- Potts, Christopher 2007b. The centrality of expressive indices. *Theoretical Linguistics* 33: 255–68.
- Poulos, George 2021. *On the Origins of Human Speech and Language*. Amazon Publishing.
- Poulos, George and Louis J. Louwrens 1994. *A Linguistic Analysis of Northern Sotho*. Pretoria: Via Afrika Limited.
- Prat, Chantel S., Debra L. Long, and Kathleen Baynes 2007. The representation of discourse in the two hemispheres: An individual differences investigation. *Brain and Language* 100, 3: 283–94.
- Prévost, Sophie 2011. *A propos* from verbal complement to discourse marker: a case of grammaticalization? *Linguistics* 49, 2: 391–413.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik 1985. *A Comprehensive Grammar of the English Language*. London, New York: Longman.
- Raymond, Geoffrey 2004. Prompting action: The stand-alone ‘So’ in ordinary conversation. *Research on Language and Social Interaction* 37: 185–218.
- Reber, A. S. 1993. *Implicit Learning and Tacit Knowledge*. Oxford: Oxford University Press.
- Reinhart, Tanya 1983. Point of view in language: The use of parentheticals. In Gisa Rauh (ed.), *Essays on Deixis*. Tübingen: Narr, pp. 169–94.
- Rennison, John R. 1996. *Koromfe*. London, New York: Routledge.
- Rett, Jessica 2011. Exclamatives, degrees and speech acts. *Linguistics and Philosophy* 34, 411–42.
- Reyna, Valerie F. and Charles J. Brainerd 1995. Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences* 7: 1–75. doi:10.1016/1041-6080(95)90031-4.
- Rhee, Seongha 2004. From discourse to grammar: Grammaticalization and lexicalization of rhetorical questions in Korean. In Gordon Fulton, William J. Sullivan, and Arle R. Lommel (eds.), *LACUS: Forum XXX: Language, Thought and Reality*. Houston: Lacus, pp. 413–23.
- Rhee, Seongha 2019. Lexicalization patterns in color naming in Korean. In Ida Raffaelli, Daniela Katunar, and Barbara Kerovec (eds.), *Lexicalization Patterns in Color Naming: A Cross-linguistic Perspective*. (Studies in Functional and Structural Linguistics, 78.) Amsterdam: John Benjamins, pp. 109–31.
- Robins, R. H. 1958. *The Yurok Language: Grammar, Texts, Lexicon*. (University of California Publications in Linguistics, 15.) Berkeley and Los Angeles: University of California Press.
- Rosch, Elenor 1973. Natural categories. *Cognitive Psychology* 4: 328–50.
- Ross, John 1970. On declarative sentences. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*. Waltham: Ginn, pp. 222–72.
- Rosten, Leo 1968. *The Joys of Yiddish*. New York: McGraw-Hill.
- Rousseau, Jean-Jacques 1990. *Essai sur l'origine des langues*. Edited by J. Starobinski. Paris: Gallimard.
- Sadock, Jerry 1969. Super-hypersentences. *Papers in Linguistics* 1, 1: 1–15.
- Sakai, Kuniyoshi L., Yoshinori Tatsuno, Kei Suzuki, Harumi Kimura, and Yasuhiro Ichida 2005. Sign and speech: Amodal commonality in left hemisphere dominance for comprehension of sentences. *Brain* 128, 6: 1407–17.
- Salmons, Joseph 1990. Bilingual discourse marking: code-switching, borrowing and convergence in some German-American dialects. *Linguistics* 29: 453–80.
- Samarin, William J. 1965. Perspective on African ideophones. *African Studies* 24: 117–21.
- Sapir, Edward 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.

- Schaden, Gerhard 2010. Vocatives: A note on addressee-management. *University of Pennsylvania Working Papers in Linguistics* 16, 1: 176–85.
- Schegloff, Emanuel A. 1968. Sequencing in conversational openings. *American Anthropologist* 70: 1075–95.
- Schegloff, Emanuel A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. and Harvey Sacks 1973. Opening up closings. *Semiotica* 7: 289–327.
- Schiffrin, Deborah 1987. *Discourse Markers*. (Studies in Interactional Sociolinguistics, 5.) Cambridge: Cambridge University Press.
- Schleicher, August [1862] 1877] *Compendium of the Comparative Grammar of the Indo-European, Sanskrit, Greek and Latin Languages. Part II: Morphology*. London, Trubner. (Translation of Schleicher, August 1862. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. II Formenlehre. Weimar).
- Schneider, Stefan 2007a. Reduced parenthetical clauses in Romance languages. In Nicole Dehé and Yordanka Kavalova (eds.), *Parentheticals*. (Linguistics Today, 106.) Amsterdam, Philadelphia: Benjamins, pp. 237–58.
- Schneider, Stefan 2007b. *Reduced Parenthetical Clauses as Mitigators: A Corpus Study of Spoken French, Italian and Spanish*. Amsterdam, Philadelphia: Benjamins.
- Schourup, Lawrence 1999. Discourse markers. *Lingua* 107: 227–65.
- Schourup, Lawrence 2003. Comment on Paul Kockelman 2003. The meanings of interjections in Q'eqchi' Maya: From emotive reaction to social and discursive action. *Current Anthropology* 44: 482–3.
- Schrock, Terrill 2014. *A Grammar of Ik (Icé-tód): Northeast Uganda's Last Thriving Kuliak language*. Doctoral thesis. Utrecht: LOT (Leiden University).
- Schwentner, Ernst 1924. *Die primären Interjektionen in den indogermanischen Sprachen mit besonderer Berücksichtigung des Griechischen, Lateinischen und Germanischen*. Heidelberg: Carl Winter's Universitätsbuchhandlung.
- Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Sherratt, Sue 2007. Right brain damage and the verbal expression of emotion: A preliminary investigation. *Aphasiology* 21, 3–4: 320–33.
- Sidtis, Diana, Gina Canterucci, and Dora Katsnelson 2009. Effects of neurological damage on production of formulaic language. *Clinical Linguistics and Phonetics* 23, 4: 270–84.
- Siemund, Peter 2015. Exclamative clauses in English and their relevance for theories of clause types. *Studies in Language* 39, 3: 697–727.
- Sloman, Steven A. 1996. The empirical case for two systems of reasoning. *Psychological Bulletin* 119: 3–22.
- Sohn, Ho-Min 1994. *Korean*. (Descriptive Grammars.) London: Routledge.
- Sohn, Ho-Min 1999. *The Korean Language*. Cambridge: Cambridge University Press.
- Sohn, Ho-Min and B. W. Bender 1973. *A Ulithian Grammar*. (Pacific Linguistics Series C, 27.) Canberra: The Australian National University.
- Sonnenhauser, Barbara and Patrizia Noel Aziz Hanna (eds.) 2013. *Vocative! Addressing between System and Performance*. (Trends in Linguistics. Studies and Monographs, 261.) Berlin: De Gruyter, pp. 1–23.
- Soukka, Maria 2000. *A Descriptive Grammar of Noon: A Cangin Language of Senegal*. Munich: Lincom Europa.

- Stange, Ulrike 2009. *The Acquisition of Interjections in Early Childhood*. Hamburg: Diplomica.
- Stange, Ulrike 2016. *Emotive Interjections in British English: A Corpus-Based Study on Variation in Acquisition, Function and Usage*. Amsterdam: Benjamins.
- Stange, Ulrike 2019. The social life of emotive interjections in spoken British English. *Scandinavian Studies in Language* 10, 1: 174–93.
- Stange, Ulrike and Damaris Nübling 2014. Interjections. In Cornelia Müller, Ellen Fricke, Alan Cienki, and David McNeill (eds.), *Body – Language – Communication*. Berlin: Mouton de Gruyter, pp. 1982–89.
- Stanovich, Keith E. 1999. *Who is Rational? Studies of Individual Differences in Reasoning*. Mahwah, NJ: Erlbaum.
- Stanovich, Keith E. 2004. *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. Chicago: University of Chicago Press.
- Stappers, Leo 1986. *Kanyok: eine Sprachskizze. Afrikanistische Arbeitspapiere, special issue*. Cologne: Schriftenreihe des Kölner Instituts für Afrikanistik.
- Stenström, Anna-Brita 2020. English- and Spanish-speaking teenagers' use of rude vocatives. In Nico Nassenstein and Anne Storch (eds.), *Swearing and Cursing: Contexts and Practices in a Critical Linguistic Perspective*. (Language and Social Life, 22.) Berlin: De Gruyter Mouton, pp. 281–301.
- Stephens, Richard, John Atkins, and Andrew Kingston 2009. Swearing as a response to pain. *NeuroReport* 20, 12: 1056–60.
- Stivers, Tanya and N. J. Enfield 2010. A coding scheme for question–response sequences in conversation. *Journal of Pragmatics* 42, 10: 2620–26.
- Stolz, Thomas 2007. *Allora*: On the recurrence of function-word borrowing in contact situations with Italian as donor language. In Jochen Rehbein, Christiane Hohenstein, and Lukas Pietsch (eds.), *Connectivity in Grammar and Discourse*. Amsterdam, Philadelphia: Benjamins, pp. 75–99.
- Taavitsainen, Irma 1995. Interjections in Early Modern English: From imitation of spoken to conventions of written language. In Andreas H. Jucker (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English*. Amsterdam, Philadelphia: Benjamins, pp. 439–65.
- Tannen, Deborah and Piyale Cömert Öztekin 1981. Health to our mouths: Formulaic expressions in Turkish and Greek. In Florian Coulmas (ed.), *Explorations in Standardized Communication Situations and Prepatterned Speech*. Volume 2. *Conversational Routine*. (Janua Linguarum. Series Maior, 96.) Berlin: De Gruyter Mouton, pp. 37–54.
- Taylor, John R. 1989. *Linguistic Categorization: Prototypes in Linguistic Theory*. Second edition, 1995. Oxford: Oxford University Press.
- Thompson, Lawrence C. and M. Terry Thompson 1992. *The Thompson Language*. Missoula: University of Montana Press.
- Tosco, Mauro 2001. *The Dhaasanac Language: Grammar, Text, Vocabulary of a Cushitic Language in Ethiopia*. Cologne: Köppe.
- Tottie, Gunnel 2011. *Uh* and *um* as sociolinguistic markers in British English. *International Journal of Corpus Linguistics* 16: 173–96.
- Tottie, Gunnel 2014. On the use of *uh* and *um* in American English. *Functions of Language* 21: 6–29.
- Tottie, Gunnel 2015. *Uh* and *um* in British and American English: Are they words? Evidence from co-occurrence with pauses. In Rena Torres Cacoulios, Nathalie Dion, and André Lapierre (eds.), *Linguistic Variation: Confronting Fact and Theory*. New York: Routledge, pp. 38–55.

- Tottie, Gunnel and Sebastian Hoffmann 2006. Tag questions in British and American English. *Journal of English Linguistics* 34, 4: 283–311.
- Traill, Anthony 1994. *A !Xóñ Dictionary*. (Quellen zur Khoisan-Forschung, 9.) Cologne: Köppe.
- Trask, R. L. 1993. *A Dictionary of Grammatical Terms in Linguistics*. London and New York: Routledge.
- Traugott, Elizabeth Closs 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred Lehmann and Yakov Malkiel (eds.), *Directions for Historical Linguistics*. Amsterdam, Philadelphia: John Benjamins. 245–71.
- Traugott, Elizabeth Closs 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65, 1: 31–55.
- Traugott, Elizabeth Closs 1995. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the International Conference of Historical Linguistics XII, Manchester 1995.
- Traugott, Elizabeth Closs 2003. From subjectification to intersubjectification. In Raymond Hickey (ed.), *Motives for Language Change*. Cambridge: Cambridge University Press, pp. 124–39.
- Traugott, Elizabeth Closs 2007. Discussion article: Discourse markers, modal particles, and contrastive analysis, synchronic and diachronic. *Catalan Journal of Linguistics* 6: 139–57.
- Traugott, Elizabeth Closs 2018. Modeling language change with constructional networks. In Salvador Pons Bordería and Óscar Loureda Lamas (eds.), *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change*. (Studies in Pragmatics, 18.) Leiden, Boston: Brill, pp. 17–50.
- Traugott, Elizabeth Closs 2022. *Discourse Structuring Markers in English: A Historical Constructionalist Perspective on Pragmatics*. (Constructional Approaches to Language, 33.) Amsterdam, Philadelphia: Benjamins.
- Traugott, Elizabeth C. and Graeme Trousdale 2014. Contentful constructionalization. *Journal of Historical Linguistics* 4, 2: 254–82.
- Tucker, Archibald N. 1994. *A Grammar of Kenya Luo (Dholuo)*, edited by C. A. Creider. Volume 1 (Nilo-Saharan, 8.1.) Cologne: Köppe.
- Tversky, A. and Daniel Kahneman 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124–31.
- van den Berg, René 1989. *A Grammar of the Muna Language*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 139.) Dordrecht: Foris Publications.
- van Dijk, Teun A. 1980. *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition*. Hillsdale, New Jersey: Erlbaum.
- Van Lancker, Diana 1988. Nonpropositional speech: Neurolinguistic studies. In A. Ellis (ed.), *Progress in the Psychology of Language*. Volume 3. London: L. Erlbaum, pp. 49–118.
- Van Lancker, Diana 1990. The neurology of proverbs. *Behavioral Neurology* 3: 169–87.
- Van Lancker, Diana 1997. Rags to riches: Our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. *Brain and Language* 57, 1: 1–11.
- Van Lancker Sidtis, Diana 2004. When novel sentences spoken or heard for the first time in the history of the universe are not enough: Toward a dual-process model of language. *International Journal of Language and Communication Disorders* 39, 1: 1–44.

- Van Lancker Sidtis, Diana 2009. Formulaic and novel language in a 'dual-process' model of language competence: Evidence from surveys, speech samples, and schemata. In Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen M. Wheatley (eds.), *Formulaic Language*. Volume 2. *Acquisition, Loss, Psychological Reality, and Functional Explanations*. (Typological Studies in Language, 83.) Amsterdam, Philadelphia: Benjamins, pp. 445–70.
- Van Lancker Sidtis, Diana and Whitney A. Postman 2006. Formulaic expressions in spontaneous speech of left- and right-hemisphere damaged subjects. *Aphasiology* 20, 5: 411–26.
- Velupillai, V. 2012. *An Introduction to Linguistic Typology*. Amsterdam: John Benjamins.
- Vendler, Zeno 1967. *Linguistics and Philosophy*. Ithaca: Cornell University Press.
- Veselinova, Ljuba N. 2006. *Suppletion in Verb Paradigms: Bits and Pieces of the Puzzle*. (Typological Studies in Language, 67.) Amsterdam, Philadelphia: Benjamins.
- Vindenes, Urd 2018. Cyclic renewal of demonstratives. *Studies in Language* 42, 3: 641–68.
- Voeltz, F. K. Erhard and Christa Kilian-Hatz (eds.) 2001. *Ideophones*. (Typological Studies in Language, 44.) Amsterdam, Philadelphia: Benjamins.
- Voorhoeve, Clemens Lambertus 1965. *The Flamingo Bay Dialect of Asmat*. Ph.D. dissertation, Leiden University.
- Waltereit, Richard 2002. Imperatives, interruption in conversation and the rise of discourse markers: a study of Italian *guarda*. *Linguistics* 40: 987–1010.
- Waltereit, Richard 2006. The rise of discourse markers in Italian: A specific type of language change. In Kerstin Fischer (ed.), *Approaches to Discourse Markers*. (Studies in Pragmatics, 1.) Amsterdam: Elsevier, pp. 61–76.
- Watanabe, Honoré 2010. Fillers and their relevance in describing Sliammon Salish. In Nino Amiridze, Boyd H. Davis, and Margaret MacLagan (eds.), *Fillers, Pauses and Placeholders*. (Typological Studies in Language, 93.) Amsterdam, Philadelphia: Benjamins, pp. 173–87.
- Weinreich, Uriel [1966] 1989. On the semantic structure of language. In William Labov and Beatrice S. Weinreich (eds.), *Uriel Weinreich: On Semantics*. Philadelphia: University of Pennsylvania Press, pp. 37–96.
- Westermann, Diedrich 1912. *A Short Grammar of the Shilluk Language*. Philadelphia: The Board of Foreign Missions of the United Presbyterian Church of N.A.; Berlin: Reimer (E. Vohsen).
- Westermann, Diedrich 1927. Laut, Ton und Sinn in westafrikanischen Sudansprachen. In Franz Boas (ed.), *Festschrift Meinhof*. Hamburg: Friederichsen, pp. 315–28.
- Westermann, Diedrich 1937. Laut und Sinn in einigen westafrikanischen Sprachen. *Archiv für Vergleichende Phonetik* 1: 154–72, 193–211.
- Wharton, Tim 2003. Interjections, language, and the 'showing/saying' continuum. *Pragmatics and Cognition* 11, 1: 39–91.
- Wharton, Tim 2009. *Pragmatics and Non-Verbal Communication*. Cambridge: Cambridge University Press.
- Wharton, Tim 2016. That bloody so-and-so has retired: Expressives revisited. *Lingua* 175–6: 20–35.
- Wichmann, Anne 2004. The intonation of *please*-requests: A corpus-based study. *Journal of Pragmatics* 36, 1521–49.
- Wichmann, Anne 2005. *Please* – from courtesy to appeal: The role of intonation in the expression of attitudinal meaning. *English Language and Linguistics* 9: 229–53.

- Wichmann, Anne 2011. Grammaticalization and prosody. In Heiko Narrog and Bernd Heine (eds.), *The Oxford Handbook of Grammaticalization*. New York: Oxford University Press, pp. 331–41.
- Wierzbicka, Anna 1992. The semantics of interjections. *Journal of Pragmatics* 18: 159–92.
- Wierzbicka, Anna 2003a. *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Berlin: De Gruyter.
- Wierzbicka, Anna 2003b. Comment on Kockelman, Paul 2003. The meanings of interjections in Q'eqchi' Maya: From emotive reaction to social and discursive action. *Current Anthropology* 44: 483–84.
- Wierzbicka, Anna 2021. 'Semantic primitives', fifty years later. *Russian Journal of Linguistics* 25, 2: 317–42.
- Wilkins, David P. 1989. *Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar*. Ph.D. dissertation, The Australian National University, Canberra.
- Wilkins, David P. 1992. Interjections as deictics. *Journal of Pragmatics* 18: 119–58.
- Wilson, Deirdre and Dan Sperber 1993. Linguistic form and relevance. *Lingua* 90, 12: 1–25.
- Wiltchko, Martina 2021. *The Grammar of Interactional Language*. Cambridge: Cambridge University Press.
- Wittgenstein, Ludwig 1953. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Blackwell.
- Woods, Rebecca 2021. Towards a model of the syntax–discourse interface: A syntactic analysis of *please*. *English Language and Linguistics* 25, 1: 121–53.
- Wright, Melissa 2011. On clicks in English talk-in-interaction. *Journal of the International Phonetic Association* 41, 2: 207–29.
- Wundt, Wilhelm 1911. *Völkerpsychologie*. First volume: *Die Sprache*. Leipzig: Engelmann.
- Yliniemi, Juha 2019. *A Descriptive Grammar of Denjongke (Sikkimese Bhutia)*. Ph.D. dissertation, University of Helsinki.
- Yngve, Victor 1970. On getting a word in edgewise. *Proceedings of the Chicago Linguistic Society* 6: 567–78.
- Zhang, M. 2016. A multidimensional analysis of metadiscourse markers across written registers. *Discourse Studies* 18, 2: 204–22.
- Zwicky, Arnold M. 1974. 'Hey, Whatsyourname.' In Michael La Galy, Robert Fox, and Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting. Chicago Linguistic Society. April 19–21, 1974*. Chicago: Chicago Linguistic Society, pp. 787–801.
- Zwicky, Arnold M. 1985. Clitics and particles. *Language* 61: 283–305.
- Zwicky, Arnold M. and Geoffrey K. Pullum 1987. Plain morphology and expressive morphology. In John Aske, Natasha Beery, Laura Michaelis, and Hana Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* 7: 330–40.



# Index of Languages

- Abbruzzese 247  
‡Akhoe Haillom 208, 409  
Akie 117*t*, 118, 124, 180–1, 202, 203, 212, 236, 239  
Albanian 117*t*  
Ambrym 117*t*, 218  
Arabic 138, 219, 284, 383–4  
Arabic, Algerian 115, 125  
Arabic, Moroccan 230  
Arabic, San'ani 219  
Arabic, Tunisian 114–15  
Arusa Maasai, *see* Maasai, Arusa  
Asmat 272  
Asturian 245 n.1  
Avar 219  
Awtuw 254  
Ayt Hadiddu 252, 253, 254, 255, 256, 257, 258*t*, 259, 260, 261, 299  
  
Bahnar 167, 170, 178, 180  
Baka 2, 152, 162  
Basque 146, 178, 181, 258  
Bulgarian 117*t*, 241, 246, 293  
  
Cambodian 152, 404  
Catalan 242, 243*t*  
Cha'palaa 208*t*  
Chinese, Standard 135, 202, 208*t*  
Croatian 241  
Czech 241  
  
Dagaare 246  
Dahalo 39 n.1  
Danish 72, 117*t*, 297, 299  
Datooga 247, 271  
Dciriku 117*t*  
Denjongke 22, 34, 134–5  
Diding'a 160, 167  
Diyari 13, 20, 117*t*  
Dutch 117*t*, 135, 270, 293  
Dyirbal 33  
  
English, Early Middle 291  
English, Old 275, 291  
Ewe 35, 111, 112, 115, 116, 117*t*, 118, 187, 222, 236, 253, 260, 275, 383  
  
Fox 243  
French 18 n.2, 68, 72, 111 n.1, 117*t*, 135, 208*t*, 218, 235 n.1, 281, 284, 287, 303, 306, 307*t*, 371, 378  
  
Gbeya 146, 178, 180, 146  
Georgian 251  
German 23, 24, 42, 65 n.1, 72, 117*t*, 187, 194, 208*t*, 216*t*, 222, 246, 258, 270, 275, 281, 285, 286, 290, 291, 301–2, 304, 307, 316*t*, 372  
Gilaki 117*t*, 118–19  
Gola 163–4  
Gooniyandi 20, 23, 33, 117*t*, 124*t*  
  
Ha 245  
Hadza 39 n.1  
Hamar 163  
Hausa 13, 155, 158, 244, 258, 316*t*, 317  
Hebrew 124–5, 219, 220  
  
Iatmul 152, 218, 317  
Icelandic 208*t*  
Ik 126–7, 284  
Indonesian 258  
Iraqw 316  
Italian 111 n.1, 117, 242, 275, 282, 287, 289 n.1, 293, 303, 304, 307, 378, 383, 398, 401, 402  
  
Japanese 8, 16, 21, 29, 36, 62 n.1, 74, 75, 76, 77, 108, 135, 144–5, 146, 150, 151, 152, 154, 155, 158–9, 161, 176, 178, 180, 230–1, 275, 301, 307, 314 n.1, 315, 318, 396  
Jarawara 115, 117*t*, 118 n.1, 220  
  
Kabiye 117*t*, 118, 149  
Kalenjin 117*t*, 118  
Kannada 263  
Kanyok 212  
Kashmiri 241  
Khoekhoe 39 n.1  
Khwe 153  
Kikongo 208  
Kisi 22–3, 73, 152, 203, 212, 328  
Kivunjo 152, 156

- Korean 5, 21, 22, 29, 34, 115, 117*t*, 132, 135, 146,  
     147, 148, 166, 167, 170, 178, 179–80, 190, 208,  
     212, 213*t*, 219, 263, 308, 327  
 Koromfe 159–60  
 Kwakiutl 244–5  
  
 Laal 40–1  
 Lahu 208*t*  
 Lakota 117*t*, 118 n.2  
 Lao 208*t*, 251  
 Latin 2 n.2, 232, 240–1, 307, 374  
 Latvian 232, 241  
 Lele 284  
 Lithuanian 232  
 Logba 117*t*, 118 n.3, 167  
 Lushai 258  
  
 Ma Manda 117, 243  
 Maale 155, 156, 166, 239, 251  
 Maasai, Arusa 3, 163, 202, 250–1, 254, 255, 256,  
     257*t*, 261  
 Macedonian 241  
 Malayalam 317  
 Manambu 32–3, 244, 255*t*  
 Mandarin, *see* Chinese, Standard  
 Mayali 2, 14, 28, 58, 91, 124*t*, 140, 215, 275,  
     399–400, 414  
 Mayan 24–5, 33, 208*t*, 212, 293, 317  
 Mparntwe Arrernte 33, 123, 216–17  
 Muna 34, 117*t*, 118 n.4, 240, 251, 255, 260, 271  
 Mundang 160  
 Murrinh Patha 208*t*  
  
 Ndebele 258*t*  
 Ndonga 33, 34  
 Nepali 117*t*  
 Nkore-Kiga 263  
 Nlaka'pamuetsin, *see* Thompson  
     (Nlaka'pamuetsin)  
 Noon 241, 253  
 Northern Sotho, *see* Sotho, Northern  
 Nyakyusa 245  
  
 Owambo 33  
  
 Paamese 21, 23, 182, 201–2, 367  
 Palula 244, 246  
 Paresi 13  
 Persian 138, 211, 213, 214, 383  
 Pitjantjatjara 260, 299  
 Polish 187, 241, 296, 297, 299, 376, 378–9  
 Portuguese 117*t*, 242, 271, 307, 378, 409  
 Punjabi 241, 246  
  
 Q'eqchi' 24–5, 33, 89, 289, 293  
 Quechua 263  
  
 Romanian 117*t*, 118 n.6  
 Runyankore 242  
 Russian 117, 242, 284, 292, 307  
  
 Sardinian 244  
 Sardinian, Logudorese 125, 244  
 Semai 146, 180  
 Semelai 316–17  
 Serbian 241  
 Shona 147, 150, 160, 161, 162–3, 172, 261, 274,  
     312  
 Sicilian 247  
 Siwu 149, 152, 154, 156–7, 159, 161, 163, 169,  
     170, 175, 208, 263, 301, 302, 309, 311, 312,  
     314–15, 400  
 Slovak 117*t*  
 Somali 219, 317  
 Sorbian, Upper 241  
 Sotho, Northern 159, 160, 162, 169  
 Sotho, Southern 34, 45, 119, 154, 181, 252, 253,  
     315, 316*t*  
 Spanish 111 n.1, 135–6, 137–8, 208*t*, 245, 248,  
     271, 283–4, 285, 286, 307, 378, 383  
 Supyire 159  
 Swahili 35, 109, 126, 138, 150, 151, 152, 160,  
     176, 187, 219, 221–2, 226–7, 257*t*, 272–4, 275,  
     279, 302, 314, 320, 367, 376 n.1, 379, 383  
  
 Tamil 36, 264  
 Thai 124  
 Thompson (Nlaka'pamuetsin) 20, 23, 117*t*, 243,  
     328, 329  
 Tigemaxo 117*t*, 205  
 Tjwao 117*t*  
 Totonac 138, 264  
 Tswana 40, 117*t*, 119, 164, 181 n.2, 255  
 Turkish 72, 138, 156, 292, 308, 383  
 Tzotzil 117*t*  
  
 Udihe 317  
 Ukrainian 241  
 Ulithian 118*t*, 245  
  
 Venda 73, 317  
 Vietnamese 258

Wampis 33, 244, 246–7, 327–8

Wintu 118*t*, 118 n.7

Wolof 40

Xhosa 20, 22, 28, 40, 42, 118*t*, 173, 174, 190, 197,  
199, 252, 260, 287, 299, 301–2, 315, 316*t*

!Xóǀ 39–40, 41, 122

!Xun 39 n.1, 42, 125–6

Yir-Yoront 20

Yucatec 317

Yurok 243

Zargulla 249, 251, 252, 254

Zulu 20, 39, 118*t*, 146, 178, 180, 243*t*, 244, 302,  
328, 329*t*

# Index of Names

- Acharya, Jayaraj 117*t*, 118 n.5  
 Ädel, Annelie 339  
 Ageneau, C. 24, 135  
 Aijmer, Karin 127, 130, 132, 276  
 Aikhenvald, Alexandra Y. 32, 120, 124, 125, 244, 255*t*  
 Akita, Kimi 16, 36, 62, 64, 68, 74, 75, 146, 152, 154, 155, 161, 177, 301, 302, 308, 314 n.1, 315, 316–17, 318–19  
 Allan, Keith 306*t*, 307  
 Alpher, Barry 20  
 Ameka, Felix K. 5, 6, 9, 12, 13, 22, 25, 35, 57, 60, 111, 112, 115–16, 117*t*, 118, 122, 123, 146, 183, 184, 185, 186, 187, 188, 191–2, 193*t*, 196, 209, 222, 223–4, 236, 248–9, 253, 260, 263 n.1, 275, 281, 298, 299, 306*t*, 319, 326–7, 365, 366, 370, 371, 373, 374, 376, 380–1, 383, 411  
 Amha, Azeb 38, 48, 155, 156, 159, 166, 239, 248, 249, 251, 252, 253, 254, 314  
 Amouzadeh, Mohammad 138, 213, 214, 276  
 Andrason, Alexander 1, 3, 8, 13, 16, 20, 22, 24, 28, 30, 32, 38, 91, 108, 113, 114, 117*t*, 118*t*, 120, 122, 125, 146, 148, 149, 160, 163, 169, 173, 179, 181, 183, 187, 190, 197, 199, 248–9, 250, 251, 252, 254, 255, 256, 257, 258–9, 260, 261, 263, 287, 295, 296, 297, 299, 301, 315, 316*t*, 366, 370, 417  
 Anstatt, Tanja 232, 241  
 Armstrong, David 34  
 Arnovick, Leslie 276  
 Atkins, John 352  
 Austin, Peter 13, 20, 117*t*  
 Baglini, Rebekah 40  
 Baldi, Sergio 150, 151, 160, 176, 273  
 Bally, C. 25, 78  
 Bannard, Colin 342  
 Barth-Weingarten, Dagmar 31  
 Baynes, Kathleen 340  
 Beeching, Kate 129, 130, 131, 132  
 Beeman, Mark 355  
 Bender, B. W. 118, 245  
 Benfey, Theodor 4, 6, 367, 371  
 Bergqvist, Henrik 47  
 Biber, Douglas 8, 12, 13, 16, 19, 24, 33, 45, 46, 80, 81 n.2, 92, 94, 96, 98, 101, 105, 109–11, 112, 114, 123, 134, 142, 143, 145, 184, 187, 190, 193, 194, 198, 199, 200–1, 204, 206–7, 209, 210, 211, 212, 213, 214, 215, 216, 217, 221, 224, 225–6, 227, 232, 233, 235–6, 237, 238–9, 242, 267, 283, 291, 304, 306*t*, 330–3, 345, 363, 371, 372, 374, 401, 406, 414  
 Bickerton, Derek 388  
 Blakemore, Diane 67, 337, 346, 404  
 Blecke, Thomas 117*t*, 205  
 Blonder, Lee Xenakis 355  
 Bloom, R. 355  
 Bloomfield, Leonard 182, 192, 209, 243, 323, 373  
 Bolinger, Dwight 23  
 Bolly, Catherine 276  
 Borchmann, Simon 183, 184, 185, 186  
 Börjars, Kersti 300  
 Borod, J. C. 355  
 Bowers, Dawn 355  
 Brainerd, Charles 340, 356, 393, 394  
 Branch, C. 353  
 Brandão, Ana Paula Barros 13  
 Braun, Bettina 202, 423  
 Brauner, Siegmund 261  
 Brenzinger, Matthias 39, 40, 41, 122  
 Brinton, Laurel J. 10, 127, 129, 140, 277, 313, 347  
 Brown, Colin M. 209, 212, 219, 355, 408, 409  
 Brown, Penelope 36, 205, 211, 212, 239  
 Buechel, Eugene 118 n.2, 177*t*  
 Bühler, Karl 25, 78, 222  
 Burridge, Kate 211, 306*t*, 307  
 Bussmann, Hadumod 148, 373, 375  
 Bynon, James 38, 248, 250, 253, 254, 255, 256, 258, 259, 260, 261*t*, 299  
 Caffi, Claudia 2 n.1, 185  
 Canterucci, Gina 343 n.1  
 Carter, Ronald 3, 184, 352  
 Cassirer, Ernst 375  
 Cecil, Matthew Jones 215  
 Chiarello, Christine 355  
 Childs, G. Tucker 3, 8, 13, 19, 22–3, 73, 108, 145, 149, 152, 158, 161, 162, 165 n.1, 167, 168, 203, 209, 212, 312, 374  
 Clancy, Patricia 215

- Claridge, Claudia 2, 129, 131  
 Clark, Eve V. 385  
 Clark, Herbert H. 25, 69–72, 73, 74–6, 77, 79, 133, 134, 135, 177, 194*t*, 271–2, 336, 337, 371, 388  
 Cleeremans, A. 360  
 Code, J. Virbel 338  
 Cole, Desmond 40, 117*t*, 119, 164, 181 n.2, 232, 255  
 Conklin, Kathy 379, 395  
 Coulmas, Florian 7, 61, 220, 221, 231, 275  
 Couper-Kuhlen, Elizabeth 31  
 Creider, Chet A. 117*t*, 118  
 Creider, Jane Tapsubei 117*t*, 118  
 Creighton, Laura A. 357 n.2  
 Croft, William 295, 347  
 Crowley, Terry 21, 23, 182, 202, 367  
 Crystal, David 374  
 Cuenca, Maria Josep 130, 183, 365
- D'Atri, Annabella 385  
 Davies, Eirlyse 63, 220, 221  
 Davis, Christopher 334  
 De Brosse, Charles 68, 371  
 De Neys, W. 358  
 Degand, Liesbeth 276  
 Dehé, Nicole 202, 346, 347  
 Dér, Csilla Ilona 31, 127  
 Detges, Ulrich 276  
 Deuchar, Margaret 323, 324*t*, 366, 371  
 Diffloth, Gérard 29, 30, 64, 146, 158, 167, 170, 174, 310, 338–9, 368, 372  
 Dik, Simon C. 337  
 Dimmendaal, Gerrit Jan 160, 172, 175  
 Dingemanse, Mark 6, 14, 16, 19, 22, 30, 31, 36, 44, 64, 66, 68, 69, 71, 72, 73, 74, 75, 77, 78, 79, 95, 116, 119, 145, 146, 147–8, 149, 150, 152, 154, 155, 156, 157, 158, 159, 161, 163, 164, 165, 166, 168, 169, 171, 172, 174, 175, 176, 177, 179, 181, 206, 207, 208, 263, 272, 300, 301, 302, 303, 308, 309–10, 311, 313, 314, 315, 316, 317, 318, 334, 368, 371, 380, 400, 412  
 Dixon, R. M. W. 12, 33, 115, 117*t*, 118 n.1, 187, 196, 220, 365, 374  
 Dlali, Mawande 8, 13, 20, 22, 28, 108, 118*t*, 120, 173, 183, 187, 190, 199, 252, 260, 299, 301, 366, 370  
 Doke, Clement M. 20, 39, 118*t*, 119, 147, 149, 154, 170, 181, 232, 243*t*, 302, 315, 328, 329*t*  
 Dorvlo, Kofi 117*t*, 118 n.3, 167  
 Dostie, Gaétane 276  
 Du Bois, John 48, 141, 183, 209–10, 389, 391, 411
- Dwyer, David 15, 16, 28, 68, 78, 80, 146, 147, 148, 152, 154, 156, 165, 168, 170, 307, 308, 312, 313, 314, 318, 340, 360*t*, 368, 379, 397
- Eastman, Carol M. 35, 109, 221–2, 279, 367, 376 n.1, 379  
 Eckardt, Regine 232  
 Ehlich, Konrad 23, 24*t*, 216*t*  
 Ekman, Paul 186, 196  
 Ellsworth, Phoebe 186, 196  
 Enfield, N. J. 31, 44, 72, 116, 119, 127, 133, 206, 207–8, 209, 210, 212, 215, 223, 251, 371, 408–9  
 Epstein, Seymour 357, 359  
 Erman, Britt 29, 276  
 Espinal, Teresa 232, 235 n.1, 238  
 Evans, Jonathan St. B. T. 47, 357, 358, 359, 361, 394  
 Evans, Nicholas 2, 6, 8, 9, 14, 19, 26, 27–8, 58, 66, 91, 124*t*, 140, 183, 184 n.1, 185, 191, 215, 275, 337, 370, 399, 414  
 Evers-Vermeul, Jacqueline 276
- Fagard, Benjamin 378  
 Fehn, Anne-Maria 1, 16, 24, 30, 108, 113–14, 117*t*, 149, 160, 179, 190, 255  
 Feldman, Harry 254  
 Ferguson, Charles A. 6, 7, 17, 32, 42, 221, 226, 383–4, 394  
 Fivaz, Derek 34  
 Florey, Margret 211  
 Floricic, Franck 18 n.1–2, 22 n.1, 34, 46 n.2, 51, 63 n.1, 68 n.1, 80 n.1, 99 n.1, 102 n.1, 111 n.1, 125*t*, 196 n.1, 219 n.1, 232, 234, 235 n.1, 242, 244, 245 n.1, 247, 282 n.1, 287, 289 n.1, 368, 374 n.1, 376 n.1, 383  
 Foolen, Ad 7, 51, 63, 249 n.1, 268, 298  
 Fortune, G. 147, 149, 150, 159, 160, 162, 274, 312  
 Fox Tree, Jean E. 25, 133, 134, 135, 337, 371  
 Frajzyngier, Zygmunt 284  
 Frank-Job, Barbara 398, 401, 402  
 Frankish, Keith 356, 357, 360*t*, 396  
 Fraser, Bruce 31, 48, 96, 97, 128, 129, 183, 233, 238, 282, 324, 325, 377  
 Frederick, S. 242, 359  
 Friesen, Wallace V. 186, 196  
 Furkó, Bálint Péter 97, 130, 337
- Gawrowski, Betram 357 n.2  
 Gehweiler, Elke 8, 80, 191, 198, 292, 304, 305, 306, 307*t*  
 Gernsbacher, M. 340  
 Gerrig, Richard J. 69, 71 n.1  
 Ghaderi, Soleiman 211, 213, 214

- Gil, David 34, 40, 218, 219, 389  
 Givón, T. 1, 2, 37 n.1, 63  
 Goddard, Cliff 57, 66, 193, 370  
 Goffman, Erving 4, 16, 35, 44, 56, 66, 187, 188, 197, 291, 365, 366, 367, 370, 371, 373, 404  
 Goldberg, Adele E. 295, 347  
 Graesser, A. C. 340  
 Greenberg, Joseph H. 39 n.1  
 Greene, S. B. 340  
 Grenoble, Leonore A. 40  
 Grice, H. Paul 408  
 Güldemann, Tom 16, 73, 75, 146, 153, 154, 178, 371, 372  
 Guma, S.M. 34, 45, 117t, 253  
 Günthner, Susanne 276  
 Guthrie, Malcolm 274  
 Gutzmann, Daniel 16, 65, 66, 81, 144–5, 333, 334, 335  
  
 Haegeman, Liliane 18 n.3, 117t, 118 n.6, 278, 348  
 Hagoort, Peter 355  
 Haiman, John 7, 13, 16, 19, 21, 22, 28, 30, 32, 35, 43 n.1, 46, 56, 57, 58, 64, 66, 68, 72, 73, 74, 79, 108, 128, 146, 149, 152, 156, 159, 164, 165, 166t, 167, 168, 170, 171, 172–3, 174, 175, 176, 177, 185–6, 232, 263, 308, 313, 317, 334, 345, 370, 371–2, 374, 375, 386, 391, 397, 404, 416  
 Halliday, Michael A. K. 45  
 Hansen, Maj-Britt Mosegaard 276  
 Harjula, Lotta 245  
 Harries, Lyndon 274t  
 Hasan, Ruqaiya 45  
 Haselow, Alexander 15, 132, 344–5, 352, 353, 429  
 Havers, Wilhelm 360  
 Hayashi, Makoto 133, 134, 135, 209, 212, 219, 366, 408, 409  
 Haywood, C. Santschi 355  
 Heilman, Kenneth M. 355  
 Heine, Bernd 17, 19, 22, 24, 43, 46, 51, 53, 71 n.1, 97, 114, 115, 117t, 118, 120, 124, 125, 126t, 127, 128, 129, 132, 133, 134, 137, 139, 140, 194, 197, 202, 203, 204, 210, 212, 227, 233, 236, 239, 268, 269, 275, 277, 278t, 281, 282, 284, 288, 295, 298, 300, 308, 309, 332, 336, 337, 345, 346, 347, 348, 352, 353, 354, 355, 357 n.1, 360t, 363, 382, 383, 386–7, 391, 400 n.1, 414  
 Heine, Steven J. 356  
 Henderson, Robert 16, 65, 66, 81–2, 144–5, 333, 334, 335  
 Hengeveld, Kees 8, 15, 112, 116, 181, 220, 232, 239, 240, 241, 242, 246, 293, 337  
  
 Henrich, Joseph 356  
 Heritage, John 190, 290, 291  
 Hewitt, B. G. 251  
 Hill, Virginia 18 n.3, 117t, 118 n.7, 278, 348  
 Hinds, John 21  
 Hockett, Charles F. 28, 38, 43, 44, 388  
 Hoffmann, Sebastian 18, 203–4  
 Hoogenraad, Robert 323, 324t, 366, 371  
 Hopper, Paul J. 46, 277  
 Huddleston, Rodney 12, 25, 184  
 Hyland, Ken 17, 339  
  
 Ifantidou-Trouki, Elly 337  
  
 Jackendoff, Ray 197, 386  
 Jacobsen, William H., Jr. 245  
 Jakobson, Roman 64, 120, 182, 222, 233, 367, 372  
 Janney, Richard W. 2, 34, 185  
 Jay, Timothy 198, 305, 306t, 419, 431  
 Jendraschek, Gerd 152, 156, 218  
 Jespersen, Otto 44, 184, 192, 371  
 Jiménez, L. 360  
 Johnstone, Barbara 339–40  
 Joseph, Brian D. 32, 380, 411  
 Jucker, Andreas H. 31, 45, 130, 283, 291  
 Jung-Beeman, Mark 353  
  
 Kahneman, Daniel 356–7, 358t, 359, 361, 362, 363, 392, 396  
 Kaltenböck, Gunther 17, 19, 22, 24, 29, 43, 51, 53, 71, 97, 98, 100 n.1, 101, 115, 127, 128, 129, 132, 133, 134, 137, 139, 140, 143 n.1, 194, 197, 204, 210, 227, 233, 268, 269, 275, 281, 282, 288, 295, 300, 308, 332, 336, 345, 346, 347, 348, 352, 353, 354, 355, 360, 363, 382, 383, 391, 400 n.1, 414  
 Kaplan, David 16, 56, 61, 63, 65–6, 67, 68, 79, 82, 88 n.1, 205, 312, 333–4, 335–6, 403, 404  
 Karani, Michael 3, 38, 163, 169, 248, 249, 250, 251, 254, 255, 256–7, 259, 261, 263  
 Karcevski, Serge 368  
 Katsnelson, Dora 343 n.1  
 Kavalova, Yordanka 346, 347  
 Kilian-Hatz, Christa 2, 6, 28, 30, 146, 147, 152, 153, 172, 174  
 Kim, Stephanie Hyeri 132  
 Kingston, Andrew 352  
 Kintsch, W. 340  
 Kirwan, B. E. R. 242  
 Kita, Sotaro 16, 29, 30, 76, 77, 79, 80, 146, 150, 151, 152, 155, 158, 159, 169 n.1, 176, 318, 396  
 Kleinknecht, Friederike 245, 284, 285–6  
 Knowles, John 29, 30

- Kockelman, Paul 24–5, 33, 55, 64, 83, 186, 187, 189, 289, 293, 370
- König, Christa 355, 382
- König, Ekkehard 117*t*, 118, 124, 202, 203, 212, 236, 239
- Koo, Hyun Jung 146, 147, 170, 180
- Koroma, Regine 163–4
- Kowal, Sabine 6, 24, 133, 134, 135, 183, 186, 193, 306*t*
- Kruglanski, A. 359
- Kryk, Barbara 44, 376, 378
- Kunene, Daniel 68, 147, 149
- Kurylowicz, Jerzy 232
- Kuteva, Tania 17, 19, 22, 24, 43, 46, 51, 53, 71, 97, 98, 101, 115, 127, 128, 129, 132, 133, 134, 137, 139, 140, 194, 197, 204, 210, 211, 217, 219, 227, 233, 246, 268, 269, 275, 278*t*, 281, 282, 284, 285, 288, 290 n.1, 293, 295, 298, 300, 308, 309, 332, 336, 337, 345, 346, 347, 348, 352, 353, 354, 355, 357 n.1, 360, 363, 382, 383, 386, 400 n.1, 414
- Lakoff, George 8
- Lambrecht, Knud 27, 234
- Langacker, Ronald W. 78
- Lébikaza, Kézié Koyenzi 117*t*, 118, 149
- Lecours, A. R. 338
- Lee-Goldman, Russell 209, 210
- Leech, Geoffrey 8, 12, 13, 16, 18, 19, 24, 33, 45, 46, 80, 81, 85, 92, 94, 96, 109, 110, 111, 112, 114, 123, 134, 142, 143, 145, 184, 186, 187, 188, 190, 193, 194, 198, 199, 200, 201, 204, 206, 207, 209, 210, 211, 212, 213, 214, 215, 216, 217, 221, 224, 225, 226, 227, 232, 233, 234, 235, 236, 237, 238, 239, 283, 291, 304, 306*t*, 323, 324*t*, 330, 331, 332, 342, 345, 366, 370, 371, 372, 374, 401, 406
- Legère, Karsten 117*t*, 118, 124, 202, 203, 212, 236, 239
- Levinson, Stephen C. 3, 21, 23, 36, 113, 205, 209, 211, 212, 219, 234 n.1, 237, 239, 399, 408, 409, 422
- Levisen, C. 197, 368, 369
- Lewis, Diana M. 131
- Ley, R. G. 355
- Li, Charles N. 202
- Lieven, Elena 342
- Liljegren, Henrik 244, 246
- Lionnet, Florian 40, 219
- Ljung, Magnus 16, 29, 99, 121, 189, 191, 195, 198, 270–1, 287, 289, 292, 303, 304–5, 306–7, 352, 372–3
- Llopis Cardona, Ana Belén 285, 286*t*
- Long, Debra L. 340
- Long, Haiping 17, 19, 22, 24, 43, 46, 51, 53, 71, 97, 98, 101, 115, 127, 128, 129, 132, 133, 134, 137, 139, 140, 194, 197, 204, 210, 211, 217, 219, 227, 233, 246, 268, 269, 275, 278*t*, 281, 282, 284, 285, 288, 290 n.1, 293, 295, 298, 300, 308, 309, 332, 336, 337, 345, 346, 347, 348, 352, 353, 354, 355, 357 n.1, 360, 363, 382, 383, 386, 400 n.1, 414
- Louwrens, Louis J. 159, 160, 162, 169, 315, 316*t*
- Lusekelo, Amani 245
- Lydall, Jean 163
- Lyons, John 222, 376
- McArthur, Tom 304
- McCarthy, Michael 184, 352, 372
- McCready, E. 144
- McGregor, William 20, 23, 117*t*
- Mackenzie, J. Lachlan 8, 11, 12, 15, 19–20, 112, 116, 142, 181, 220, 232, 239, 240, 241, 242, 246, 288, 289, 293, 305, 307, 337, 367, 370, 409
- McKoon, G. 340
- MacWhinney, Brian 391
- Malchukov, Andrej 391
- Markó, Alexandra 132
- Martin, Samuel 108
- Martinović, Martina 39 n.1, 40
- Maschler, Yael 125, 220, 338, 340
- Matisoff, James A. 67, 225
- Matras, Yaron 137
- Matthews, Peter H. 374, 375
- Matutu, Haile 22, 181, 183
- Mauranen, A. 17, 339
- Mauri, Caterina 119, 121
- Meinard, Maruszka Eve Marie 19, 80, 96, 100, 142, 148, 149, 150, 181, 249, 371, 374, 375
- Michaelis, Laura A. 51
- Milioni, Georgia 287, 289 n.1
- Milner, B. 353
- Miltersen, Ehm Hjorth 299
- Mitchell, Alice 245, 247, 271
- Mofokeng, S.M. 119, 154, 170, 181, 232, 315
- Möhligh, Wilhelm J. G. 117
- Molinu, Lucia 125*t*, 242, 244
- Moore, Erin 211
- Moravcsik, Edith 391
- Morris, H. F. 242
- Mortensen, Sune Sønderberg 407
- Moscovitch, M. 355
- Moshi, Lioba 15, 16, 28, 68, 78, 80, 146, 147, 148, 152, 154, 156, 165, 168, 170, 307, 308, 312, 313, 314, 318, 340, 360*t*, 368, 379, 397
- Müller, Friedrich Max 1, 4, 367
- Munaro, Nicola 18 n.3, 37, 62, 80, 278, 348, 350

- Mutz, Katrin 276  
Mwinlaaru, Isaac Nuokyaa-Ire 246
- Narrog, Heiko 46, 231  
Nassenstein, Nico 352  
Nespoulous, J. L. 338, 340  
Nevalainen, Terttu 231  
Newman, Paul 13, 155, 158, 244, 316*t*, 317, 318  
Noel Aziz Hanna, Patrizia 23, 27, 120, 233, 241, 284, 336, 348  
Noora, Azam 138, 276  
Nordbustad, Frøydis 316*t*  
Norde, Muriel 276, 300  
Norenzayan, Ara 356  
Norrick, Neal R. 12, 14 n.1, 37, 62, 134, 142, 143, 182, 186, 190, 193, 199, 217, 239, 287–8, 289, 325, 326*t*, 361, 365, 370, 400, 402  
Nübling, Damaris 6, 13, 23, 24*t*, 30, 182, 187, 192, 194, 198, 270, 275, 307, 366, 370, 436  
Nuckolls, Janis B. 171  
Nuyts, Jan 91 n.1, 188
- Ocampo, Francisco 276  
O'Connell, Daniel C. 6, 24, 133, 134, 135, 183, 186, 193, 306*t*  
Ogi, Naomi 8  
Onodera, Noriko Okada 132  
Osman, Magda 359, 360  
Over, David E. 357, 359, 394  
Öztek, Piyale Cömert 13, 220
- Padilla Cruz, Manuel 79  
Parker, Steve 119, 218  
Paton, W. F. 117*t*, 218  
Pawley, Andrew 29, 30, 221, 341–2, 343, 360, 363, 379, 394–5, 405  
Peirce, Charles Sanders 69, 70, 71, 77  
Peña, Jaime Germán 33, 327–8  
Pennington, Ryan 117, 243  
Phiri, Admire 1, 16, 24, 30, 108, 113–14, 117*t*, 149, 160, 179, 190, 255, 420  
Pinker, Steven 303, 304, 305, 306*t*, 352, 405  
Pitkin, Harvey 118*t*  
Poletto, Cecilia 348  
Pons Bordería, Salvador 285, 286*t*  
Poplack, Shana 139  
Postman, Whitney A. 343 n.1  
Potts, Christopher 64–5, 67, 68, 81, 82, 404  
Poulos, George 39, 159, 160, 162, 169, 315, 316*t*, 389  
Prat, Chantel S. 340  
Prévost, Sophie 277  
Pullum, Geoffrey K. 5, 6, 12, 25, 184, 368
- Quirk, Randolph 18, 85, 184, 186, 188, 193, 221, 234, 370
- Rasmussen, T. 353  
Ratcliff, R. 340  
Raumolin-Brunberg, Helena 231  
Raymond, Geoffrey 263  
Reber, A. S. 359  
Reinhart, Tanya 346  
Rennison, John R. 159–60  
Rett, Jessica 51  
Reyna, Valerie F. 340, 356, 393, 394  
Rhee, Seongha 115, 146, 147, 166, 167, 170, 180, 219  
Robins, R. H. 243  
Rosch, Elenor 8  
Ross, John 348  
Rosten, Leo 67  
Rousseau, Jean-Jacques 385, 418
- Sacks, Harvey 95, 212, 223  
Sadock, Jerry 348  
Sakai, Kuniyoshi L. 353  
Salmons, Joseph 222  
Samarin, William J. 149  
San Roque, Lila 47  
Sansò, Andrea 119, 121  
Sapir, Edward 4, 365, 366, 374  
Schaden, Gerhard 234, 235, 237, 238  
Schegloff, Emanuel A. 95, 212, 223, 237, 405  
Schiffrin, Deborah 48, 96, 97, 128, 129, 325  
Schleicher, August 232  
Schmitt, Norbert 379, 395  
Schneider, Stefan 346, 360  
Schourup, Lawrence 31, 127, 186  
Schrock, Terrill 126–7  
Schwentner, Ernst 192  
Searle, John R. 192  
Shah, Sheena 39, 40, 41, 122  
Sherratt, Sue 355  
Sidtis, Diana, see Van Lancker Sidtis, Diana  
Siemund, Peter 51  
Simon-Vandenberg, Anne-Marie 130  
Singer, M. 340  
Sloman, Steven A. 262, 357  
Sohn, Ho-Min 5, 21, 22, 34, 117*t*, 118*t*, 148, 166, 190, 213*t*, 327  
Sohn, Sung-Ock 132, 245  
Sonnenhauser, Barbara 23, 27, 120, 233, 284, 336  
Souza, Miguel 285, 286*t*  
Sperber, Dan 337  
Stange, Ulrike 2 n.1, 6, 62 n.2, 66, 122, 185, 186, 191, 192, 198, 209, 298, 306*t*, 319, 326–7, 366



- Stanovich, Keith E. 357, 358, 359, 361  
 Stappers, Leo 212  
 Stenström, Anna-Brita 248, 279, 284  
 Stephens, Richard 352  
 Stivers, Tanya 209, 212, 219, 408, 409  
 Stokoe, William 34  
 Stolz, Thomas 137–8  
 Storch, Anne 352  
 Swaab, Tamara Y. 355  
 Syder, F. 379
- Taavitsainen, Irma 189  
 Tannen, Deborah 220  
 Taylor, John R. 8  
 Thompson, Lawrence C. 20, 23, 117*t*, 243, 328, 329  
 Thompson, M. Terry 20, 23, 117*t*, 243, 328, 329  
 Thompson, Sandra A. 202, 215  
 Torreira, Francisco 31, 44, 72, 116, 119, 207, 208, 371  
 Tosco, Mauro 21, 23, 117*t*, 118, 171, 212  
 Tottie, Gunnel 18, 133, 134, 135, 203, 204  
 Trabasso, T. 340  
 Traill, Anthony 40, 41, 122  
 Trask, R. L. 374, 375  
 Traugott, Elizabeth Closs 10, 17, 46, 91 *n*.1, 128, 129, 188, 277, 295, 296 *n*.1, 313 *n*.1  
 Trousdale, Graeme 295, 296  
 Tucker, Archibald N. 276  
 Tversky, A. 359
- van den Berg, René 34, 117*t*, 118*t*, 240, 251, 255, 260, 271  
 van Dijk, Teun A. 337, 340  
 Van Lancker Sidtis, Diana 343, 344, 353, 354, 355, 356*t*, 360*t*, 382  
 Velupillai, V. 13  
 Vendler, Zeno 148, 151, 273  
 Veselinova, Ljuba N. 115, 125  
 Vincent, Nigel 300  
 Vindenes, Urd 300
- Virbel, J. 338  
 Voeltz, F. K. Erhard 6, 146, 147, 419  
 Vogel, Marianne 198, 270  
 Voorhoeve, Clemens Lambertus 272
- Waltereit, Richard 276, 282–3, 378  
 Wamitila, Kyallo Wadi 150, 151, 160, 176, 273  
 Warren, Beatrice 29  
 Watanabe, Honoré 133  
 Weinreich, Uriel 9  
 Westermann, Diedrich 166*t*, 168  
 Wharton, Tim 4, 6, 18, 33 *n*.1, 43, 44, 56 *n*.1, 57, 60, 61, 63, 66, 78, 79–80, 105–6, 185, 186, 298, 335–6, 365, 366, 367, 370, 373  
 Wichmann, Anne 227, 346, 377  
 Wierzbicka, Anna 6, 17, 35, 57, 61, 67, 90 *n*.1, 122, 185, 187, 188, 191, 365, 394  
 Wilcox, Sherman 34  
 Wilkins, David P. 6, 13, 17, 25, 26–7, 33, 34, 35, 39, 43, 44, 46, 57, 60–1, 62, 92–3, 121, 123, 124, 125, 183, 184, 186, 187, 196, 209, 217, 220, 260, 263, 281, 298, 299, 306*t*, 326*t*, 327, 365, 366, 370  
 Wilson, Deirdre 275, 337  
 Wiltshko, Martina 8, 47, 60 *n*.1, 92, 106, 200, 201, 204, 209, 210, 211, 214, 215, 278, 302, 338, 348, 349–51, 364, 369, 406, 416  
 Wittgenstein, Ludwig 8, 82  
 Woods, Rebecca 227–9  
 Wright, Melissa 39, 136, 137  
 Wundt, Wilhelm 192
- Yliniemi, Juha 22, 34, 135  
 Yngve, Victor 214, 401  
 Yoon, Kyung-Eun 133, 134, 135, 209, 212, 219, 366, 408, 409
- Zhang, M. 17, 339  
 Zwicky, Arnold M. 12, 14, 17, 19, 22, 23, 98, 111, 112, 114, 234, 237, 239, 323–4, 368, 370, 374, 375

# Index of Subjects

- adjacency pair 212 n.1, 406
- adjunct 18, 20, 70, 84–5, 86, 90, 98 n.1, 100–1, 143, 214, 250, 275, 296, 406
- adverb 20–1, 22, 25, 28, 34, 45, 77, 83, 121, 129, 135, 137, 139, 147, 149, 151, 154–5, 157*t*, 159 n.1, 165, 171, 172, 174, 177, 214, 227, 269, 296, 308, 310, 312, 314, 315*t*, 317, 318, 320, 327, 339, 373, 375, 397
- Aktionsart 148, 151
- alignment 48, 141, 183, 210, 389
- analytic mode 343, 356, 386
- analytic processing 360*t*
- andative 251–2
- animal call, conative 3, 34, 248, 249, 250, 254, 256, 259, 261 n.1
- animal directive 3 n.1–3, 34, 40, 41–2, 119, 169, 194, 248–61, 263–4, 298, 299, 389
- aphasia 353, 354
- aphasic 354–5, 356*t*, 368
- appropriateness condition 17
- argument configuration 93, 181, 219 n.1, 262
- argument structure 55–107, 110, 112–13, 121, 122, 130–1, 140, 143, 144, 149–50, 151, 173–4, 188–9, 196, 201, 207, 213, 224–5, 229, 235, 250, 262, 280–2, 290–3, 295*t*, 320, 417
- article, definite 242, 267, 387
- article, indefinite 387
- attention signal 2, 13, 19, 32*t*, 35, 51, 72, 82 n.1, 88*t*, 89*t*, 90, 98, 105, 111, 112, 113–19, 123, 208, 235, 280–1, 283, 285–6, 290, 294, 320, 324, 326, 328, 329, 361, 378*t*, 399, 415
- backchannel 40, 48, 209, 211, 215, 217, 349–50, 401–2
- body movement 34, 46, 73, 107, 161, 212, 263, 397, 407
- borrowing 139, 256, 257, 380, 383, 439
- brain lateralization 352–3
- camouflaging 303–5, 306*t*, 307, 320
- click 32–4, 39–42, 119, 136–7, 160–1, 164, 165 n.1, 177, 179, 218–19, 251–2, 253*t*, 254 n.1, 255–6, 258
- click, alveolar 33 n.1, 34, 39, 40, 136, 160, 179, 251, 255–6, 258–9, 262, 298, 329, 370, 374, 386, 389
- click, bilabial 39, 40, 41, 136, 212, 255
- click, dental 32, 33, 34, 39, 40, 136, 160, 179, 219, 255, 298, 370, 374
- click, lateral 39, 40, 136, 219, 253, 329
- click, naked 39, 40, 41, 218, 270, 389
- click-only interactive 39, 40, 41, 137, 389
- cognitive system 356, 357
- cohesion 43, 45, 46
- common ground 25, 349, 364
- competing motivations 391
- conative animal call, *see* animal call, conative
- conceptual displacement, *see* displacement
- confirmational 92, 200–1, 348, 350, 406
- conjunction 38, 218, 284, 325, 373, 374, 375
- construction grammar 295, 347
- context extension 247–9, 282, 287, 294, 297, 300, 310, 311
- continuer 209, 402
- cooperative principle 408
- cooptation 192, 257, 266–77, 278, 279, 284, 285, 294, 299, 300, 305, 307, 314, 320, 346, 415
- decategorialization 121, 226, 278, 279*t*, 297, 309, 310, 311, 318, 319
- decision making 358, 363, 394, 412
- definition, prototypical 7–8, 168
- definition, working 110, 111, 120, 128, 142, 147, 183, 200, 211, 222, 233, 249
- deictic 7, 9, 25–8, 35–6, 51, 62, 87, 93, 104, 107, 115, 126, 165, 249, 252*t*, 269, 334, 384, 397
- deictic motion 115, 249, 281
- deixis 9, 27, 43, 45, 66, 67, 104, 313
- depiction 16, 34, 55–6, 68–71, 72, 73–7, 79, 82–3, 93, 95, 107, 109*t*, 148, 171–6, 178, 179, 188, 267, 308, 336, 363, 379, 388, 393, 412–14
- depiction, expressive 83, 93, 267, 393, 413
- depiction, interactive 71, 74–6
- derivation 12, 13, 14, 20, 168, 183, 268, 300, 311, 314, 315*t*, 334 n.1, 337
- descriptive meaning 16, 56, 57, 61, 62, 63, 65, 66*t*, 67, 75, 79, 81, 85, 93, 105, 162, 171, 249, 374 n.1, 413
- desemanticization 226, 278*t*, 279, 281, 283, 293, 297, 309, 310, 311, 313, 319
- determiner 26
- diasystem 376, 383

- dimension, affecto-imagistic 16, 70, 396
- dimension, analytic 15, 16, 78, 80, 308, 312, 340, 360*t*, 368, 379
- dimension, expressive 15, 78, 80, 308, 340, 360*t*, 368, 379
- direct-access hypothesis 77, 83, 107, 109, 270, 336, 351, 382, 393, 396
- directive 2, 3, 9, 10*t*, 19*t*, 20, 27, 32*t*, 33, 34, 40, 41, 42, 47, 49*t*, 51, 73–4, 76, 88*t*, 89*t*, 90, 91, 92, 98, 108, 109*t*, 110*t*, 111, 114–16, 119–27, 139–40, 164, 169, 176, 185, 187, 194, 196, 197, 211, 223, 233, 239, 243, 248–64, 280–3, 291, 293–4, 298–9, 320, 323*t*, 324*t*, 326*t*, 327, 328–9, 331, 338, 347*t*, 356*t*, 374, 375, 378*t*, 385–6, 387*t*, 389, 399, 415, 416, 418
- discourse act 11, 12, 181
- discourse gloss 339, 400 n.1
- discourse grammar 8, 14, 43, 53, 71*t*, 110, 181, 220, 332–3, 348, 349, 360*t*, 363, 369, 383, 391, 398, 414
- discourse-level grammar 17, 103, 106, 268–9, 276, 277, 294, 314, 337–40, 344, 348, 349, 351, 355, 369, 414
- discourse marker 10, 11, 19, 25–6, 39, 48, 88*t*, 89*t*, 96–7, 98, 105, 109, 115, 129–32, 136–8, 140, 162, 163, 190, 196, 205, 227, 269, 280*t*, 281–6, 289–91, 294, 308, 320, 324, 329, 332, 361, 376–8, 382–3, 400–1, 407, 415
- discourse marker, constituent anchored 97, 131, 132
- discourse model 340
- discourse organization 45, 156, 206, 299, 348
- discourse particle 127, 195, 378
- discourse structuring marker 129
- discreteness 44–5
- discursization 276
- dispersal call 251, 258, 259, 389
- displacement 28, 31, 38, 386, 388
- dual-process framework 336–60, 362, 368, 369
- edge-feature 37
- emotive stratum 182, 367
- engagement 47
- epithet 50*t*
- erosion 227, 278*t*, 279, 286, 297, 304, 309, 311, 318, 319
- etymologically opaque 135, 145, 219, 256, 267, 386, 387, 418
- evaluation 48, 91 n.1, 141, 183, 389
- evaluative 10*t*, 19*t*, 48, 49*t*, 51, 88*t*, 89*t*, 109*t*, 110*t*, 141–5, 176, 181, 264*t*, 326, 329, 331, 334, 356*t*, 378*t*, 389
- exclamative 51, 118 n.7, 141, 144, 372
- experienter 87, 91, 122, 188, 296, 297, 417
- expletive 11, 12, 13, 20, 23, 50*t*, 110*t*, 190, 193, 194, 195, 197–9, 271, 279, 292, 296, 303, 304, 306, 321, 328, 329, 331, 332*t*, 343, 352, 370, 373
- expletive, moderated 50*t*, 193, 194, 198
- explicit system 58, 357, 358, 359, 360
- expressive meaning 55, 58, 63, 64, 65, 66*t*, 67, 68, 79, 80, 81, 82, 85, 162, 180, 207, 249, 273*t*, 336, 376, 388, 415, 416
- expressive reinforcement 278, 288, 300–3
- expressiveness 16, 64, 301, 308
- falsetto voice 33
- filler 26, 49*t*, 50 n.1, 127, 128, 133, 134, 135, 136, 138, 195 n.1, 197, 284, 289, 343, 355, 366, 371
- foregrounding 16, 64, 106, 154 n.1, 172, 301, 309 n.1, 310, 311, 319
- formulaic speech 343–4, 360, 379
- fossil 297, 302, 385–9
- framing structure 361, 401
- free form 310, 373
- functional discourse grammar 8, 14, 181, 220
- fuzzy-trace theory 340, 341, 356, 393, 394
- Gestalt psychology 340–1
- gesture 16, 27, 34–6, 43, 46, 56 n.1, 70–1, 73–4, 77, 79, 83, 90, 161–2, 184, 188, 210, 218, 223, 251, 336, 370, 372, 380, 386, 388, 392*t*, 397, 407, 416–17
- gesture, linguistic 16, 35, 370
- gesture, physical 35
- gesture, vocal 35, 184, 188, 370
- gist representation 394
- grammaticalization 43, 46, 48, 128, 139, 140, 147, 154, 155, 157, 164, 165, 168, 170, 174–5, 177, 185, 196, 204, 226–9, 232, 245, 259, 266–7, 276–99, 300, 302, 307–20, 332, 350–1, 413, 415
- greeting 3, 13, 36, 50, 61, 110, 210, 212, 221, 222, 223, 224–6, 230, 249, 275, 293, 302, 323, 328*t*, 331, 332*t*, 342, 371, 383, 398, 399, 406–7, 416
- greeting formula 13, 36, 226, 275
- H argument 90–1, 101, 102, 123, 124, 131, 132, 143, 187, 188, 250, 281, 291, 292
- hesitation marker 130, 131, 133, 328, 355
- holistic mode 343, 386
- holistic processing 341, 360*t*, 405
- holophrase 12, 20, 157, 168, 309, 367, 384–5, 394, 409
- holophrastic stage 384–5
- hypoalgesic effect 352
- hypocorism 242*t*, 243*t*

- icon 69, 70–1
- ideophone 2, 5, 6, 8, 10, 13–16, 19–22, 28, 29, 30, 32–5, 36, 41, 43 n.1, 44–5, 46, 47, 49*t*, 57–8, 62 n.1, 63–4, 66, 67, 68, 69, 71–8, 80, 88*t*, 89*t*, 91, 95–6, 107, 108, 109, 110*t*, 140, 145–81, 184, 194*t*, 199, 201, 229, 263–4, 271–4, 297, 299–303, 307–21, 326, 328, 329*t*, 332, 334, 338, 340, 345, 350, 355, 362, 368, 371, 372, 374–5, 378 n.1, 379, 380, 382, 385–7, 395 n.1, 396–7, 403–5, 412, 413, 415
- illocution 240
- illocutionary force 16, 64 n.1, 181, 344, 348
- imagery, *see* sensory imagery
- imperative 49*t*, 103, 119, 120–2, 123, 124–7, 140, 181, 226, 228, 229 n.2, 232, 234, 243–4, 247*t*, 257, 281, 282–3, 294, 325, 328, 347*t*, 359, 378
- imperative, solidarity 281, 285
- implicit system 357, 378*t*
- imprecation 120, 285
- index, temporal 25
- indexical 9, 24–7, 38, 74, 77, 81, 83, 89, 93, 94, 104, 134, 172–4, 178, 187, 205, 250, 259, 269, 313, 334, 336, 362, 386, 388, 413
- indexicality 9 n.1, 38, 173, 174, 313
- inflection 12–14, 20, 168, 169, 183, 240–1, 300, 311, 314, 315*t*
- insert 8, 12, 19, 24, 80, 109–10, 142, 178, 184, 190, 202, 211, 232, 236, 272, 273*t*, 303, 330–3, 371, 372, 383 n.1, 403–5
- Interactional Spine Hypothesis 348, 350, 351
- interactive 1–2, 7–14, 19, 20, 23, 27–8, 35–48, 52, 55–8, 63, 64 n.1, 66–7, 71, 74–9, 81, 85–94, 100–13, 115–16, 120, 123–4, 128–30, 137, 142, 148, 161, 165 n.2, 174, 176, 177, 181, 185, 192, 200, 202, 204, 207, 210–11, 215, 222, 224, 228, 233, 240, 247–300, 302, 306, 308, 312, 314, 320, 322, 327–36, 338–52, 355–7, 360–4, 368, 369, 376–418
- interactive, primary 37, 165 n.2, 192, 208, 219, 294 n.1, 387, 418
- interactive, secondary 165 n.2, 257, 266–7, 268 n.1, 387
- interjection 1–4, 9, 11–14, 17, 19*t*, 22–7, 30, 33, 37, 39, 45, 48, 51, 53, 56–63, 66–7, 74–8, 86, 88*t*, 89, 90 n.1, 95, 99–102, 105, 108–9, 113, 118, 122, 124, 125, 130, 140, 143, 164, 173, 174, 181–99, 207, 208, 236, 245, 260, 271, 275, 276, 279, 280–301, 320, 321, 326, 327, 335, 341, 361, 367, 371, 376–8, 382, 394, 397, 398, 399–402, 408, 409, 412, 415
- interjection, cognitive 51, 80, 100, 108, 109*t*, 122, 185, 187 n.1, 327
- interjection, conative 51, 115, 118, 122, 187, 196, 248, 327
- interjection, emotive 2, 3, 51, 64, 66, 86, 95, 99, 100, 108, 109*t*, 122, 124, 164, 173, 176, 178, 182, 184, 185, 186, 187, 191, 198, 270, 292, 294, 327, 370
- interjection, expressive 5, 15, 16, 55, 56, 58, 63–8, 78, 85, 93, 106, 174, 181, 184, 187, 196, 197, 207, 220, 222, 270, 272, 275–6, 301, 303, 319, 333–5, 338, 368, 371, 376, 379, 403, 413, 415
- interjection, laughter 197, 287
- interjection, phatic 196, 222, 223, 327
- interjection, pragmatic 195, 196, 289
- interjection, primary 37, 44, 49*t*, 56, 111, 186, 189, 190, 191, 192, 193, 194, 223, 289, 290, 297, 299, 325, 327, 331, 373–4, 386, 402
- interjection, secondary 37, 49*t*, 123, 188, 189, 190, 191–9, 217, 221, 270, 287, 288, 297, 327, 331, 332*t*, 370, 386, 402
- interjection, volitive 90 n.1, 122, 185
- interjective 11, 119, 133, 170, 181, 232, 296, 328, 329*t*, 338
- invective 14 n.1, 287
- kinaesthesia 148
- kinship term 198, 238*t*, 243–7, 292
- language acquisition 364, 384–5, 390, 417
- language contact 52, 124, 137–9, 220, 383–4, 390, 417
- language evolution 38, 39, 385–90
- language periphery, *see* periphery of language
- linguaging 338
- leave-taking 406
- left hemisphere damage 344, 354
- lexicalization 259, 313
- liminal sign 36
- linearization 76, 268, 344, 346
- loanword 244, 257*t*
- macrostructure 337, 339, 352
- marginalia 364, 366, 380, 381–3, 390, 411
- maxim of quantity 408
- meaning, conceptual 62, 65, 80, 268, 271
- meaning, descriptive 16, 56, 57, 61, 62, 63, 65, 66*t*, 67, 75, 79, 81, 85, 93, 105, 162, 171, 249, 374 n.1, 413
- meaning, procedural 79, 80, 106, 337
- metadiscourse 339
- metadiscursive 17
- metalanguaging 338
- metatextual 17, 109*t*, 128–9, 168, 345–6
- microstructure 337, 352
- mimetic 16, 29, 62 n.1, 74, 76, 77, 146, 151, 155, 158–9, 318, 396

- mimicry type 248  
 modalizing discourse 338  
 mode, descriptive 77  
 mode, mimetic 16, 146  
 modifier construction 152, 153*t*, 154, 155–7,  
 171, 274, 309, 310, 314  
  
 naked click, *see* click  
 nasalization 34, 137, 254  
 natural semantic metalanguage 17, 57, 62,  
 106, 365  
 negation 4, 28–31, 40, 46, 173, 174–5, 202, 218,  
 219, 280, 286–8, 294*f*, 298, 311, 321, 367,  
 371, 397  
 neurolinguistic 52, 73, 322, 342, 345, 352–6,  
 360*t*, 412  
 non-clausal unit 45, 330, 331, 333  
 novel speech 343, 344, 354, 360*t*  
 nursery term 186, 260, 261, 299  
  
 onomatopoeia 19, 41, 44, 64, 96, 148–50,  
 160, 380  
 other-initiated repair 31, 44, 116, 206–8, 216,  
 258, 371, 389  
  
 paralinguistic 250, 266, 416  
 paralinguistic communication 250, 366, 416  
 paraphrase 27, 56–83, 85, 88, 94, 97–100, 102,  
 105, 106, 112, 114, 121, 122, 130–2, 143,  
 145, 149–50, 162, 188, 189, 201, 207, 213,  
 224–5, 229, 233, 235, 250, 298, 335, 337, 341,  
 393–4, 404  
 parenthetical 22, 29, 154*t*, 309, 310, 336, 346,  
 347 n.1, 348  
 particle 51, 127, 136, 161, 195, 202, 235, 241,  
 277, 323, 336, 339, 344, 370–1, 373, 375, 378,  
 380, 408  
 performance, depictive 16  
 performative hypothesis 348  
 periphery, left 18, 37, 105, 113, 132, 143, 152,  
 190, 202, 213, 237, 362, 378, 395, 405  
 periphery, right 152, 201, 202, 213, 269, 378, 405  
 periphery of language 42, 366, 380, 411, 412  
 phenomime 148  
 phonotactic constraint 32, 33, 263  
 phoric marker 11, 339, 356*t*  
 politeness formula 32, 42, 63 n.2, 221, 383  
 positioning 48, 141, 183, 389  
 pragmatic marker 129, 227, 325  
 pragmatic particle 195  
 pragmatization 276, 277  
 preposition 121, 373, 374, 375  
 problem solving 358  
 proceduralization 276  
  
 propositional discourse 7, 15, 16, 47, 51, 56, 60,  
 80, 82, 87, 106, 209, 245, 295, 338–40, 342–5,  
 348–52, 355, 369, 391, 404, 405  
 prototype 8, 9, 48, 53, 148, 259, 261 n.1, 318  
 psychomime 148  
  
 quasi-word 6, 366, 370, 373  
 question tag 26, 202, 407  
 quotative construction 152–7, 178 n.1, 309, 311,  
 314, 317, 318  
  
 reduplication 152, 169, 262, 263, 301, 302  
 referential discourse 338  
 repair marker 44, 116, 206, 207, 208, 216, 258,  
 371, 389  
 response cry 186, 291, 293, 375  
 response elicitor 19, 26, 46, 59, 72, 88–92, 99,  
 105, 111, 200–6, 210, 328, 351, 378, 403, 407  
 response signal 2, 11, 19, 23, 26, 30, 38, 40, 73–4,  
 76, 88–90, 92, 102, 105, 115, 190, 197, 210–18,  
 225, 287, 288, 323*t*, 328*t*, 361, 378*t*, 401–2,  
 406–8, 416  
 right hemisphere damage 353–5  
 ritual 3, 6, 210, 221, 224, 225, 249, 279, 326*t*,  
 366, 370, 398, 399, 406  
 ritualized act 366, 370  
  
 S argument 20, 89, 100, 122, 149, 173, 176,  
 188, 417  
 scope, semantic-pragmatic 97, 131, 269  
 semantic decomposition 27, 62, 93, 365  
 semantics of Meanings 82 n.2, 334  
 semantics of Use 65, 66*t*, 82 n.2, 344  
 semi-word 6, 16, 44, 176, 366, 370, 372, 373, 375  
 sensory imagery 36, 75, 78, 109*t*, 148, 161, 165,  
 176, 179, 181  
 sentence grammar 6, 7, 10, 12, 15, 17, 21, 23, 26,  
 28, 29, 31, 32, 36–8, 40–7, 55, 56, 61, 64, 65 n.2,  
 67–8, 76–9, 82–7, 92, 94, 101–9, 120, 129, 135,  
 141, 145, 155, 159–60, 164–5, 167, 169–79,  
 192–3, 201, 221, 227–8, 231, 233, 237, 240–1,  
 244–5, 254–60, 263, 266–72, 275–9, 283,  
 287–9, 292, 294–300, 308, 311, 313–21, 325,  
 330, 333*f*, 334–41, 343, 345–6, 348, 349–57,  
 360–3, 368–9, 375, 379, 381–8, 391–7, 400–17  
 sentence-level grammar 339  
 sequence of activation 395  
 sex gender marking 240, 241  
 sibilant 33, 258, 259  
 situation of discourse 9, 12, 16–17, 25, 26, 27,  
 28, 38, 45, 64, 65, 70, 73, 76, 77, 81, 83, 86, 87,  
 89, 90, 92, 93, 94, 101, 104, 107, 116, 128 n.1,  
 131, 144, 173–4, 188, 189, 195, 204, 222, 225,  
 229, 230, 245, 247, 268, 269, 271, 277, 320,

- 333, 336, 339, 345–6, 351, 362, 379, 382, 384, 388, 392*t*, 393, 396, 407, 414
- social formula 19*t*, 26, 36, 48, 73, 82 n.2, 88*t*, 89*t*, 90, 98 n.1, 101, 109, 164, 220–30, 235, 275, 279, 293, 302, 324, 325, 328–9, 333, 342, 377, 378*t*, 398, 399, 400, 401, 407, 415, 416
- social psychology 52, 322, 357, 359, 360, 361, 382, 392, 395, 412, 414
- soliloquy 91, 99, 102, 124, 131, 143, 187, 189, 213, 292
- sound symbolism 5, 166, 167, 168
- speech act 63, 145, 221, 341, 342, 348
- speech act formula 342
- speech event 38, 124, 173, 174
- stancetaking 48, 141, 183, 209, 386, 389, 411
- stand-alone word 111, 228, 229, 235, 237, 255, 280, 291, 330, 331, 371
- sub-cortical circuitry 368
- suppletive 125, 126*t*
- swear word 11, 13, 29, 185, 198, 199, 270, 271, 301, 304, 323, 371, 380
- symbol 25, 44, 69*t*, 70–1, 75, 76, 77, 228, 255 n.2
- synchronization of minds 47, 349
- syntactic detachment 185, 299
- system 6, 38, 39, 47, 52, 58, 78, 133, 158, 232, 240, 315, 323, 347, 356–62, 364, 365, 367, 368, 369, 376, 377, 379, 386, 387, 388, 389, 390, 391, 394, 395, 396, 397, 411, 414
- T argument 92–4, 96–100, 102, 111–15, 123, 130–2, 139, 143–4, 149–50, 153–6, 164, 171, 173, 176, 180–1, 189, 195, 201–2, 206, 213–15, 224, 225, 235, 237, 241, 273, 280, 282, 287, 290, 295*t*, 350
- taboo word 198, 199, 327, 372, 380
- tag 3, 202, 203, 407
- theme argument 91
- thetical, constructional 346, 347
- thetical, formulaic 127, 346, 347, 354, 355, 360
- thetical, instantaneous 346, 347
- thetical grammar 51, 332, 333*f*, 345–8, 352–4, 360*t*
- truncation 240, 242, 272
- truth condition 65, 66*t*, 334, 393, 416
- truth-conditional 7, 16, 333, 335
- turn taking 281, 283, 381
- usage-based theory 53
- use-conditional 16, 334, 335
- valence 55, 86, 90, 99, 100, 107, 123, 138, 143, 178, 196, 214, 220, 224, 280, 413
- venitive 251, 252*t*
- verbatim representation 340–1, 356, 392*t*, 393–4
- vocal gesture 35, 184, 188, 370
- vocative 7, 10, 14, 19*t*, 21, 23, 27, 28, 38, 47, 48, 50*t*, 51, 88*t*, 89*t*, 90, 97, 98, 101, 109, 110–18, 120, 128, 139, 140, 164, 177–8, 181–2, 211, 226, 231–48, 250, 256, 264*t*, 270, 271, 279, 280, 282, 283–6, 291–4, 301, 302, 320, 323, 325, 326, 327, 329, 331, 336, 345, 347*t*, 348, 356*t*, 377, 378*t*, 385, 401, 406
- voice-gesture model 35, 74, 161, 162, 397, 417
- whispering 22, 32, 33, 177
- whistle 250, 252, 253, 255, 256, 259, 262
- wild animal 248, 252, 262
- word 12, 15, 16, 21, 23, 29, 34, 44, 64, 65, 67, 76, 119, 131, 132, 136, 146, 147, 150, 153 n.1, 158, 159 n.1, 169 n.1, 183, 191, 192, 193, 195, 198, 207, 208, 216, 221, 226, 242, 244, 268, 271, 274, 275, 300, 301, 304, 313, 314, 315*t*, 319, 321, 329*t*, 333, 369–77, 379, 384, 386, 395, 396, 405, 415, 419
- working memory 340, 358















