

Oxford
LINGUISTICS

Pronouns

D. N. S. Bhat

Oxford Studies in Typology and Linguistic Theory

PRONOUNS

Oxford Studies in Typology and Linguistic Theory

Series Editors: Ronnie Cann, *University of Edinburgh*, William Croft, *University of New Mexico*, Martin Haspelmath, *Max Planck Institute Leipzig*, Nicholas Evans, *University of Melbourne*, Anna Siewierska, *University of Lancaster*.

Published

Classifiers: A Typology of Noun Categorization Devices

Alexandra Y. Aikhenvald

Auxiliary Verb Constructions

Gregory D.S. Anderson

Pronouns

D. N. S. Bhat

Subordination

Sonia Cristofaro

The Paradigmatic Structure of Person Marking

Michael Cysouw

Indefinite Pronouns

Martin Haspelmath

Anaphora

Yan Huang

The Emergence of Distinctive Features

Jeff Mielke

Applicative Constructions

David Peterson

Copulas

Regina Pustet

The Noun Phrase

Jan Rijkhoff

Intransitive Predication

Leon Stassen

Co-Compounds and Natural Coordination

Bernhard Wälchli

Published in Association with the Series

The World Atlas of Language Structures

edited by Martin Haspelmath, Matthew Dryer, Bernard Comrie, and David Gil

In Preparation

Imperatives and Commands

Alexandra Y. Aikhenvald

Reciprocals

Nicholas Evans

Reference in Discourse

Andrej A. Kibrik

Double Object Constructions

Maria Polinsky

Predicative Possession

Leon Stassen

Sign Languages

Ulrike Zeshan

PRONOUNS

D. N. S. BHAT

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

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 in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai

Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata

Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi

São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States

by Oxford University Press Inc., New York

© D. N. S. Bhat 2004

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2004

First published in paperback 2007

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, 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 book in any other binding or cover
and you must impose the same condition on any acquirer

A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

(Data available)

Typeset by SPI Publisher Services Ltd, Pondicherry, India

Printed in Great Britain

on acid-free paper by

Biddles Ltd., King's Lynn, Norfolk

ISBN 978-0-19-926912-9 (Hbk.)

978-0-19-923024-2 (Pbk.)

1 3 5 7 9 10 8 6 4 2

Contents

<i>Abbreviations</i>	x
<i>Preface</i>	xi
1 Introduction	1
1.1 'Pronoun' as a category	1
1.1.1 Defining pronouns	2
1.1.2 Two types of 'pronouns'	4
1.2 Basis of the distinction	6
1.2.1 Differentiating characteristics	9
1.2.2 Possibility of gradation	12
1.3 Position of third person pronouns	13
1.4 'Free-pronoun' and 'bound-pronoun' languages	15
1.4.1 Two kinds of agreement	16
1.4.2 Disparity between pronominal systems	18
1.4.3 Obligatoriness	22
1.4.4 Stability	25
1.4.5 Nature of the distinction	28
1.4.6 Only a speculation	29
1.5 Languages without personal pronouns	30
1.6 Structure of the monograph	31
Part I: Personal Pronouns	35
2 Relation with the Referent	37
2.1 Introduction	37
2.2 Dissociation from the referent	38
2.2.1 Referentiality of personal pronouns	40
2.2.2 File cards for referents	42
2.2.3 Use of alternative devices	43
2.3 Notion of definiteness	47
2.3.1 Evidence for definiteness	48
2.3.2 Evidence against definiteness	48
2.3.3 Personal pronouns as determiners	50
2.4 Association with the nominal category	52
2.4.1 Inability to take modifiers and complements	53
2.4.2 Association with numerals	55
2.5 Conflicting characteristics	56
2.6 Summary	57

3	Coreference and Non-Coreference	58
3.1	Logophoric non-coreference	58
3.1.1	Third person 'logophors'	62
3.1.2	Anaphoric pronouns	66
3.1.3	Contrast between logophoric and anaphoric pronouns	67
3.2	Anaphora among personal pronouns and proforms	74
3.3	Distinctions in the predicate	78
3.3.1	Reflexive meaning	79
3.3.2	Switch-reference	84
3.3.3	Reciprocal meaning	85
3.3.4	Modal distinctions	87
3.4	Summary	89
4	Association with Grammatical Categories	91
4.1	Introduction	91
4.2	Association with the category of number	91
4.2.1	Homogeneous and heterogeneous plurals	93
4.2.2	Plurality and conjunction	95
4.2.3	Dichotomous association	99
4.2.4	Combinations of pronouns	100
4.2.5	Minimal and augmented forms	101
4.2.6	Inclusory conjunction	103
4.2.7	Hierarchy of number marking	105
4.3	Association with the category of gender	109
4.3.1	Restricted occurrences	109
4.3.2	Explanations for irregularity	110
4.3.3	Politeness distinctions	111
4.4	Association with the category of case	112
4.4.1	Sagittal case marking	113
4.4.2	Prominence of speech act participants	115
4.4.3	Direct-inverse marking	116
4.4.4	Constraint on the occurrence of bound themes	117
4.5	Summary	118
5	Conflicting Characteristics	120
5.1	Introduction	120
5.2	Prominence of speech roles	121
5.2.1	Prominence of the speaker	122
5.2.2	Prominence of the addressee	124
5.2.3	Hierarchy of nominal categories	125
5.3	Personal pronouns as Janus-like	128
5.4	Basis of differentiation	130
5.5	Summary	131

6	The Position of Third Person Pronouns	132
6.1	Introduction	132
6.2	'Two-person' and 'three-person' languages	134
6.2.1	Nature of the affinity	135
6.2.2	Distinction between third person pronouns and demonstratives	138
6.3	Gender distinctions	138
6.4	Person-oriented and distance-oriented demonstratives	140
6.4.1	Exceptions to the correlation	143
6.5	Distinctions in other areas of grammar	145
6.6	Position of language families	147
6.6.1	Two-person languages	147
6.6.2	Three-person languages	148
6.7	Functional distinctions	149
6.8	Summary	150

Part II: Proforms 151

7	The Structure of Proforms	153
7.1	Introduction	153
7.2	Paradigms of proforms	153
7.2.1	Order of constituents	157
7.2.2	Identifying the constituents	161
7.3	Complex structures	163
7.3.1	Complexity in the expression of functional distinctions	164
7.3.2	Complexity in the expression of general concepts	169
7.3.3	Effects of grammaticalization	171
7.3.4	Use of peripheral concepts as basic constituents	172
7.4	Summary	173
8	Constituent Elements of Proforms	175
8.1	Introduction	175
8.2	Functional distinctions	175
8.2.1	Distinctions among demonstratives	177
8.2.2	Distinctions among interrogative-indefinites	182
8.2.3	Distinctions among relative-anaphors	183
8.3	Neutralization of functional distinctions	186
8.3.1	Combining all the three main functions	186
8.3.2	Combining demonstratives with interrogative-indefinites	187
8.3.3	Combining demonstratives with relative-anaphors	189
8.3.4	Combining interrogative-indefinites with relative-anaphors	189

8.4	General concepts	190
8.4.1	Word-class distinctions	190
8.4.2	Distinctions among different word classes	192
8.4.3	Representation of general concepts	195
8.5	Overlap between pronominal functions and general concepts	196
8.6	Summary	198
9	Characteristics of Proforms	200
9.1	Introduction	200
9.2	Denoting a general concept	200
9.2.1	Linguistic and extra-linguistic identifications	202
9.2.2	Persistence of indefiniteness	207
9.2.3	Introducing new entities	209
9.3	Using proforms for different purposes	210
9.3.1	Indefiniteness of interrogatives	211
9.3.2	Range of occurrence	212
9.4	Associating with different categories	214
9.4.1	Notion of pragmatic reference	215
9.4.2	Notion of semantic reference	220
9.4.3	Notion of scope	222
9.5	Summary	224
10	Interrogative–Indefinite Puzzle	226
10.1	Introduction	226
10.1.1	Puzzle about the interrogative–indefinite relationship	227
10.1.2	As part of a system of proforms	229
10.1.3	Semantic differences	231
10.1.4	Solution to the puzzle	232
10.2	Request for information	234
10.2.1	Use of interrogative intonation	234
10.2.2	Use of interrogative particles	235
10.2.3	Use of interrogative mood	238
10.2.4	Derived interrogatives	238
10.3	Restricting the question to a constituent	240
10.3.1	Shifting the pronoun to the sentence-initial position	240
10.3.2	Shifting to other focus positions	242
10.3.3	Use of focus particles	242
10.3.4	Use of focus constructions	243
10.4	Combining the two notions	244
10.4.1	Use of distinct intonations	245
10.4.2	Use of distinct question particles	245
10.5	Non-marking and double-marking	246
10.6	Summary	248

11	Other Related Puzzles	250
11.1	Introduction	250
11.2	Derivation of indefinite pronouns	250
11.2.1	Differentiating languages	252
11.2.2	Non-differentiating languages	255
11.2.3	Possible counter-examples	257
11.2.4	Distinction between ‘even’ and ‘at least’	258
11.2.5	Additional distinctions	260
11.2.6	Grammaticalization of marked indefinites	261
11.3	Puzzle about indirect ‘questions’	262
11.4	Interrogative–relative puzzle	266
11.5	Summary	271
12	Concluding Remarks	272
12.1	Definitions for pronouns	272
12.2	Plurality and conjunction	274
12.3	Pragmatic and semantic identity and reference	275
12.4	Interrogation and indefiniteness	276
	Appendix: List of 225 languages used as a sample	277
	<i>References</i>	292
	<i>Index of Authors</i>	307
	<i>Index of Languages</i>	312
	<i>Index of Subjects</i>	316

Abbreviations

1	first person	Imp	imperative
2	second person	Incl	inclusive of 2
3	third person	Int	interrogative
Abl	ablative	Loc	locative
Abs	absolute	Log	logophoric
Acc	accusative	M	masculine
Acco	accompanitive	N	neuter
Ana	anaphoric	Neg	negative
Ass	associative	Nom	nominative
Aux	auxiliary	NRef	non-referential
C	common gender	Obj	object
Comp	complementizer	Perf	perfect
Cond	conditional	Pl	plural
Dat	dative	Poss	possessive
Decl	declarative	Pres	present
Def	definite	Prog	progressive
Dem	demonstrative	Rec	reciprocal
Det	determiner	Ref	referential
DS	different subject	Refl	reflexive
Emph	emphatic	Rel	relative
Erg	ergative	Sg	singular
Excl	exclusive of 2	SS	same subject
F	feminine	Subj	subject
Fut	future	Super	superlative
Gen	genitive		

Preface

I had an opportunity, in the year 2000, to spend three months as a Visiting Fellow at the Research Centre for Language Typology, La Trobe University, Melbourne, Australia, thanks to an invitation from Bob Dixon, and in the year 2001 I had another opportunity to spend six months as a Guest Scientist at the Max Plank Institute of Evolutionary Anthropology, Leipzig, Germany, thanks to an invitation from Bernard Comrie. During these two visits, I was able to gather data on pronouns from a large number of languages belonging to various parts of the world. The present monograph is the final outcome of these two visits. I am thankful to both Bob Dixon and Bernard Comrie for providing me with these opportunities.

This study is actually a continuation of a report entitled ‘Wh-words’ (unpublished) that I had prepared earlier in the year 1989 as part of a project of producing a typological study of Indian languages. In that report, I had examined the relationship between the various uses of the so-called ‘wh-words’ among Indian languages. The above-mentioned visits have allowed me to expand the scope of that study both from the point of view of the pronouns covered and also from that of the languages covered. Part of this study has appeared as an article (Bhat 2000) in *Linguistic Typology*.

I hope to have uncovered several interesting characteristics of pronouns through this cross-linguistic study. For example, I have been able to show that the notion of (in)definiteness that is associated with pronouns is quite different from the one that is associated with ordinary noun phrases. The two notions differ not only in their location (speaker vs. hearer) but also the level to which they belong. The former is semantic whereas the latter is only pragmatic. Another interesting characteristic, which is restricted to only some of the pronouns like demonstratives, interrogatives, and indefinites (called ‘proforms’), is the occurrence of a dual structure, which has been used here for establishing a better understanding of the functioning of indefinite and interrogative pronouns and the relationship between the two. Personal pronouns, on the other hand, show several unique characteristics of their own that can be derived directly from an assumption that their function is to denote speech roles rather than to identify the individuals who perform those roles.

There is, on the other hand, one deficiency that affects the first part of this study that I am, unfortunately, not in a position to remove. We need to differentiate between languages in which the primary function of personal pronouns is carried by their *bound* forms, and the ones in which it is carried out by their *free* forms. Any satisfactory characterization of personal pronouns has to take into

account this typological distinction. However, the relevance of this distinction for my study occurred to me at the very end of my stay in Leipzig, and once I returned to India there was no way for me to pursue that line of study further. I have no access to grammars of the relevant languages, especially since I have now retired from service. I therefore have to leave the job of pursuing it further to more fortunate scholars who have continuing access to excellent libraries such as that of the Max-Planck Institute in Leipzig.

I wish to acknowledge my indebtedness to Bernard Comrie who has gone through the first part of this monograph, and to Edith Moravcsik who has gone through the whole monograph, for their many useful comments. I have also been helped by the comments of the two reviewers of Oxford University Press. Writing this book has been one of the most exciting experiences of my research career. I hope that I am able to share some of this excitement with others through this book.

D.N.S.B.

Mysore

1 March 2003

1 Introduction

1.1 ‘Pronoun’ as a category

The term ‘pronoun’ is generally used for referring to several different sets of words such as personal pronouns, demonstratives, interrogatives, indefinites, relatives, correlatives, etc. Defining and delimiting these into a category of words, however, has been rather problematic. Traditionally, pronouns are defined as words that ‘stand for nouns’, but most linguists find this definition to be unsatisfactory. This is mainly because personal pronouns do not ‘stand’ for any nouns as such, whereas demonstrative or interrogative pronouns can ‘stand for’ adjectives, adverbs, or even verbs (see §1.1.1). On the other hand, attempts to establish alternative definitions for pronouns have not been satisfactory either. Grammarians are therefore forced to retain the traditional definition as the only workable definition.

I wish to argue in this chapter that the main reason for this failure to establish a satisfactory definition for pronouns is the fact that words that are generally included under the category of pronouns do not together form a single category. Most importantly, personal pronouns are quite different from the rest of the pronouns. There are several differentiating characteristics that occur between them that derive primarily from the fact that the two have entirely different functions to perform in language (see §1.2). Because of this difference, one can hardly find any characteristic that can be regarded as common to both the sets. Hence, it is also rather difficult to formulate a definition that can be applied to both of them.

Grammarians also find it rather difficult to decide as to which lexical items to include in the category of pronouns and which to exclude from it. Even while dividing pronouns into subgroups like personal pronouns, demonstratives, and interrogatives, grammarians are troubled by this problem. Questions such as (i–iii) are generally left unanswered or answered only arbitrarily.

- (i) Are third person pronouns ‘personal’ or demonstrative?
- (ii) Are we to regard pronominal adjectives and adverbs as ‘pronouns’?
- (iii) Are we to include words like *such*, *other*, and *one* under pronouns?

The absence of a satisfactory characterization of the category of pronouns and its various sub-categories, as mentioned above, is apparently the main reason for this difficulty. Another reason, as suggested by some linguists, is that the classical theory of categorization is rather unsuitable for grouping words into word classes. The use of an alternative theory such as ‘Prototype Theory’ rather than

the classical theory, could resolve this problem to a certain extent (see Rosch 1978: 37, Hopper and Thompson 1984, Lakoff 1987: 47, Bhat 1994: 13).

1.1.1 *Defining pronouns*

Applying the characteristic of ‘standing for nouns’ to pronouns is problematic in several different ways. First of all, it is not generally made clear what the notion ‘stand for’ denotes and why it should be applicable to pronouns only and not to other kinds of expression. It is possible, for example, to regard a general term like *human* as standing for several more specific terms like *man*, *woman*, *boy*, *girl*, etc. In what sense do pronouns stand for nouns and these general terms do not? On the other hand, there is clearly some sense in which the relation between pronouns and nouns is different from the one that occurs between general terms and specific ones. Another point, mentioned by J. Lyons (1968) and several others is that it is not just ‘nouns’ that pronouns stand for, but rather ‘noun phrases’.

The notion of standing for ‘nouns’ can also be troublesome because, excepting personal pronouns, all others may involve the categories of adjectives, adverbs, and even verbs. For example, the category of demonstrative pronouns includes, in most languages, not only nominal forms but also adjectival and adverbial ones. Some languages have even verbal demonstratives. This is also true of interrogatives, indefinites, and relatives. Some grammarians regard only words that belong to the nominal category as pronouns, and relegate others to the category of adjectives or adverbs, in order to maintain the definition of pronouns as words that stand for nouns. However, this has the effect of splitting the various paradigms of pronouns, such as those of demonstratives, interrogatives, indefinites, etc. into two or more sets and placing them in different word classes. It also forces the linguists to assign single forms like *this* to different word classes like nouns and adjectives (Jespersen 1924: 84). Some linguists like Quirk, Greenbaum, Leech, and Svartvik (1985: 76) use the term ‘pro-form’ instead of ‘pro-noun’ in order to account for this diversity.

On the other hand, the notion of ‘standing for’ something else is completely unsuitable for characterizing first and second person pronouns. This is evident from the fact that any other noun or pronoun that we try to use instead of these pronouns would fail to provide the crucial kind of meanings that they are meant to denote.

- (1) a. *I am reading a book.*
- b. *The speaker is reading a book.*

Notice that the sentence (1b) cannot carry out the function of (1a). The pronoun *I* occurring in (1a) indicates that the agent of the action (of reading) denoted by that sentence is the speaker of that sentence, whereas the noun phrase *the speaker* occurring in (1b) cannot do so. It can only refer to the speaker of some other sentence and not the speaker of (1b).

There is a similar kind of functional distinction occurring between other pronouns like demonstratives, interrogatives, indefinites, etc. on the one hand and nouns on the other, which also makes it rather difficult to view the former as 'standing for' the latter. For example, demonstratives denote objects that are not actually named, but are pointed out. It is true that in some languages some of the demonstratives may be used for referring back or standing for some other expression, but this can only be regarded as an *extension* of their deictic use (see J. Lyons 1977: 670). The primary function of other types of pronoun like indefinites and interrogatives is also rather different from that of being words that stand for some other expression. They may be anaphoric in some usages but in several others they would be introducing new entities into the discourse whose identity is unknown and is being questioned, or entities whose identity is being kept hidden or merely left unspecified. In these latter types of usages, it would be difficult to think of them as 'standing for' some other expression.

On the other hand, there are languages in which a distinction is made between demonstrative pronouns on the one hand and anaphoric pronouns on the other, with the latter being specialized for the function of 'standing for' some other expression. Sinhalese (Indo-Aryan), for example, has the demonstrative *eyaa* 'that person (anaphoric)' contrasting with *arəya* 'that person (remote)', with the former being restricted to anaphoric use. The language has a reflexive pronoun *taman* which can also have only anaphoric usage, but unlike *eyaa* 'that person (anaphoric)', it is locally bound (Gair 1998: 113). Such anaphoric and reflexive pronouns and the so-called relative pronouns of familiar languages are perhaps the only items in the category of pronouns that are used primarily as words that stand for some other expression.

Wales (1996: 4) suggests the possibility of regarding pronouns as words that function primarily as useful 'shorthand' referring expressions that are alternatives either to ostensibly more explicit or descriptive forms, or to forms that are more cumbersome. Wales points out that pronouns generally have a low semantic content, which derives from the fact that their primary function is to avoid repetition. They may occur in their full form as referring expressions or in their truncated form as affixes while being parts of other expressions like adjectives and verbs.

This suggestion of Wales is meant for describing personal pronouns, but as I will be pointing out later on in this chapter (§1.2), the 'low semantic content' of first and second person pronouns is meant for allowing them to function as 'shifters' rather than as shorthand referring expressions. In order to satisfactorily indicate the speech roles of 'being the speaker' and 'being the addressee' respectively, personal pronouns tend to be dissociated from the actual persons who perform those speech roles. Languages use other devices like appositive noun phrases or vocatives along with personal pronouns in order to identify these latter individuals that function as the referents of those pronouns, in some of the exceptional contexts (see §2.2.3–4).

The notion of 'shorthand expression' would be more appropriate for the remaining pronouns like demonstratives, interrogatives, and indefinites than to personal pronouns. However, only some of them have the function of 'reference'. They may be adjectives, adverbs, or even verbs as mentioned earlier. Further, 'avoiding repetition' is only one of the reasons for using these second types of pronoun (i.e. pronouns other than those of first and second person). In the case of indefinite and interrogative pronouns, for example, a speaker would be forced to use a general term because of his inability to provide more specific information about the entity under consideration. This is also true of pronouns used in exclamatory sentences. They need to be general because the speaker does not want to be more specific about the quality or quantity that is being expressed. There are also uses of pronouns that are meant for suppressing information from the addressee. Thus, there are several reasons for preferring to use general terms rather than specific ones in the formation of these pronouns. The use of different sets of pronouns is an important strategy that languages utilize in order to differentiate between these different purposes that underlie the use of such general terms (see §1.2).

1.1.2 Two types of 'pronouns'

It is evident from the foregoing that it would be impossible to formulate a definition that can take care of all pronouns. We cannot find even a single characteristic that is shared by all of them, excepting perhaps a vague notion of 'being very general'. The traditional term 'pronoun' cannot therefore be regarded as representing a lexical category. Most importantly, we need to differentiate between (i) personal pronouns (especially of first and second person) on the one hand and (ii) the remaining pronouns on the other, on the basis of distinct functions that they have to perform in language. Such a differentiation can form the basis for establishing two different definitions, of which one can be applied to personal pronouns and the other one to the rest of the pronouns. However, we may not be able to include all the various types of form that are generally regarded as 'pronouns' by different grammarians under the latter category. We may have to make it more compact on the basis of the primary function that the pronouns have to perform in language.

The need to differentiate between two different types of 'pronouns' has been recognized by several grammarians on the basis of the fact that the languages that they are describing show several morpho-syntactic differences between the two sets of words. For example, Dixon postulates 'pronouns' and 'deictics' as two distinct categories for Yidiny, an Australian language (1977: 123) and also for Boumaa Fijian, a Pacific (Austronesian) language (1988: 58). The former category includes only personal pronouns whereas the latter category includes all the remaining pronouns. The division is based upon certain morpho-syntactic differences occurring between the two. In Yidiny, for example, pronouns show a

nominative–accusative case marking, whereas deictics and other nominals show an absolutive–ergative case marking. Pronouns also differ from others in not having any locative case forms (Dixon 1977: 167, 169).

Showalter (1986) uses the terms ‘interlocutory’ and ‘substitutive’ (following Houis 1977) for a similar division of pronouns in a study of Lyle, a language belonging to the Gur branch of Niger-Congo family. Parker (1986) also makes a similar distinction in the case of Mandani, a Benu-Congo language. Both of them include only first and second person pronouns in the set of ‘interlocutory pronouns’. Third person pronouns, which also function as demonstratives, are included in the set of ‘substitutives’. The two are differentiated from one another by the fact that only interlocutory pronouns participate in conversational exchange. Further, interlocutory pronouns are invariable of class while substitutives vary according to the class of the noun that is being pronominalized. The latter also have the additional function of specification as demonstratives, interrogatives, and relators, which interlocutory pronouns do not. This distinction between interlocutory pronouns and substitutives is also used by Marchese (1986) in a description of the pronouns of Godie, a Kru language.

Pitkin (1984: 230) divides the pronouns of Wintu, an American Indian language of northern California, into two subgroups, with first and second person pronouns falling into one subgroup and the rest of the pronouns into another subgroup. The two subgroups differ in their case and aspect and also in their derivational potential. First and second person singular pronouns show no aspect distinctions and are never used syntactically as demonstratives. Fortune (1955), on the other hand, makes a distinction between ‘absolute’ and ‘qualificative’ pronouns for Shona, a Bantu language, in which the third person pronouns are included in the group of absolutes rather than that of qualificatives. This is done in spite of the fact that the third person pronouns show distinctions for 22 noun classes and are similar to qualificatives on this point. The latter are considered to include demonstratives, enumeratives, quantificatives, possessives, and relatives. They are said to differ from the former (absolutes) in that the former do nothing more than indicate a certain thing or collection of things.

The present monograph is a detailed examination of this need to make a distinction between two different types of ‘pronouns’. I propose to use the terms ‘personal pronoun’ and ‘proform’ for representing this distinction, with the former term including first and second person pronouns and the latter term including all the other types of pronouns. The position of third person pronouns in this regard, however, needs to be examined separately, as they appear to belong to the system of personal pronouns in some languages and to that of demonstratives in others (see Ch. 6). I propose to use the term ‘pronoun’ as a cover term for referring to both these types of words (personal pronouns and proforms), even though, as mentioned earlier, there may not be any basis for the establishment of such a super-category, other than the fact that it has the backing of an extended grammatical tradition. I also propose to use the term ‘pronominal’

as a cover term for representing the characteristic of being a personal pronoun in the first part of this monograph and of being a proform in the second part, instead of coining new terms like ‘proformal’.

1.2 Basis of the distinction

There is an important functional distinction between personal pronouns and proforms, which, I believe, is the basis of most of the differences that occur between them. The function of first and second person pronouns is primarily to indicate the two principal speech roles, namely that of ‘being the speaker’ and ‘being the addressee’ respectively. Since these two personal pronouns are generally associated with case markers, they are also able to connect the speech roles with case roles, such as the agent, patient, experiencer, beneficiary, etc. However, personal pronouns generally do not provide any basis for identifying the actual individuals who perform those speech roles.

- (2) a. *I gave a book to John.*
- b. *John gave me a book.*

Notice that in (2a) the first person pronoun is able to indicate that the speaker is the agent in the action of giving, whereas in (2b), it is able to indicate, instead, that the speaker is the recipient of that action. The position of the pronoun in these sentences (pre-verbal vs. post-verbal) and the two distinct case forms that it assumes (*I* vs. *me*) provide the basis for establishing this relation between the speech role and the two case roles. However, the pronoun provides no help whatsoever for identifying the actual person that has uttered the sentences (2a) or (2b). In order to identify that person, one will have to be present when the sentences are uttered. This is also true of the second person pronoun, as shown by (3a–b).

- (3) a. *You may go home.*
- b. *Who gave this book to you?*

The identity of the person that the speaker had in mind while using the second person pronoun *you* in (3a) or (3b) could be obtained only by being present when the sentences were uttered and by observing the person whom the speaker was facing while uttering them. The pronoun *you* provides no basis whatsoever for this identification. It only tells us that *whichever* person is intended to be the addressee by the speaker is expected by him to perform the activities concerned, namely ‘going home’ in the case of (3a) and ‘providing an answer for the question’ in the case of (3b).

Proforms, on the other hand, are general terms that are used for carrying out different functions like (i) identifying the participants of an event by locating them with reference to the spatio-temporal location of the speech act participants

(or indicating the location, time, manner, etc. of the event itself in a similar fashion), (ii) referring back (or forward) to other expressions that occur in the utterance or in previous utterances, or (iii) indicating the scope of a question, negation, or exclamation. They are generally made up of two different elements of which one indicates the function of the proform (demonstrative, interrogative, indefinite, or anaphoric (relative)) and the other one denotes their category or scope (person, thing, place, manner, quality, quantity, etc.). In the somewhat shortened paradigms of the proforms of Lezgian (Daghestanian: Haspelmath 1993: 188), given in (4), for example, the columns represent the function of the proforms and rows represent their scope.

(4)	Proximate	Distal	Interrogative	Indefinite
Nominal	<i>im</i>	<i>am</i>	<i>him</i>	<i>sam</i>
Place	<i>inag</i>	<i>anag</i>	<i>hinag</i>	<i>sanag</i>
Source	<i>inaj</i>	<i>anaj</i>	<i>hinaj</i>	<i>sanaj</i>
Goal	<i>iniz</i>	<i>aniz</i>	<i>hiniz</i>	<i>saniz</i>
Location (in)	<i>inra</i>	<i>anra</i>	<i>hinra</i>	<i>sanra</i>
Location (on)	<i>inal</i>	<i>anal</i>	<i>hinal</i>	—
Manner	<i>ik'</i>	<i>ak'(a)</i>	<i>hik'(a)</i>	<i>sak'(a)</i>
Quality	<i>iġtin</i>	<i>aġtin</i>	<i>hiġtin</i>	—
Quantity	<i>iq'wan</i>	<i>aq'wan</i>	<i>hiq'wan</i>	—

We can analyse these proforms as consisting of an initial element (*i-*, *a-*, *hi-*, or *sa-*) that indicates their function and a general term (*-m*, *-nag*, *-naj*, *-niz*, *-nra*, etc.) that denotes their scope. There are a few additional initial elements like *at'a-* 'remote', *ha-* 'anaphoric', *aġa-* 'up' and *wini-* 'down' and also a few additional general terms like *-nhe* 'presentational' in the language that are not shown in the table.

The number of paradigms that occur in the system of proforms differs from language to language. In most of the Australian languages, for example, there is only a single paradigm of proforms that is used for denoting both the interrogative as well as the indefinite meanings, as shown by the paradigms of proforms given in (5), which belong to Yidiny (Dixon 1977).

(5)	Gloss	Definite pronouns			Interrogative/ Indefinite pronouns
		Near	Far	Very far	
	Animate	<i>yipdu</i>	<i>ɲupdu</i>	<i>yupdu</i>	<i>wapdu</i>
	Inanimate	<i>yingu-</i>	<i>ɲungu-</i>	<i>yungu-</i>	<i>wapi</i>
	Adjectival	<i>yiju</i>	<i>ɲuju</i>	<i>yuju</i>	<i>wapa</i>
	Place/time	<i>yingu</i>	<i>ɲungu</i>	<i>yungu</i>	<i>wapda</i>
	Manner	<i>yingu:rup</i>	<i>ɲungu:rup</i>	—	<i>wapda:rup</i>
	Source	<i>yingum</i>	<i>ɲungum</i>	—	<i>wapɖum</i>
	Type	<i>yinariɲ</i>	<i>ɲunariɲ</i>	<i>yunariɲ</i>	<i>wapda:riɲ</i>

In Lezgian, on the other hand, there are two distinct paradigms of which one is used for denoting the interrogative meaning and the other one for denoting the indefinite meaning, as shown in (4). Similarly, English has only a single paradigm that is used in interrogative sentences as well as relative clauses. Sanskrit, on the other hand, has two distinct paradigms (in addition to two demonstrative paradigms), as shown in (6), of which one is used in interrogative sentences (7a) and the other one in relative clauses (7b) (Speijer 1884: 204).

(6)	Gloss	Demonstrative		Relative	Interrogative
		Proximate	Remote		
	Person (M)	<i>ayam</i>	<i>saḥ</i>	<i>yaḥ</i>	<i>kaḥ</i>
	Person (F)	<i>iyam</i>	<i>sa:</i>	<i>ya:</i>	<i>ka:</i>
	Thing	<i>idam</i>	<i>tad</i>	<i>yad</i>	<i>kad/kim</i>
	Place/Goal	<i>atra</i>	<i>tatra</i>	<i>yatra</i>	<i>kutra/kva</i>
	Source	<i>itaḥ</i>	<i>tataḥ</i>	<i>yataḥ</i>	<i>kutaḥ</i>
	Time	<i>ida:ni:m</i>	<i>tada:ni:m</i>	<i>yada:/yarhi</i>	<i>kada:/karhi</i>
	Quality	<i>i:dṛśaḥ</i>	<i>ta:dṛśaḥ</i>	<i>ya:dṛśaḥ</i>	<i>ki:dṛśaḥ</i>
	Amount	<i>iyat</i>	<i>ta:vat</i>	<i>ya:vat</i>	<i>kiyat/kiyant</i>
	Number	—	<i>tati</i>	<i>yati</i>	<i>kati</i>
	Manner	<i>ittham</i>	<i>tatha:</i>	<i>yatha:</i>	<i>katham</i>

- (7) a. *ko bhava:n*
 who (Int) you
 ‘Who are you?’
- b. *ya a:ryaḥ tam pṛccha*
 who (Rel) gentlemam him ask
 ‘Ask that (person) who is a gentleman’

Languages may use a single paradigm for carrying out all these functions. For example, Lyle, belonging to the Gur branch of Niger-Congo family (Showalter 1986) is reported to use a single set of proforms in the function of demonstratives, interrogatives, and relatives (see §7.3.1 for details). In (8), for example, the same proform functions both as an interrogative (8a) and a demonstrative (8b). The only difference between the two is that the high tone of the interrogative contrasts with the falling tone of the demonstrative. Showalter (1986: 202) suggests, however, that the high tone, which is apparently optional, can be regarded as the ‘intonation’ of the constituent question in which the interrogative pronouns occur (Showalter 1986: 202).

- (8) a. *jì kèé/kèê*
 house which/this
 ‘Which house?’
- b. *jì kèê*
 house which/this
 ‘This house’

Languages that combine together two different functions and represent them through a single paradigm would generally be using some other device for indicating the distinctions that occur between those functions. For example, there are several languages that use the same set of proforms in interrogative and indicative sentences, but they use additional devices like having a question particle, interrogative verb, or interrogative intonation for differentiating between the function of interrogativity and that of indefiniteness (see §10.2 for details).

There is a similar kind of variation among languages regarding the number of rows that occur among the paradigms of proforms. It derives from the fact that different languages use different sets of categories (word classes) or general concepts (like person, thing, place, time, manner, amount, type, etc.) among their proforms. For example, some may have distinct nominal and adjectival proforms, whereas some may use the same set of proforms in both these functions. Some may have special verbal proforms and some may not. Similarly, some languages may have a single set of proforms for denoting location whereas others may differentiate between source, goal, and location and use distinct sets of proforms for representing them. However, the former may use additional markers for differentiating between different types of locations (see §7.4 for details).

There is, thus, an important functional distinction between personal pronouns and proforms. The former are used for denoting speech roles and for indicating their involvement in the events or states that the sentences in which they occur describe, whereas the latter are used for locating the participants of events, or the events themselves with reference to the speech context, or for indicating the scope of questions, negations, etc. This distinction in the functioning of these two types of pronoun gives rise to several differences in the characteristics that get associated with them. And this is the reason why we find it rather difficult to establish a single definition that can cover both these types of pronoun.

1.2.1 Differentiating characteristics

I will be describing in detail the various characteristics of personal pronouns and proforms separately in the following chapters. My purpose in the present section is merely to indicate some of the most important differences that occur between these two types of ‘pronouns’ in order to justify my decision to describe them in two distinct parts of this monograph. All these differences derive directly from the functional distinction that separates personal pronouns from proforms.

Languages tend to dissociate personal pronouns from any of the identifying characteristics of their referents. For example, modifiers and complements that help to differentiate between different types of referent are not generally associated with personal pronouns. Gender distinctions are also generally restricted to third person pronouns. They are rarely shown in the case of first and second person pronouns. Even when gender distinction is shown in second person or first

person plural, it may be used for denoting speech role distinctions rather than distinctions among the referents (see §4.3 for details). This is also true of the association of these pronouns with the category of number. It is used for denoting different combinations of speech roles rather than the plurality of their referents. First person plural pronoun, for example, denotes a combination of first and third person or of first and second person and not a plurality of speakers (see §4.2).

This tendency to dissociate personal pronouns from their referents gets reflected in the fact that languages use different kinds of alternative device for indicating the identity of their referents in some of the contexts in which such an indication is necessary. For example, several languages use appositive noun phrases for identifying the referents of personal pronouns, while the pronouns themselves are allowed to remain unchanged (see §2.2.3). Thus, an individual who is taking an oath is identified in that oath not by directly modifying the first person pronoun that refers to that individual, but rather by placing the name of that person in apposition to the pronoun. The pronoun itself is not burdened with this identifying information.

(9) *I, John Smith, hereby declare that . . .*

The reason for this dissociation of personal pronouns from aspects of their referents is quite evident. As a conversation progresses, the speech roles of being the speaker and being the addressee shift constantly among two or more individuals who are participating in that conversation. In order to indicate, consistently, the speech roles of being the speaker and being the addressee, personal pronouns need to remain insensitive to such shifts. That is, they should remain constant when the speech roles remain the same even if the participants change, whereas when the speech roles change, they must change even if the participants remain the same.

- (10) a. John: *'I need to go to the city urgently.'*
 b. Bill: *'You can take my car.'*
 c. Mary: *'But you had promised to take me to the movies.'*

Notice that the first and second person pronouns occurring in (10a–c) consistently indicate the speaker and the addressee respectively. They are insensitive to any of the shifts that occur among the referents of those pronouns. For example, the first person pronoun has three different referents in these three sentences and the second person pronoun also has different referents in (10b) and (10c). It would not be possible for these pronouns to remain insensitive to changes occurring among their referents in this fashion if they are directly associated with information concerning their referents. For example, if the second person pronoun occurring in (10a) were to indicate any of the identifying characteristics of its referent, namely John, it would not have been possible for Mary to use that pronoun unaltered in (10c) as it has a different referent, namely Bill in that sentence. Thus the tendency to dissociate personal pronouns from information

concerning their referents derives directly from their primary function of denoting speech roles (see §2.2).

There is no comparable restriction concerning the association of proforms with information regarding their referents. As determiners, for example, proforms can be directly associated with noun phrases that include different types of modifier (as in (11a) of English) or complements (as in (11b) of Kannada (Dravidian) given below). As nominals, they can be modified directly in order to indicate distinctions of gender and number, deictic distinctions, etc., all of which have the primary function of providing information about their referents. Their function in language apparently does not make it necessary to dissociate them from all this information regarding their referents, and on this point they are clearly different from personal pronouns.

- (11) a. *This beautiful pen is a gift from Mary.*
 b. *ni:vu tand-a i: pustaka canna:g-ide*
 you brought-Rel this book good-is
 'This book, which you have brought, is good'

Another interesting difference between personal pronouns and proforms concerns their association with the notion of definiteness. There are two different meanings in which linguists use the term 'definite'. In the case of third person pronouns (or noun phrases), it is used for denoting the speaker's assumption regarding the ability of the *addressee* to identify their referents. In the case of proforms, on the other hand, it is used for indicating the *speaker's own* ability to do so (§9.3). For example, a speaker would use an indefinite noun phrase like *a car* even for referring to a car that he has purchased, if his addressee has not heard about it earlier. Evidently, the speaker himself would not have any difficulty in identifying the referent of that noun phrase. Its 'indefiniteness' depends upon his inability to assume that his addressee can identify it. On the other hand, a speaker would use an interrogative pronoun when he *himself* is unable to identify an individual or object. In fact, he would be expecting, while using an interrogative pronoun, that his addressee would be in a position to provide the necessary information.

Another important difference between these two uses of the term 'definite' is that in the case of noun phrases definiteness gets established merely through 'verbal' or 'linguistic' identification, whereas in the case of proforms there is generally a need to establish a more substantial 'extra-linguistic' identification. The indefiniteness of a noun phrase comes to an end the moment it is used in a sentence, whereas the indefiniteness of an interrogative (or indefinite) pronoun persists so long as the speaker feels that he does not possess sufficient information about its referent for establishing a proper 'extra-linguistic' identification (see §9.2.1–2 for details).

Notice, however, that in both these usages, the term 'definiteness' denotes primarily the identifiability of the *referents* of an expression. Since first and

second person pronouns are generally dissociated from any of the identifying characteristics of their referents, questions concerning the identifiability or non-identifiability of their referents (either for the speaker or for the addressee) would be irrelevant for their use. Hence the notion of definiteness in either of the above two meanings would not have any relevance as far as first and second person pronouns are concerned. This irrelevance of the notion of definiteness appears to manifest itself in the form of certain puzzling usages of first and second person pronouns (see §2.3.1–2).

As I have pointed out earlier (§1.2), the stems of proforms are generally made up of two different elements, namely a ‘pronominal’ element that indicates their function in language, and a general term that indicates their scope. It is true that while functioning as determiners, proforms may have only a single element in their stem, namely the pronominal element, but the nouns that occur with that element would then be functioning as the general term (second element). The stems of personal pronouns, on the other hand, generally involve only a single element, namely the pronominal one that denotes speech roles. This is another interesting distinction between the two types of pronoun, which also derives from the distinction occurring in their primary function in language.

1.2.2 *Possibility of gradation*

A question has been raised as to whether this distinction between personal pronouns and proforms constitutes a distinction between two different categories as suggested above or whether the two only form part of a *continuum* of nominals (Comrie, personal communication). There are some interesting characteristics that appear to support the latter contention. For example, some of the proforms resemble personal pronouns rather than nouns in their number and case marking. This is especially true of proforms that denote human beings. In English, for example, third person pronouns *he*, *she*, and *they* and the interrogative pronoun *who* resemble personal pronouns in showing distinct accusative and genitive case forms. There are certain other characteristics shown by these forms in other languages, like the distinction between ergative and accusative, that make it possible to establish a ‘hierarchy of nominal categories’ in which personal pronouns occupy the top position (see §5.2.3 for details). Similarly, first and second person pronouns manifest some differences that make one of them more similar to third person pronouns or demonstratives (see §5.2.1–2).

A closer examination of these aspects of gradation reveals, however, that they derive from an interesting conflict that occurs in our use of personal pronouns. As I have pointed out in the preceding sections, the fact that personal pronouns need to be dissociated from their referents in order to indicate, consistently, the speech roles of ‘being the speaker’ or ‘being the addressee’, is primarily responsible for most of the characteristics that differentiate them from other nominals, including proforms. They cannot, however, completely escape from the fact that they do

possess referents of their own. This latter factor gives rise to certain characteristics that conflict with the former set of characteristics, and to that extent they become less consistent in their primary function. The appearance of a gradation between personal pronouns and proforms is mainly due to the occurrence of these latter characteristics.

For example, first and second person pronouns do not generally show any gender distinctions. This is apparently due to the fact that gender marking would associate personal pronouns directly with their referents and this would adversely affect their primary function of denoting speech roles. However, there do occur some languages (primarily those of the Semitic family) that allow second person pronouns to show gender distinction. Several other languages show gender distinction in the non-singular forms of all persons (see §4.3). This cross-linguistic variation in the gender marking of personal pronouns makes it appear that they form a gradation along with third person pronouns, proforms, and other nominals. Notice, however, that gender marking derives from their *association* with their referents, which goes against their primary function of denoting speech roles.

In view of this conflict occurring among the characteristics of personal pronouns, we may perhaps have to make use of 'Prototype Theory' as the basis of our categorization of personal pronouns (and also of proforms), as I have suggested earlier in this chapter (see §1.1). We may regard first person singular pronoun as the most prototypical among personal pronouns, as it shows almost all the characteristics that derive from the primary function of this category. Other pronouns of this category are less prototypical and therefore show fewer of those characteristics. They also tend to show some of the characteristics that belong to the neighbouring categories like proforms and nouns, and this gives rise to the appearance of a gradation among these categories.

1.3 Position of third person pronouns

Another point that needs to be examined in this connection is the position of third person pronouns in this division of pronouns into personal pronouns and proforms. Should we regard them as part of the system of personal pronouns or should we relegate them to the system of proforms? Several linguists have advocated the latter alternative, on the basis of the claim that only first and second person pronouns denote individuals who participate in the speech act. The referents of third person pronouns are seen as 'non-persons' whose position in the speech act can only be described negatively with respect to first and second person pronouns (J. Lyons 1977: 638). This claim is supported by the tendency shown by several languages to dispense with third person pronouns in favour of demonstrative pronouns.

In Kannada (Dravidian), for example, third person pronouns are identical with remote demonstratives. They do not share with first and second person pronouns

any of the characteristics that derive from their specialized function of denoting speech roles rather than the individuals who perform those roles. For example, the plural markers that occur with third person pronouns are the same as the ones that occur with nouns and are different from the ones that occur with first and second person pronouns; third person pronouns show gender distinctions and distinctions for endophoric and exophoric usage (§3.2), unlike first and second person pronouns, and so on. In view of these points, there would not be any difficulty, as far as Kannada is concerned, in regarding third person pronouns as belonging to the system of proforms rather than that of personal pronouns.

On the other hand, there do occur several languages like Angami (Tibeto-Burman: Giridhar 1980), Mundari (Austroasiatic: Osada 1992), etc. in which third person pronouns are quite different from demonstratives. Unlike the latter, they share several characteristics with first and second person pronouns, such as the occurrence of distinct plural markers and case markers. Hence, in such languages, there is some basis for claiming that the third person pronouns properly belong to the system of personal pronouns rather than that of proforms or demonstratives. This dichotomous treatment of third person pronouns either as related to demonstrative pronouns or as forming a distinct personal pronoun appears to form the basis of an interesting typological distinction among languages. We may regard languages that have third person pronouns of the former type as 'two-person' languages, and the ones that have personal pronouns of the latter type as 'three-person' languages, as I point out in detail in the sixth chapter.

One interesting distinction that can be correlated with this typological distinction concerns the occurrence of gender distinctions. Languages that manifest gender distinctions among their third person pronouns tend to be two-person languages, whereas the ones that do not show any gender distinctions among their third person pronouns tend to be three-person languages (§6.3). Another cross-linguistic difference that can be correlated with this typological distinction occurs in the system of demonstratives. Anderson and Keenan (1985: 282) suggest differentiating between 'person-oriented' and 'distance-oriented' deictic systems, with the former involving first, second, and third persons as distinct deictic centres ('near 1', 'near 2', and 'near 3' or 'away from both 1 and 2') and the latter involving only the speaker as the deictic centre (proximate and remote or proximate, medial, and remote). It appears that distance-oriented deictic systems are generally preferred by two-person languages whereas person-oriented deictic systems are preferred by three-person languages (§6.4). Changes that affect languages also appear to support these correlations between the two-person/three-person typology on the one hand and distinctions among demonstratives, gender-marking among third person pronouns, etc. on the other (§6.6).

It is possible to regard this typological distinction between two-person and three-person languages as representing a distinction in the world-view of the speakers of these languages. Speakers of two-person languages, for example, may be viewing the speech act as basically a two-way affair, taking place

between a speaker and his addressee. Speakers of three-person languages, on the other hand, may be viewing it as a group affair, taking place among three or more individuals, with the speech roles, denoted by the pronouns of first, second, and third person, shifting among the individuals of that group. It would be interesting to see whether there are any cultural distinctions that can be correlated with this linguistic distinction.

1.4 'Free-pronoun' and 'bound-pronoun' languages

Another important typological distinction that we need to make among languages is between those in which the primary function of personal pronouns is carried out by independent personal pronouns that occur as arguments on the one hand, and the ones in which it is carried out, instead, by bound pronouns (clitics or affixes) that occur as part of the predicate on the other. In the former case, agreement markers that occur in the predicate would not have to perform the primary function of personal pronouns, namely to denote the involvement of speech roles in the event or state that the predicate denotes, and hence they may not manifest the prototypical characteristics of personal pronouns. In the latter case, on the other hand, it would be the independent personal pronouns rather than agreement markers or clitics that would be free of this function and hence the independent pronouns of such languages may not show any of the characteristics that they manifest in languages of the former type (see below, §1.4.1).

Several languages manifest disparity between the system of independent personal pronouns and that of bound forms (clitics or agreement markers), but this disparity, as it occurs among bound-pronoun languages, appears to be the opposite of the one that occurs among free-pronoun languages (§1.4.2). Secondly, while the occurrence of agreement markers is rather non-obligatory and irregular in free-pronoun languages, the occurrence of independent pronouns is non-obligatory in bound-pronoun languages (§1.4.3). Another interesting difference between these two types of language concerns the diachronic changes that affect personal pronouns or agreement markers. In free-pronoun languages, it is the agreement marker that appears to show instability, whereas in bound-pronoun languages independent pronouns are rather unstable (§1.4.4). All these characteristics, which differentiate between these two types of languages, appear to derive from the primary functional distinction mentioned above that forms the basis of this typology.

Unfortunately, however, this interesting typological distinction appears to be affected by the occurrence of languages in which both these types of pronominal expression are able to carry out the primary function of personal pronouns. Such languages may show complementation between the two types of expression, or they may redundantly use both of them together. In fact, the two distinct types of personal pronoun system mentioned above appear to form two ends of a

continuum, with the systems occurring in individual languages being more or less close to one or the other of the two ends of the continuum (§1.4.5).

1.4.1 Two kinds of agreement

Traditionally, pronouns that occur as independent words and function as arguments of sentences are considered to be rather different from agreement markers that occur with verbs and denote distinctions of person. Only the former are considered to be essential for the functioning of language. Jespersen (1922: 335), for example, regards the latter as superfluous. Further, the categories of person and number are considered to be 'nominal' categories, and their association with verbs is considered to be 'determined' by distinctions occurring in the former case (J. Lyons 1968: 242).

This traditional view is apparently based upon the fact that, in most of the familiar languages, speech act participants are indicated primarily by independent personal pronouns. Case markers that indicate the nature of their participation in events and states are also attached to these independent pronouns. Agreement markers that occur in the verb, on the other hand, have only the function of partially repeating this information. They may or may not occur in a language and can easily be lost, unlike independent pronouns. They may provide information only about the most prominent participant, like the subject or agent of a clause but not about other (oblique) participants. Further, they are generally combined with other elements that have more important functions to perform, such as denoting the tense, aspect, and mood of the verb. Paradigms of different tenses and moods may differ in showing or not showing personal distinctions or in showing only a few distinctions. These characteristics of agreement markers, in contrast to the characteristics shown by independent personal pronouns, indicate clearly that agreement markers have only a minor function to perform, namely that of supporting personal pronouns.

On the other hand, there are several other (non-familiar) languages in which the agreement markers or bound pronouns occurring in the predicate are more prominent and obligatory as compared to independent pronouns. The latter are used only for emphasis or contrast in these languages. Case role distinctions are expressed through distinctions occurring in agreement markers rather than through affixes occurring with independent pronouns. Linguists who worked on such languages feel that the traditional view, regarding agreement markers as having only a supportive role, needs to be replaced by a view according to which the two types of pronominal markers are considered to be *equivalent* ways of person marking (Cysouw 2001: 13). Thus, Anderson and Keenan (1985: 260) claim that the deictic function of first and second person elements is independent of whether the pronominal form is an independent word, a clitic, or simply an inflectional affix. This is also apparently the view of Givón (1984: 353), who argues that stressed independent pronouns, unstressed/clitic pronouns, and verb agreement markers

constitute both a functional-synchronic and diachronic cline. They form a hierarchy as shown in (12).

(12)	<u>strong</u>	<u>weak</u>	<u>clitic</u>	<u>bound</u>	zero
	pronouns			agreements	

More recently, however, several linguists have pointed out that there is a need to differentiate between two types of bound pronouns or agreement markers. For example, Jelinek (1984: 44) argues that, in languages like Warlpiri, clitic pronouns occurring with auxiliaries function as arguments, and the core case relations are expressed by distinctions occurring in them rather than in independent pronouns or noun phrases. The latter expressions function only as adjuncts and are non-obligatory. The latter also do not take any case markers.

There are also languages like Straits Salish that possess no independent pronouns as such. Affixes and clitics occurring with the predicate denote all the core arguments in them. Even for indicating oblique arguments that are identified only with respect to person and number, Straits Salish makes use of a set of deictic roots that function syntactically as third person pronouns which, however, can also function as predicates or adjuncts (Jelinek and Demers 1994: 714).

Bresnan and Mchombo (1987: 741) also consider it necessary to differentiate between two kinds of verbal agreement, which they call 'grammatical agreement' and 'anaphoric agreement'. In grammatical agreement, there is generally a noun phrase that bears the argument relation with the verb and the agreement marker (verbal affix) has only the function of representing, redundantly, the person, number and gender of that noun phrase. In anaphoric agreement, on the other hand, the verbal affix functions as an incorporated pronominal argument of the verb and the coreferential noun phrase has only a non-argument function—either as an adjunct or as a topic or focus of the clause or discourse structure. Bresnan and Mchombo (1987) refer to the occurrence of certain additional characteristics that can be correlated with this distinction between grammatical agreement and anaphoric agreement. In Chicheŵa (a Bantu language) in which both these types of agreement marker are reported to occur, for example, only an anaphoric agreement marker can have a non-local agreeing adjunct, that is, one that occurs in a different clause.

Baker's (1996) notion of 'polysynthetic parameter' is also primarily based upon a similar distinction among agreement systems. Languages that show obligatory morphological marking of arguments in the verb (called 'morphological visibility condition'), coupled with the occurrence of non-configurational syntax, are considered to be polysynthetic. Baker (1996: 11) points out that the idea of describing certain languages as having all the core arguments obligatorily represented in the predicate itself is also a traditional one, with a long history in Amerindian linguistics. It goes back to at least as far as Wilhelm von Humboldt's analysis of Aztec in the 1830s, according to Foley (1991: 228). Baker (1996: 14),

however, regards the agreement markers to be ‘reinforcing’ the conventional argument phrases rather than replacing them, apparently as an effort to maintain the Binding Theory.

I would like to suggest that this distinction in the nature of verbal agreement is one of the characteristic differences that can be correlated with the typological distinction between ‘free-pronoun’ and ‘bound-pronoun’ languages, mentioned above. It can be derived directly from the functional distinction that forms the basis of that typology. Languages that manifest the polysynthetic parameter or the ones that show ‘anaphoric’ rather than ‘grammatical’ agreement also have the primary function of personal pronouns carried out by personal markers occurring in the verb rather than by independent pronouns. They can be regarded as ‘bound-pronoun’ languages. On the other hand, languages that have ‘grammatical’ agreement markers also have the primary function of personal pronouns carried out by independent pronouns. They are ‘free-pronoun’ languages.

1.4.2 Disparity between pronominal systems

There are apparently several other differentiating characteristics that can be correlated with this typological distinction between ‘free-pronoun’ and ‘bound-pronoun’ languages. For example, several languages show disparity between the system of independent pronouns on the one hand, and the system of their bound forms on the other, but the kind of disparity that occurs among free-pronoun languages appears to be the *opposite* of the one that occurs among bound-pronoun languages. This difference appears to derive from the fact that these two types of languages use different types of pronouns (bound vs. free) for carrying out the function of denoting speech roles.

Several Indo-Aryan languages manifest gender distinction among agreement markers but not among independent personal pronouns. In Marathi (Indo-Aryan), for example, first and second person pronouns do not show any gender distinction when they occur as independent pronouns (see 13a), but when they occur as agreement markers with the verb, they show a two-fold masculine–feminine gender distinction as shown in (13b), which represents the past forms of the verb *as* ‘to be’ (Bernsten and Nimbkar 1982: 42, 57).

- | | | | |
|------|----|--------------|--------------------|
| (13) | | Singular | Plural |
| a. | 1 | <i>mi</i> | <i>amhi</i> (Excl) |
| | | | <i>apəŋ</i> |
| | 2 | <i>tu</i> | <i>tumhi</i> |
| b. | 1M | <i>hoto</i> | MF <i>hoto</i> |
| | 1F | <i>hote</i> | |
| | 2M | <i>hotas</i> | MF <i>hote</i> |
| | 2F | <i>hotis</i> | |

Sentences (14a–b) exemplify the occurrence of this differentiation between personal pronouns and agreement markers in Marathi (Kashi Wali, personal communication).

- (14) a. *tu ghari gel-aa-s asa Raam mhaṇaala*
 you home went-MSg-Clitic so Ram said
 ‘You went home, so Ram said’
- b. *tu ghari gel-i-s asa raam mhaṇaala*
 you home went-FSg-Clitic so Ram said
 ‘You went home, so Ram said’

Notice that the second person verbal forms, *gelaas* ‘went (MSg)’ and *gelis* ‘went (FSg)’, occurring in the embedded clauses of (14a–b) show MF gender distinction, but the pronoun *tu* ‘you’ occurring in the subject position of these clauses does not show any such distinction.

This disparity in gender marking leads to an interesting distinction in the control of verbal agreement. While the person (and number) distinction is controlled by the pronoun that occurs in the subject position, gender distinction is controlled by the noun phrase that provides the basis for the identification of the *referent* of that pronoun. This can be exemplified with the help of sentences that involve verbs of saying. For example, there are two different kinds of ant, of which one is called *mungi*, a word in feminine gender, and the other one is called *mungla*, a word in masculine gender. In a story that reports a conversation among these ants there would be gender distinction in the case of first person verbs but not in that of the corresponding personal pronouns, as shown in (14c,d) (Kashi Wali, personal communication).

- (14) c. *mi ghari jaat-e asa mungi mhaṇaa-l-i*
 I home go-Pres.1FSg so ant (FSg) said-Past.3Fsg
 ‘I go home so ant said’
- d. *mi ghari jaat-o asa mungla mhaṇaa-l-aa*
 I home go-Pres.1MSg so ant (MSg) said-Past.3MSg
 ‘I go home so ant (of a different kind) said’

Nama Hottentot (Central Khoisan: Hagman 1977) also shows disparity between personal pronouns and agreement markers regarding both gender marking and number marking. In the case of personal pronouns, there is no gender distinction, and further there is only a two-fold singular–plural number distinction. The latter distinction is also restricted to first person (15a). In the case of agreement markers, on the other hand, there is a three-fold singular–dual–plural number distinction in all three persons, and also a two-fold masculine–feminine gender distinction in the singular (and dual), and a three-fold masculine–feminine–common distinction in the plural (15b). This latter distinction, however, does not occur in first person singular (Hagman 1977: 42, 44).

(15) a. *Personal pronouns*

	Singular	Plural
1	<i>tií</i>	<i>síí</i>
2	<i>saá</i>	
3	<i>//ĩĩ</i>	

b. *Agreement markers*

	Singular	Dual	Plural
1M	<i>ta</i>	<i>kxm̩</i>	<i>ke</i>
1F	„	<i>ñt/ĩm̩</i>	<i>se</i>
1C	—	„	<i>tá</i>
2M	<i>ts</i>	<i>kxò</i>	<i>ko</i>
2F	<i>s</i>	<i>rò</i>	<i>so</i>
2C	—	„	<i>tù</i>
3M	<i>p/i</i>	<i>kxà</i>	<i>ku</i>
3F	<i>s</i>	<i>rà</i>	<i>tì</i>
3C	—	„	<i>ñt/ĩn̩</i>

However, Nama Hottentot allows these agreement markers to be attached to personal pronouns, as in *saáts* ‘you (MSg)’, and also to nouns that occur in apposition to these pronouns, as in *saá kxòets* ‘you (MSg) the person’.

The situation appears to be exactly the reverse among bound-pronoun languages. They show several distinctions in their independent personal pronouns that are absent among agreement markers. For example, *Tukang Besi* (Austronesian: Donohue 1999: 113) shows a two-fold singular–plural distinction in the case of pronominal affixes that occur with the verb; the distinction, however, is restricted to first and second person. There is also a paucal form, used in first person only, involving a lot of overlap. Free pronouns, on the other hand, show the distinction between dual, paucal, and plural in all persons. Donohue notes that pronominal affixes in this language are used obligatorily, whereas free pronouns are rarely used.

Tinrin (Melanesian: Osumi 1995: 37, 170) is similar in that it has dual and plural forms for all persons, including the inclusive–exclusive distinction in the first, in the case of independent personal pronouns, but in that of reduced forms of pronouns that occur with the verb, the language has only two non-singular forms, namely first person inclusive and third person.

(16) a. *Independent pronouns*

	Singular	Dual		Plural	
		Excl	Incl	Excl	Incl
1	<i>nro</i>	<i>haru</i>	<i>komu</i>	<i>hari</i>	<i>kevi</i>
2	<i>nrii</i>		<i>kou</i>		<i>wiri</i>
3	<i>nri</i>		<i>nrorru</i>		<i>nrorri</i>

b. *Subject pronouns*

	Singular	Plural
1	<i>u</i>	<i>ri</i> (inclusive)
2	<i>ke</i>	—
3	<i>nrâ</i>	<i>rri</i>

Tamazight of the Ayt Ndhir (Berber: Penchoen 1973: 25, 27) shows a two-fold gender distinction in second and third persons in the case of free pronouns, whereas in that of agreement markers it shows this distinction only in third person.

(16) c. *Independent pronouns*

	Singular	Plural
1	<i>nəkk</i>	<i>nukni</i>
2M	<i>šəgg</i>	<i>kʷənni</i>
2F	<i>šəmm</i>	<i>kʷənim̄i</i>
3M	<i>nətta</i>	<i>nihni</i>
3F	<i>nəttat̄</i>	<i>nihənt̄i</i>

d. *Subject affixes*

1	-x	n-
2	<i>t̄...d̄</i>	<i>t̄...m̄</i>
3M	y-	-n
3F	<i>t̄-</i>	<i>-nt̄</i>

This difference in the kind of disparity that occurs between the personal pronouns and agreement markers of free- and bound-pronoun languages appears to be even more striking in the category of case. Traditionally, case marking is considered to be a property of nominal expressions. Nouns and pronouns are described as taking different case markers in order to indicate either directly or indirectly the various case roles in which their referents participate. This point can be exemplified with the help of the Kannada (Dravidian) sentences (17a–b) in which case distinctions are indicated primarily through affixes that occur with noun phrases.

- (17) a. *ra:ju tanna pustakav-annu hari-ge to:risida*
 Raju self's book-Acc Hari-Dat showed
 'Raju showed his book to Hari'
- b. *ra:ju hariy-annu mane-ge kaḷuhisida*
 Raju Hari-Acc house-Dat sent
 'Raju sent Hari to (his) house'

In both (17a) and (17b), the agent is left unmarked, but it is differentiated from the patient by marking the patient with the accusative suffix and the two are differentiated from the goal by marking the goal with the dative suffix. Agreement markers that occur with the verb, on the other hand, can indicate only one of

the core arguments, namely the agent of volitional verbs and the patient of non-volitional verbs. There are some verbal forms like the negative and the infinitive that show no personal agreement whatsoever.

The case-marking disparity that occurs among bound-pronoun languages, on the other hand, is the opposite of the disparity found in free-pronoun languages. In bound-pronoun languages, case is associated with clitics or agreement markers either by placing the relevant pronominal elements in different positions in the predicate or by making distinctions in the form of the elements used (or both). It is also possible to have the case markers attached to these pronominal elements. Nichols (1986: 100) uses the terms ‘dependent-marking’ (nominal marking) and ‘head-marking’ (verbal marking) for describing these two kinds of case-marking strategy. Sentence (18) of Abkhaz (Hewitt 1979: 36 as quoted in Nichols 1986: 61) exemplifies the latter strategy.

- (18) *a-xàc'a a-phš a-š° q'š Mφ-M|š-My-Hte-yt'*
 the-man the-woman the-book it-to.her-he-gave-Finite
 ‘The man gave the woman the book’

There are three different noun phrases in (18), but they remain unmarked for case. The nature of their participation in the action of ‘giving’ is expressed by the relative position of the cross-referring pronominal elements that occur with the verb.

Kamera (Central Malayo-Polynesian: Klammer 1998: 62, 67) is another language of this latter type. It has different sets of pronominal clitics showing distinctions for nominative, accusative, dative, and genitive. Its independent pronouns (and also noun phrases), however, do not show any case role distinctions. They are used only non-obligatorily when the relevant participant is to be emphasized or disambiguated.

- (19) *na tau wútu na-palu-ka nyungga*
 Article person be.fat 3Sg.Nom-hit- 1Sg.Acc I
 ‘The big man hit me’

The first person occurs as a patient in (19), but this fact is indicated through the use of the accusative form of the clitic *-ka*, while the independent pronoun *nyungga* ‘I’ occurs in its invariable form.

1.4.3 Obligatoriness

The main criterion that I have used for differentiating between free- and bound-pronoun languages depends upon the obligatoriness and consistency of personal markers. In free-pronoun languages, it would be the independent personal pronouns that are obligatory and consistent. Agreement markers may or may not occur in these languages, and further, even when they occur, they may be inconsistent in that different types of paradigm may show different types of

personal distinctions. They may also show mergers ('syncretisms') that do not have any functional explanation.

For example, the agreement markers occurring in most of the Indo-Aryan languages show a distinction between two different types of paradigms called personal and nominal, with the latter showing distinctions for gender and number but not for person (Masica 1991: 259). Some languages like Marathi show personal distinctions in the auxiliary of such verbs, but even here the distinction is restricted to second person (Berntsen and Nimbkar 1982: 259).

- (20) a. *Personal agreement* (future paradigm of the verb *ja-* 'to go')

	Singular	Plural
1	<i>jañ</i>	<i>jau</i>
2	<i>jašil</i>	<i>jal</i>
3	<i>jail</i>	<i>jatil</i>

- b. *Nominal agreement* (perfect paradigm of the verb *kər-* 'to do')

M	<i>kela</i>	<i>kele</i>
F	<i>keli</i>	<i>kelya</i>
N	<i>kelə:</i>	<i>keli</i>

Forms of the paradigm (20b) denote only gender and number distinctions that are meant for denoting agreement with the object. They occur with an auxiliary that differentiates between second person singular subject and all other subjects by having the form *-es* in the former case and *-e* in the latter case.

- (21) a. *mi kam kelə:-e*

I work(N) done(N)-have
'I have done the work'

- b. *tu kam kelə:-es*
you work(N) done(N)-have(2Sg)
'You have done the work'

- c. *tyani ĵvari perli-e*
he millet(F) sown(F)-have
'He has sown millet'

Even in the case of paradigms that do show personal distinctions, there is no consistency. For example, past habitual paradigm differs from future paradigm (given earlier in (20a)) in showing gender distinctions in addition to person and number distinctions, but in the case of first person, the plural form is identical with the masculine singular form. There is also syncretism between second plural and third singular masculine and between third singular feminine and third plural neuter. Agreement markers occurring in this paradigm are shown in (22) (Berntsen and Nimbkar 1982: 79).

(22)	<u>Singular</u>			<u>Plural</u>		
	M	F	N	M	F	N
1	<i>o</i>	<i>e</i>	—	<i>o</i>	<i>o</i>	—
2	<i>as</i>	<i>is</i>	—	<i>a</i>	<i>a</i>	—
3	<i>a</i>	<i>i</i>	<i>ə:</i>	<i>e</i>	<i>ya</i>	<i>i</i>

Additional irregularities occur in other paradigms. For example, imperfect habitual shows syncretism between first and third person singular forms, whereas perfect intransitive shows syncretism between second and third person singular forms (Berntsen and Nimbkar 1982: 88, 93). There are apparently diachronic explanations for these and other syncretisms that occur in the marking of personal distinctions through agreement affixes in this language, but there do not appear to be any functional explanations other than that the agreement affixes do not have the primary function of indicating the involvement of speech roles in the event or state that the predicate denotes. They only repeat the information that is already provided by personal pronouns and hence it is not necessary for them to be consistent.

Bound-pronoun languages, on the other hand, use agreement markers very consistently and regularly for denoting personal distinctions. For example, Baker (1996: 20) points out that all polysynthetic languages have full and obligatory agreement paradigms for both subject and object. The languages also allow argument-drop and at least some degree of freedom in word order. Baker points out further that these languages do not have any infinitives (verbal forms that lack agreement inflection for subject) of the type found in English, apparently because the occurrence of such forms would violate the condition that verbal forms must obligatory mark the core arguments.

While agreement markers are used obligatorily and consistently in these bound-pronoun languages, independent personal pronouns are used only optionally, and have only a supportive role. Their function is merely to emphasize (or contrast) the involvement of speech roles in the events or states that the predicates denote. For example, Tutelo (Siouan: Oliverio 1997: 63) has three sets of pronominal prefixes that occur in the verb and indicate the actor, patient, and the dative case roles.

(23)	<u>Actor prefixes</u>		<u>Patient prefixes</u>		<u>Dative prefixes</u>	
	Singular	Plural	Singular	Plural	Singular	Plural
1	<i>wa-</i>	<i>wqk/wq</i>	<i>wi-</i>	<i>wai-</i>	<i>wi-</i>	<i>wqk-/wq-</i>
2	<i>ya-</i>		<i>yi-/hi</i>		<i>yi-</i>	
3	(zero)		<i>i-/ (zero)</i>		(zero)	

- (24) *wa-yi-ku:-wa*
 1Sg.Agent-2Dat-give-realis
 ‘I give (it) to you’

- (25) *wai-yqt-o-ya-ste:ka*
 1Pl.Patient-heart-Loc-2Agent-good
 'You loved us'

Independent pronouns are formed in this language by attaching the patient prefixes to the stems *-ma* 'disjunctive' or *-sá* 'emphatic'. The former can also occur with the prefix *ik^há-* to provide restrictive meaning or as a verb to provide meanings like 'It is him'. That is, these independent pronouns are used only when special meanings like disjunction, emphasis, or contrast are to be indicated.

There are also some bound-pronoun languages that do not have any independent personal pronouns as such. In Cayuvava (Key 1967: 34), for example, the class of words that may be translated as personal pronouns involve the use of the auxiliary verb *āre* 'is, exists' along with the affix *hi* and one of the direct object markers. They can occur as auxiliary verbs as well.

- (26) a. *are-hi-ai* 'I am'
 b. *are-hi-a* 'you are, It is you'

1.4.4 Stability

Traditionally, personal pronouns are considered to form part of the basic vocabulary of language. Goddard and Wierzbicka (1994: 37) claim, for example, that the pronouns *I* and *you* are well-established semantic primitives and are attested as distinct lexical elements in every known human language. Personal pronouns are also generally considered to involve un-analysable stems unlike other pronouns such as demonstratives that involve the combination of deictic elements with certain general terms. It appears, however, that this stability and un-analysability are the characteristics of independent personal pronouns that occur among free-pronoun languages; they do not appear to be the characteristics of such pronouns that occur among bound-pronoun languages. In the latter case, independent personal pronouns are generally derived by attaching personal affixes to certain stems like emphatic markers or through reduplication.

In Warekena (Maipuran: Aikhenvald 1998: 322), for example, independent personal pronouns are formed with the help of an emphatic suffix *-ya* attached to the cross-referring personal prefixes. Similarly, in Jacalteco (Mayan: Craig 1977: 100), first and second person pronouns are formed by prefixing *ha-* (which has the function of clefting elsewhere) to the absolute set of clitic pronouns. In Tzutujil (Mayan: Dayley 1985: 61), on the other hand, independent pronouns for first and second person are formed by reduplicating the absolute person markers. They are used to mark contrastive information, and are not required in non-contrastive situations.

- (27) a. *Absolute personal markers*
- | | Singular | Plural |
|---|------------|----------------|
| 1 | <i>in-</i> | <i>oq-</i> |
| 2 | <i>at-</i> | <i>ix-</i> |
| 3 | ϕ | <i>ee-/ee?</i> |
- b. *Independent personal pronouns*
- | | | |
|---|-------------|---------------|
| 1 | <i>inin</i> | <i>ojoj</i> |
| 2 | <i>atet</i> | <i>ixix</i> |
| 3 | <i>jaa?</i> | <i>ja?ee?</i> |

In Bella Coola (Salishan: Nater 1984: 112), on the other hand, independent personal pronouns are reported to be, strictly speaking, stative verbs ('to be me') that can be nominalized like other verbs. There are also languages like Mbay (Central Sudanic, Nilo-Saharan: Keegan 1997: 62) that are reported to have no independent pronouns as such. Subject, object, and possessive affixes that are attached to verbs, prepositions, and nouns carry out the function of such pronouns. Further, certain nouns like *yáá* 'thing', *dèē* 'person', *tàa* 'speech', and *lòo* 'place' may be used in a general pronominal sense, translating as 'something' and 'somebody'. For example, *yáá* 'thing' may refer to an unspecified thing or action ('something'), whereas *tàa* 'speech' may refer to something said.

The relationship between independent pronouns and bound forms of pronouns (clitics or affixes) also shows a similar type of disparity. When the two are not derivationally related, it would be the bound forms of free-pronoun languages and independent pronouns of bound-pronoun languages that would be resulting from rather unusual sources. Helmbrecht (1996: 131) provides evidence from East Caucasian languages that supports such a claim as far as free-pronoun languages are concerned. Personal agreement is not a central typological feature of these languages, but some of them have developed personal distinctions among their agreement markers. The sources of these markers, however, are rather disparate. For example, Lak, belonging to the Lak-Dargva group, has developed agreement markers for representing a distinction between speech-act participants (first and second persons) and non-speech-act participants as shown in (28), quoted from Žirkov (1995: 95). There is no number distinction in third person.

- (28)
- | | Singular | Plural |
|-----|----------------------------------|------------|
| 1,2 | <i>-ra</i> | <i>-ru</i> |
| 3 | <i>-r/-ri/-ϕ</i> | |

Helmbrecht refers to a hypothesis put forth by Burchuladze (1979) and other authors regarding the diachronic source of these personal markers. According to this hypothesis, these agreement markers represent old auxiliaries that were fused with the participial form of the main verb. There is an additional set of

person markers occurring in Lak that is used predominantly with the past tense verb. It makes no number distinction even in first and second person. The markers that belong to this latter set are *-jaw* for first and second person and *-ja* for third person. It is hypothesized that these markers have developed from old participial suffixes that had a slot for a class marker. This class marker got petrified in this suffixal position and changed its agreement function (Helmbrecht 1996: 131).

According to Jacquesson (2001: 129), there is a distinction between first and non-first (2, 3) persons among the verbal forms of Ide, a dialect of Mishmi (Tibeto-Burman), but the suffixes used for this purpose have nothing to do with pronouns. They appear to be modal in origin. Different sets of suffixes are used in present and past forms of the verb.

(29)	Present	Past
1	<i>-zi</i>	<i>-la</i>
2, 3	<i>-ga</i>	<i>-hiba</i>

Personal pronouns of Mishmi, according to Sastry (1984: 80) are *hã* '1Sg', *nín* '1Pl', *nyú* '2Sg' and *ánè* '2Pl'.

It appears that the primary function of these agreement markers is merely to differentiate between different persons rather than to indicate speech roles. It is apparently possible, therefore, to use any device that is available for making this differentiation. Havyaka Kannada (Dravidian) shows a development of agreement marking that appears to support such a claim. While Standard Kannada does not make any distinction between inclusive and exclusive plurals in first person, Havyaka Kannada has distinct pronominal forms, *eŋgo* 'we (excl)' and *na:vu* 'we (incl)' for representing this distinction. The latter form goes back to the Proto-Dravidian inclusive pronoun, which, as I point out (§4.2.4), is made up of a second person element prefixed to first person plural form. Havyaka Kannada has a distinct personal agreement marker for representing the exclusive meaning, namely *-yõ*, but for showing agreement in the case of inclusive meaning, it uses the third person neuter singular affix. The strategy used here is apparently to employ a verbal affix that is most distinct from the exclusive one with which the inclusive meaning needs to be primarily differentiated. But the result of the use of this strategy is that the inclusive agreement marker fails to show any derivational relationship with the inclusive pronoun in this language.

- (30) a. *eŋgo ondu sinema no:ḍidde-yõ*
 we (Excl) one cinema seen-1Excl
 'We (Excl) have a cinema'
- b. *na:vu ondu sinema no:ḍid-du (nemp-idd-o)*
 we (Incl) one cinema seen-3NSg (memory-have-Int)
 'We (Incl) have a cinema, do you remember?'

It has been claimed by some linguists that diachronically clitics develop from independent personal pronouns. Givón (2001: 400) suggests the rise of pronominal systems to be proceeding from stressed independent personal pronouns to unstressed/clitic anaphoric pronouns to anaphoric pronominal affixes to obligatory pronominal agreement. However, the points mentioned above appear to indicate that the direction of development may depend upon language type. In free-pronoun languages, clitics and affixes may develop from independent pronouns, but in bound-pronoun languages, the direction of diachronic development is probably the reverse: independent pronouns derive from clitics or affixes through the attachment of emphatic markers to them.

1.4.5 *Nature of the distinction*

As mentioned earlier, this typological distinction between free-pronoun and bound-pronoun languages does not appear to be sharp and clear-cut. Just like other typological distinctions, it forms a gradation with prototypical free-pronoun languages forming one end of the gradation and prototypical bound-pronoun languages forming the other end. Remaining languages show the characteristics of both these types of languages to different degrees and fall in-between the two. For example, Bresnan and Mchombo (1987: 752) report that in Chicheŵa (a Bantu language), only the object marker shows unambiguously the anaphoric agreement. The subject marker is ambiguous between grammatical agreement and anaphoric agreement. That is, the characteristic of being a bound-pronoun language is shown clearly only in the case of object marking in Chicheŵa.

Irish (McCloskey and Hale 1984: 491) is another language in which there is an absolute complementarity between the appearance of person-number morphology on the verb and the appearance of independent phonologically expressed subject. If the subject does receive independent phonological expression, the 'analytic' form of the verb, which is unmarked for person and number, is to be used. If, on the other hand, the 'synthetic' form, which is marked for person and number is used, then the subject may not receive independent phonological expression. McCloskey and Hale (1984: 493) point out that in sentences that contain a synthetic verbal form, the 'inflectional subject' behaves for syntactic purposes exactly as if it were an overt pronoun. For example, the element *féin*, which occurs with pronouns to denote reflexive or emphatic sense, can also occur with the inflectional subject.

- (31) *An gcuirfeá féin isteach ar an phost sin*
 Int put (Cond) self in on that job
 'Would you yourself apply on that job?'

McCloskey and Hale (1984: 518) point out further than while Irish forbids the appearance of an independent pronoun in the argument position of constructions that contain a synthetic verbal form, Welsh permits an independent pronoun to

coexist with person-number morphology in the verb. Sonora Yaqui (Southern Uto-Aztecan) is another language in which there is complementarity between free and bound occurrences of pronouns. According to Dedeick and Casad (1999: 277), the language generally does not use the subject clitic in the verb for marking first or second person when it is marked by an independent pronoun. In Jacaltec (Mayan: Craig 1977: 100) also, person is marked by either an independent pronoun or a clitic pronoun. It may, however, be marked by both of them in emphatic contexts.

The distinction between dependent-marking and head-marking grammars, proposed by Nichols (1986) appears to be correlatable with this distinction between free-pronoun and bound-pronoun languages. In a head-marking language, for example, the dependent is considered to be optional, occurring only for emphasis, focus, or disambiguation. It stands in a roughly appositive relation to the pronominal marker that occurs with the head constituent. Case markers are attached to the pronominal elements that occur with the head rather than to the dependent constituent (Nichols 1986: 107–8). These are also the characteristics of independent personal pronouns occurring in bound-pronoun languages. Nichols (1986: 72) reports that most of the languages in her sample are predominantly either head-marking or dependent-marking, but probably no language is exclusively of one or the other type. There are also languages that are double-marking (having markers on the head as well as the dependent) or split-marking (having some head-marked and some dependent-marked patterns).

1.4.6 *Only a speculation*

What I have been suggesting in the foregoing sections (§1.4.1–5) regarding the possibility of establishing a distinction between free-pronoun and bound-pronoun languages, and the characteristics that can form the basis of such a distinction, is only a *speculation* that needs to be carefully examined against relevant and exhaustive data. Unfortunately, this idea occurred to me at the very end of my study of pronouns, and circumstances do not permit me to proceed further with a detailed study of this speculation.

I am aware of the fact that the distinction is of crucial importance for any characterization of personal pronouns. For example, the evidence provided by the independent pronouns of bound-pronoun languages would not be relevant for such a characterization. On the other hand, while the agreement markers of free-pronoun languages are mostly irrelevant for the characterization of personal pronouns, those of bound pronoun languages would be very relevant because we can expect them to show all the characteristics that derive from the function of denoting speech roles. Hence a proper characterization of personal pronouns would have to restrict itself to free pronouns in the case of free-pronoun languages and bound pronouns in that of bound pronoun languages.

There are several other problems that an exhaustive study of personal pronouns will have to face in connection with this possibility of making a distinction between free-pronoun and bound-pronoun languages. For example, I mentioned earlier (§1.3), that a typological distinction can be made between ‘two-person’ and ‘three-person’ languages, depending upon whether the third person pronouns are included in the system of demonstratives (proforms) or of personal pronouns. I suggested that this distinction can be correlated with a distinction between (i) languages in which speech act is regarded as a ‘two-way’ affair taking place between two individuals (with the roles of being the speaker and being the addressee shifting between them), and (ii) languages in which speech act is regarded as ‘group affair’ taking place among several individuals (with the roles of being the speaker, being the addressee, and being neither of the two shifting among those individuals; see §6.3 for details).

It is evident that this typology cannot be applied to bound-pronoun languages by using the criteria mentioned above, namely the presence versus absence of affinity between third person pronouns and demonstratives. The reason is that we cannot characterize bound pronouns (clitics or agreement markers) as not forming part of the system of personal pronouns on the basis of their affinity with demonstratives. We may have to use some other basic criterion such as the occurrence of unmarked (zero-marked) third person for establishing a comparable typological distinction in the case of these languages.

As I have mentioned earlier, however, due to my inability to proceed further with an examination of this particular aspect of personal pronouns, I have to leave all these tantalizing problems of their characterization unexplored.

1.5 Languages without personal pronouns

A question has been raised as to whether some of the Southeast Asian languages like Burmese, Thai, and Japanese can be regarded as not possessing any personal pronouns. These languages use different nouns in place of pronouns in order to indicate social status, politeness, etc. (see §4.3.4). For example, Bradley (1993: 158) points out that in Burmese the current first person pronouns are derived by adding the suffixes *fɔ* ‘male’ and *má* ‘female’ to the word *cun* ‘slave’, and the current second person pronouns are derived from religious titles. The male form, in this latter case, is probably a contracted version of *θekhin phəyà* ‘Bodhisattva’, and the female form derives from the word *shin* ‘master’.

Some linguists consider these languages as not possessing any pronouns as such, whereas others regard the nominal forms that are used in place of pronouns to be functioning as first or second person pronouns (see Diller 1994). It appears to me that a definite answer to this question can be provided only through an examination of the use of those nouns from the point of view of their ability to denote speech roles as distinct from the individuals who perform those roles. That

is, we need to find out answers to several questions of the following type. Do these languages require noun phrases to be used in matrix sentences, in addition to the nominal expressions that are used as personal pronouns, for identifying the speaker and the addressee, as in the case of sentences like (8a–c) given in §2.2.3? Do these languages allow the nominal expressions that occur as personal pronouns to take modifiers and complements such that they can establish the identity of the participants by themselves? Do they require noun phrases to be placed in apposition in an oath-taking context as in (9)? Depending upon the answers that we obtain to these questions, we would be able to say whether the nominal expressions that occur in the place of personal pronouns in these languages are to be regarded as personal pronouns or merely as nominal expressions.

It is possible, however, that these nominal expressions show a gradation in manifesting the differentiating characteristics of personal pronouns in these languages. Sugamoto (1989) makes such a claim regarding the personal pronouns of Japanese, based upon an examination of some of their characteristics. According to Hinds (1986: 241) also, personal pronouns of this language differ from the pronouns of other languages like English (i) in having nominal origins, (ii) in being terms of occupation or status titles, (iii) in being very large in number, with different forms being selected depending upon sex, age, perceived social status, and emotional correlation, and (iv) in showing most of the nominal characteristics like occurring after demonstratives and being modified by adjectives or relative clauses. Hinds concedes, however, that the nominal expressions show some characteristics of their own, which makes it necessary to regard them as forming a distinct ‘pronominal’ category.

Even in other languages, there are some special contexts in which the replacement of a personal pronoun by a nominal is considered to be necessary. They also appear to retain the force of first and second person pronouns to a certain extent in such uses. For example, while talking to a child during its language-acquiring period, the mother might use the word ‘mother’ in place of the first person singular pronoun and the name of the child in place of the second person singular pronoun (Wales 1996: 56).

(32) *Mummy will do that for Timmy.*

This replacement is apparently necessitated by the fact that language-acquiring children find it rather difficult to cope with the shifting nature of indexicals like *I* and *you*. That is, the nouns *mummy* and *Timmy* in (32) are used not as first and second person pronouns, but only as nouns. They do not function as ‘shifters’. In fact, they are meant for avoiding the use of shifters.

1.6 Structure of the monograph

I have argued in this chapter that there is a need to differentiate between pronouns of first and second person on the one hand and the rest of the pronouns, called

'proforms', on the other. There is a crucial functional distinction between the two, which makes it practically impossible to establish a definition that can be applicable to both of them. First and second person pronouns are used to indicate the speech roles of 'being the speaker' and 'being the addressee' respectively of the clause in which they occur. They can relate these speech roles with case roles, but generally they are not used to connect those speech roles with the actual individuals who perform those roles. This is because they need to transcend the individual (referential) distinctions that occur among the speakers and addressees in order to indicate, consistently, those two speech roles.

Proforms, on the other hand, are general terms that are used for carrying out different functions like identifying the participants of an event by locating them with reference to the location of the speaker or the addressee, relating them with previously identified participants, or denoting the scope of a question, negation, or exclamation. They are prototypically made up of two distinct constituent elements, namely a general term and a pronominal element. The latter indicates the function in which the former term is used. They can belong to different categories like nominal, adjectival, adverbial, and verbal, and are different from personal pronouns on this point as well: personal pronouns can only be nominal.

Since these two types of pronoun show very few characteristics that are common to both, it has become necessary to describe them separately under two different parts in this monograph. In Part I, I describe personal pronouns, and point out how they acquire several unique characteristics because of their primary function of denoting speech roles. They get dissociated from their referents (Ch. 2), take special markers for denoting non-coreference—unlike third person pronouns and demonstratives that take special markers for denoting coreference (Ch. 3), get associated with the category of number for denoting conjunction rather than plurality, and with the category of gender for denoting social distinctions rather than distinctions concerning the identity of their referents (Ch. 4). However, the fact that they do have referents of their own makes it possible for them to manifest certain additional characteristics that conflict with the former set of characteristics (Ch. 5). The last chapter (Ch. 6) of this first part deals with a typological distinction among languages, which depends upon whether third person pronouns function as part of the system of personal pronouns or of proforms (demonstratives) in those languages.

This characterization of personal pronouns is primarily based upon free-pronoun languages, that is, languages in which the primary function of denoting speech roles is carried out by independent personal pronouns. We can expect several of these characteristics to be shown by bound-pronoun languages among their agreement markers or clitics occurring with the verb, but unfortunately, I do not have sufficient data on languages of this latter type to establish this point. I am also not in a position to gather such data. It is possible as well that languages of this latter type show additional unique characteristics among their bound personal markers that derive from their primary function of denoting speech

roles. A detailed study of these languages from this particular perspective is yet to be undertaken.

Part II of this monograph provides a characterization of proforms. It describes the general structure of proforms (Ch. 7), the nature of their constituent elements (Ch. 8), and some of their important characteristics (Ch. 9), such as, for example, the occurrence of a distinct variety of reference and (in)definiteness. I point out how these basic structures and characteristics derive directly from their primary function in language. The affinity between interrogative and indefinite pronouns, which is considered to be rather puzzling by several linguists, is accounted for in the next chapter (Ch. 10). The final chapter summarizes the findings of this study (Ch. 11).

Before proceeding further with this characterization of the distinction between personal pronouns and other types of pronouns ('proforms'), I wish to emphasize an important point concerning my basic approach to the study of lexical or grammatical categories and processes. I believe that a distinction needs to be made between primary and secondary (or extended) functions of various categories and processes that occur in a language. We can account for their prototypical characteristics only on the basis of their primary function. When used in one of the extended functions, they assume additional characteristics that are relevant for that function and also lose some of the characteristics that derive from their primary function, as the latter are irrelevant for the extended function. However, they may still retain some of these latter characteristics in their extended use, in spite of the fact that the characteristics have no relevance for that extended function. We cannot therefore provide a satisfactory explanation for the various characteristics that lexical or grammatical categories and processes manifest in a language if we do not differentiate between their primary and secondary (or extended) functions or uses.

For example, the primary function of logophoric pronouns is to differentiate between speech roles (or speech act participants) that occur in endophoric contexts on the one hand, and the ones that occur in exophoric speech contexts on the other (see §3.1). They manifest several characteristics that derive directly from this primary function. For example, logophoric pronouns translate as first and second person pronouns in some languages, but as third person pronouns in others. Some languages show first person agreement for logophoric pronouns that indicate an endophoric speaker as distinct from an exophoric speaker, whereas some show third person agreement for the same. Some even show a distinct agreement marker. It is also possible to establish a hierarchy of logophoric trigger predicates (see (27) in §3.1.3) with the help of which cross-linguistic differences concerning the kind of matrix predicates that allow logophoric pronouns to occur in their embedded clauses can be described.

All these aspects of logophoric pronouns can be accounted for satisfactorily only on the basis of their primary function of differentiating between endophoric and exophoric speech roles. Several languages extend the use of these pronouns

to anaphoric contexts in which no speech act participant is involved, but even in such languages we need to regard these pronouns as having the above-mentioned use as their primary function. If they are viewed as merely anaphoric (or reflexive), we would have no explanation for many of the characteristics that they manifest in these extended uses as well, that is, characteristics that originally derived from their primary function.

I have provided justification for this approach on the basis of a study of lexical categories, especially the category of adjectives, in Bhat (1994). The present monograph can be considered to provide additional justification for that approach on the basis of a study of pronouns.

PART I

Personal Pronouns

This page intentionally left blank

2 Relation with the Referent

2.1 Introduction

First and second person pronouns are rather different from other nominals in their relation with their referents. Languages tend to dissociate them from their referents by not allowing them to take any modifiers or complements whose function is to indicate the identifying properties of their referents. Similarly, languages generally do not allow personal pronouns to occur with definite or indefinite articles, whose function is to indicate the identifiability or non-identifiability of their referents. On the other hand, in contexts in which there is a need to provide information regarding the identity of their referents, languages tend to use alternative devices like attaching appositive noun phrases or vocative nouns to personal pronouns. They are thereby able to avoid the direct association of personal pronouns with information concerning their referents. Languages may also use some of the redundant markings of personal distinctions, like the use of agreement markers, for providing information regarding the identity of the referents of personal pronouns (§2.2).

This tendency to dissociate personal pronouns from their referents has the effect of making the notion of definiteness or indefiniteness rather irrelevant for their description. There are actually two distinct meanings in which linguists have been using the term ‘definite’ and ‘indefinite’. When used with reference to noun phrases, the term ‘indefiniteness’ generally indicates the addressee’s inability to identify the referent (or the speaker’s assumption regarding that inability) whereas, when used with reference to proforms, it denotes the speaker’s own inability to do so. There are also other additional distinctions between these two uses, such as, for example, the ‘level’ (linguistic or extra-linguistic) at which the identification cannot be established (§5.2). However, the notion of definiteness or indefiniteness would not generally be relevant for first and second person pronouns in either of these two meanings because their functioning does not directly involve the identification of their referents. Some of the uses of first and second person pronouns appear to reflect this irrelevance of the notion of definiteness or indefiniteness for their occurrence (§2.3).

Notice, however, that this dissociation of personal pronouns from their referents is only a preferred characteristic and not an absolute necessity. It is helpful in making the pronouns function as efficient shifters. There are two main aspects of referents, namely their gender and number, which tend to get represented in personal pronouns in spite of the fact that such a representation makes the pronouns less efficient in their primary function of being shifters. In fact,

personal pronouns appear to behave like the two-faced Janus in this respect. Number marking is mainly used for denoting the number of speech roles rather than that of their referents in the case of personal pronouns. Gender marking also tends to involve distinctions among speech roles, as seen, for example, in their use for denoting social distinctions. However, both can also be used for providing information regarding the referents. Further, there is an enormous amount of cross-linguistic variation on both these aspects, which appears to represent the various types of compromises that languages have arrived at as solutions for this conflict (§2.4).

2.2 Dissociation from the referent

The tendency to dissociate personal pronouns from their referents derives from the fact that the primary function of personal pronouns is to indicate the involvement of speech roles in the event or state that the sentences in which they occur describe. In order to carry out this function satisfactorily, they need to be insensitive to any shifts that occur among their referents. The nature of communication is such that the speech roles of ‘being the speaker’ and ‘being the addressee’ (along with that of ‘being a third person’ (non-speaker/non-addressee)) shift constantly among two or more individuals as a conversation progresses. In order to indicate, consistently, the involvement of speech roles in an event or state, it is therefore necessary for personal pronouns to remain unaffected by changes that take place in their referents. Further, even when their referents remain unchanged, they need to change if the speech roles that they represent undergo change.

- (1) a. Mary: *‘I want to go home early today.’*
 b. John: *‘I want to come with you.’*
 c. Bill: *‘You have to finish your work before you go.’*

The referent of the first person pronoun *I* shifts from Mary to John in its occurrence in (1b) but the pronoun remains unchanged. On the other hand, the referent of *I* occurring in (1a) and that of *you* occurring in (1b) are the same, namely Mary, but the speech roles that they represent are different: it is ‘being the speaker’ in the former case and ‘being the addressee’ in the latter case. This latter shift in the speech role is represented by a change in the pronoun in (1b).

There is a shift of the referent between the second person pronoun occurring in (1b) and the one occurring in (1c) (from Mary to John) whereas there is no shift of referents between the two occurrences of the same pronoun in (1c). The second person pronoun, however, has remained unchanged in all these contexts (1b–c), indicating clearly that it is insensitive either to the shift or the non-shift of its referents.

The point to be noted here is that any direct association of first and second person pronouns with information concerning their actual referents would make them less efficient in their primary function of denoting speech roles. For example, if the first person pronoun occurring in (1a) were to indicate, directly, the characteristic of its referent (Mary) being a female, it would not be possible for John to use the same pronoun in (1b). Similarly, if the second person pronoun occurring in (1b) were to indicate John's love for Mary, it would not be possible for Bill to use the same pronoun in (1c). The dissociation of personal pronouns from any of the properties of their referents makes them highly efficient as indicators of speech roles.

Jespersen (1923: 123) suggests using the term 'shifter' for personal pronouns in order to describe this property of having a meaning that differs according to the situation, but he applies it even to words like *father*, *mother*, *enemy*, etc., as these words also have a meaning that changes according to the situation in which they are used. This resemblance between personal pronouns and these common nouns derives from the fact that both of them denote the 'roles' of their referents. However, in the case of personal pronouns, the primary function is the denotation of those roles, whereas in that of common nouns the primary function is to identify their referents. Common nouns make use of the roles (or set of characteristics) only as an aid in their primary function. This is evidenced by the fact that common nouns are unlike personal pronouns in being sensitive to shifts occurring among their referents. For example, unlike personal pronouns, they can be used with modifiers and complements whose primary function is to reduce the referential scope of their head nouns by indicating some of the characteristics of their referents. They are also used with definite or indefinite articles whose function is primarily to indicate the continuation or non-continuation (or shift) of their referents. The use of both these types of devices with common nouns can be seen in (2a-c).

- (2) a. *A woman was standing with a small boy near the bus stop.*
 b. *The boy asked the woman to give some money to him.*
 c. *A boy asked the woman to give some money to him.*

The use of the modifier *small* with the noun *boy* in (2a) has the effect of reducing the scope of the noun phrase in which the noun *boy* functions as the head. It provides additional information regarding the referent of that noun, and thereby makes it easier for the addressee to identify that referent. On the other hand, the use of an indefinite article in that noun phrase and also with the noun *woman* in (2a) has the function of indicating that, in those cases, new referents are being introduced. This is also true of the noun phrase *a boy* occurring in (2c). It introduces a referent that is different from the one denoted by the noun *boy* in (2a) and (2b). In contrast to this, the use of the definite article with the noun *boy* in (2b) and the noun *woman* in (2b) and (2c) indicates that the same referents that were introduced earlier by those nouns are intended by the speaker in these cases.

Jakobson (1971) and other later writers restrict the use of the term 'shifter' to deictic words, that is, words whose meaning is more crucially dependent upon the speech context than that of common nouns. These expressions differ from common nouns in that they generally do not occur with articles whose primary function is to indicate the continuation or non-continuation of their referents. Demonstratives like *this* and *that*, adverbials like *yesterday* and *today*, and tense markers like present and past are regarded as shifters, in addition to personal pronouns, by these scholars. For example, sentences (3a–b) indicate that the demonstratives *this* and *that* are similar to personal pronouns in being insensitive to shifts or non-shifts of their referents.

- (3) a. John: *I like this table₁.*
 b. Bill: *I think this table₂ is better than that one₁.*

There is, on the other hand, an important difference between personal pronouns and these other types of expressions as far as their 'shifting' nature is concerned. The latter are dependent upon the former for the establishment of their meaning, whereas the former (especially the first person pronouns) are independent of any other expression. Further, as pointed out by Parret (1980: 97), the first person pronoun differs from other shifters ('demonstratives') in not requiring any demonstration. If any demonstration is supplied with a first person pronoun, it is either irrelevant or is meant for stylistic or rhetorical emphasis. 'The speaker refers to himself when he uses *I* and no amount of pointing to someone else in his environment can reverse this reference' (see also Cornish 1999: 54). This centrality of personal pronouns among deictics might be the reason why languages do not generally allow them to be directly associated with any of the identifying properties of their referents while other deictic expressions like demonstratives are freely associated with such properties (as can be seen, for example, in the use of expressions like *this red box*).

2.2.1 Referentiality of personal pronouns

Some scholars view this dissociation of personal pronouns from their referents as implying that they are non-referential. For example, Benveniste (1971: 218) argues that the instances of the use of *I* do not constitute a class of reference since there is no 'object' definable as *I* to which these instances can refer in identical fashion. He considers the first person pronoun (along with other 'indicators') to be an empty sign that is non-referential with respect to 'reality'. It becomes full as soon as a speaker introduces it into an instance of his discourse. He suggests, '*I* can only be identified by the instance of discourse that contains it and by that alone'. It is solely a 'reality of discourse'.

Anscombe (1975) compares the use of first person pronouns with that of proper names and finds the two to be different on several points. For example, a personal pronoun is of no use for introducing people to one another, or for calling

someone. While it might be used as a signature, one would be quite dependent on other clues to the identity of the signatory. The real difference, according to Anscombe (1975: 54), lies in the fact that each one uses the name *I* only to speak of himself. Anscombe concludes that *I* is neither a name nor another kind of expression whose logical role is to make a reference (1975: 60).

There is, however, a need to differentiate between the dissociation of person pronouns from their referents on the one hand, and the absence of any referents that may occur in the use of definite or indefinite noun phrases on the other. The latter depends upon the intention of the speaker, and is represented through formal distinctions in some languages. For example, Givón (1984: 412) points out that in Bemba (Bantu), nouns are marked for the referential/non-referential distinction. They take a nominal prefix of the form VCV—in order to denote referentiality and a prefix of the form CV—to denote non-referentiality.

- (4) a. *a-à-fwaaya* *ici-tabo*
 he-Past-want Ref-book'
 'He wanted a (specific) book/the book'
 b. *a-à-fwaaya* *ci-tabo*
 he-Past-want NRef-book
 'He wanted some (non-specific) book'

Chesterman (1991: 26) suggests that, even in English, the contrast between the use vs. non-use of articles can be regarded as denoting this basic distinction between reference vs. non-reference (see §9.4.1). However, English allows its noun phrases to be used non-referentially even when they are associated with a definite or indefinite article. This has led to the possibility of its noun phrases that contain an article to show ambiguity between referential and non-referential meanings. The occurrence of such an ambiguity in the case of a definite noun phrase can be seen in (5), which is from Donnellan (1978).

- (5) *The murderer of Smith is insane.*

Donnellan points out that in (5), the definite noun phrase *the murderer of Smith* can have either (i) a referential use, in which case it would have been used by the speaker to refer to a specific person who is known to have murdered Smith, or (ii) an attributive (non-referential) use, in which case (5) would be a statement about whoever it is that has murdered Smith (with the speaker having no specific person in mind).

There can be a similar distinction in the use of first and second person pronouns as well, which would also depend upon the intention of the speaker. For example, while describing how a particular activity is to be performed, a speaker may use a first or second person plural pronoun without actually associating it with a set of individuals.

- (6) a. *We heat the milk first and then add the dough to it.*
 b. *You heat the milk first and then add the dough to it.*

The pronoun *we* of (6a) and *you* of (6b) may not be intended to actually refer to any specific individuals as such but they would still indicate different speech roles in the sense that they associate the description with a speaker's or addressee's viewpoint.

The characteristic of personal pronouns mentioned earlier (dissociation from their referents) differs from this particular characteristic of non-referentiality in that the former derives from the function of personal pronouns in language and not from the intention of the speaker. It derives from the fact that denoting speech roles rather than the individuals who enact those roles is the primary function of first and second person pronouns. If these pronouns are burdened with information that is specific to the individuals who enact those roles, they would fail in their primary function of being the same for all speakers or addressees. They would no more be able to function as 'shifters' (see §2.2).

2.2.2 *File cards for referents*

There is an interesting suggestion, made by some linguists (see Du Bois 1980, Reinhart 1981, Heim 1988), that the notion of reference can be described in terms of a system of file cards. Understanding a discourse is like keeping a file in which each discourse referent is represented by a numbered file card. Whenever a new discourse referent is introduced, a new card is added to the file. On the other hand, whenever something is said about a discourse referent that has been introduced earlier in the discourse, it is entered in a file card that already exists in the file.

Erteschik-Shir (1997: 18) suggests that file cards for the speaker (first person) and the hearer (second person) exist permanently on top of the file and are ready to be used as topics. This would, however, be feasible only if the conversation takes place between two persons who are familiar with one another, but not in contexts in which it takes place among a group, as, for example, in a party or a conference. One will have to pick out the relevant 'cards' from the (memory) file and place them on top as different known persons start speaking, or insert new cards, as, for example, when unknown or unfamiliar persons introduce themselves or are introduced by others as they start speaking.

What is interesting in the present context is that these file cards do not identify the speaker as a 'speaker' or the addressee as an 'addressee', but rather as a specific individual through his or her own personal name or description. Consider, for example, the utterances made by the speaker A to his addressee B, as described by Erteschik-Shir (1997: 20).

(7) *I have a dog. It is brown.*

Erteschik-Shir (1997: 20) represents (7) with the help of two file cards of which one is headed by A_1 and the other one by dog_2 . The first card mentions ' A_1 has a dog' and the second one mentions 'The dog_2 (that A_1 has) is brown'. Notice that the first file card is not headed by the first person pronoun *I* or by the word

'speaker', but rather by a marker that identifies a specific individual as A_1 . That is, the file cards do not appear to use any information that is provided by personal pronouns. As far as the referential identity of the speaker and the addressee are concerned, they depend upon information that is provided by the speech context or by noun phrases occurring elsewhere in the discourse. Personal pronouns are completely unhelpful in preparing these file cards.

On the other hand, there would be a need to make additional distinctions in these file cards, concerning factors such as the *source* of information that forms the basis of a given instance of identification. They may have to indicate whether the identity of a referent that a given file card provides is only 'linguistic', based upon someone's statement, or extra-linguistic, based upon the perception or conviction of the person who 'keeps' those file cards (see §9.2.1). There are some languages that use distinct sets of personal pronouns, called 'logophoric', for indicating that the responsibility for the truth of a statement is only with the person whose statement is being reported and not with the reporter himself (Sterling 1993: 266). Some languages indicate this distinction through the use of subjunctive or irrealis mood in the case of such reports (see §3.1 for details).

Thus the relevance of personal pronouns, as far as these file cards that keep a record of referents is concerned, is restricted to the denotation of the *reliability* of their identification. They have no relevance as far as the actual identification of those referents.

2.2.3 Use of alternative devices

Having the primary function of being 'independent' shifters, I believe, is the main reason why first and second person pronouns are distanced from their referents. There are clearly other factors that make it possible for them to remain unaffected by changes occurring in their referents. For example, in a prototypical speech context, the speaker and the addressee would be fully familiar with one another and hence there would be no need for either of them to use identifying expressions while referring to himself or herself as the speaker or to the other person as the addressee. However, this could not be the *reason* why personal pronouns are dissociated from the characteristics of their referents.

This point is supported by the occurrence of different types of non-prototypical contexts in which there is a need to associate personal pronouns with expressions that identify their referents. For example, when a person cannot see the speaker of an utterance, nor recognize his voice, he will have to ask the speaker to identify himself through the use of his personal name or a description. This would happen when a person is trying to contact someone by telephone or when he is speaking to a person standing on the other side of the door. Merely using a personal pronoun, as, for example, saying *It's me*, would be completely unhelpful in such contexts. However, languages do not allow identifying expressions to be directly associated with personal pronouns even in such non-prototypical

contexts. Instead, they generally make use of alternative devices with the help of which the necessary identifying information can be provided indirectly, and thereby the dissociation of personal pronouns from their referents can be maintained (see below (10)).

When sentences that contain first and second person pronouns are used in written texts like novels and dramas, it would be necessary to specify the speaker and the addressee with the help of personal names or other descriptive expressions. This is especially true of contexts in which the identity of the speaker and the addressee shifts among three or more individuals as the reported conversation progresses (see (1a–c)). The device generally used in such contexts is to indicate the identity of the speaker and the addressee through matrix clauses as in (8a). When the conversation is only between two individuals, it may not be necessary to specify the referents in each case, but still, there would be a need to indicate shifts that occur among speech role performers through the use of distinct paragraphs and distinct sets of quotation marks as in (8b–c).

- (8) a. *'Shall we go to the seashore?' John asked Mary.*
- b. *'Sorry John, I have a headache.'*
- c. *'If we go there, perhaps your headache will go away.'*

Languages use such devices even in the spoken form in order to indicate the identity of the speaker or the addressee indirectly, while maintaining the dissociation of personal pronouns from their referents. For example, while taking an oath, it is considered to be necessary to identify the person who is taking the oath in the oath itself. In order to satisfy this requirement, the personal name of the oath-taker is placed next to the first person pronoun.

- (9) a. *'I will not enter the house of Mary Smith,' declared John Smith.*
- b. *'I, John Smith, hereby declare that I will not enter the house of Mary Smith.'*

Notice that the first person pronoun in (9a) merely indicates that the oath-taker is the speaker of (9a). The proper noun, *John Smith*, which identifies the oath-taker by naming him, occurs in the matrix sentence. In (9b), which is a proper oath, it is placed next to the pronoun and made part of that oath. However, the proper name occurs only in apposition to the first person pronoun; the two do not form a single phrase and hence the dissociation of that personal pronoun from its referent is maintained.

A speaker may use this device for identifying himself in a telephone conversation, as shown in the Kannada sentence (10). This device is used quite frequently by Kannada speakers while telephoning a stranger who would not be in a position to recognize the speaker by his voice.

- (10) *na:nu ra:ju ma:tana:duvudu*
 1Sg Raju speaking
 'This is Raju speaking'

Another situation of the above type arises in the use of the plural forms of personal pronouns. When a speaker uses one of these forms, his addressee might not be able to identify the persons, other than the speaker or himself, who might have been intended by the speaker as its referent. In such contexts, the speaker might specify the intended individuals by placing identifying noun phrases in apposition to personal pronouns.

- (11) a. *We teachers must resolve this problem.*
 b. *You students need not wait until four.*

It is also possible to use this construction while referring to one (or a few) of the persons that belong to a group, instead of referring to the whole group, with the help of the second person pronoun. The noun phrase that occurs in apposition to the pronoun would help the hearers to identify the person or persons to whom the message or order is intended. The sentences (12a–b) of Havyaka Kannada (Dravidian), exemplify this usage.

- (12) a. *ni:nu a: mu:le-li ku:dugonḍ-ipp-o:nu hera ho:gu*
 2Sg that corner-Loc sitting-be-3MSg outside go(2Sg)
 ‘(The person among) you, who is sitting in the corner, go outside’
 b. *ningo mu:ru jana itla:gi banni*
 2Pl three person this.side come(2Pl)
 ‘Three of you come to this side’

According to Roberts (1987: 210), pronoun–noun constructions of the above type are used in Amele (a language of Gur family) for giving clarification as to the identity of the pronominal element. Roberts considers the nominal or noun phrase to be occurring in apposition to the pronoun in such constructions. It is separated from the pronoun by a slight pause and has its own intonational peak. The appositive nominal can itself have a further pronoun following it.

Erromangan, a Southern Vanuatu language, uses appositive constructions of the above type that contain personal pronouns only in the written register but not in speech. According to Crowley (1998: 184), a noun referring to a person’s position in a society or kinship relationship with the addressee can be apposed to a pronoun referring either to the writer of the message or to the recipient.

- (13) *kik tavsogi ra nur igko*
 2Sg teacher Possessive place here
 ‘you, the teacher of this place’

This constraint apparently results from the fact that such appositive uses of noun phrases are more relevant in written register than in the spoken register.

According to Haiman (1980: 226), first and second person pronouns cannot be used in Hua (Gorokan, Papuan) in appositive constructions of the type *we linguists* with common nouns, but in the case of kin terms and personal names, their

use in apposition is obligatory. They follow the possessive suffix, if there is one as shown in (14).

- (14) *nomo* *dima-ga*
 maternal.uncle your-I
 ‘I, your maternal uncle’

Notice that (14) consists of the possessive pronoun *dima* ‘your’ and appositive pronoun *ga* ‘first person’. If personal pronouns are regarded as denoting speech roles, kin terms and person names that occur in apposition with them can be considered as expressions that identify the persons who enact those roles. The need to identify this particular type of persons may have contextual or cultural significance.

Another strategy used by some languages for specifying a person out of a group of persons while using a second person singular pronoun is the use of the vocative. In Havyaka (Kannada, Dravidian), for example, a noun in the vocative case is placed in front of the pronoun rather than after it, as can be seen in the sentence (15).

- (15) *appā, ni:nu manege batt-ey-o:*
 father, 2Sg home come-2Sg-Int
 ‘Father! Are you coming home?’

The noun used as a vocative is separated from the rest of the sentence by intonation. Havyaka also allows the second person pronoun to be left unspecified in such situations, but the verb continues to show agreement with that pronoun.

- (16) *appā, manege batt-ey-o:*
 father home come-2Sg-Int
 ‘Father! Are (you) coming home?’

Blake (1994: 9) points out that vocatives do not appear as dependents in constructions, but rather stand outside constructions and are inserted parenthetically. They are unlike other cases in that they do not mark the relation of dependents to heads. Even languages like Yapese (Austronesian), which do not morphologically mark their nouns for case, are reported to have special addressee (vocative) forms for personal names. That is, the marking of the vocative appears to be unrelated to the marking of nouns for case. This puzzling aspect of the vocative can be accounted for by the claim that their main function is to provide referential identity to second person pronouns which, by themselves, are incapable of identifying their referents.

What is interesting about these constructions is that they very clearly indicate the general tendency of languages to avoid directly associating personal pronouns with identifying information regarding their referents. Whenever there is a need to provide such information, languages appear to use devices that keep personal pronouns dissociated from it.

2.3 Notion of definiteness

As I have mentioned earlier, linguists use the terms ‘definite’ and ‘indefinite’ in two different meanings that are quite distinct from one another. When used with reference to noun phrases or the articles that occur in them, the notion of definiteness depends upon the identifiability of a referent for the addressee, or the speaker’s assumption regarding his ability to do so. On the other hand, when used with reference to proforms like demonstratives or indefinite-interrogatives, it depends upon the identifiability of a referent for the speaker himself. For example, an interrogative pronoun is considered to be ‘indefinite’ in the sense that the speaker himself is incapable of identifying its referent. In fact, the speaker would be assuming that his addressee would be in a position to identify it (which is the reason why he puts the question to the addressee). This is clearly the *reverse* of the assumption that he holds while using an indefinite noun phrase.

There are several additional differences between the two notions of definiteness and indefiniteness. For example, the definiteness of noun phrases is only ‘verbal’ or ‘linguistic’; the addressee concedes it readily the moment a referring expression is used in an utterance. On the other hand, the definiteness of a proform is more than ‘verbal’; it needs to be established on a firmer footing. An addressee may persist in viewing a referent as ‘indefinite’, and may continue to use an indefinite or interrogative pronoun for referring to it, so long as its identity at the relevant non-verbal or extra-linguistic level is not established to his satisfaction. That kind of ‘persistence’ is generally unavailable in the case of the indefiniteness of noun phrases (see §9.2.1).

The characteristic that is common to both these varieties of definiteness (or indefiniteness) is that they are concerned with the identifiability (or non-identifiability) of a referent. First and second person pronouns are unconcerned with this factor, because they need to be insensitive to any shifts or non-shifts that occur among their referents in order to function satisfactorily as shifters. They are generally distanced from any identifying properties of their referents, as we have seen in the previous section. This unique characteristic of personal pronouns makes it rather difficult to associate them directly with either of the above-mentioned two types of definiteness or indefiniteness.

Linguists generally regard personal pronouns to be ‘inherently’ definite. This is a long-standing tradition among grammarians. As pointed out by C. Lyons (1999: 26), personal pronouns are often referred to as ‘definite pronouns’ in contrast to indefinite pronouns like *one* and *someone*. However, justifying this assumption in the case of first and second person pronouns has turned out to be rather difficult because there are certain uses of these pronouns that clearly contradict this assumption. The situation appears to be rather complex, involving a dichotomous behaviour.

2.3.1 *Evidence for definiteness*

Languages like English, which use a definite or an indefinite article obligatorily with nouns that have a specific reference, generally do not allow any article to occur with first and second person pronouns. In order to account for this constraint, it has been claimed that personal pronouns are similar to proper names (which also do not generally occur with articles) in being ‘inherently’ definite. Postal (1970) adduces certain usages of first and second person pronouns in order to support this claim.

- (17) a. *Big as the boy was, he couldn't lift it.*
 b. **Big as a boy was, he couldn't lift it.*
 c. *Big as I am, I couldn't lift it.*
- (18) a. *The best of the men here*
 b. **The best of some men*
 c. *The best of us*
- (19) a. *Which of the men*
 b. **Which of some men*
 c. *Which of us*

Notice that the first person pronoun resembles definite noun phrases rather than indefinite ones in (17–19) and hence one can assume that they are definite.

On the other hand, as I have suggested above, there is also the possibility of deriving this non-occurrence of definite and indefinite articles with personal pronouns from the fact that they are insensitive to shifts or non-shifts that occur among their referents. In fact, they need to be insensitive to such shifts in order to function as shifters, and the use of articles with them, whose primary function is to indicate the shift or non-shift of referents, could conflict with their function as shifters.

2.3.2 *Evidence against definiteness*

There are also certain other factors that appear to go against the claim that personal pronouns are definite. For example, C. Lyons (1999: 312) refers to the occurrence of an ‘intriguing agreement phenomenon’ in English, involving first person plural pronoun and a possessive or a reflexive pronoun.

- (20) a. *Some of us like our beer chilled.*
 b. *Some of us like their beer chilled.*
- (21) a. *Some of us disgraced ourselves last night.*
 b. *Some of us disgraced themselves last night.*

As pointed out by Jespersen (1943: 128), the expressions *some of us* and *most of us* are, strictly speaking, in the third person, but it is excusable if the

speaker includes himself and uses the first person plural in such contexts. He describes this ‘indefinite’ use of first person pronoun as involving a ‘plural of approximation’. There is actually a meaning distinction between these two usages, as pointed out by C. Lyons (1999: 312): The speaker is included in the (a) sentences and excluded in the (b) sentences. Lyons suggests that in both these cases *some* denotes an ‘indefinite’ subset, in spite of the inclusion of the speaker in the (a) sentences. What is problematic here is that the first person plural pronoun *our* has indefinite reference that is identical with that of the phrase *some of us*.

The pronoun *our* follows the indefinite noun phrase in English, but in Havyaka (one of the coastal dialects of Kannada, Dravidian) it *precedes* such a noun phrase (see (22a) given below). Further, it can be followed by indefinite as well as definite noun phrases, which provides the basis for claiming that the plural form of first person pronoun can function as having either definite or indefinite reference. That is, the pronoun can represent either a definite or an indefinite set in which the speaker is included, and the distinction between these two possibilities is represented by the noun phrase that occurs in an appositive construction with the personal pronoun (see §2.2.3) and not by the pronoun itself. This is true of the second person plural pronoun as well.

- (22) a. *eṅgo kelavu jana pe:ṭe-ge ho:vutt-eyō*
 1Pl(Excl) some person market-Dat go-1Pl (Excl)
 ‘Some of us are going to the market’
- b. *niṅgo kelavu jana pe:ṭe-ge ho:g-i*
 2Pl some person market-Dat go-2Pl(Imperative)
 ‘Some of you go to the market’
- (23) a. *eṅgo i: na:lku jana pe:ṭe-ge ho:vutt-eyō*
 1Pl(Excl) this four person marker-Dat go-1Pl (Excl)
 ‘*These four of us are going to the market’
- b. *niṅgo i: na:lku jana pe:ṭe-ge ho:g-i*
 2Pl this four person market-Dat go-2Pl (Imperative)
 ‘*These three of you go to the market’

Notice that the plural forms of first and second person pronouns, *eṅgo* ‘we (Pl. Excl)’ and *niṅgo* ‘you (Pl)’ are in apposition with an indefinite noun phrase in (22a–b) and with a definite noun phrase in (23a–b). Since these noun phrases identify the referent of the pronoun that occurs before them, the pronoun itself may have to be regarded as ‘indefinite’ in (22a–b) and ‘definite’ in (23a–b).

The sentences given in (22a–b) also appear to go against the claim of C. Lyons (1999: 315), namely that an indefinite plural that includes a person is never directly encoded in a noun phrase.

2.3.3 *Personal pronouns as determiners*

The non-occurrence of definite article with personal pronouns, in spite of their apparent definiteness, has led to the claim that personal pronouns are not words that stand for nouns, but rather are determiners or articles that stand adjacent to (specified or unspecified) nouns. This claim was originally put forth by Postal (1970), but since then, it has been conceded by several linguists (see, for example, Radford 1993: 109, C. Lyons 1999: 303).

The occurrence of expressions like *we linguists*, in which a personal pronoun is followed by a noun or a noun phrase, is considered by some linguists to support the above-mentioned claim (see C. Lyons 1999: 141). Personal pronouns are considered to occur as determiners in such constructions. On the other hand, linguists like Rigter (1980), and Diessel (1999: 67) argue that the constructions involve apposition and not modification or determination. Further, the motivation for the use of this construction, as pointed out in §2.2.3, makes it evident, I believe, that the personal pronouns do not determine the referential scope of the associated nominals; rather, they occur as shifters that cannot themselves identify the individuals concerned, and the appositive noun phrases are meant for carrying out this latter function.

Certain languages like English show a constraint in their use of this construction, namely that only the plural forms of first and second person pronouns can occur in that construction. The singular form of second person pronoun can occur in exclamatory contexts like *You idiot!*, which is probably an idiomatic shortening of a sentence like *You are an idiot!*, but the first person singular form cannot occur even in such contexts. On the other hand, languages like German allow both singular as well as plural pronouns to be associated with such noun phrases (C. Lyons 1999: 141). It would be rather difficult to account for this cross-linguistic difference in the range of occurrence of this construction if we are to claim that the pronouns function as determiners.

According to Davies (1981: 157), in Kobon (Kalam family of Papua New Guinea), such pronoun–noun constructions are preferred to a pronoun alone in the case of second person dual and plural forms and are common in the case of first person dual and plural forms. On the other hand, according to Hagman (1977: 43), Nama Hottentot (Central Khoisan Group) makes it obligatory (i) to use first and second person pronouns and (ii) not to use third person pronouns, when noun phrases are inflected for person, number, and gender.

- (24) a. *tií kxòe-ta*
 I person-1Sg
 ‘I, the person’
 b. *saá kxòe-ts*
 you person-2MSg
 ‘You, the male person’

These and other similar constraints concerning the use of this construction (such as, for example, its restriction in the written form in Erromangan (see §2.2.3)) can be accounted for on the basis of the claim that the noun phrases occurring in them are meant for establishing the referential identity of the pronoun. For example, the addressee would generally have difficulty in identifying the participants only when dual or plural forms of these personal pronouns are used. Languages like English restrict the use of identifying noun phrases to such potentially ambiguous contexts, whereas others like German extend it to non-ambiguous contexts as well.

C. Lyons (1999: 144) points out that such constructions are not possible in Spanish even in the case of first and second person plural pronouns. However, Spanish uses an alternative device for denoting the identity of such pronominal referents: it places the noun phrase in the subject position but indicates the relevant speech role with the help of the agreement markers occurring in the verb. As in the case of English, this possibility is available only in the case of plural reference (C. Lyons 1999: 144).

- (25) a. *Los estudiantes trabajan mucho*
 the students work-3Pl much
 ‘The students work hard’
 b. *Los estudiantes trabajamos mucho*
 the students work-1Pl much
 ‘We students work hard’
 c. *Los estudiantes trabajáis mucho*
 the students work-2Pl much
 ‘You students work hard’

According to Saltarelli (1988: 210), pronoun–noun combinations of the above type, occurring in Basque, are to be regarded as appositive because both the pronoun and the noun must bear the same case marking. On the other hand, when determiners are used with a noun in a noun phrase, the case markers occur only with the determiner that follows the noun. The pronouns also differ from determiners in preceding the noun rather than following it.

- (26) a. *hain emakume lodi-ek*
 so woman fat-Pl.Erg
 ‘such fat women’
 b. *liburu berri hori-ek oso interesgarriak dira*
 book new those-Pl.Erg very interesting are
 ‘Those books are very interesting’

In the Havyaka dialect of Kannada (Dravidian) also, both the pronoun as well as the noun that occur in such constructions show case markers (27a). On the

other hand, determiners or adjectives that precede a noun in a noun phrase cannot take any case markers (29b).

- (27) a. *niŋgo-ge makko-ge mada:lu baɫsutte*
 you-Dat children-Dat first serve
 ‘I will serve (food) to you children first’
 b. *saɲɲa makko-ge mada:lu baɫsutte*
 small children-Dat first serve
 ‘I will serve (food) to small children first’

Another aspect of this construction, which also appears to go against the claim that the pronouns function as determiners, is the possibility of the noun phrase occurring with determiners in such constructions. In Hausa (Chadic), for example, the noun occurring in constructions of the above type can be followed by a definite article or a demonstrative pronoun (Newman 2000: 155).

- (28) a. *mū mālāman-nān*
 we teacher-these
 ‘we teachers’
 b. *sū mutānē-n*
 they men-the
 ‘they, the men’

In Mupun (Chadic), on the other hand, the use of the definite article after a noun occurring in such constructions is obligatory. According to Frajzyngier (1993: 90), this definite article can only be anaphoric. The noun phrase to which it is attached must have an antecedent occurring either in the same sentence or in one of the preceding sentences. That is, the occurrence of an article obligatorily after the noun or noun phrase in such constructions implies that the noun phrase is anaphoric. This supports the assumption that the noun or noun phrase is appositional to the pronoun, and its referent is established independently by that pronoun. The noun phrase only provides information for identifying that referent.

2.4 Association with the nominal category

Personal pronouns are generally considered to belong to the nominal category. Their ability to occur in the argument position of sentences and their ability to take gender, number, and case markers are considered to form the basis of their inclusion in the nominal category. However, as I will be pointing out in the fourth chapter, the association of first and second person pronouns with these categories (especially gender and number) is for entirely different purposes as compared to the association of other nominals with them. It cannot therefore be regarded as supporting the inclusion of personal pronouns in the nominal category.

On the other hand, there are several other characteristics shown by personal pronouns that are generally considered to go against their inclusion in the nominal category. For example, personal pronouns do not normally take modifiers and complements. They also do not occur with determiners like definite or indefinite articles or demonstratives. These constraints clearly derive from the functional irrelevance of modifiers, complements, and determiners to first and second person pronouns (§2.4.1). Similarly, the association of personal pronouns with numerals is rather different from the association of proforms or nouns with them. In the former case numerals function as appositive phrases that do not directly modify the pronouns. They provide additional information regarding the identity (in terms of number) of their referents, but since personal pronouns are not directly concerned with the identity of their referents, their association with numerals can only be indirect (§2.4.2).

As I have pointed out in the previous chapter (§1.5), a number of languages prevalent in South Asia are reported to replace their personal pronouns with sets of expressions that are of nominal origin, like kinship terms, religious titles, deprecatory terms for oneself, or honorary terms for the addressee, and so on. The occurrence of nominals in such replacements might be regarded as indicating the ‘nominal’ character of personal pronouns. Notice, however, that the nominals tend to lose their nominal character and to take on, gradually, the characteristics of personal pronouns in such uses. This interesting characteristic of these replaced nominals appears to derive from the primary function of personal pronouns.

2.4.1 Inability to take modifiers and complements

Most grammarians have recognized the non-occurrence of any modifiers or complements with personal pronouns as one of the important and distinguishing characteristics of personal pronouns. We can consider this characteristic also as resulting from the primary function of personal pronouns. Notice that in the case of common nouns, modifiers and complements have the function of restricting the scope of reference so as to make it easier for the addressee to identify the referent. They associate additional information with the noun to which they are attached and thereby make the description of the referent richer.

- (29) a. *Give me a stick.*
 b. *Give me a long, pointed stick.*
 c. *Give me a stick that won't break easily.*

As more and more modifiers and complements are attached to a noun, it becomes more and more easy for the addressee to identify the kind of object that the speaker has in mind while using the expression.

Pronouns of first and second person are not generally associated with modifiers and complements because identifying the referent does not form part of their

primary function. In fact, they need to be *dissociated* from their referents in order to carry out their primary function of denoting speech roles. They need to remain unchanged as the reference shifts among different individuals (see §2.2). Their non-occurrence with modifiers and complements very clearly derives from this unique function that they have to perform. If modifiers or complements are attached to them, they would become less efficient or inefficient in their primary function.

It has been suggested by some linguists that there do occur some languages in which personal pronouns are allowed to take modifiers and complements. However, a closer examination of these languages reveals that the expressions that are attached to these pronouns are not actually modifying them but are only appositional to them. They do provide additional information about the referents of the pronouns, but they do so by staying at a distance. For example, Iraqw (Cushitic: Mous 1993: 113) allows its personal pronouns to be modified by adjectives, numerals, and relative clauses. But the function of these appears to be appositional rather than modificational. In the case of relative clauses, for example, personal pronouns must obligatorily be followed by demonstratives.

- (30) *aní-w-í* *goo'ûm* *a-ga* *hlaqáat*
 1Sg-M-Dem write 1Sg-Perf tired (1Sg)
 'I, who am writing, am tired'

It can perhaps be claimed that the pronoun denotes the speech role in these constructions and the demonstrative identifies the referent for that speech role. Similarly, Li and Thompson (1981: 134) report that in Chinese pronouns take modifiers only in certain restricted contexts, as for example, when one is appraising or evaluating oneself. They are not meant for identifying the referent.

- (31) *kělián* *de* *wǒ*
 pitiful Nom I
 'Poor me!'

Personal pronouns differ from nouns in the kind of modifying affixes or particles that they take. In the case of personal pronouns, the affixes generally involve notions like 'only', 'alone', 'emphatic', 'all of ...', 'the whole lot of ...', etc. None of them have the function of providing additional information about the referents that can be used by the addressee for identifying their referents. That is, they do not conflict with the primary requirement of personal pronouns, namely that they must be dissociated from their referents in order to denote speech roles.

Another characteristic that differentiates personal pronouns from nouns is that they do not take any determiners. Linguists have used this characteristic either to claim that they are inherently definite and therefore do not require the attachment of a definite article, or that they themselves are definite articles (see §2.3.3). It is,

on the other hand, possible to account for this constraint on the basis of their primary function in language. Definite and indefinite articles are generally used in languages in order to indicate the nature of identification that a given noun phrase performs. Since first and second person pronouns do not identify their referents, they need not be associated with such distinctions.

Languages that do attach articles with personal pronouns are found to make use of that association for entirely different purposes. For example, Mupun (Chadic) is reported to use the definite article *nə* with personal pronouns, but its function is to add emphasis to those pronouns. In the case of nouns, on the other hand, it has the function of denoting that their referents have already been introduced earlier in the discourse (Frajzyngier 1993: 171).

- (32) *wur nə*
 3M Def
 ‘he himself’

2.4.2 Association with numerals

As mentioned earlier, languages generally do not allow modifiers to occur with personal pronouns, apparently because attaching them would make personal pronouns functionally less efficient. However, most languages allow numerals to occur with them. Some grammarians consider these to be representing exceptional uses of modifiers, but actually the numerals occur as appositional phrases rather than as modifiers in these constructions. They specify the exact number of referents that the pronouns have left unspecified. They do not actually ‘constrain’ that number. This claim is supported by the fact that the position in which numerals occur in the case of personal pronouns is generally quite different from the one in which they occur in the case of nouns.

In Kannada, for example, numerals precede nouns in a noun phrase, whereas they follow personal pronouns.

- (33) a. *ibbaru huḍugaru*
 two boys
 b. *na:vu ibbaru*
 we two

In the case of noun phrases, the position before a noun is a modifying position, whereas the one after the noun is an appositional one. This is shown by the contrastive use of both numerals (34a–b) as well as adjectives (35a–b) in these two types of contexts.

- (34) a. *i: mu:ru pustaka nanage be:ku*
 this three book me want
 ‘I want three of these books’

- b. *i: pustaka mu:ru nanage be:ku*
 this book three me want
 'I want these books, which are three in number'

- (35) a. *dodḍa huḍugaru ellige ho:gidda:re*
 big boys where gone
 'Where have the big boys gone?'

- b. *huḍugaru dodḍ-avaru ellige ho:gidda:re*
 boys big-MPI where gone
 'Where have the boys, the big ones, gone?'

This is also true of English and several other languages. In Awa Pit (Barbacoan family), for example, numerals precede nouns but follow personal pronouns (Curnow 1997: 86).

2.5 Conflicting characteristics

This dissociation of personal pronouns from their referents is not an absolute necessity for their functioning. Languages appear to distance personal pronouns from their referents mainly because they can be more efficient as shifters if they are not burdened with information regarding the identity of their referents. On the other hand, the direct association of such information with personal pronouns, such as, for example, through affixes, would apparently lead to an economy of effort. There can clearly be a conflict between these two advantages, and languages appear to choose different compromise solutions.

For example, while most languages disallow gender distinctions to be marked on first and second person pronouns, some allow them to be shown in the case of non-singular forms, and some allow them even in the case of the singular form in second person. Similarly, some languages use number distinction in the case of first and second person pronouns only for denoting speech role distinctions, whereas several other languages extend their use for denoting the number of speech act participants (referents) as well. The existence of an enormous amount of cross-linguistic variation in the occurrence of both these categories with personal pronouns indicates clearly that languages face a conflict while associating them with personal pronouns. I discuss these points in detail in a following chapter (see §4.2–4).

Personal pronouns show similar kind of conflict in their association with the notion of definiteness and indefiniteness as well, as we have seen in the previous section (§2.3). Their dissociation from their referents makes it rather difficult to view them as definite or indefinite, whereas the fact that they have readily identifiable referents makes it possible for them to behave

like ‘inherently’ definite expressions. The characteristics that personal pronouns manifest in this regard clearly indicate this dichotomous nature of their association with the notion of definiteness (see §5.3 for details).

2.6 Summary

Languages generally show a tendency to dissociate first and second person pronouns from their referents. This tendency derives from their primary function of occurring consistently as shifters, that is, as expressions that denote speech roles irrespective of the actual participants who perform those roles. Generally, the identity of speech act participants would be readily observable, and hence there would be no need to indicate that identity, but there do occur some contexts in which a speaker would have to provide some identifying information either about himself or about his addressee. Languages generally possess alternative devices like the use of appositive noun phrases, vocative forms, matrix sentences, etc. which allow personal pronouns to remain dissociated from their referents even in such non-prototypical contexts.

Personal pronouns are generally considered to be ‘definite’, but the fact that they are generally distanced from any identifying information regarding their referents makes it rather difficult to associate them with the notion of definiteness. Notice that the distinction between definiteness and indefiniteness depends primarily upon the identifiability of the referent of an expression.

3 Coreference and Non-Coreference

3.1 Logophoric non-coreference

Pronouns of first and second person differ from those of third person in the way in which their use in language gives rise to problems of coreference. In the case of first and second person pronouns, problems arise primarily in the so-called logophoric contexts, that is, contexts in which a speaker depends upon some other person or persons as the source for his statements. In such contexts, there can be a conflict between a first person pronoun used (in that statement) by the speaker of the source statement on the one hand, and by the current speaker (reporter) on the other. Similarly, in the case of second person pronouns, there can be a conflict between the pronoun used for referring to the source addressee on the one hand and the present addressee (addressee of the report) on the other.

We may regard this as a conflict between endophoric and exophoric speech contexts, that is, between speech contexts that occur ‘inside’ linguistic expressions and the ones that occur outside. In the case of written texts, the problem of differentiating between these two types of speech contexts can generally be resolved by placing the clause that is being ascribed to some other person inside inverted commas (as direct speech). In the case of the spoken form, on the other hand, there might not be any such device for differentiation, and hence ambiguities may arise. This point can be exemplified with the help of the sentences (1a–b) of Havyaka, one of the coastal dialects of Kannada (Dravidian), which does not differentiate between direct and indirect speech.

- (1) a. *en-na ello:ru-de hogaḷuttavu he:ḷi ra:ju enna-tre*
me-Acc all-Emph praise that Raju me-with
he:ḷiddā
tell (Perfect)
(i) ‘Raju₁ has told me₂, “Everybody praises me₁”’
(ii) ‘Raju₁ has told me₂ that everybody praises me₂’
- b. *ni:nu be:ga bar-ekku he:ḷi avu hariya-tre he:ḷiddavu*
you early come-must that they Hari-with tell (Perfect)
(i) ‘They have told Hari₁, “You₁ must come early”’
(ii) ‘They have told Hari₁ (asked him₁ to tell you₂) that you₂ must go early’

In (1a), the referent of *enna* ‘my’ may be Raju, the speaker of the embedded clause, or the speaker of the whole sentence (i.e. the reporter of that clause). Similarly, in (1b), the referent of *ni:nu* ‘you’ can be Hari, the addressee of the

embedded clause, or, alternatively, it can be the addressee of the whole sentence (addressee of the reporter).

Languages like Havyaka Kannada use the same set of pronouns for denoting endophoric as well as exophoric speech act participants, whereas languages like English differentiate between the two by using their third person pronouns for denoting the endophoric speech act participants. While the former alternative gives rise to ambiguity between endophoric and exophoric speech act participants as shown in (1a) and (1b), the latter gives rise to a different type of ambiguity, namely between an endophoric speech act participant and a non-participant (of a speech act).

- (2) a. *John₁ told me that he_{1,2} had killed a tiger.*
 b. *I told John₁ that he_{1,2} had killed a tiger.*

The third person pronoun *he* occurring in (2a) can be coreferential either with the endophoric speaker, John, or any other third person (exophoric) referent. Similarly, the same pronoun (*he*) in (2b) can be coreferential either with the endophoric addressee, John, or any other third person referent.

A number of West African languages are reported to use a distinct set of pronouns, called 'logophoric' pronouns, in order to differentiate between the original speaker (or the source) of information on the one hand, and the current speaker (reporter of that information) on the other. These languages are able to remove both the ambiguities mentioned above through the use of such pronouns. Most of these languages use only a single logophoric pronoun that differentiates between endophoric and exophoric speakers. Donno so (Dogon: Culy 1994) is a language of this type. Some languages, on the other hand, have an additional logophoric addressee pronoun that differentiates between endophoric and exophoric addressees. Mupun, a Chadic language (Frajzyngier 1993) is of this type. Languages may also use the same pronoun for both these purposes, as, for example, Mundani, a Grassfields Bantu language (Parker 1986).

Sentences (3a–b) from Mupun (Frajzyngier 1993: 107, 111) exemplify the occurrence of a distinction between an exophoric speaker pronoun and an endophoric (logophoric) speaker pronoun, whereas (5a) from the same language, given later, exemplifies the occurrence of an endophoric (logophoric) addressee pronoun.

- (3) a. *wu sat nə n-nas wur*
 3MSg say that 1Sg-beat 3MSg
 'He said that I beat him'
 b. *wu sat nə di nas an*
 3MSg say that Log1.MSg beat 1sg
 'He₁ said that he₁ beat me'

In (3a), the first person pronoun *n-* can only refer to the exophoric speaker (the speaker of the whole sentence), whereas in (3b), the logophoric pronoun *di* can

only refer to the endophoric speaker (the speaker of the embedded sentence). There is no ambiguity in Mupun that is comparable to the one occurring in (1a–b) of Havyaka, or to the one that is comparable to the ambiguity occurring in (2a–b) of English, because it uses a distinct set of pronouns for denoting endophoric and exophoric speakers.

When the source of information is the same as the current speaker (reporter), Mupun uses the ordinary first person pronoun as seen in (4) (Frajzyngier 1993: 111).

- (4) *n sat n-wur nə an a nga kwat*
 1Sg say to-3M Comp 1Sg be man hunt
 ‘I told him that I am a hunter’

Mupun makes a similar type of differentiation between endophoric and exophoric addressees through the use of a distinct logophoric addressee pronoun when the two are distinct from one another as can be seen in (5a) (Frajzyngier 1993: 113).

- (5) a. *n-sat n-wur nə gwar ji*
 1Sg-say to-3Sg that Log2 come
 ‘I told him₁ that he₁ should come’

We may contrast (5a) with (5b) given below in which the Log2 pronoun has been replaced by the corresponding third person pronoun.

- (5) b. *n-sat n-wur nə wur ji*
 1Sg-say to-3Sg that 3Sg come
 ‘I told him₁ that he₂ should come’

It is also possible to replace the log1 pronoun of sentence (3a), given earlier, by a third person pronoun in order to provide a disjoint third person reference as in (5c).

- (5) c. *wu sat nə wu nas an*
 3M say that 3M beat 1sg
 ‘He₁ said that he₂ beat me’

These possibilities have led to the view, generally held by linguists, namely that the logophoric pronouns denote the *coreference* between those pronouns and their third person antecedents. That is, logophoric pronouns are viewed as removing the endophoric–exophoric ambiguity that affects third person pronouns in familiar languages like English (see 2a–b). For example, in the English gloss given in (5c), the two instances of third person pronoun *he* can either be coreferential or non-coreferential. In Mupun, on the other hand, the third person pronoun *wu* occurring in (5c) can only have disjoint reference, in view of the fact that when there is coreference, it is obligatory to use a logophoric pronoun. This is also true of (5b) as compared to (5a).

However, regarding logophoric pronouns as having the function of denoting third person coreference is rather misleading because it misrepresents their basic function, which is to denote the *non-coreference* between endophoric and exophoric speech act participants. The primary purpose of their use is to remove the endophoric–exophoric ambiguity that affects the use of first and second person pronouns. Their ability to remove the endophoric–exophoric ambiguities that affect the use of third person pronouns in languages like English as seen in (5c) and (5b)), on the other hand, can only be regarded as an indirect or incidental one. Languages actually use an entirely different type of pronoun (called anaphoric pronoun) for this latter purpose, as I point out below (§3.1.2).

There are, thus, three possible ways in which languages respond to the problem of denoting the distinction between the speech roles of two different speech contexts. (i) They may use distinct first and second person pronouns (logophoric and non-logophoric) as in Mupun described above. Alternatively, (ii) they may use the same set of first and second person pronouns for denoting the speaker and the addressee of both the speech contexts as in Havyaka, or (iii) they may restrict the use of their first and second person pronouns for denoting the speech roles of the current speech context and use ordinary third person pronouns for denoting the speech roles of other speech contexts as in English.

There is a similar kind of cross-linguistic difference among languages that possess distinct logophoric pronouns in their manifestation of verbal agreement for those pronouns. Most of these languages are reported to show third person agreement for logophoric pronouns, in contrast to having first person agreement for pronouns that have exophoric speakers as their referents. This is similar to the third alternative given above. On the other hand, there are some languages that manifest first person agreement for logophoric as well as first person pronouns. This is true of Donno So, a Dogon language, as can be seen in (6), which is from Culy (1994*b*) as quoted by Huang 2000: 173).

- (6) *Oumar minnε inyemε m̃ gendεzem gi*
 Oumar field Log Poss regard.Prog.1Sg said
 ‘Oumar₁ said that (he₁) would look at his₁ field’

The third possibility of having an agreement marker that is different from both first and third persons also occurs in some languages. For example, Ngiti (Central Sudanic, Nilo-Saharan) has a logophoric pronoun *ndi*, which has an agreement marker *nd-* that is different from that of first person (*m-*), second person (*ny-*), and also of third person (zero) (Lojenga 1994: 192). This resembles the first alternative mentioned above.

Languages also manifest a similar type of cross-linguistic variation regarding the diachronic sources that they utilize for deriving logophoric pronouns. For example, Efik has its logophoric pronoun derived from third person plural pronoun, whereas Sango has it derived from first person pronoun. There are also languages like Mundani that use the reflexive pronoun (which is different from

their first and third person pronouns) as the source of their logophoric pronoun (Huang 2000: 190).

3.1.1 Third person 'logophors'

It is evident from the preceding discussion that there would be a need to have distinct logophoric pronouns only for denoting the endophoric speaker and the addressee. This is because there would be a need to differentiate only between these two endophoric speech act participants from their corresponding exophoric speech act participants. There would not be any need to have logophoric third person pronouns because third person is not generally viewed as a distinct speech act participant. However, it has been claimed that in the case of languages like Gokana (Cross-river, Niger-Congo: Hyman and Comrie 1981), Lele (East Chadic: Wiesemann 1986a: 445) and Yag Dii (Bohnhoff 1986), logophoric marking occurs in the case of all three persons. I would like to suggest that the function of these marked pronouns is to denote anaphoric coreference rather than logophoric non-coreference.

Gokana is reported to mark pronouns of all three persons as logophoric by attaching a suffix to the verb (Hyman and Comrie 1981: 20, 22–3).

- (7) a. *m̀̀n kɔ m̀̀n d̩*
 I said I fell
 'I said that I fell'
- b. *m̀̀n kɔ m̀̀n d̩-ɛ*
 I said I fell-Log
 'I said that I fell' (dispreferred to 7a)
- (8) a. *ò̀d kɔ ò̀d d̩*
 you said you fell
 'You said that you fell'
- b. *ò̀d kɔ ò̀d d̩-ɛ*
 you said you fell-Log
 'You said that you fell'
- (9) a. *à̀d kɔ à̀d d̩*
 he said he fell
 'He₁ said that he₂ fell'
- b. *à̀d kɔ à̀d d̩-ɛ*
 he said he fell-Log
 'He₁ said that he₁ fell'

There are two important differences between the use of this 'logophoric' verbal marker, on the one hand, and the logophoric speaker and addressee pronouns of Mupun, described earlier, on the other. (i) The pronominal distinction

that is represented by the marked pronouns of Gokana belongs to the current (reporting) speech context, whereas the distinction that the two logophoric pronouns represent in Mupun belongs to the reported speech context. Notice that the two logophoric pronouns of Mupun are translated by *third* person pronouns in English, whereas the three marked pronouns of Gokana are translated by pronouns that belong to three different persons in English.

This difference also occurs between Gokana and Angas (Chadic: Burquest 1986: 92), another language that is reported to have logophoric pronouns only for first and second persons. The logophoric pronouns of Angas are also translated with the help of the third person pronouns of English and not with the help of its first and second person pronouns, as can be seen in (10a–b).

- (10) a. *músa ló téné dýí mét kàsúwá*
 Musa told that Log1 go market
 ‘Musa₁ said that he₁ will go to the market’
- b. *músa ló mBulus téné gwā jè*
 Musa told Balus that Log2 come
 ‘Musa₁ told Balu₂ to come (that he₂ will come)’
- c. *músa ló mBulus téné kó jè*
 Musa told Balus that he come
 ‘Musa₁ told Balus₂ that he₃ will come’

The second difference is that, in Mupun, the source addressee is marked by the second person logophor in order to indicate that it is *different* from the current (reporter’s) addressee (as shown in (5a)). This is also true of Angas, as shown in (10b). In Gokana, on the other hand, the source addressee is marked as ‘logophor’ only if it is *the same* as the current addressee (see 8b). Hyman and Comrie (1981: 22) point out that a sentence like (11), which is comparable to sentence (5a) of Mupun (repeated below), is ungrammatical in Gokana, if the pronoun is marked for logophoricity. That is, a pronoun that denotes the source addressee cannot be marked by the second person ‘logophor’ in Gokana if it is not coreferential with the current (reporter’s) addressee.

- (11) **m̀m̀ kó nè lébàrè kɔ̀ àè d̀-è*
 I said to Lebare that he fell-Log
 ‘I said to Lebare₁ that he₁ fell’
- (5) a. *n-sat n-wur nə gwar ji*
 1Sg-say to-3Sg that Log2 come
 ‘I told him₁ that he₁ should come’

That is, the logophoric addressee pronoun is used in Mupun and Angas in order to indicate *non-coreference* between the source addressee and the current addressee, whereas in Gokana it is used in order to indicate the *coreference* between the two.

Lele (East Chadic), another language that is reported to have ‘logophoric’ pronouns for denoting all three persons, appears to be similar to Gokana rather than to Mupun and other West African languages on these points. In fact, Lele is said to have a whole series of such pronouns that are derived from ordinary subject pronouns by prefixing the indirect speech introducer *nā* (Wiesemann 1986a: 445).

(12)	Subject pronouns		Logophoric pronouns	
	Singular	Plural	Singular	Plural
1	<i>ng</i>	<i>ni</i>	<i>nāng</i>	<i>nāni</i>
2M	<i>gi</i>	<i>ngu</i>	<i>nāgi</i>	<i>nāngū</i>
2F	<i>me</i>	„	<i>nāme</i>	„
3M	<i>-di</i>	<i>-ge</i>	<i>nāy</i>	<i>nāgē</i>
3F	<i>-du</i>	„	<i>nādū</i>	..
1 + 2	<i>nga</i>		<i>nānga</i>	
1 + 2 + 3	<i>ngagu</i>		<i>nāngāngū</i>	

Wiesemann (1986a: 448) points out, however, that there is no need to resort to the participants of the original speech act for describing the occurrence of these logophoric pronouns. That is, the pronouns are similar to those of Gokana in denoting *coreference* between endophoric and exophoric speech act participants. Their main function is to indicate that the marked pronoun has an endophoric antecedent. They are quite different from the non-coreferential logophoric pronouns of Mupun and other similar languages.

Yag̃ Dii (Adamawa: Bohnhoff (1986: 112) is another language that is reported to have logophoric pronouns for all three persons. There is a whole series of pronouns, called *bi*-series, which is used in subordinate clauses in this language. However, Bohnhoff notes that the only condition regarding its antecedent is that it needs to be the subject of the matrix clause. It does not seem to be limited to contexts containing a performative verb, nor to a desiderative context; nor do pragmatic–semantic notions as source–receiver of the information seem to govern the use of the series. These pronouns are considered to be logophoric only because the subordinate clauses in which they occur may have been derived from underlying quotes. That is, the pronouns are similar to those of Gokana and Lele in marking coreference rather than non-coreference between endophoric and exophoric speech act participants.

As I will be pointing out in the next section, languages use an entirely different device, called ‘anaphoric pronoun’, for removing the endophoric–exophoric ambiguity that affects the use of third person pronouns. Gokana and Lele have apparently extended such a device to all three pronouns. It would therefore be less confusing if we regard these languages as having extended anaphoric pronouns rather than logophoric pronouns.

Another interesting point that supports this claim is that, in Gokana, this ‘logophoric’ marking is obligatory only in the third person. In the second person,

it is optional but preferred, whereas in the first person it is dispreferred (Hyman and Comrie 1981: 22–3). Similarly in Yag Dii, the first person forms of the logophoric *bi*-series are only ‘occasionally’ differentiated from ordinary first person pronouns (Bohnhoff 1986: 113).

In view of these points, the person hierarchy that has been suggested by Hyman and Comrie (1981: 33) for the likelihood of logophoric marking, given in (13) (which is referred to by Wieseemann 1986a: 450, Sterling 1993, Culy 1994, and others), cannot be correct.

- (13) *Person hierarchy for logophoric pronouns*
3rd > 2nd > 1st

The preferred pronoun for logophoric marking is the endophoric first person, which may be represented in other languages either by the exophoric first person pronoun or by the exophoric third person pronoun. Endophoric second person is rarely marked for logophoricity, and the marking of endophoric third person is apparently absent. This latter factor has probably made it possible for some languages to use the exophoric third person pronoun for representing the endophoric first person.

We cannot, however, represent these cross-linguistic variations with the help of a hierarchy in which first person occurs at the top, because in Pero, a Chadic language, there is a logophoric addressee pronoun but not a logophoric speaker pronoun. Frajzyngier (1989: 122) refers to the former as a pronoun used in reported speech.

- (14) a. *Non-reported speech*
Singular Plural
2M *kàì/kè* 2MF *cínù*
2F *cì*
- b. *Reported speech*
Singular Plural
2M *péemò* 2MF *péémè*
2F *péejè*

The pronouns shown in (14b) occur when A is reporting an utterance of B addressed to C in which a second person pronoun was used. However, in the place of the first person pronoun that was used by B, Pero uses its *third* person pronoun because it does not have any special reporting pronoun for first person. Notice that Pero is an exception to the hierarchy given in (13) as well.

Culy (1997: 856) refers to the possibility of languages having an ‘anti-logophoric’ pronoun, that is, a pronoun that *cannot* have the logophoric trigger as its antecedent. Adiukrou is reported to have such a pronoun, which is a blend of the corresponding direct discourse pronoun with the reporting subject pronoun. Culy exemplifies the use of this pronoun with the help of (15) that he quotes from Hill (1995: 97).

- (15) *li dad eke ow'n im Dabu*
 3Sg said that AntiLog went Dabou
 'He₁ said that he₂ went to Dabou'

If the language is using these marked forms for all three persons, we may perhaps have to regard them as similar to those of Gokana and Lele (with the marker being used for non-coreference inside the sentence) rather than those of Mupun and other West African languages, which denote non-coreference between the roles of two different speech acts.

3.1.2 Anaphoric pronouns

The problem of coreference that arises in the case of third person pronouns derives from an entirely different source as compared to that of first and second person pronouns. In most languages, third person pronouns can have their reference established either by the speech context, called exophoric reference, or by an expression occurring in the same sentence (or in one of the previous sentences), called endophoric reference.

- (16) a. *John₁ gave his_{1,2} book to Mary.*
 b. *John₁ plays violin well but he_{1,2} can't sing.*

Notice that the third person pronoun *his* occurring in (16a) can have its reference determined by the expression *John* occurring in the same sentence. Alternatively, however, it can also have its reference determined by the speech context, independently of the referent of the expression *John*. That is, it can either be coreferential with *John* or non-coreferential with it. This is also true of the third person pronoun *he* occurring in (16b).

In the use of first and second person pronouns, on the other hand, there is no comparable problem of coreference because two different uses of either of these pronouns would generally be obligatorily coreferential, and further, their referent would also be the same as the one that is externally determined. The only exception to this is the case of reported speech (logophoric context), involving two or more speech contexts, described in the previous section.

- (17) a. *I₁ gave my₁ book to Mary.*
 b. *You₁ play violin well, but you₁ can't sing.*

Notice that the ambiguity between endophoric and exophoric references, which affects third person pronouns as seen in (16a and b), does not affect the first and second person pronouns *me* and *you* occurring in (17a) and (17b) respectively. There are certain exceptional uses of second person pronouns that do involve non-coreference (as, for example, in the sentence *you₁ take this book and you₂ take that one*) but this is not generally perceived as giving rise to ambiguity as in the case of third person pronouns. Several languages make use of a distinct

pronoun or set of pronouns, called ‘anaphoric pronouns’ (see §3.2 for a different use of the term ‘anaphoric’), in order to differentiate between exophoric and endophoric references of third person pronouns. For example, Kannada (Dravidian) uses the pronoun *ta:nu* (plural *ta:vu*) for specifying endophoric (anaphoric) reference as distinct from exophoric reference.

- (18) a. *ra:ju avana pustakav-annu hari-ge to:risida*
 Raju₁ his book-Acc Hari-Dat showed
 ‘Raju₁ showed his_{1,2} book to Hari’
- b. *ra:ju tan-na pustakav-annu hari-ge to:risida*
 Raju Ana-Gen book-Acc Hari-Dat showed
 ‘Raju₁ showed his₁ (own) book to Hari’

In (18a), the ordinary third person pronoun *avana* ‘his’ can have its reference dependent on the noun *ra:ju* or it can have it determined independently by the speech context. On the other hand, the anaphoric pronoun *tanna* occurring in (18b) can only have its reference determined by the noun *ra:ju*. It cannot have its reference determined independently (or externally) by the speech context.

Languages differ in several ways regarding the use of their anaphoric pronouns. There are some languages in which anaphoric pronouns appear to be in no way different from third person pronouns except for the fact that they must have an antecedent occurring either in the same sentence or in one of the preceding sentences. On the other hand, there are other languages that show different kinds of constraints regarding the use of their anaphoric pronouns. In Kannada, for example, the antecedent of the anaphoric pronoun *ta:nu* must generally be present in the same (simple or complex) sentence and further, if there is an agent in the sentence, it alone can be the antecedent.

- (19) *ra:ju hariy-annu tanna maney-alli hedarisida*
 Raju Hari-Acc Anaph’s house-Loc threatened
 ‘Raju₁ threatened Hari₂ in his₁ (Raju’s) house’

In (19) the anaphoric pronoun *tanna* can only have Raju as its antecedent and not Hari or someone else. Several other South Asian languages like Marathi (Indo-Aryan) also have anaphoric pronouns whose occurrence is constrained in various ways.

In spite of the occurrence of such differences, it is evident that all these anaphoric pronouns share the primary function of differentiating between endophoric and exophoric references of *third person* pronouns. They need to be differentiated from logophoric pronouns on this point, as I show in the next section.

3.1.3 Contrast between logophoric and anaphoric pronouns

There is clearly some similarity between logophoric pronouns and anaphoric pronouns deriving from the fact that both have the function of differentiating

between endophoric and exophoric meanings, and further, both are connected with the notion of reference. However, there are also several differences between the two. The most important and crucial one among them concerns the primary function for which they are used in language. Logophoric pronouns are used primarily for differentiating between endophoric and exophoric *speech contexts*, by specifically marking the participants of the endophoric speech context as distinct from those of the exophoric speech context. As such, they are primarily concerned with the occurrence of first and second person pronouns. On the other hand, anaphoric pronouns are used primarily for differentiating between endophoric and exophoric uses of *linguistic expressions*. Since this latter distinction affects primarily third person pronouns, they are primarily concerned with the occurrence of third person pronouns and not of first and second person pronouns.

Languages that possess only one of these pronouns (anaphoric or logophoric) may appear to be using that pronoun in the function of the other pronoun in contexts in which the two overlap. For example, Kannada (Dravidian) has an anaphoric pronoun but not any logophoric pronoun. It can, however, remove the ambiguity between endophoric and exophoric speakers, through the use of its anaphoric pronoun in logophoric contexts.

- (20) a. *ra:ju nana-ge ondu huli ka:ṇis-idey-endu he:ḷida*
 Raju me-Dat one tiger appear-is-that said
 ‘Raju₁ said that I have seen a tiger/he₁ has seen a tiger’

- b. *ra:ju tan-age ondu huli ka:ṇis-idey-endu he:ḷida*
 Raju Ana-Dat one tiger appear-is-that said
 ‘Raju₁ said that he₁ has seen a tiger’

The use of first person pronoun in (20a) is ambiguous between endophoric and exophoric speakers, whereas the use of an anaphoric pronoun in (20b) has the effect of making the sentence unambiguous. However, the primary function of the anaphoric pronoun *tan-* ‘self’ occurring in (20b) is to remove a different kind of ambiguity, namely the one seen in the English translation of (20b), between endophoric and exophoric uses of the ordinary third person pronoun *he*. The removal of the former ambiguity is only the indirect effect of its use. This latter point is supported by the occurrence of contexts (21–2) in which Kannada allows only first and second person pronouns to be used in complement clauses. These pronouns cannot be replaced by the anaphoric pronoun *tan-* ‘self’ in spite of the fact that the above-mentioned ambiguity between endophoric and exophoric speakers or addressees affects the first and second person pronouns occurring in them.

- (21) *na:nu pa:s-a:gidd-e:ne-endu ra:ju nan-na hattira he:ḷid-a*
 1Sg passed-have-1Sg-that Raju 1Sg-Gen with said-3Sg
 ‘Raju₁ has told me that he₁ has/I have passed’

- (22) *ni:vu be:ga bara-be:ku anta ra:ju he:l-idd-a:ne*
 2Sg early come-must that Raju told-has-3Sg
 ‘Raju has told me that I/you must come early’

Notice that in (21) the first person pronoun *na:nu* ‘I’ can be coreferential either with the speaker of the reported sentence, namely Raju (endophoric), or with the speaker of the whole sentence (exophoric). But one cannot use the anaphoric pronoun *ta:nu* in place of *na:nu* in order to remove that ambiguity. This is also true of the second person pronoun *ni:nu* ‘you (Sg)’ occurring in (22). It can be coreferential either with the addressee of the embedded sentence (which is the speaker of the whole sentence) or with the addressee of the whole sentence, and yet the anaphoric pronoun cannot be used in its place.

Languages may also extend the use of their anaphoric pronouns to logophoric contexts. For example, the anaphoric pronouns (also called ‘long-distance’ reflexive pronouns) of some of the East Asian languages like Chinese generally require their antecedents to be the subject of the matrix clause. However, in the case of logophoric contexts (i.e. contexts that involve the ‘source’ of the proposition or the ‘experiencer’ of the mental state that is being described), they may also allow a non-subject antecedent (Huang 2000: 192).

- (23) *ta ting tongshi shuo ziji tishang le jiaoshou*
 3Sg hear colleague say self promote Particle professor
 ‘He₁ hears from the colleague₂ that he_{1,2} has been promoted to a professor’

On the other hand, languages that possess only logophoric pronouns but not any anaphoric pronouns may extend the use of their logophoric pronouns to non-logophoric contexts in which a distinction between endophoric and exophoric speech roles does not actually exist. For example, Babungo, a Grassfields Bantu language, uses its logophoric pronoun even in the case of embedded sentences in which the pronoun cannot have an endophoric speaker (or a source of information) as its antecedent (Schaub 1985: 111).

- (24) a. *ŋwó gí lāa yì táa jwí*
 he say that Log Fut come
 ‘He₁ said that he₁ would come’
 b. *lámíbí jwì lāa yì jwì ndì fá*
 Lambi₁ come that Log₁ come take thing
 ‘Lambi has come in order to collect something’

Notice that in (24a) the logophoric pronoun *yì* has been used for specifying an endophoric speaker, whereas in (24b), it has no such function. It only has the function of denoting anaphoric coreference, and thereby removing the ambiguity between endophoric and exophoric third person referents.

It may be possible to regard some of the languages that have *extended* the use of their anaphoric pronouns to logophoric contexts as having *changed* their anaphoric pronouns into logophoric ones. Similarly, some of the languages that have extended the use of their logophoric pronouns to anaphoric contexts may be regarded as having changed their logophoric pronouns into anaphoric ones. It is not possible, however, to identify anaphoric pronouns with logophoric ones as claimed by some scholars (see Huang 2000: 190) because we would not then be able to account for the prototypical characteristics of either of these two types of pronoun. In the case of logophoric pronouns, for example, the prototypical characteristics can be accounted for only on the basis of their primary function of differentiating between endophoric and exophoric speech roles, whereas in the case of anaphoric pronouns the prototypical characteristics can be accounted for only on the basis of their primary function of differentiating between endophoric and exophoric (third person) referents.

The defining characteristics that have been used for claiming that a given language has logophoric pronouns are entirely different from the ones that have been used for claiming that a particular language has anaphoric pronouns. These distinct sets of characteristics very clearly derive from the functional distinction mentioned above. The most important concept that underlies the use of logophoric pronouns is the notion of the 'source' of information. Generally, the source of a piece of information that is contained in a statement would be the speaker himself. When, however, the speaker is not the source of information, there would be a need to specify that source. The fact that the source is other than the speaker and the addressee can form the basis for referring to it by a third person pronoun. However, if the source is viewed as the 'speaker' of a statement, there would be a need to differentiate it from an ordinary third person pronoun. Logophoric pronouns are thus partly like third person pronouns and partly like first person pronouns.

This dichotomous nature of logophoric pronouns gets reflected in the characteristics that they manifest in different languages. As I have pointed out earlier (§3.1), languages that do not have a distinct set of logophoric pronouns may use either their first and second person pronouns for this purpose, or alternatively, they may use their third person pronouns. This cross-linguistic variation is comparable to the variation occurring among the agreement markers of languages that do possess a distinct set of logophoric pronouns. They may have a distinct set of agreement markers for these pronouns, but the ones that do not have any such markers may use either first and second person agreement, or third person agreement. The diachronic source of logophoric pronouns also shows a similar kind of cross-linguistic variation. Logophoric pronouns may derive from first and second person pronouns or from third person pronouns. They may also have other derivational sources.

In the case of anaphoric (long-distance reflexive) pronouns, on the other hand, the most important concept that is used for characterizing them is the notion of

coreference with a linguistic expression. A third person pronoun can have its reference generally determined by the speech context, but it is possible, in the case of most languages, to use these pronouns coreferentially with previously used linguistic expressions. This double use of third person pronouns leads to ambiguity in some contexts. Anaphoric pronouns are used in order to remove this ambiguity, by specifying the coreferential use. While the problem of 'source' is primarily concerned with first and second person pronouns, that of coreference is concerned only with third person pronouns. It has no relevance as far as first and second person pronouns are concerned. This distinction between logophoric and anaphoric pronouns gets reflected in the fact that the former are generally restricted to first and second person pronouns (with the first person logophor being generally termed as a 'third person' pronoun) whereas the latter are restricted to third person pronouns.

Another interesting point, concerning the relationship between a logophoric pronoun and its antecedent, also appears to support this claim. Sells (1987: 449) points out that logophoric pronouns do not require antecedents with identical extensions, but instead require that the antecedent be *included* in the set denoted by the pronoun. The sentence given in (25), which is an Ewe sentence quoted by Sells from Clements (1975), exemplifies this point.

- (25) *kofi kpɔ be yèwo-do go*
 Kofi see Comp Log.Pl-come out
 'Kofi saw that they (including Kofi) had come out'

Notice that the logophoric plural pronoun *yèwo* has *kofi* as its antecedent which, however, is only included in it. Sells points out further that the opposite situation, in which a singular logophoric pronoun has a plural antecedent is not possible. Reflexive pronouns used in logophoric contexts, on the other hand, appear to require identical extensions (see also Wiesemann (1986a: 442) for a similar opinion regarding reflexive pronouns).

This possibility of the use of logophoric pronouns resembles the so-called 'plural of approximation' of first and second person pronouns, described in the previous chapter (§2.3.2). In English, for example, the expression *some of us* is, strictly speaking, in the third person, but the speaker may include himself and use a coreferential first person plural with it or exclude himself and use a coreferential third person plural with it (Jespersen 1943: 127, C. Lyons 1999: 312).

- (26) a. *Some of us like our beer chilled.*
 b. *Some of us like their beer chilled.*

As I will be pointing out in the next chapter (§4.2.1), this unique characteristic of first and second person pronouns derives from the fact that their association with the category of number involves the notion of conjunction rather than that of plurality. Logophoric pronouns appear to resemble first and second person

pronouns in manifesting this notion of conjunction when they are associated with the category of number.

Languages that make use of logophoric pronouns differ from one another in the kind of predicates that can function as 'logophoric triggers' in them, that is, predicates whose complement clauses can contain logophoric pronouns. In order to describe this variation, linguists have established a hierarchy of logophoric trigger predicates, such as (27) (Sterling (1993: 259).

(27) Communication > thought > psychological state > perception

It is claimed that the occurrence of any of these types of predicates as 'logophoric triggers' in a language would imply that the predicate types occurring to its left in the hierarchy would also function as logophoric triggers in that language. For example, Igbo allows only verbs of communication as logophoric trigger predicates, whereas Ewe allows all four types. Taburi excludes only verbs of perception, whereas Mandang excludes psychological state verbs as well (Huang 2000: 184).

One common characteristic of all these logophoric trigger predicates is that they involve a 'source' of information as distinct from the 'reporter' of information. It is true that a distinction between endophoric and exophoric speech roles actually occurs only in the case of the predicate type that occurs to the extreme left of the hierarchy, namely 'communication', but other predicate types that occur to its right can also be thought of as indirectly involving such a distinction. Further, the hierarchy clearly indicates the centrality of that distinction.

When logophoric pronouns are compared with third person pronouns, their use in a given context appears to be obligatory. That is, the use of a logophoric pronoun denotes coreference and that of the corresponding third person pronoun denotes disjoint reference. Sentences (28a–b) of Donno So (Culy 1997: 848) exemplify this point.

- (28) a. *Oumar Anta inyemeñ waa be gi*
 Oumar Anta Log (Acc) seen Aux said
 'Oumar said that Anta had seen him (Oumer)'
- b. *Oumar Anta woñ waa be gi*
 Oumar Anta 3Sg-Acc seen Aux said
 'Oumar said that Anta had seen him (some other person)'

Notice that in (28a), the logophoric pronoun *inyemeñ* 'him' indicates that the person seen is Oumar, who is the speaker of the embedded clause (source of information). In (28b), on the other hand, the third person pronoun *woñ* 'him' that is used in its place indicates that the person seen is someone other than Oumar.

On the other hand, the use of anaphoric pronouns is generally considered to be non-obligatory. Speakers can use either an anaphoric pronoun or the corresponding third person pronoun in most of the contexts in which the former can be used. The

Kannada sentences (18a–b) given earlier exemplify this non-obligatoriness of anaphoric pronouns. There are apparently only very few languages like Mupun (Chadic: Frajzyngier 1993: 83) in which no overlap is allowed to occur between the uses of anaphoric (endophoric) and exophoric pronouns. Some linguists have suggested that this obligatoriness vs. non-obligatoriness is one of the important differences between logophoric and anaphoric pronouns (Sterling 1993, Culy 1997: 848). Comrie (personal communication) points out, however, that in Tuburi (an Adamawa language of Chad) and Gbaya, the use of logophoric pronouns (even when compared with that of third person pronouns) is non-obligatory.

I wish to suggest, on the other hand, that the use of logophoric pronouns appears to be obligatory in most languages only because they have been compared with the use of third person pronouns. If we compare them with the use of the corresponding *first person* pronouns, we would find the occurrence of logophoric pronouns to be non-obligatory in roughly the same way in which the occurrence of anaphoric pronouns is considered to be non-obligatory. This latter comparison is actually the correct one because, as I have suggested earlier, logophoric pronouns are used in order to disambiguate the use of first and second person pronouns, whereas anaphoric pronouns are used in order to disambiguate the use of third person pronouns. Thus, in any given context in which a logophoric pronoun can be used, one can generally use a first person pronoun as well, with the former use being regarded as involving ‘indirect speech’ and the latter involving ‘direct speech’. In several languages, especially in the spoken form, the distinction between these two ways of reporting a given statement is not very clear-cut (see, for example, Schaub 1985: 5 on Babungu). This is also apparently true of Mupun (Chadic). According to Frajzyngier (1993: 109), the distinction between direct and indirect speech in Mupun is ‘marked’ with the help of logophoric pronouns.

- (29) a. *wu sat nə an nas ha*
 3MSg say that 1Sg-Fut beat 2MSg
 ‘He₁ said “I_{1,2} will beat you”’
- b. *wu sat nə dī nas an*
 3MSg say that LogMSg beat 1sg
 ‘He₁ said that he₁ beat me’

There is apparently an ambiguity in (29a), as the first person pronoun can have an endophoric antecedent if the sentence is regarded as involving direct speech, whereas if it is regarded as involving indirect speech it will have an (‘obligatory’) exophoric antecedent. The ‘non-obligatory’ use of logophoric pronoun as in (29b) removes this ambiguity just as the non-obligatory use of an anaphoric pronoun in the Kannada sentence (18b) given earlier removes a similar type of ambiguity occurring in the use of third person pronouns.

I had suggested earlier in this section that logophoric pronouns differ from anaphoric pronouns in representing non-coreference rather than coreference. It is

quite possible to regard both of them as representing non-coreference, as both of them have the function of *differentiating* between endophoric and exophoric uses of pronouns. That is, we can regard the two as complementing one another, with logophoric markers affecting first and second person pronouns and anaphoric markers affecting third person pronouns. However, the latter are generally perceived as representing coreference rather than non-coreference. In fact, languages that possess only anaphoric pronouns may extend their use for denoting, rather redundantly, first and second person coreference as well, and further, they may even differentiate between these three types of coreference by associating the anaphoric pronoun with personal suffixes or pronouns of all three persons.

For example, the anaphoric (long-distance reflexive) pronoun of Kashmiri (Indo-Aryan: Wali and Koul 1997: 123, 126–7) can be used for denoting coreference in the case of all three persons.

- (30) a. *bɪ ba:vɪ ma:la:-yi panun bo:y ə:nas manz*
 I show Mala-Dat Refl's brother mirror in
 'I show my brother to Mala in the mirror'
- b. *tsɪ chakh panɪni kəmi:zɪ chala:n*
 you are Refl's shirts washing
 'You are washing your shirts'
- c. *aslamən thəv kita:b panɪni beni nishi*
 Aslam kept book Refl's sister near
 'Aslam kept a book near his sister'

Notice, however, that in the case of these anaphoric pronouns, the denotation of first and second person coreference (or non-coreference) is redundant. In the case of logophoric pronouns, on the other hand, the denotation of first and second person non-coreference is of crucial importance. We have seen how failure to recognize this point has misled linguists, as in the case of languages like Gokana and Lele. There is thus clearly a need to differentiate between anaphoric coreference and logophoric non-coreference.

3.2 Anaphora among personal pronouns and proforms

There is a striking disparity in the use of the term 'anaphoric pronoun' between theoretical writings on this topic on the one hand, and the grammars of various individual languages on the other. In the former case the term is used mainly for denoting the expressions that remove the ambiguity between exophoric and endophoric uses of *personal pronouns* (especially that of third person pronouns), whereas in the latter case it is used mainly for denoting the expressions that remove this ambiguity from similar uses of *demonstrative pronouns* (proforms). There is clearly a difference between these two types of expression, in spite of the fact that in several languages, third person pronouns are either identical or

derivationally related to demonstratives. The use of the same term in both these cases appears to have masked this difference.

Most of the theoretical discussions of anaphoric pronouns or the notion of anaphora are concerned with the problem of coreference that affects the use of personal pronouns. For example, Huang (2000: 2) considers anaphors to be of two main categories, namely noun phrase (or noun) anaphors and verb phrase anaphors. The latter involve devices like ellipsis, gapping, sluicing, and stripping, and not any overt marking as such, and hence, as far as overt anaphoric pronouns are concerned, the study restricts itself to noun phrase or noun anaphors. Further discussions of this topic deal only with personal pronoun anaphors. This is also true of the Binding Theory of Chomsky (1981), which forms the basis of most of the theoretical discussions of anaphora; it is restricted to personal pronoun anaphors, especially the ones that are locally bound.

On the other hand, most grammarians use the term 'anaphoric pronoun' for denoting an entirely different kind of expression, namely the one that removes the ambiguity resulting from the double use (endophoric–exophoric) of *demonstrative* pronouns. For referring to pronouns that indicate personal pronoun anaphora, they use rather invariably, the term 'reflexive' (pronoun or affix). In my collection of data from the grammars of 225 languages, I find altogether 37 instances in which the languages are reported to have an 'anaphoric pronoun', and in the case of *all* of them, the pronoun is part of the system of demonstratives, and is mainly used for removing the ambiguity resulting from the exophoric and endophoric uses of demonstrative pronouns. It may be either a distinct pronominal element that contrasts with other demonstrative elements on this point, or alternatively, it may be an affix that is attached to demonstratives (or to ordinary nouns) for indicating their endophoric use.

For example, Maricopa (Yuman: Gordon 1986: 55) has the demonstrative root *aany* 'anaphoric' that contrasts with three other demonstrative roots, namely *da* 'near', *va* 'medial', and *aas* 'distant, out of sight'. In contrast to this, West Greenlandic (Inuit: Fortescue 1984: 254) has an anaphoric prefix *ta-* that can be attached to any demonstrative in order to change it into an anaphoric pronoun. Hunzib (Daghestanian: Van den Berg 1995: 61) is similar in having the prefix *yo-/yə* that can also be attached to demonstrative pronouns for a similar purpose.

I have been using the term 'anaphoric pronoun' in the previous sections of this chapter in the sense of 'personal pronoun anaphora'. I will be describing demonstrative anaphors ('anaphoric proforms') in greater detail in part two (Chs. 8 and 9) of this monograph. I have pointed out in the first chapter (§1.2.1) several important characteristics that differentiate between personal pronouns and proforms. We can expect at least some of these characteristics to be relevant for differentiating between these two types of anaphoric pronouns. For example, we generally find proforms to be made up of two different elements, namely a pronominal element that indicates their function and a general term that indicates their scope (§7.1). Anaphoric proforms (demonstratives) generally resemble

these other proforms in having a dual structure as shown by the proforms of different languages given in (31).

(31)		Anaphoric	Proximate	Remote
	Sinhalese (Indo-Aryan: Gair 1998: 113)	<i>eyaa</i>	<i>meyaa</i>	<i>aṛəyaa</i>
	Lezgian (Nakh-Daghestanian: Haspelmath 1993: 188)	<i>ham</i>	<i>im</i>	<i>am</i>
	Nivkh (Paleosiberian: Gruzdeva 1998: 26)	<i>kud</i>	<i>tyd</i>	<i>hyd</i>
	U. Kaapor (Tupi-Guaraní: Kakumasu 1986: 353)	<i>ame'ẽ</i>	<i>kom'ẽ</i>	<i>peme'ẽ</i>
	Godoberi (Daghestanian: Kibrik 1996: 42)	<i>ho = w</i>	<i>ha = w</i>	<i>hu = w</i>
	H. Oromo (Owens 1985: 87)	<i>xáaní</i>	<i>xuni</i>	<i>suní</i>
	Basque (Saltarelli 1988: 213)	<i>berau</i>	<i>hau</i>	<i>hura</i>

Third person anaphors (long-distance reflexives), on the other hand, generally involve a single element stem. For example, Kannada has a 'long-distance' anaphoric pronoun that resembles personal pronouns in showing allomorphic distinctions and in having a distinct plural marker.

(32)	First	Second	Anaphoric	Third	'king'
Singular	<i>na:nu</i>	<i>ni:nu</i>	<i>ta:nu</i>	<i>avanu</i>	<i>ra:ja(nu)</i>
Sg. Oblique	<i>nan-</i>	<i>nin-</i>	<i>tan-</i>	<i>avan-</i>	<i>ra:jan-</i>
Plural	<i>na:-vu</i>	<i>ni:-vu</i>	<i>ta:-vu</i>	<i>ava-ru</i>	<i>ra:ja-ru</i>
Pl. Oblique	<i>na-m-</i>	<i>ni-m-</i>	<i>ta-m-</i>	<i>ava-r-</i>	<i>ra:ja-r-</i>

Diachronically, the anaphoric *ta:nu* appears to have been, originally, a third person pronoun that got restricted to anaphoric use, with the remote demonstrative replacing it in the exophoric use. The latter also has an optional anaphoric use. Other Dravidian languages like Tamil and Telugu also have similar single stem forms that function as anaphoric personal pronouns.

Reinhart and Reuland (1993: 658) refer to an observation made by Faltz (1977), namely that when anaphors are complex expressions, they are universally local, whereas the long-distance type is universally simplex. They give examples like Dutch *zich*, Norwegian *seg*, and Italian *sé* to illustrate the latter claim. Several languages use terms denoting body parts as anaphoric (reflexive) pronouns that take personal markers when they are locally bound (have an antecedent in the same clause). This constraint regarding the underived nature of long-distance anaphors, however, is only a tendency rather than an absolute rule, as has been pointed out by some linguists like Huang (2000: 96) and Levinson (2000: 311). In Chinese, for example, a complex form like *ta ziji* ('pronoun' + self) can also occur as a long-distance reflexive (i.e. in addition to the simple form *ziji*).

The occurrence of a distinction between local and non-local anaphors is an important characteristic of personal pronoun anaphors. It does not appear to have any relevance as far as anaphoric demonstratives (proforms) are concerned. Further, the denotation of personal distinctions is also a characteristic of personal pronoun anaphors. In the case of some languages, the use of these pronouns is restricted to human referents, which again is a characteristic that they derive from being part of the system of personal pronouns. Levinson (2000: 312) points out that long-distance anaphors differ from ordinary (personal) pronouns in having something to do with emphatic contrast, or protagonist's perspective, subjective point of view, etc. Anaphoric demonstratives do not appear to provide any such connotations.

Another interesting difference between personal pronouns and proforms that gets extended to these two types of anaphoric pronouns is that the former (personal pronouns) can only occur as a nominal expression whereas the latter (proforms) can be nominal, adjectival, adverbial, or even verbal. For example, according to Heath (1984: 269), Nunggubuyu (Australian) has four different demonstrative roots, *ya*:- 'proximate', *da*:- 'immediate', *yuwa*:- 'distant', and *ba*:- 'anaphoric' that can occur with noun class suffixes to form demonstrative pronouns. They can also be used in forming a variety of adverbs. There are also distinct predicate forms for these demonstratives.

Similarly, Sinhalese (Indo-Aryan) has an anaphoric stem *ee* that functions like a demonstrative, as shown in (33). It contrasts with three different demonstrative stems, namely *mee* 'this (near the speaker)', *oya* 'that (near the addressee)', and *arə* 'that (distant from both)'. It can occur in a whole paradigm of pronominal forms, and resembles the three demonstrative stems (and also other proforms like interrogatives) in this respect (Gair 1970, 1998: 113).

(33)		Near 1	Near 2	Distal	Anaphoric	Interrogative
	Person	<i>meyaa</i>	<i>oyaa</i>	<i>arəya</i>	<i>eyaa</i>	<i>kauru</i>
	Animal	<i>meeka</i>	<i>ooka</i>	<i>arəka</i>	<i>eeke</i>	<i>kooka</i>
	Inanimate	<i>meekə</i>	<i>ookə</i>	<i>arəkə</i>	<i>eeke</i>	<i>kookə</i>
	Place	<i>mehee</i>	<i>ohee</i>	<i>arəhee</i>	<i>ehē</i>	<i>kohee</i>
	Manner	<i>mehemə</i>	<i>ohomə</i>	<i>arəhemə</i>	<i>ehemə</i>	<i>kohumə</i>

Among the three deictic demonstratives given in (33), the first two can also be used as anaphoric pronouns, but Gair (1998: 113) points out that in such a use, they generally also give the corresponding deictic meanings. The anaphoric *e*-, on the other hand, is neutral in this respect. Gair points out further that the distal *a*- and the anaphoric *e*- are specialized in function in a near complementary fashion. The distal pronoun is rarely used anaphorically, whereas the anaphoric pronoun is never used in spatial terms. The latter is restricted to the anaphoric use. Sinhalese also has a reflexive pronoun (anaphoric personal pronoun) *taman* 'self' that is generally restricted to third person and has only 'human' connotation.

Eastern Pomo, a Hokan language, is similar to Sinhalese in having an 'anaphoric' personal pronoun as well as an anaphoric demonstrative pronoun. Its

anaphoric demonstrative stem *‘u*, contrasts with the spatial demonstrative *khu*, in that the latter requires the object to be pointed out. The language also has an anaphoric pronoun *hí* contrasting with third person pronoun. The latter is derived by adding gender markers to the stem *mí*. (McLendon 1975: 124, 161).

As I will be pointing out in detail in the ninth chapter (§9.2.1–2), third person pronouns and proforms represent two entirely different notions of definiteness. The identity of the referent that needs to be established in the case of third person pronouns is only ‘linguistic’ whereas in the case of proforms like demonstratives it needs to be more substantial (extra-linguistic). It is possible that there is a similar distinction between anaphoric personal pronouns and anaphoric proforms. There is apparently a need to examine the relevance of this distinction (and also others) more systematically for understanding the exact nature of these two varieties of anaphoric pronouns.

3.3 Distinctions in the predicate

Some of the distinctions that occur in the predicates of sentences, such as reflexive, switch-reference and reciprocal, involve, though secondarily, coreference between two of their arguments, or between one of their arguments and the argument of a different predicate that is connected with it. Pronouns of first and second person differ from those of third person in the way in which these distinctions affect (or do not affect) them. For example, the distinctions that are connected with the reflexive and switch-reference predicates generally do not have any effect upon first and second person pronouns in the sense that the coreference that the verbs are expected to impose upon them is already present in the use of those pronouns. This difference between first and second person pronouns on the one hand and third person pronoun on the other gets reflected in the fact that in the case of some languages, the pronominal reflexive device is restricted to third person (§3.3.1).

We may regard the coreference that is connected with a reflexive predicate as complementary to the one that is connected with a switch-reference predicate. In the former case, the arguments that are regarded as coreferential (or non-coreferential) belong to a single predicate, whereas in the latter case, they belong to different predicates (§3.3.2). Reciprocal predicates, on the other hand, do not involve any coreference as such. They only involve sets of arguments that reciprocate each other’s case functions. However, this reciprocation gives rise to an ‘appearance’ of coreference (§3.3.3).

In addition to these meaning distinctions (which involve coreference in one form or the other), there are also other meaning distinctions that are connected with the predicate and affect the form and occurrence of personal pronouns. The most important one among them is modal distinction, especially between realis and irrealis. Several languages use distinct sets of personal pronouns depending

upon the modal distinctions that occur in the predicate (§3.3.4). It could be seen that in all these cases, the distinctions occurring in the predicate have influence on personal pronouns only to the extent that they have any influence on speech roles.

3.3.1 Reflexive meaning

The term ‘reflexive’ is generally used for denoting coreference between two different noun phrases, apparently due to the fact that the language that formed the main basis for its study, namely English, happened to have a marker of coreference (a pronominal reflexive device) for denoting it. As the term ‘reflexive’ clearly describes, however, the notion that underlies the use of this device is primarily a distinction between actions that affect the performer himself on the one hand, and the ones that affect some other person or object on the other. There are different ways in which an action can affect its performer. It may make him function as its patient, experiencer, or beneficiary in addition to being its performer. It is also possible for the reflexive predicate to merely indicate that an action has been carried out (or a process has taken place) without the intervention of any individual other than the performer.

Several languages represent the reflexive–non-reflexive distinction by modifying the verb, and thereby emphasize the fact that it is basically a verbal notion (see Bhat 1978: 9). For example, Manipuri (Tibeto-Burman) attaches the suffix *jə* to the verb in order to indicate that the verb involves reflexive meaning (Bhat and Ningomba 1997: 234).

- (34) a. *məhaknə miŋseŋ-də yeŋŋi*
 he mirror-Loc saw
 ‘He saw (something) in the mirror’
 b. *məhaknə miŋseŋ-də yeŋ-jəy*
 he mirror-Loc saw-Refl
 ‘He saw himself in the mirror’

The use of a reflexive verb implies, in most contexts, that two of the arguments of the predicate are coreferential. However, Manipuri may allow one of the coreferential arguments to be left unspecified, as in (34b), or specify it with the help of an ordinary third person pronoun, as in (34c).

- (34) c. *məhaknə mabu miŋseŋ-də yeŋ-jəy*
 he him mirror-Loc saw-Refl
 ‘He saw himself in the mirror’

There are some languages that possess two different devices, one verbal and the other one pronominal, of which the verbal device is used for denoting reflexive meaning and the pronominal one is used for denoting anaphoric coreference. There are contexts in which only one or the other of these two meanings

occurs, but there are also contexts in which the two meanings overlap. In the case of contexts in which the two meanings overlap, some of these languages are able to leave either of the two meanings unspecified, or alternatively, they may specify both of them. For example, sentences (35a–b) and (36) of Kannada (Dravidian) exemplify contexts in which only one of the two meanings occur, whereas sentences (37a–c) exemplify an overlapping context in which the three alternatives mentioned above (specifying the reflexive meaning, specifying the anaphoric coreference, or specifying both) have been utilized.

- (35) a. *ra:ju kurciy-alli kulita*
 Raju chair-Loc sat
 ‘Raju sat on the chair’
- b. *ra:ju kurciy-alli kulitu-konḍa*
 Raju chair-Loc sat-Refl
 ‘Raju sat on the chair (on his own)’
- (36) a. *ra:ju hari-ge avana pustakav-annu to:risida*
 Raju Hari-Dat his book-Acc showed
 ‘Raju₁ showed his_{1,2} book to Hari₁’
- b. *ra:ju hari-ge tan-na pustakav-annu to:risida*
 Raju Hari-Dat Ana-Gen book-Acc showed
 ‘Raju₁ showed his₁ book to Hari₂’
- (37) a. *ra:ju tana-ge anna baḍisida*
 Raju self-Dat food served
 ‘Raju served food to himself’
- b. *ra:ju anna baḍisi-konḍa*
 Raju food served-Refl
 ‘Raju served food to himself’
- c. *ra:ju tana-ge anna baḍisi-konḍa*
 Raju self-Dat food served-Refl
 ‘Raju served food to himself’

Notice that in (35) reflexive meaning is possible, as shown by the use of the verbal reflexive marker *konḍ* in (35b), but anaphoric meaning is not possible. On the other hand (36) exemplifies a sentence in which anaphoric meaning is possible, as shown by the occurrence of *tan-* ‘self’ in (36b), but reflexive meaning is not possible. Sentence (37) exemplifies a context in which both these meanings are possible and therefore there is an overlap. In (37a), the anaphoric coreference has been specified through the use of the anaphoric pronoun *tan*, but the reflexive meaning has been left unspecified. It can, however, be inferred from the fact that two of the arguments of the sentence are coreferential. In (37b), on the other hand, the reflexive meaning has been specified through the use of the verbal reflexive marker *konḍ-*, but the anaphoric coreference has been left unspecified. As

in the previous case, the meaning can be inferred from the use of the former device. Both these meanings have been specified in (37c) through the use of both the verbal reflexive marker as well as the anaphoric pronoun.

It is rather difficult to differentiate between these two types of meanings if we use the same term 'reflexive' (or 'anaphora') for denoting both coreference as well as verbal reflexivity. Since the term anaphora is generally used for denoting coreference, it would be helpful if we restrict the use of the term 'reflexive' for denoting verbal reflexivity as I have suggested (Bhat 1978: 12). That is, endophoric reference needs to be split into reflexive and anaphoric, with the former being restricted to simple clauses (locally bound). This need to differentiate between the reflexivity of the predicate on the one hand and anaphora (or coreference) of nouns or noun phrases on the other has been recognized recently by Reinhart and Reuland (1993: 658). They propose to differentiate between what they call 'SELF-anaphors' and 'SE-anaphors' of which the former have the function of turning a predicate into a reflexive one (by imposing identity on two of its arguments) whereas the latter has the function of denoting coreference. Only the latter (SE-anaphor) is considered to be capable of participating in long-distance binding.

As I have mentioned earlier, several languages use a verbal device for denoting the verbal reflexive meaning (see Huang 2000: 163), whereas several others use, instead, a nominal or pronominal device for the same purpose. For example, Tarma Quechua (Adelaar 1977) has reflexive verbs that are derived by adding *ku* to the verb. On the other hand, English attaches the marker *-self/selves* to the pronoun that denotes one of the coreferential arguments of a reflexive predicate. The pronoun is in the genitive in first and second persons and in the accusative in third person. Similarly, Cubeo (Tucanoan: Morse and Maxwell 1999: 80) attaches the word *baxu* 'self' to possessive forms of personal pronouns to derive their reflexive forms. Languages may also use bound forms of personal pronouns that are attached to certain invariable words for this purpose. For example, Bagirmi (Central Sudanic: Stevenson 1969: 45) adds the personal affixes to the word *kow* 'self, life' in order to derive reflexive pronouns.

There are some languages like Hungarian (Kenesei, Vago, and Fenyvesi 1998: 271) and Harar Oromo (Owens 1985: 170, 187) that are reported to use both these types of 'reflexive' devices. It is possible that in the case of some of them the two have distinct functions to serve, as in the case of Kannada described earlier. The latter (Kannada) is also generally described as using both verbal and pronominal 'reflexive' devices but, as I have pointed out above, the two have entirely different functions to perform. There would at least be a better possibility of uncovering such languages if we differentiate between 'reflexive' devices whose primary function is to mark the reflexivity of the predicate on the one hand, and 'anaphoric' devices whose primary function is to indicate the coreference of arguments on the other.

There are also some languages that do not possess any specific device for denoting reflexive meaning. In first and second persons, the occurrence of the

same pronoun for denoting two different arguments of the predicate implies, in these languages, that the predicate is reflexive. For example, Pirahã, an Amazonian language belonging to the Mura family, does not have any special pronominal or verbal device for denoting reflexive meaning. The meaning is expressed as a normal transitive configuration (Everett 1986: 215).

- (38) a. *hi hi xibáobahá*
 3 3 hit
 (i) 'He hit himself'
 (ii) 'He hit him'
- b. *ti ti xibáobahá*
 1 1 hit
 'I hit myself'

Everett notes that (38a) may be interpreted as involving either reflexive meaning or disjoint reference depending upon the available pragmatic information.

Since this ambiguity between coreference and disjoint reference occurs only in the case of third person pronouns, several languages restrict the use of their reflexive pronouns only to such contexts. For example, Reuland (2000: 11) points out that in the reflexive sentences of Dutch (and many other Germanic and Romance languages), pronouns of first and second person are used as they are, whereas pronouns of third person are necessarily replaced by the corresponding 'reflexive' pronoun. (This latter pronoun is actually 'anaphoric', as it is used in other non-reflexive contexts as well.)

- (39) a. *Ik voelde mij wegglijden*
 I felt me slide.away
 'I felt myself slide away'
- b. *Jij voelde je wegglijden*
 you felt you slide.away
 'You felt yourself slide away'
- c. *Hij voelde zich/*hem wegglijden*
 he felt himself/*him slide.away
 'He felt himself slide away'

On the other hand, several languages have developed special nominal or pronominal devices that can be used for denoting reflexive meaning. These differ from the anaphoric pronouns mentioned above in that they can be used in the case of all three persons. For example, Chinese attaches personal pronouns to its anaphoric pronoun *ziji* 'self' to indicate reflexive meaning (Li and Thompson 1981: 137, Zhu 1997: 74).

- (40) a. *lìsì zài zéběi tā zìjì*
 Lishi Durative blame 3Sg self
 'Lishi is blaming herself'

- b. *Zhangsan renwei ziji meiyong*
 Zhangsan thinks self useless
 'Zhangsan thinks that he is useless'

Notice that in (40a) the anaphoric pronoun *ziji* 'self' occurs with the third person pronoun to indicate reflexive meaning, whereas in (40b) it occurs alone to indicate anaphoric meaning (coreference).

According to Zhu (1997: 71), this complex reflexive pronoun, which contains a personal pronoun and *ziji* 'self' (as in (40a)), can only be used locally, very much like the reflexive pronouns of English. According to Huang (2000: 96), however, such complex forms can occur optionally as long-distance reflexives as well.

Languages that have a verbal reflexive device show a tendency to extend its use for indicating intransitive or passive connotations, whereas the ones that have a pronominal device show an entirely different kind of tendency, namely to extend it for denoting emphatic and anaphoric (long-distance reflexive) connotations (Bhat 1978: 26). In spite of this latter possibility, there is a need to differentiate between a prototypical reflexive device, which is clause-bound, and a prototypical anaphoric device that can have an antecedent occurring in a different clause. Some of the constraints that are shown by the latter can be accounted for only by assuming that it has an entirely different function to perform in language as compared to the clause-bound reflexive device. The function of the anaphoric device, as I have suggested (Bhat 1978: 52), is to remove two kinds of ambiguity that affect the use of *third person* pronouns: (i) ambiguity between endophoric and exophoric references and (ii) ambiguity between different endophoric antecedents. The function of the reflexive device, on the other hand, is to indicate a distinction in the predicate irrespective of the person to which its arguments belong.

Since the pronouns of first and second person are unaffected by either of the two ambiguities mentioned above, we can expect an anaphoric device to be relevant only in the case of third person pronouns, but not in that of first and second person pronouns. A reflexive device, on the other hand, would be relevant for all three persons as it indicates a distinction in the predicate. Hence, if the use of a given pronominal device is restricted to third person, we can assume that it is basically an anaphoric device, whereas if it is unrestricted by personal distinctions, but restricted by the notion of a predicate (local binding), we can assume that it is a reflexive device.

This point is supported by the fact that in the case of languages that use a verbal reflexive device, coreference among the arguments is not an obligatory condition. In Kannada (Dravidian), for example, a reflexive verb can be intransitive, and may denote the accidental occurrence of an event.

- (41) a. *koḍa oḍeyitu*
 pot broke
 'The pot broke'

b. *koḍa oḍedu-koṇḍitu*

pot broke-Refl

'The pot broke' (without the conscious intervention of an agent)

Another important difference between reflexive and anaphoric devices is that the scope of the former is restricted to a clause as it represents a meaning distinction that occurs in the predicate. The scope of the latter (anaphoric device), on the other hand, is constrained only by the notion of a sentence, as it is primarily meant for differentiating between endophoric and exophoric references in the use of third person pronouns. Its scope may even extend to other preceding sentences, but a reflexive device cannot have such extensions.

3.3.2 *Switch-reference*

Another device that is used for denoting a distinction occurring in the predicate is switch-reference. It is concerned with the notion of continuity or discontinuity in the use of a predicate. In Diyari (Australian), for example, subordinated clauses can be marked by either of a set of subordinating markers, depending upon whether their subject is coreferential or non-coreferential with the subject of the main clause (Austin 1981: 316).

- (42) a. *nhulu nganthi pardakarna warrayi, thanali thayi-lha*
 he meat brought Aux they eat-Implicated (SS)
 'He brought meat for them (i.e. him and others) to eat'

- b. *nhulu nganthi pardakarna warrayi, thanali*
 he meat brought Aux they
thayi-rnanthu
 eat-Implicated (DS)
 'He brought meat for them (others) to eat'

Notice that in (42a), the verb *thayi* 'eat' of the subordinated clause occurs with the subordinating marker *lha* that has 'same subject' connotation, whereas in (42b) it occurs with the subordinating marker *rnanthu* that has 'different subject' connotation.

Grammarians generally describe this device as involving a distinction between coreference and disjoint reference, but most of the 'switch-reference' languages use only a *verbal* device for this purpose. Further, the distinction is restricted to a clause, as it is necessary to specify coreference or non-coreference in the case of each clause, as, for example, when a series of clauses are subordinated to a main clause. That is, languages appear to use this device in order to differentiate between two different types of predicate of which one is perceived as continuing the characterization of a given subject or topic and the other one as not continuing it. Mithun (1999: 270) points out that in some of the languages that are reported to be using a switch-reference device, there are contexts in which the

device appears to be used in order to differentiate between the continuity and discontinuity of events rather than that of referents. It is therefore necessary, I think, to regard switch-reference as basically a feature of the predicate rather than of its arguments.

3.3.3 Reciprocal meaning

Reciprocity is another concept that is primarily concerned with the predicate rather than with the argument. It resembles reflexive meaning in some respects but differs from it in others. The most important difference between the two is that the notion of 'coreference', when it occurs, is a real one in the case of reflexive meaning, whereas in the case of reciprocal meaning, it is only illusory.

- (43) a. *John beats Mary.*
 b. *John beats himself.*
 c. *John and Mary beat each other.*

Notice that the reflexive sentence (43b) indicates coreference between the agent and the patient (with *John* replacing *Mary* as the patient, while continuing as the agent), whereas the reciprocal sentence (43c) denotes only the conjoining of the two arguments *John* and *Mary*. Both of them continue in the respective case roles that they had in (43a), but they additionally assume the role of the other individual. However, this does not give rise to 'real' coreference as shown by (43d), which is a paraphrase of (43c).

- (43) d. *John beats Mary and Mary beats John.*

The coreference that occurs in the reciprocal sentence (43c) is only an 'apparent' one. It derives from two different facts: (i) the sentence combines together two different events in which the same set of participants, namely *John* and *Mary*, are involved; (ii) the involvement of these two participants in the second event is the *reverse* of their involvement in the first event. That is, in the conjoined sentence, both the participants appear as being involved in *both* those roles, and this makes it possible to regard the two roles as being represented by the same set of participants (see Langacker 1976: 48, Bhat 1978: 41). It is true that all reciprocal sentences cannot be derived from conjoined sentences, but none of them can be regarded as involving real coreference.

Several languages express the concept of reciprocity with the help of a verbal affix. For example, Manipuri (Tibeto-Burman) attaches the suffix *nə* to the verb in order to indicate reciprocal meaning (Bhat and Ningomba 1997: 237).

- (44) a. *tombə-nə cawbə-də phuy*
 Tomba-Nom Chaoba-Dat beat
 'Tomba beat Chaoba'

- b. *tombə-gə cawbə-gə phu-nəy*
 Tomba-Gen Chaoba-Gen beat-Rec
 ‘Tomba and Chaoba beat one another’

According to Lichtenberk (2000b: 57), languages may use a reciprocal marker that is connected either with a reflexive marker or with a collective marker, but not with both of them. For example, many Oceanic languages use the same construction for denoting reciprocal and collective senses, but a different one for denoting the reflexive meaning. On the other hand, most Australian languages use the same construction for denoting reciprocal and reflexive meanings. This duality in the diachronic source of reciprocal markers derives from the fact that reciprocal meaning involves two distinct notions. It prototypically requires the involvement of a plurality of participants, and languages that emphasize this aspect of reciprocal meaning use a device for representing it that is also used for denoting collective meaning. On the other hand, reciprocal meaning also involves the notion of an illusory coreference as described earlier, and languages that emphasize this latter aspect of reciprocal meaning use a device for representing it that is also used for denoting reflexive meaning.

Languages may also use a pronominal device for denoting reciprocal meaning, as some languages do in the case of reflexive meaning (see §3.3.1). However, there is an important difference between reflexive and reciprocal meanings, as far as the actual type of pronominal devices that are used for denoting them are concerned. While the pronominal devices that occur in the case of reflexive meaning can be regarded as anaphoric and definite, in the case of reciprocal meaning, the devices can only be regarded as ‘anaphoric’ but not as definite. That is, languages generally use an *indefinite* expression as a reciprocal pronoun. For example, English does not use its personal pronouns for deriving its reciprocal expressions, as it does in the case of its reflexive device. Instead, it uses a complex form that is made up of the ‘indefinite’ pronouns, *each*, *other*, *one*, and *another*.

- (45) a. *The students are helping each other.*
 b. *The boys looked at one another.*

This is true of several other languages that use a pronominal device for denoting the reciprocal meaning. For example, Dravidian languages like Kannada derive their reciprocal pronouns through the use of the numeral for ‘one’.

- (46) *ra:ju mattu hari obbar-ann-obbaru hogaḷidaru*
 Raju and Hari one-Acc-one praised
 ‘Raju and Hari praised one another’

Similar derivations of pronominal reciprocal devices have been reported to occur in Marathi and other Indo-Aryan languages as well. Hunzib (Daghestanian: Van den Berg 1995: 133) also has a similar derivation; it reduplicates the cardinal

number *hās* ‘one’ in the appropriate case. This tendency appears to derive from the fact, mentioned earlier, that the coreference that is involved in the case of reciprocal meaning is only a fictitious one. The reciprocal pronouns select an indefinite individual from the group that the subject identifies, and indicate that that indefinite individual is related with another indefinite individual through the relation of ‘reciprocity’. It is apparently for this reason that the English reciprocal, unlike its reflexive, does not show any person distinctions either, even though the pronoun is applied to all three persons.

- (47) a. *We were praising one another.*
 b. *You were praising one another.*

3.3.4 Modal distinctions

I have suggested (Bhat 1999: 91) that a three-fold typological distinction can be made among languages depending upon whether they give prominence to the categories of tense, aspect, or mood. We may designate languages as ‘tense-prominent’, ‘aspect-prominent’, or ‘mood-prominent’ based on this classification. One of the characteristics that can be associated with languages that give prominence to mood is that they tend to possess distinct sets of personal pronouns that are to be used in a sentence depending upon the mood of its predicate (Bhat 1999: 135). This correlation appears to derive from the fact that modal distinctions represent either (i) distinctions in the speaker’s assessment of the reliability of information or (ii) distinctions in the illocutionary force of a speech act. Both are directly related to speech contexts and speech roles and hence their influence on the form and occurrence of personal pronouns can perhaps be regarded as supporting the claim that the primary function of personal pronouns is to indicate speech roles.

There are several languages in which personal pronouns have two distinct sets of forms of which one occurs in realis sentences and the other one in irrealis sentences. For example, Wari’, a Pacaas Novos language of western Brazil (Everett and Kern 1997: 329) makes a distinction between realis and irrealis sentences through the use of distinct sets of pronominal clitics. There are two distinct sets of realis clitics that are used to differentiate between future and non-future realis verbs, which contrast with a single set of irrealis clitics. There is also a set of tenseless clitics that is used when tense is marked elsewhere in the verb.

- (48) a. *cao’* ‘*iri*’
 eat 1 (Incl.Realis.Present)
 ‘We are eating’
 b. *’om* *ta querec* *xi’*
 not Realis.future.see 1 (Incl)
 ‘We (Incl) will not see’

- c. *tomi'* *ha'* *xixi'*
 speak pay.attention 1 (Incl.Irrealis)
 'We should obey you'

According to Chafe (1985), Caddo (Iroquian) makes a distinction between realis and irrealis verbs by attaching distinct sets of personal prefixes to them. The irrealis verb is used in yes-no questions, negations, obligatories, conditionals, and several others like simulative *dùy* 'as if', infrequentative *wàs* 'seldom', admiring *hùs* 'surprise', etc. Similar distinctions occur in several other languages like Manam of the Oceanic group (Lichtenberk 1983) and Amele of the Gum family (Roberts 1987) as well.

There are other mood-prominent languages that use distinct sets of personal pronouns depending upon the illocutionary force of a sentence or clause. For example, Supyire, a language of the Gur family (Niger-Congo) makes a distinction between declarative and non-declarative sentences. The latter includes sentences like commands, prohibitions, questions, blessings, vocatives, and exclamations. There are two distinct sets of forms for first and second person pronouns in this language, of which one is used in declarative sentences and the other one in non-declarative sentences (Carlson 1994: 153).

(49)	Declarative		Non-declarative	
	Singular	Plural	Singular	Plural
1	<i>mùì</i>	<i>wùu</i>	<i>na</i>	<i>wu</i>
2	<i>mu</i>	<i>yìì</i>	<i>ma</i>	<i>yì</i>

- (50) a. *u a mùì pyà*
 he Perf me (Decl) see
 'He has seen me'
- b. *ku kan na à*
 it give me (Non-Decl) to
 'Give it to me!'

- (51) a. *mu a mùì kánhá*
 you (Decl) Perf me (Decl) tire
 'You have annoyed me'
- b. *ma pa.*
 you (Non-Decl) come
 'Come!'

According to Everett (1986: 246), Pirahã (Mura language family) makes a distinction between first and second person pronouns occurring in imperative and hortatory sentences on the one hand, and the ones occurring in other types of sentences on the other.

- (52) a. *gói pii oái-pí pii igopái hái*
 2Imp water fetch water bring certain
 'Go fetch water. Bring water!'

- b. *gí ti xibáobáhá*
 2 1 hit
 'You hit me'

Nicobarese (Munda family) differentiates between interrogative and non-interrogative sentences by using distinct sets of personal pronouns. The interrogative set is used in the subject position of either interrogative or exclamatory sentences (Braine 1970: 145).

- (53) a. *haʔén əc*
 late 1 (Int)
 'Am I late?'
 b. *fohñə misɔ.kə cin*
 hit.out mosquito 1 (Non-Int)
 'I hit the mosquito out'
- (54) a. *sítih kanɔ.ŋeváh tíʔ əm tɔ ɔ*
 how ability (Perf) do you (Int) to it
 'How did you catch it?'
 b. *fě.lə cɔ.n man*
 kill John you (Non-Int)
 'Kill John!'

Notice that distinct first (*əc/cin*) and second (*əm/man*) person pronouns are used in the (a) and (b) sentences of (53) and (54) respectively.

3.4 Summary

First and second person pronouns differ from third person pronouns and also from proforms and nouns concerning the kind of 'coreferential ambiguity' that affects their use. Since the primary function of first and second person pronouns is to denote speech roles, their use may involve coreferential ambiguity between the speech roles that belong to two different speech contexts, as, for example, between those of the reported speech context and the reporting speech context. A given personal pronoun may represent a speech role that belongs to the reported speech context, which may or may not be coreferential with the corresponding speech role of the reporting speech context. Languages have been reported to use special 'logophoric' pronouns in order to indicate that a reported speech role is *non-coreferential* with the corresponding reporting speech role.

In the case of third person pronouns, on the other hand, coreferential ambiguity can occur within a single speech context, depending upon whether the pronoun

has an endophoric antecedent or an exophoric antecedent. Languages generally resolve this ambiguity by using special 'anaphoric' pronouns that indicate that the antecedent is anaphoric (denoted by a linguistic expression). That is, they are used for denoting the *coreference* between the two expressions. Notice that this type of ambiguity does not occur in the case of first and second person pronouns.

There is a need to differentiate between these logophoric and anaphoric distinctions that affect pronominal reference on the one hand, and reflexive, reciprocal, switch-reference, and modal distinctions that affect the meaning of the predicate on the other. Languages may use either pronominal or verbal devices for representing these latter meaning distinctions of the predicate, and depending upon the type of device used, they may also show different types of extensions. However, the prototypical characteristics of these latter devices cannot be properly accounted for if we do not differentiate them from logophoric and anaphoric devices whose primary function concerns arguments and not predicates.

4 Association with Grammatical Categories

4.1 Introduction

Personal pronouns are generally considered to be similar to nouns in their ability to associate with grammatical categories like number, gender, and case. This assumption, however, is only superficially correct. In their actual association with these categories, pronouns of first and second person are rather different from proforms and nouns. Number marking, for example, involves primarily conjunction in the case of first and second person pronouns, whereas in that of proforms and nouns it involves plurality. The crucial difference between these two concepts is that, in conjunction, differences that occur among the entities that are numbered are emphasized, whereas, in plurality, they are suppressed (§4.2). The association of personal pronouns with the categories of gender (§4.3) and case (§4.4) also involves certain unique characteristics that are absent in the association proforms and nouns with those categories. I wish to suggest that most of these unique characteristics of personal pronouns derive from their primary function of denoting speech roles rather than the individuals who enact those roles.

4.2 Association with the category of number

The association of personal pronouns with the category of number is generally described in the same way in which the association of other nominals with that category is described. Both are considered to involve notions like dual and plural, and other such number distinctions. However, there are several points on which personal pronouns differ from other nominals in their association with the number category. In the case of nominals, for example, there is generally a one-to-one relationship between the singular forms and the corresponding non-singular (plural) forms (as in *boy-boys*, *tree-trees*, *stone-stones*, etc.), whereas in the case of personal pronouns, the relationship is more complex. Several languages show two different non-singular forms, called ‘inclusive’ and ‘exclusive’ for their first person pronoun. The paradigm of Malayalam (Dravidian: Asher and Kumari 1997: 258) personal pronouns given in (1) exemplifies this point. Notice that there are only two singular pronouns but corresponding to these, there are three different plural pronouns. (Third person pronouns of Malayalam are the same as demonstratives.)

(1)	Singular	Plural	
1	<i>naan</i>	1 Excl	<i>naṇṇa!</i>
		1 Incl	<i>naam</i>
2	<i>nii</i>		<i>niṇṇa!</i>

The distinction occurring in first person plural in (1) involves the inclusion vs. exclusion of the addressee, which means that it is not merely a distinction in plurality. It indicates a distinction between two different kinds of *combinations* in which the first person is included; one of these combinations includes second person, whereas the other one excludes it.

Even in the case of languages in which there is no disparity in the actual number of singular and non-singular forms, the non-singular forms of personal pronouns differ in their connotation from those of other nominals. As pointed out by several linguists like Jespersen (1924: 192), Benveniste (1971), and J. Lyons (1968: 277), the terminology used for describing the non-singular forms of first and second person pronouns is misleading because the forms do not stand in the same relationship to singular forms as *boys*, *cows*, etc. do to *boy*, *cow*, etc. The word *boys* indicates several boys, but the pronoun *we* does not indicate several speakers; instead, it indicates one speaker (specifically, the speaker of the sentence in which it occurs) and one or more non-speakers. Jespersen (1924: 192) points out that even when a body of men, in response to ‘who will join me?’ answer ‘we all will’, it means in the mouth of each speaker nothing but ‘I will and all the others will (I presume)’. The plural of second person is somewhat different in that it may indicate several addressees in some contexts, but generally, it is used to indicate an addressee and one or more other persons.

This unique characteristic of personal pronouns does not appear to get extended to other pronouns (proforms) or nouns. For example, it is quite possible to think of the word for ‘father’ as having two different plural forms of which one includes ‘mother’ and the other one excludes. Nevertheless, languages do not appear to express such an inclusive–exclusive distinction through their number system. Why is this distinction restricted to the number system of personal pronouns?

There do occur instances in which a dual form of a noun may indicate the inclusion of a specific individual. For example, Sanskrit has the dual form *ma:tarau* ‘mother (Dual)’, which has the meaning ‘mother and father’. Grammarians regard it as a shortened form of the *dvandva* compound *ma:ta:-pitar-au* ‘mother-father-Dual’. The compound also has the shortened form *pitarau* ‘father (Dual)’, with the meaning ‘father and mother’. However, neither of these dual forms is contrasted with a form that *excludes* reference to father in the former case and mother in the latter case. That is, we have an ‘inclusive’ form here, but that form is not contrasted with an exclusive form as in the case of personal pronouns.

Jespersen (1924: 191) considers the pronoun *we* to be involving a ‘plural of approximation’ where several individuals are comprised in the same form though not belonging exactly to the same kind. It is similar, according to him, to certain other plural forms like *sixties* which indicates sixty-one, sixty-two, and so forth till sixty-nine or *the Vincent Crummleses* which indicates Vincent Crummles and his family. However, there are contexts in which the non-singular forms of personal pronouns do not involve this notion of ‘approximation’. For example, when they are used for denoting only two persons, there would not be any approximation as such, especially if they have the inclusive meaning. This is also true of the distinct first and second person ‘dual’ forms that some languages possess. On the other hand, Huddleston and Pullum (2002: 1465) consider *we* and *I* to be ‘distinct lexemes, not plural and singular forms of a single lexeme’.

Cysouw (2001: 66) suggests using the term ‘group’ rather than ‘plurality’ for referring to the association of personal pronouns with the category of number, especially when it is used for denoting combinations like $1 + 2$, $1 + 3$, $1 + 2 + 3$ and $2 + 3$. However, the notion of ‘group’ has apparently misled him, because he includes $3 + 3$, which is in no way different from the plural forms of other pronouns and nouns, in his category of ‘groups’; further, he does not include pronouns that specify the actual numbers, like dual and trial in this category of ‘group’, even though they resemble other non-singular personal pronouns in being different from the non-singular forms of other pronouns and nouns.

The crucial point on which personal pronouns differ from other pronouns and nouns in this association with the category of number is that, in the former case, there is a need to specify the nature of the individuals that are included under its reference, whereas, in the latter case, distinctions occurring among the referents are to be disregarded. The first person plural pronoun, *we*, for example, specifies that it includes, minimally, the speaker as one of its referents, and that its other referents may be the addressee and/or other persons. The notion of ‘approximation’ concerns these other persons that get included under its reference, and not the speaker. The inclusive–exclusive distinction that is associated with personal pronouns also requires the specification of a particular type of individual that occurs (or does not occur) as their referent. The non-singular forms of other pronouns and nouns like *they* and *boys*, on the other hand, do not require their referents to be differentiated from one another in this fashion.

4.2.1 Homogeneous and heterogeneous plurals

The case of Mao Naga (Tibeto-Burman: Giridhar 1994: 114) is rather interesting in this context. It differentiates between (i) a plural that indicates similarity among the individuals that are grouped together (called ‘homogeneous’ plural) and (ii) a plural that indicates differences among them (called ‘heterogeneous’ plural). The normal plural suffix that is attached to nouns or pronouns for

denoting plurality in this language is *khru*. It may indicate either a homogeneous group or a heterogeneous group. This suffix is contrasted with another plural suffix, namely *ta*, which indicates, specifically, a homogeneous part of a heterogeneous group. (These two suffixes are followed, as seen in the examples given below, by one of the individuating suffixes like *hi* ‘visible’, *ʔi* ‘invisible, definite’, and *sü* ‘invisible, indefinite’.)

- (2) a. *laruübvü-khru-ʔi* ‘books (of different types)’
 b. *larübvü-ta-ʔi* ‘books (copies of the same book out of a collection of books of different types)’
- (3) a. *ayi madi-khru-ʔi pika*
 me green.one-Pl₁-Ind give
 ‘Give me the green ones’
 b. *ayi madi-ta-ʔi pika*
 me green.one-Pl₂-Ind give
 ‘Give me the green ones only (out of a collection of things of different colours)’

Giridhar (1994: 116) suggests that only *khru* is a genuine plural marker. He considers *ta* to be related to the word *oʔa* ‘kind, type, species’. In some dialects, the two occur with the same consonant. Further, when the suffix occurs with the singular marker *na* as in *hi-na-ta-hi* ‘this-one-Pl₂-this’: ‘these-one kind-but many in number’, it clearly has this latter connotation, namely ‘kind’.

What is interesting in the present context is that when these two suffixes are attached to personal pronouns, they indicate the inclusive–exclusive distinction (Giridhar 1994: 139).

- (4) a. *a-khru-müy iniu*
 1-Pl₁-person village
 ‘our (Incl) villages’
 b. *a-ta-müy iniu*
 1-Pl₂-person village
 ‘our (Excl) village’

Notice that the distinction here is not merely between homogeneity and heterogeneity as in the case of nouns. It indicates, specifically, a distinction between the inclusion vs. exclusion of the addressee.

It is also possible to use these two plural suffixes with the second person pronoun in Mao Naga. In this case, they appear to indicate a distinction between the inclusion vs. exclusion of the third person. Giridhar (1994: 141) points out that *ni-khru-müi* ‘2-Pl₁-person’ (2.Incl.PI) would mean that all its referents are present; it cannot refer to anyone who is not present (i.e. third person). On the other hand, *ni-ta-müi* ‘2-Pl₂-person’ (2.Excl.PI) could refer to persons who are not present as well. The first person inclusive form *a-khru-müy* ‘1-Pl₁-person’

also appears to have this special connotation of indicating only the persons who are actually present.

4.2.2 Plurality and conjunction

I believe that this interesting difference between personal pronouns and other pronouns and nouns can be described more effectively by associating the former with the notion of ‘conjunction’ and the latter with the notion of ‘plurality’. These are two entirely different devices that languages utilize in order to indicate the involvement of two (or more) persons or things in a particular case role. Conjunction emphasizes the *differences* that occur between the persons or things that are conjoined by it, whereas plurality emphasizes the *similarity* that occurs among them. In fact, the differences are to be *disregarded* when a plural marker is used.

- (5) a. *John, Bill, and Tom have gone home.*
- b. *The boys have gone home.*

Notice that the use of conjunction in (5a) depends upon the specification of *John*, *Bill*, and *Tom* as three distinct individuals, whereas the use of plurality in (5b) depends upon their non-specification as distinct individuals, and their specification as just *boys*. The noun *boys* disregards the differences that occur among them and emphasizes the common characteristic of ‘being boys’ that is shared by them. One cannot use conjunction if no difference is indicated, either explicitly or implicitly, whereas one cannot use plurality if differences cannot be disregarded. We cannot have **he and he* as a conjunction because no difference between the two conjoined entities can be perceived, whereas *John and John* is possible only if the context indicates that the two personal names denote different individuals.

In the case of the ‘plural’ forms of personal pronouns, however, this basic requirement of plurality is completely disregarded. As we have seen earlier, the ‘dual’ and ‘plural’ forms of these pronouns are used for denoting conjunction (i.e. in contexts in which no similarity is involved) in exactly the same way in which conjoint phrases like *I and you*, *I and he/they*, or *you and he/they* might be used. The first person plural pronoun *we* of English represents the combination of a speaker and his addressee, or a speaker and a third person; there cannot be any *similarity* between the referents that get included in either of these two sets (i.e. between the speaker and his addressee or between the speaker and a third person), at least not one that can be indicated by a ‘first person’ pronoun, in the way in which the word *boy*, mentioned earlier, represents the common characteristics of the individuals that can be referred to by it. Hence, these denotations of first person plural pronoun can only be regarded as involving the notion of conjunction and not that of plurality. The second person plural pronoun, however, is somewhat different, as mentioned earlier; it generally represents an addressee and some other person or persons (conjunction), but occasionally it can also represent a group of addressees (plurality).

There are several languages in which the first person dual or plural shows two different forms, namely exclusive and inclusive (see the Malayalam example given in (1)). These forms are generally described as being part of the number system. But the forms represent only a distinction in the *type* of individuals that get included or excluded, and not a distinction in the *number* of individuals that are referred to by those expressions. The proper device for representing such a distinction is clearly conjunction and not plurality.

We can account for the above-mentioned unique characteristic of personal pronouns on the basis of the assumption that the primary function of personal pronouns is to indicate speech roles and not the persons who enact those roles. In order to carry out this function satisfactorily, they need to be dissociated from their referents (see §2.2). If we associate speech roles (rather than the performers of speech roles) with the category of number, the resulting notion can only be conjunction, involving different combinations of speech roles. It cannot be thought of as resulting in the notion of 'plurality' because there does not appear to be any context in which an indication of several 'instances' of the same speech role would be useful.

If the non-singular forms of personal pronouns are used for denoting the combination of speech roles, we can expect the possible number of such forms to be *different* from that of singular forms in at least some cases. For example, when a language has only two singular personal pronouns denoting the first and second persons, there would be a need to have only a single non-singular personal pronoun, namely the one that denotes the combination 1 + 2. On the other hand, when there are representations for three different persons, namely first, second, and third, there can be three or more combinations, such as 1 + 2, 1 + 3, 2 + 3, and 1 + 2 + 3. Thus, the disparity that occurs among languages between singular and non-singular personal pronouns can nicely be accounted for by the claim that the association of personal pronouns with the category of number leads to the notion of conjunction.

There are some languages in which the number distinction is restricted to personal pronouns. For example, Chalcatongo Mixtec (Macaulay 1996: 81) has two distinct personal pronouns, namely first and second, corresponding to which it has a *single* non-singular form that has the meaning '1 + 2'.

(6)	Singular	Non-singular
1	<i>ru'ù</i>	<i>žó'ó</i>
2	<i>ro'o</i>	

Third person pronouns of this language, are general nouns, such as, for example, *čàà* 'man' (for 3M) and *ñã'ã* 'woman' (for 3F), but the language does not express plurality in the case of these nouns or in that of any other noun. In order to indicate the plurality of the subject, it uses the prefix *ká* that is attached to the verb.

If we regard this language as having a number category that is restricted to first person, it would imply that the language has a nominal number category that is

unusually defective, as it is restricted to dual meaning and also to a single word in the language, namely the first person pronoun. On the other hand, if we consider the language as expressing conjunction and not duality among its personal pronouns, the paradigm given in (6) would appear to be perfectly formed. There can only be a single conjunctive form (1 + 2) in a system that consists of only two personal pronouns, namely 1 and 2, and the language correctly possesses a single conjunctive form for representing it.

Kwakiutl (Wakashan: Boas 1911: 527) is similar in that it also appears to be defective if the former way of describing the number category is applied to it. It has singular forms for all three persons, but non-singular form for first person only. There are actually two different non-singular forms in first person, namely an inclusive and an exclusive, as shown in (7).

(7)	Singular	Inclusive	Exclusive
1	<i>En</i>	<i>Eu^εɬu</i>	<i>Ens</i>
2	<i>Es</i>		
3	(<i>ē</i>)		

If we regard these forms as representing the possible conjunctions of persons, rather than as the 'plural' forms of first person pronoun, the paradigm would appear to be less defective. Notice that the three personal pronouns can involve three possible conjunctions, namely 1 + 2, 1 + 3, and 2 + 3. The language has representations for two of them, inclusive (1 + 2) and exclusive (1 + 3). The absence of 2 + 3 can be accounted for by the fact that the third person pronoun is rather different in this language, as it is almost always used with the demonstrative pronoun. It is also not specified in the verb when the subject is to be represented. Concerning this problem of the 'plurals' of Kwakiutl personal pronouns, Boas suggests that the two 'are not conceived as plurals' but merely as inclusive and exclusive forms.

A similar system is reported to occur in Acehnese (Durie 1985, quoted by Corbett 2000: 64). It also has a first person pronoun restricted to singular use, and two additional forms used for denoting exclusive and inclusive meanings. Second and third person pronouns, and also nouns, do not show any number distinction. In languages of this type, the primary function of number marking is clearly the denotation of conjunction (emphasizing the differences among the entities that are put together rather than similarities) and hence there cannot be any language-internal basis for regarding them as having 'dual' or 'plural' marking among the pronouns. This is also true of other languages in which number distinction occurs only among personal pronouns. (See Helmbrecht, forthcoming, for additional languages of this type.)

In the case of most languages, on the other hand, number markers occur with personal pronouns as well as with other nominals, but the actual markers that occur with first and second person pronouns are rather different from the ones that occur with other nominals. Many languages use suppletive forms for

denoting number distinction in the case of first and second person pronouns and regular affixes in that of proforms and nouns. In English, for example, first person plural is derived through suppletion (*I, me* vs. *we, us*) whereas the plural forms of nouns are generally derived through the addition of a plural suffix (*tree* vs. *tree-s*). There are several languages of this type, of which some are shown in (8).

(8)	1Sg	1Pl	2Sg	2Pl
Bukiyip (Arapesh: Conrad 1991: 11)	<i>yek</i>	<i>apak</i>	<i>nyak</i>	<i>ipak</i>
Chukchee (Bogoras 1922: 719)	<i>ğ̃um</i>	<i>mu'ri</i>	<i>g̃it</i>	<i>tur'i</i>
Taba (Malayo-Polynesian: Bowden 1997: 215)	<i>yak</i>	<i>tit</i>	<i>au</i>	<i>meu</i>
Supyire (Gur: Carlson 1994: 151)	<i>m̃i</i>	<i>w̃u</i>	<i>mu</i>	<i>ỹi</i>
Kobon (Kalam: Davies 1981: 154)	<i>yad</i>	<i>hon</i>	<i>ne</i>	<i>k̃ol</i>
Luvale (Bantu: Doke 1949: 64)	<i>àmi</i>	<i>ètu</i>	<i>òve</i>	<i>ènu</i>
Maybrat (Papuan: Dol 1999: 68)	<i>tuo</i>	<i>amu</i>	<i>nuo</i>	<i>anu</i>
Lezgian (Daghestanian: Haspelmath 1993: 184)	<i>zun</i>	<i>čun</i>	<i>wun</i>	<i>k̃ün</i>
Warao (Isolate: Romero-Figeroa 1997: 64)	<i>ine</i>	<i>oko</i>	<i>ihi</i>	<i>yatu</i>
Bagirmi (Nilo-Saharan: Stevenson 1969: 32)	<i>ma</i>	<i>je</i>	<i>i</i>	<i>se</i>

In some of the languages given in (8), there are additional distinctions like dual-plural and inclusive-exclusive; these are not shown in (8). The first person plural form given for languages in which there is an additional inclusive-exclusive distinction is the inclusive one. In my sample of 225 languages (see §6.1), about 60 per cent show suppletive forms for first person plural. This is less frequent (only about 40 per cent) for second person plural, but still, compared to proforms and nouns, the occurrence of suppletive plural forms among personal pronouns is very prominent.

There are also several languages in which the non-singular number marker used with personal pronouns is different from the one that is used with proforms and nouns. For example, Kannada (Dravidian) uses the suffix *-vu/m* with first and second person pronouns and *-ru* 'human' and *-gaḷu* 'non-human' with nouns and proforms. Turkish (Lewis 1967: 67) uses the marker *-iz* with first and second person pronouns and *-lar* with third person pronouns (demonstratives) and nouns. Even languages that use the same number marker with all nominals generally use either suppletive forms or altered stems with their first and second person pronouns. For example, Maori (Eastern Polynesian: Bauer 1999: 366) is regular in its use of the dual suffix *-ua* and the plural suffix *-tou*, but the first person pronoun itself has distinct stems in singular and non-singular forms.

(9)	Singular stem	Non-singular stem
1	<i>au</i>	<i>maa</i> (Excl), <i>taa</i> (Incl)
2	<i>koe</i>	<i>koor/kou</i>
3	<i>ia</i>	<i>raa</i>

I would like to suggest that this differentiation in the number markers used, coupled with a differentiation in the connotation of these markers (conjunction in the case of personal pronouns and plurality in the case of other nominals) needs to be regarded as representing a distinction in the category of number itself.

4.2.3 Dichotomous association

As I will be pointing out in the next chapter (§5.3), personal pronouns appear to be Janus-like in their association with some of the categories like number, gender, and definiteness. They show some characteristics that appear to derive from their primary function of denoting speech roles, but the fact that they do have referents of their own appears to form the basis of certain other characteristics that conflict with these characteristics.

In their association with the category of number, for example, their primary function of denoting speech roles would make it sufficient for them to have ‘non-singular’ forms that denote the possible combination of speech roles, such as 1 + 2, 1 + 3, and 2 + 3, but many languages have additional non-singular forms whose function is clearly to indicate the ‘plurality’ of their referents. Generally these forms represent the plurality of third person referents, but they may also indicate the plurality of second person referents.

Languages may differentiate between these two types of ‘number’ distinctions by using distinct devices for representing them or they may allow them to be represented ambiguously with the help of a single set of number markers. Languages may also restrict the association of personal pronouns with the category of number for indicating only one of these notions, namely conjunction, as we have seen in the case of languages like Chalcatongo Mixtec, Kwakiutl, and Acehnese (§4.2.1).

Zewen (1977: 50) points out, for example, that in Marshallese (Micronesian), personal pronouns differentiate between two different types of forms, called plural and ‘collective’. The emphatic singular and plural forms are shown in (10).

(10)	Person	Singular	Plural
	1	<i>ŋa</i>	<i>kim</i> (Excl)
	1 + 2		<i>kij</i> (Incl)
	2	<i>kwe</i>	<i>kom</i> ^o
	3	<i>e</i>	<i>ir</i>

According to Zewen (1977), these emphatic personal pronouns (and also the non-emphatic object pronouns) can take the numeric morphemes *ro* ‘dual’, *jil*

‘trial’, *eaŋ* ‘quatral’, and *wij* ‘universal plural’ to form ‘collective’ forms. It is perhaps possible to assume that in this language, ‘plural’ forms denote speech role combinations (conjunction), except in the case of *ir* ‘3PI’, whereas the forms derived by adding numeric morphemes to plural forms indicate number distinctions occurring among their referents. We can then regard the former as only ‘implying’ the relevant number distinctions (that occur among their referents) while specifying different combinations of speech roles.

4.2.4 *Combinations of pronouns*

What I am suggesting here is that number marking represents two different notions in the case of personal pronouns, namely conjunction in the case of some forms like 1 + 2, 1 + 3, and 2 + 3 and plurality in the case of others (like 2 + 2 and 3 + 3). Some languages like Mixtec, Kwakiutl, and Acehnese have markers only for conjunction whereas others have markers for both conjunction as well as plurality. Such a claim is supported by the occurrence of suppletive forms and also distinct number markers for denoting the conjunctive type of number distinction in several other languages (see §4.2.2). It is also supported by the occurrence of conjoined forms for representing some of the non-singular pronouns, especially the ones that denote 1 + 2 in several other languages.

For example, Kurukh (Dravidian: Hahn 1908) has a system of personal pronouns in which the first person plural pronoun is used for denoting the exclusive (1 + 3) meaning, whereas for denoting the inclusive (1 + 2) meaning, first and second person pronoun are combined together, as shown in (11). This is also true of Malto, another Dravidian language (Mahapatra (1979: 75).

(11)	Singular	Plural (Excl)	Plural (Incl)
1	<i>e:n</i>	<i>e:m</i>	<i>na:m</i>
2	<i>nī:n</i>	<i>nī:m</i>	

According to Krishnamurti (1968), this system goes back to Proto-Dravidian for which one can reconstruct (i) **yān* and **nīn* as the first and second person pronouns respectively, having **yām* and **nīm* as their respective plural forms, and (ii) **n-(y)ām* as the inclusive (1 + 2) form, in which the initial element **n-* represents the second person.

(12)	Singular	Plural (Excl)	Plural (Incl)
1	<i>*yān</i>	<i>*yām</i>	<i>*n(y)ām</i>
2	<i>*nīn</i>	<i>*nīm</i>	

The Proto-Dravidian forms **yān* and **yām* have changed into *e:n* and *e:m* respectively in Kurukh. This system has undergone different kinds of modifications in other Dravidian languages, such as, for example, the loss of the inclusive–exclusive distinction, as in Standard Kannada, or the occurrence of a nominal plural marker in the formation of the exclusive form, as in Tulu (Bhat 1967).

(13)	Singular	Plural (Excl)	Plural (Incl)
1	<i>e:ni</i>	<i>eŋkɭu</i>	<i>na:vu</i>
2	<i>i:</i>	<i>niŋkɭu</i>	

Among the Tulu pronominal forms given in (13) the inclusive form is retained (except for changing *-m* to *-vu*), but the exclusive form and also the second person plural form have their plural marker *-m* replaced by one of the nominal plural markers.

According to Bradley (1993: 178), an unambiguous way of indicating the inclusive meaning in Lahu (Tibeto-Burman) is to conjoin the second and first person pronouns (*nɔŋáhi* 'we (incl) plural', literally 'you-I-plural'). On the other hand, Nama Hottentot (Central Khoisan) forms its first person inclusive dual and plural forms by attaching the first person dual and plural suffixes to the second person singular pronoun (Hagman 1977: 44).

- (14) a. *saá-kxm* 'we (Dual, Incl, M)'
2Sg-1DualM
b. *saá-ke* 'we (Pl, Incl, M)'
2Sg-1PlM

This is also true of Plains Cree (Algonquian: Wolfart 1973: 39) in which the inclusive form is derived by adding the plural marker of first person (or rather a marker that partly represents first person and partly represents third person) to the second person stem, whereas the exclusive form is derived by adding the plural marker of third person to the same (second person) stem.

(15)	Singular	Inclusive	Exclusive
1	<i>nīya</i>	<i>nīyanān</i>	
2	<i>kīya</i>	<i>kīyānaw</i>	<i>kīyawāw</i>
3	<i>wīya</i>		<i>wīyawāw</i>

According to Davies (1981: 158), it is possible, in Kobon (Kalam family) to replace non-singular pronominal forms by pronominal forms that specify more precisely the identity of the referents. This appears to be an alternative device used by the language for indicating, more clearly, the underlying conjunctive meaning.

- (16) *hel* '1Dual' (i) *ne yad* '1Sg 2Sg' or *yad ne* '2Sg 1Sg'
(ii) *nipe yad* '3Sg 1Sg' or *yad nipe* '1Sg 3Sg'

By choosing either (i) or (ii), a speaker is able to remove the ambiguity that occurs in the use of the dual form *hel*.

4.2.5 Minimal and augmented forms

There is an interesting problem of analysis that has been noticed in the case of the non-singular forms of personal pronouns in some of the languages of

Arnhem Land (Australia) and Philippines. For example, the representation of the pronouns of Rembarrnga (McKay 1978, as referred to in Corbett 2000: 166) provides a disjointed occurrence of the 'dual-trial' affix *-bbarrah*, as can be seen in (17).

(17)		Singular	Dual	Trial	Plural
	1Excl	<i>ngunu</i>	<i>yarr-bbarrah</i>	—	<i>yarru</i>
	1Incl	—	<i>yukku</i>	<i>ngakorr-bbarrah</i>	<i>ngakorru</i>
	2	<i>kut</i>	<i>nakor-bbarrah</i>	—	<i>nakorru</i>
	3M	<i>navu</i> MF	<i>barr-bbarrah</i>	—	<i>barru</i>
	3F	<i>ngadu</i>			

Notice that the affix *-bbarrah* occurs as a 'dual' marker in the case of all pronouns excepting 1Excl; in the latter case, it occurs as a 'trial' marker. Further, the concept of 'trial' is restricted to this particular pronoun. In order to provide a more systematic account of number marking, McKay proposes to regard the inclusive dual form *yuk~~ku~~* '1 + 2' as a 'minimal' form, along with the singular pronouns 1, 2, and 3, and to consider the remaining dual forms along with the trial form of *yuk~~ku~~*, namely *ngakorr-bbarrah*, as 'unit augmented' and the plural forms as 'augmented'. This results in a better arrangement of the forms into paradigms (especially from the point of view of the suffix *-bbarrah*), as shown in (18).

(18)		Minimal	Unit augmented	Augmented
	1Excl	<i>ngunu</i>	<i>yarr-bbarrah</i>	<i>yarru</i>
	1Incl	<i>yukku</i>	<i>ngakorr-bbarrah</i>	<i>ngakorru</i>
	2	<i>kut</i>	<i>nakor-bbarrah</i>	<i>nakorru</i>
	3M	<i>navu</i> MF	<i>barr-bbarrah</i>	<i>barru</i>
	3F	<i>ngadu</i>		

McKaughan (1972) refers to the occurrence of a similar problem in Maranao, a Philippines language. It has four different plural forms corresponding to two singular forms. If we use the traditional approach for describing these forms, we obtain a defective paradigm as seen in (19).

(19)		Singular	Dual (Incl)	Plural (Excl)	Plural (Incl)
	1	<i>ako</i>	<i>ta</i>	<i>kami</i>	<i>tano</i>
	2	<i>ka</i>	—	<i>kano</i>	—

McKaughan (1972) suggests that the pronoun *ta* may be regarded as [—Plural] so that the pronouns can be represented as showing a compact system.

(20)	Person	—Plural	+Plural
	1	<i>ako</i>	<i>kami</i>
	2	<i>ka</i>	<i>kano</i>
	1 + 2	<i>ta</i>	<i>tano</i>

Similar suggestions to shift the 1 + 2 form to the column that contains singular pronouns have been made for languages like Tarma Quechua (Adelaar 1977: 218), Popolocan languages (Veerman-Leichsenring 2000: 322) and some additional American Indian languages (Mithun 1999: 218). Some, like Adelaar (1977) describe the 1 + 2 form as the ‘fourth person’. However, these analyses appear to be rather ad hoc because one cannot escape from the fact that the 1 + 2 form is unlike singular forms in that its reference is non-singular. If, on the other hand, we regard the inclusive form as representing a conjunction, its separation from dual and plural forms would be less ad hoc. Most of the Australian languages are two-person languages (see §6.2) in which third person pronoun is primarily a demonstrative. There are, therefore, only two distinct personal pronouns in these languages, and hence they need only a single form, namely ‘1 + 2’, to represent the conjunction of their personal pronouns.

4.2.6 Inclusive conjunction

Another type of construction that is considered to be rather puzzling occurs (i) in the case of conjunctive coordination that involves personal pronouns and also (ii) in that of comitative arguments that occur with personal pronouns. In most languages, an action performed jointly by either first or second person with a third person (or by first person with a second person) can be expressed in either of two ways. (i) One may use a coordinated noun phrase in which the two expressions are conjoined together such that they function as a single argument (21a) or (ii) one may use one of the participants as a distinct comitative argument (21b). One may also use a plural pronoun as in (21c).

- (21) a. *I and John went to the market.*
 b. *I went to the market with John.*
 c. *We went to the market.*

There are several languages that require the personal pronoun occurring in such constructions to be in its ‘dual’ or ‘plural’ form. That is, instead of denoting the ‘number’ of its own referents, the pronoun needs to indicate the number of the whole phrase (or of two different phrases put together). (22) of Toqabaqita, an Oceanic language, shows this characteristic in a coordinate construction (Lichtenberk 2000a: 2), whereas (23) of Erromangan (Vanuatu: Crowley 1998: 180) shows it in a sentence that contains a comitative (accompanitive) argument.

- (22) *kamareqa doqora-ku meki lae mai*
 1 (Dual-Excl) brother-1Sg will 1 (Dual-Excl) go
 qusungadi
 at.tomorrow
 ‘I and my brother will come tomorrow’

- (23) *kim-ndal r-apamu-hai-me kimu-velom*
 2PI-Acco PI-brother-2Sg-PI 2PI-came
 'You and your brothers came'

Notice that in (22) only two individuals are involved in the action of 'coming', but the pronoun *kamareqa* 'we (Dual-Excl)', occurring as one of the constituents of the coordinated noun phrase, is in dual and not in singular as in the English gloss. The pronoun appears to indicate the number of the whole coordinated noun phrase, which is dual, rather than the number of its own referent, which is only singular. Constructions of this type have been reported to occur throughout the Austronesian language family and also in several other languages that belong to Australian, Indo-European, Polynesian, and Mayan families (Haspelmath, forthcoming).

The occurrence of a dual or plural personal pronoun in these constructions, apparently for denoting a *single* individual, has led to a dispute among linguists, especially because other languages use only a singular pronoun in such constructions. Some consider the inclusory pronouns to be basically 'singular' whereas others argue that they are in no way different from dual and plural pronouns occurring in other contexts. For example, Schwartz (1988: 243) argues that the plurality of inclusory pronouns is internal to the construction, meaning, apparently, that it is not the same as the plurality that is expressed by agreement markers occurring with the verb. Schwartz (1988: 243) suggests that an explanation for this constraint may have to do with the heterogeneity of first and second person non-singular pronouns. That is, these non-singular pronouns represent heterogeneous groups that consist of a speech act participant and one or more persons who are not of the same status, whereas non-singular forms of other nouns represent a homogeneous grouping.

On the other hand, Lichtenberk (2000a: 8) argues that the inclusory pronoun encodes a set of which the noun phrase that occurs with it encodes a proper subset. That is, the inclusory pronoun, according to him, is the head of the construction and the included noun phrase is its modifier, an adjunct. The former indicates the plurality of the construction that is also expressed by the agreement marker occurring in the verb.

The solution suggested by Schwartz appears to be superior in that it can account for some of the constraints that affect the occurrence of inclusory constructions. For example, there is a preference for first and second person pronouns to occur as inclusory pronouns, with some languages like Tagalog restricting the construction to only these two pronouns (Schwartz 1988: 241). It is extended to third person pronouns in some languages, but the use of ordinary nouns as inclusory plurals is extremely rare. Haspelmath (forthcoming) mentions one such case, namely that of Margi, a Chadic language (C. Hoffmann 1963: 57).

- (24) *kàmbóràwázhá-'yàr àgá màlà gə̀ndà*
 Kamburawazha-Ass.PI with wife of.him
 'Kamburawazha and his wife'

Notice that this exceptional non-pronominal inclusory construction also involves a heterogeneous combination.

There are certain other types of inclusory constructions that could be problematic for an analysis that regards inclusory pronouns as in no way different from plural pronouns. For example, Tinrin (Melanesian: Osumi 1995:41) uses an inclusory pronoun along with a comitative noun phrase, but in addition to this the sentence also contains a non-inclusory (singular) subject marker.

- (25) a. *u fi komu nrî pwere numea*
 1Sg go 1Du (Excl) 3Sg to Noumea
 'I go to Noumea with him'

The construction containing the inclusory pronoun and the comitative noun phrase can be shifted to the sentence-initial position or to the final position.

- (25) b. *komu nrî u fi pwere numea*
 1Du (Excl) 3Sg 1Sg go to Noumea
 c. *u fi pwere numea komu nrî*
 1Sg go to Noumea 1Du (Excl) 3Sg

We can resolve this puzzle, of inclusory constructions with the help of the above-mentioned assumption (see §4.2.3), namely that number distinctions occurring among personal pronouns can represent either (i) a combination of speech roles (conjunction) or (ii) a combination of the individuals who perform those roles (plurality). In sentences that contain inclusory conjunction, the non-singular forms of personal pronouns indicate the conjunction of speech roles, whereas the plurality of the individuals concerned is represented not by the pronoun itself, but rather by the conjoining particle that occurs with it or by the agreement marker that occurs in the verb. Languages that do not have any inclusory constructions apparently do not consider it necessary to specify the conjunction of speech roles in such coordinated constructions or constructions involving a comitative argument.

4.2.7 Hierarchy of number marking

Corbett (2000: 55–132) discusses in detail the proposal of setting up an Animacy Hierarchy for number marking, put forth earlier by Smith-Stark (1974), in which the three personal pronouns are considered to occupy the top position.

- (26) 1 > 2 > 3 > kin > human > animate > inanimate

The placement of the first person pronoun on top of this hierarchy is based upon the occurrence of two languages, namely Kwakiutl (Boas 1911) and Acehnese (Durie 1985) mentioned above (§4.2.2), in which number marking is restricted to first person. On the other hand, the placement of second person pronoun in the second position is based upon the occurrence of two other languages, namely

Asmat (Papuan: Voorhoeve 1965: 142) and Guaraní (Tupi: Gregores and Suárez 1967: 140) in which number marking is restricted to first and second persons. The pronouns of Asmat are given in (27).

(27)		Singular	Plural
	1	<i>nor</i>	<i>nar</i>
	2	<i>or</i>	<i>car</i>
	3	<i>ar</i>	—

It appears to me that this positioning of first and second person pronouns on top of the number hierarchy in (26) is rather odd because, generally, the most prototypical members of a group are placed on the top position of a hierarchy. For example, in an accusative-ergative case-marking hierarchy, first and second person pronouns rightly belong to the top because their occurrence with accusative marking is the most prototypical one. In the case of the category of number, on the other hand, personal pronouns do not constitute prototypical instances because the notion of plurality has an extended use among them—a use that actually conflicts with its primary characteristic of disregarding the differences that occur among individuals, as I have pointed out in a previous section (see §4.2.2).

Corbett (2000) has some difficulty in justifying this animacy hierarchy. For example, he finds a particular variety of number-marking, called Associative plural, to be occurring only with personal names, kin terms, and certain other human nouns but not with personal pronouns. That is, Associative plural occupies the middle portion of the hierarchy and therefore forms an exception to (26). As an example of this possibility, Corbett refers to the case of Hungarian (based on Edith Moravcsik, personal communication) in which personal names, kin terms, and nouns denoting professions can occur either with a plural suffix *-ok* or an associative suffix *-ék*, whereas personal pronouns have only plural forms.

(28)	<i>jános</i>	‘John’	<i>jános-ok</i>	‘Johns’ (more than one called John)
			<i>jános-ék</i>	‘John and his associates’
	<i>apa</i>	‘father’	<i>apa-ék</i>	‘father and his group’
	<i>tanító</i>	‘teacher’	<i>tanító-ék</i>	‘teacher and his group’

Following a suggestion made by Moravcsik, Corbett examines the possibility of regarding personal pronouns also as representing Associative plural rather than ordinary plural. However, this would make the occurrence of ordinary plural exceptional for the hierarchy, because personal pronouns would, according to this reinterpretation, show Associative plural rather than ordinary plural. As a result, there would now be a gap in the occurrence of ordinary plural, as it would not extend up to the top (namely personal pronouns).

In order to avoid this latter problem, Corbett (2000: 107) proposes to regard ‘Associative plural’ as not forming part of the category of number. He considers such a proposal to be supported by the occurrence of a post-base (derivational

marker) in Central Alaskan Yupik (Jacobson 1995), namely *-nku* 'and family'. This post-base can be attached only to proper names, but it can be followed either by the dual suffix *-k* or the plural suffix *-t* (*cuna-nku-k* 'cuna and his friend', *cuna-nku-t* 'cuna and his family/friends'). The dual and plural markers apparently indicate the total number of persons that are denoted by the expression. Two other types of distinctions that are associated with the notion of number also appear to be exceptions to the animacy hierarchy. They are Distributives and Collectives, which occur with nouns but not with personal pronouns. Corbett proposes to regard these latter distinctions also as not belonging to the number category.

One can actually extend the argument that Corbett puts forth, in order to regard the Associative category as not belonging to the number category, to non-singular forms of personal pronouns that involve either suppletion or special number markers and denote different combinations of speech act participants. Such non-singular forms are similar to Associative forms in their ability to take dual and plural markers. We have already seen this situation occurring in the case of languages that have 'augmented' forms rather than plural forms (§4.2.5). Corbett (2000: 169) considers these cases as involving an alternative way of organizing the morphology of person and number, but does not say anything regarding its similarity with the Associative plural. Another type of situation occurs in Nivkh (Palaeo-Siberian: Gruzdeva 1998: 26). It has a 1 + 2 dual form, inclusive and exclusive first plural forms, and second and third person plural forms. The formation of these non-singular forms involves either suppletion or irregular affixation.

(29)		Singular	Dual	Plural
	1Excl	<i>n'i</i>		<i>n'yŋ</i>
	1Incl		<i>megi</i>	<i>mer</i>
	2	<i>či</i>		<i>čin</i>
	3	<i>if</i>		<i>imŋ</i>

All the non-singular forms given in (29) can occur with the plural suffix *-gu*. Similarly, Imonda (Papuan: Seiler 1985: 44) has a single suppletive form for denoting inclusive meaning, and all its personal pronouns, along with this inclusive one, can take *-id* 'men' to denote plural meaning.

Exceptions to the animacy hierarchy of number marking also result from the fact that there are some languages in which dual occurs among some nouns but not among personal pronouns. According to Plank (1989: 297) several languages of the Semitic family, and also Irish and possibly Polish, manifest such an occurrence of the dual. Plank points out, however, that these are diachronically unstable. On the other hand, Hopi appears to have newly developed a dual that is restricted to some nouns.

Plank refers to an interesting distinction proposed by Humboldt (1830) between languages that restrict the dual (i) to pronominals, (ii) to nominals

(in order to represent natural pairs), and (iii) languages that allow it (dual) to occur with nominals as well as pronominals. In the last case (iii), the category would be of an all-pervasive type. Plank (1996) is a more detailed study of these exceptional duals. He finds 26 languages (in a sample of 205) to have a dual that occurs only with some nouns, and in 12 of these the dual does not occur with personal pronouns. Among the remaining languages, in which dual is restricted to personal pronouns and to some nouns, the latter (nouns) include primarily animate nouns.

Thus, the animacy hierarchy, as represented in (26) is defective not only in placing non-prototypical members on top, but also in having several exceptional cases to deal with. In view of these problems, it is perhaps better to regard the hierarchy as involving two different criteria rather than a single one, and to split it into two different levels. We can regard the notion of conjunction as moving down from the top of the hierarchy, and plurality as moving down from third person pronouns. Dual will have to be regarded as representing (i) conjunction in the case of personal pronouns (especially when it is represented by distinct number markers or by suppletion) and (ii) number distinction in other cases. The restricted use of duality to denote 'animate pairs' can perhaps be viewed as an extension of its use for denoting conjunction.

(30) a. 1 > 2 >

Conjunction →

3 > kin > human > animate > inanimate

Plurality →

It is now possible to regard different types of plurals like associative, collective, and distributive, which do not extend up to personal pronouns, as occurring in the second level of the hierarchy starting from the third person pronoun or kin terms. They are connected with the notion of plurality in the sense that they do not differentiate between individual members that are regarded as associates or collections, or entities that are distributed over a particular location. On the other hand, the notion of inclusory conjunction can be added to the first level of hierarchy as it forms part of the notion of conjunction. Its preferred occurrence is with first and second persons but it may be extended to third person and also to nouns involving differentiation.

(30) b. 1 > 2 >

Conjunction (animate dual) →

Inclusory construction →

3 > kin > human > animate > inanimate

Plural →

Associative →

Collective →

Distributive →

4.3 Association with the category of gender

Gender distinction is generally absent in first and second person pronouns. Most of the languages that manifest gender or noun class distinctions among their pronouns restrict their occurrence to third person pronouns and to proforms like demonstratives. In my sample of 225 languages, for example (see §6.1), 62 show gender (or noun class) distinction in the third person, of which only 10 extend it to second person singular but none to first person singular. There is one language, namely Iraqw (Cushitic: Mous 1993: 112) in which the distinction occurs in second person singular pronoun, but is not shown in third person.

(31)	Singular	Plural
1	<i>aníng</i>	<i>atén</i>
2M	<i>kúung</i>	MF <i>kuungá'</i>
2F	<i>kíng</i>	
3	<i>inós</i>	<i>ino'ín</i>

An example for the more common occurrence of gender distinction that is restricted to third person pronouns is Kannada, a Dravidian language (32). It shows a three-fold masculine–feminine–neuter gender distinction in third person pronouns (which is two-fold in plural) and demonstratives, but there is no gender distinction in first and second person pronouns.

(32)	Singular	Plural
1	<i>na:nu</i>	<i>na:vu</i>
2	<i>ni:nu</i>	<i>ni:vu</i>
3M	<i>avanu</i>	3MF <i>avaru</i>
3F	<i>avaḷu</i>	
3N	<i>adu</i>	<i>avu</i>

There is also an interesting correlation between gender marking among third person pronouns on the one hand, and the distinction between ‘two-person’ and ‘three-person’ languages, as I point out in the sixth chapter (§6.3). The majority of languages that show gender distinction in third person are two-person languages, that is, languages in which third person pronouns are regarded as part of the system of demonstratives rather than that of personal pronouns.

4.3.1 Restricted occurrences

Languages that extend gender distinction to second person pronouns are very few and most of them belong to the Afro-Asiatic language family. In my sample of 225 languages, only 10 (out of 80 that show gender distinction) extend it to second person singular. Eight of these belong to the Afro-Asiatic family, whereas one (Khmu: Premrirat 1987) belongs to Mon-Khmer, one (Ambulus: Wilson 1980) to the Ndu family (Middle Sepik stock, Papuan) and one (Tunica: Haas 1940)

is an isolate (originally spoken in Louisiana). In Pero, a Chadic language, for example, there is masculine–feminine gender distinction in second and third person singular, but not in first person or in plural (Frajzyngier 1989: 122).

(33)		Singular		Plural
	1	<i>nè</i>		<i>minù</i>
	2M	<i>kè</i>	MF	<i>mà</i>
	2F	<i>cì</i>		
	3M	<i>cákkà</i>	MF	<i>cínù</i>
	3F	<i>tè</i>		

There are two additional languages, namely Nunggubuyu (Heath 1984: 242) and Spanish (Green 1988: 95), in which gender distinction gets extended to second person only in non-singular (dual and trial only in Nunggubuyu). These languages extend the distinction to first person as well in these contexts. However, languages that extend gender distinction to first person are rather rare. Further, such languages extend it to first person only in non-singular but not in singular, as in Ntifa, a Berber language of central Morocco (Loust 1918: 211, quoted in Plank and Schellinger 1997: 65).

(34)		Singular		Plural
	1MF	<i>nkí(n)</i>	M	<i>núkni</i>
			F	<i>núkěnimti</i>
	2M	<i>kií(n)</i>		<i>kúnni</i>
	F	<i>kemmí(n)</i>		<i>kúnimti</i>
	3M	<i>ntá</i>		<i>nútni</i>
	F	<i>ntat</i>		<i>nútěnti</i>

The pronouns of first and second person also show certain additional idiosyncratic characteristics in their gender marking that keep them apart from proforms and nouns. For example, the markers that are attached to them for indicating gender distinctions may be different from the ones that are used in proforms and nouns, as for example, in Afro-Asiatic languages (Lipiński 1997: 301, Gensler MS).

4.3.2 *Explanations for irregularity*

This irregularity occurring in the association of first and second person pronouns with the category of gender resembles to a certain extent the irregularity that occurs in their association with the number category that we have examined in the previous section (§4.2). And as in the previous case, it is possible that the purpose for which gender distinction is used in first and second person pronouns is also different from the one for which it is used in proforms and nouns. In the latter case, the distinction is used primarily for removing some of the ambiguities that occur in their reference or coreference. That is, its use is primarily associated with

the function of identifying the referents or establishing coreference with the proper antecedents. In the case of first and second person pronouns, on the other hand, gender distinction appears to be used primarily for indicating social distinctions or for complying with social requirements. That is, it appears to represent distinctions in the speech roles themselves, rather than in the performers of those roles.

The irregular occurrence of gender distinction among personal pronouns can also be due to factors that have nothing to do with the primary function of those pronouns. For example, Plank and Schellinger (1997: 69) point out that one of the diachronic sources of restricted gender distinctions appearing in first and second person pronouns is the grammaticalization of adjectives (nouns) and numerals that occurs in apposition with pronouns. In Spanish, for example, simple plural forms of first and second person pronouns were replaced by complex structures that contained the relevant pronoun and the adjective meaning 'other'. These adjectives brought gender distinction along with them, as can be seen in the second person plural forms, *alteros* '2MPI' and *alteras* '2FPI'. Lithuanian, on the other hand, has grammaticalized its appositive constructions that involved the numerals for 'two'. These constructions showed gender distinction and hence the language has come to possess a system of personal pronouns in which gender distinction in first and second person occurs only in dual (Plank and Schellinger 1997: 71).

(35)	Singular	Dual	Plural
1MF	<i>às</i>	M <i>mùdu</i> F <i>mùdvi</i>	MF <i>mēs</i>
2MF	<i>tù</i>	M <i>jùdu</i> F <i>jùdvi</i>	MF <i>jūs</i>
3M	<i>jīs</i>	<i>jiēdu</i>	<i>jiē</i>
F	<i>jì</i>	<i>jiēdvi</i>	<i>jūs</i>

4.3.3 Politeness distinctions

The occurrence of politeness distinctions primarily among the pronouns of first and second person in several East Asian languages like Thai, Burmese, Japanese, etc. also appears to support the view that distinctions occurring in these pronouns are meant for establishing distinctions in the speech roles that they represent rather than for identifying the persons who perform those roles. The diachronic sources of these distinctions also appear to support such a view.

According to Nguyen (1997: 123), for example, Vietnamese has the first person pronouns *tôi* (derived from the noun denoting 'servant') used in contexts of modesty and respect and *tao* 'term of arrogance used by superiors'. It also has several second person singular pronouns such as *anh* 'used with a young man', *mày* 'term of arrogance', *ông* 'used with a gentleman', *bà* 'used with a lady', *cô* 'used with a young lady', and *bác* 'a polite term'. Many of these pronouns also

function as kinship terms, like *ông* 'grandfather', *bà* 'grandmother', *cô* 'father's younger sister', and *bác* 'father's older brother'. These can be used for denoting the relevant kinship relations as well.

Similar distinctions have been reported to occur in other East Asian languages as well. In Japanese, for example, several different nominal forms are used for denoting first and second persons, with the choice depending upon stylistic (formal–non-formal), performative (humble, differential), and other sociolinguistic factors. Thai has an even larger set of forms involving distinction of status, intimacy, and non-restraint, closely linked with other factors like politeness or respect, and also assertiveness. Some of these are based on objective facts whereas others are subjective, expressing the speaker's attitudes (Cooke 1968).

Notice, however, that all these distinctions are directly connected with speech roles. That is, the speech communities regard the speech roles of 'being a polite speaker' as distinct from that of 'being an impolite speaker', or the speech role of 'being an honoured addressee' as distinct from that of 'being an addressee of the same or lower status'. The various forms occurring as first and second person pronouns are clearly meant for denoting these distinct sub-types of speech roles. They do not appear to have the function of providing information for identifying the individuals who perform those roles.

There is an interesting distinction between first and second person pronouns, on the one hand, and those of third person, on the other, in the kind of lexical items that replace them. Bradley (1993) points out that the former tend to get replaced by denominals whereas the latter tend to get replaced by pronominals (demonstratives). In Burmese, for example, the current first person pronouns are derived from the word *cun* 'slave' + the royal or respectful suffix *fɔ̌* for male speaker form and the female suffix *má* for the female speaker form. The current second person pronouns, on the other hand, are derived from religious titles. The male form is probably a contracted version of *θəkhin phəyā* 'Bodhisattva' and the female form derives from the word *shin* 'master' (Bradley 1993: 158). The tendency to replace third person pronouns by demonstratives, on the other hand, occurs in several Tibeto-Burman languages as well.

This differentiation, however, is probably true of 'two-person' languages (see §4.2) but not of 'three-person' languages. The latter may use nominal stems rather than demonstratives for deriving their third person pronouns. For example, Yosondúa Mixtec (Farris 1992: 136), a three-person language, derives its third person pronouns from general nouns like *nāhā* 'woman', *kĩtĩ* 'animal', *yivĩ* 'person', etc.

4.4 Association with the category of case

The association of personal pronouns with the category of case is generally considered to be the most important characteristic that supports their inclusion

under the nominal category. However, pronouns of first and second person show their own unique characteristics even in this association with the category of case. There are actually some points in which they resemble other proforms and nouns in this association but others in which they differ. We can regard the former as deriving from the fact that they share with proforms and nouns the necessity to associate themselves with case roles. They need to do this in order to relate the speech roles that they represent with case roles (see §2.1). Their resemblance on this point is primarily with nouns and proforms that denote human beings.

On the other hand, they differ from other nominals on two important points. (i) There are several languages in which the involvement of first and second persons in an event or state is indicated jointly with the help of a single element called 'sagittal', instead of using distinct markers for the agent and the patient. It is possible that this special type of case marking derives from the primary function of personal pronouns suggested earlier (§1.2), namely that of denoting speech roles rather than the individuals who perform those roles (§4.4.1). (ii) Most languages give greater prominence to speech act participants as compared to other participants. As a result, first and second person pronouns occur at the top of the hierarchy of nominals as far as case marking and certain other aspects of grammar are concerned. In some languages, personal pronouns are the only ones that are marked for case. Proforms and nouns may take the adpositions or affixes that denote distinctions of location and other similar peripheral aspects of the sentence but are otherwise unmarked for case (§4.4.2). There are also several languages in which the pronominal elements that indicate first and second persons are kept in the prominent position irrespective of the nature of their participation. These languages use other devices like marking the verb as inverse or passive for indicating this latter distinction (§4.4.3). There are also languages that do not allow the use of *bound* forms (i.e. non-emphatic forms) for first and second person pronouns, if they are in a lower position in a case hierarchy (§4.4.4).

Since the notion of 'prominence' is a relative concept, the differences that derive from the second factor mentioned above, unlike the differences that derive from the primary function of personal pronouns described earlier, give rise to a gradation rather than a clear-cut distinction. For example, as I will be pointing out in the next chapter (§5.2.3), several languages differentiate between personal pronouns and nouns by using the accusative system in the former case and ergative system in the latter case. However, the distinction involves, cross-linguistically, a hierarchy of nominal categories with the cut-off point for the accusative or ergative marking falling on different places on the hierarchy for different languages.

4.4.1 Sagittal case marking

We have seen earlier (§4.2) how personal pronouns differ from other nominals in not having a one-to-one correlation between singular and plural forms.

The correlation is more complex, with one singular form relating with two plural forms (inclusive and exclusive), two singular forms (of first and second person) relating to one plural (inclusive) form, three singular forms relating to four plural forms, etc. There is a similar kind of disparity between first and second person pronouns on the one hand and other nominals on the other in their association with the category of case as well. Several languages manifest very complex systems of case marking in which the complexity is restricted to personal pronouns. This complexity derives from the fact that case markers are used among these pronouns in order to indicate the *relation* between speech roles rather than the actual case roles in which the speech roles (or the individuals who enact those roles) occur.

This interesting characteristic of case marking is shown by languages in which pronominal elements occur as affixes or clitics in the predicate rather than as independent arguments. Such languages may place the affixes at two different positions in the predicate or keep them together in a single location. In Kambera (Central Malayo-Polynesian: Klamer 1998: 62), for example, subject markers precede the verbal stem whereas object markers follow the stem.

- (36) *na-palu-ka*
 3Sg-hit-1Sg
 'He hit me'

On the other hand, there are languages like Maricopa (Hokan), in which the two affixes occur together. In such languages, there is a possibility of the two affixes merging together into a single un-segmentable element. In Maricopa, for example, the verbal form that has a second person participant as the subject and a first person participant as the object allows the pronominal marker that occurs before the verbal stem to be segmented into a subject prefix and an object prefix, but the one that has a first person subject and second person object does not. Maricopa does not have any marker for third person (Gordon 1986: 18).

- (37) a. *ʔny-m-wik-k*
 2 > 1-help-Realis
 'You helped me'
 b. *ny-wik-k*
 1 > 2-help-Realis
 'I helped you'

Languages may show very complex systems of pronominal affixes that derive from such fusions of subject and object markers. In Oneida (Iroquoian: Abbott 2000: 20), for example, such fused affixes differentiate between three numbers (with inclusive-exclusive distinction in first person), and four genders in third person. Abbott lists 58 different prefixes of this type for this language. Postal (1979) provides a detailed analysis of a similar set of pronominal prefixes occurring in another Iroquoian language, namely Mohawk. What is interesting in

the present context is that this tendency to fuse subject and object affixes or clitics into un-segmentable elements occurs most predominantly in the case of first and second person pronouns.

According to Jacquesson (2001: 138), this denotation of $1 > 2$ ('I verb you') or $2 > 1$ ('you verb me') by a single morpheme may involve a problem of selection, with either the morpheme for 1, or for 2, being selected for that purpose, according to the hierarchy at work in the language (see §5.2). There are also languages that use special morphemes for this purpose. Jacquesson suggests that such morphemes do not encode persons separately but a specific and oriented *relationship* between persons. The term 'sagittal' was coined by Hegège (1982) for referring to this special morpheme; it emphasizes the fact that the movement from one person to another is here more central than the persons themselves (sagittal is from Latin *sagitta* 'arrow'). The occurrence of such sagittal morphemes make it clear, according to Jacquesson (2001: 139), that persons are not only like facts and figures but also like relations, and person systems are like structures.

Notice, however, that this unique characteristic is restricted to personal pronouns, and this restriction can be accounted for more satisfactorily on the basis of the important functional difference that exists between personal pronouns and other nominals; the former denote speech roles rather than the individuals who perform those speech roles, whereas the latter denote (or refer to) the individuals (§1.2).

4.4.2 Prominence of speech act participants

First and second person pronouns manifest several other differentiating characteristics in their association with the category of case. These differences appear to have an entirely different basis, namely the greater prominence that speech act participants receive as compared to other participants. For example, several languages use ergative case marking for some nominals and accusative case marking for others. The cut-off point for the use of either of these two markers differs from language to language. This cross-linguistic difference appears to derive partly from the prominence of speech act participants and partly from the possibility of establishing a 'hierarchy' of nominal categories, as I point out in the next chapter (§5.2.3).

Even in the case of languages in which the same set of case markers is used among personal pronouns as well as nouns, there are generally some constraints that affect the occurrence of case markers with personal pronouns. For example, Manipuri (Tibeto-Burman: Bhat and Ningomba 1997: 70) requires a special marker *-non* is to be added to personal pronouns before adding the locative case marker. Kannada (Dravidian) extends such a constraint about the use of locative case suffix to all nominals that denote human beings. It denotes possession rather than actual location when attached to words that denote human beings whereas when attached to other nouns it denotes the actual location. In order to indicate

the actual location in the former case, one has to use more complex constructions like *nanna kayyalli* ‘in my hand’ as shown by (38c).

- (38) a. *nann-alli ondu pustaka ide*
 I-Loc one book is
 ‘I possess (have) a book’
 b. *marad-alli ondu hakki ide*
 tree-Loc one bird is
 ‘There is a bird on the tree’
 c. *nanna kayy-alli ondu pustaka ide*
 my hand-Loc one book is
 ‘There is a book with me’

Constraints regarding the use of the locative with expressions having human reference have been reported to occur in several other languages. They also appear to involve a hierarchy of nominals. For example, according to Dixon (1977: 169), locatives occur in Yidiny (Australian) predominantly with inanimates, less with non-human animates, and not at all with humans (including personal pronouns). Dixon points out that as the continuum of inanimate–human is gradually ascended, so the chances of locative being used decreases, and the probability of dative increases proportionately.

4.4.3 *Direct–inverse marking*

Some languages have developed an interesting way of maintaining the prominence of first and second person participants while representing, in the predicate, their occurrence in different case roles. Actions in which the speaker or the addressee occurs as an agent are marked ‘direct’ and the ones in which one of them occurs as the patient are marked ‘inverse’. In both these instances, however, the pronominal markers that represent the speaker or the addressee are placed in the prominent (initial) position. In Algonquian languages, for example, predicates that involve first or second person as a participant would have first or second person prefix, irrespective of whether the participant is an agent or a patient. However, predicates with (i) first or second person agent and third person patient would be differentiated from the ones that have (ii) first or second person patient and third person agent by marking the former as direct and the latter as inverse. Sentences (39a–b) of Cree (Algonquian: Wolfart 1996: 412) exemplify this usage.

- (39) a. *ni-wa.pam-a.w*
 1Sg-see-Direct.3
 ‘I see him’
 b. *ni-wa.pam-ik*
 1Sg-see-Inverse.3
 ‘He sees me’

The first person prefix *ni-* is placed before the verb in (39b) in spite of the fact that the speaker is not the agent but the patient. However, the use of the inverse marker *ik* in (39b) is able to indicate this difference between (39a) and (39b). Goddard (1967) reconstructs the suffixes **a* ‘direct’ and **ekw* ‘inverse’ for Proto-Algonquian.

This use of the direct–inverse distinction for maintaining the prominence of speech act participants can be regarded as somewhat similar to the use of the passive device for keeping the speaker or the addressee in a prominent (subject) position. Mithun (1999: 226) refers to this latter use in Picurís (Kiowa-Tanoan: Zaharlick 1982) that resembles, even more prominently, this direct–inverse distinction. In that language, passive is obligatory when a third person acts on first or second. On the other hand, when a first or second person acts on the third, the non-use of the passive is obligatory. It is only when both the arguments are third person that one can use either active or passive.

Apurinã (Maipur/Arawak) makes a direct–inverse distinction among its relative clauses. Facundes (2000: 582) points out that in that language, a restrictive relative clause of transitive verbs takes the inverse marker *keru* or *kero* if the subject of the clause is third person and the object is first or second person singular.

- (40) a. *sutowakoro n-atama-ta-kunu kona pita atama-ta*
 women 1Sg-see-Verbalizer-Rel (Direct) not 2Sg see-Verbalizer
 ‘The women whom I saw you didn’t see’
- b. *kuku n-atama-ta-keru apo-pe*
 man 1Sg-see-Verbalizer-Rel (Inverse) arrive-Perf
 ‘The man who saw me has arrived.’

In (40a), the relative clause has the agent of that clause as the initial element, whereas in (40b), the initial element indicates the patient. In both these clauses, prominence has been given to first person in contrast to third person.

4.4.4 Constraint on the occurrence of bound themes

The tendency among languages mentioned in the previous section (§4.4.3) is apparently to place the bound occurrences of first and second person pronouns in the most prominent position. There is a comparable tendency that is manifested by several languages regarding the use of weak pronouns (clitics or affixes) for denoting the case roles of theme and recipient. Haspelmath (MS) points out that when clusters of such weak pronouns are used for denoting the roles of theme and recipient, the theme may not be first or second person when the recipient is third person. The Spanish sentences (41a–b) exemplify this constraint regarding the use of clitics. The language apparently regards the recipient to be higher in the case hierarchy as compared to theme.

- (41) a. *Miguel me lo recomendó*
 Miguel 1Sg 3Sg recommended
 'Miguel recommended it to me'
- b. **Miguel me le recomendó*
 Miguel 1Sg 3Sg recommended
 'Miguel recommended me to him'

Languages use pronominal forms that are not weak or cliticized in order to indicate the meanings of sentences like (41b). The Spanish sentence (41c) exemplifies this alternative.

- (41) c. *Miguel me recomendó a él*
 Miguel 1Sg recommended to 3Sg
 'Miguel recommended me to him'

4.5 Summary

I have pointed out in this chapter (and in other chapters of this first part) a number of characteristics that can be ascribed to first and second person pronouns, in contrast to other pronouns and nouns. I have argued that most of these characteristics derive from the primary function of these pronouns, namely that of denoting speech roles. As I have suggested earlier (§2.2), first and second person pronouns tend to be dissociated from the actual individuals who perform those speech roles because in conversations among two or more individuals, the performers of speech roles constantly shift among different individuals; a speaker of one sentence would become the addressee of the next sentence and the speaker again in the next. If personal pronouns are to indicate speech roles consistently, they need to be unaffected by these shifts. That is, they cannot be indicators of speech roles and identifiers of speech act participants at one and the same time. Languages generally use personal pronouns for performing only the former function.

This inability of personal pronouns to identify the actual persons who perform speech roles gives rise to several of their interesting characteristics. For example, in some of the contexts in which such persons need to be identified, languages have to use different types of supporting devices like the use of a noun in the vocative, the use of appositional noun phrases, the use of matrix clauses, and so on. Personal pronouns do not directly take any markers or attachments that have the function of assisting the addressee in identifying their referents. They do not take definite or indefinite articles or determiners for this purpose. They also do not occur with modifiers and complements. Affixes that are meant for providing information regarding their referent(s) are also generally kept away from them. On the other hand, some of the affixes that do have such a function are

found to alter their function when used with personal pronouns. For example, gender and number markers, when used with personal pronouns, appear to indicate distinctions in the speech roles rather than in the individuals who perform those roles. Definite articles indicate emphasis rather than definiteness.

The fact that personal pronouns have speech act participants as their referents gives rise to the remaining characteristics that differentiate them from other nominals. Languages give greater prominence to speech act participants as compared to other types of participants. This tendency gets reflected in the occurrence of various types of preferential treatments that first and second person pronouns receive, as, for example, in case marking, verbal agreement, and so on. Personal pronouns also show conflicting characteristics because of their dichotomous association with reference and referent, as I point out in the next chapter. Many of these unique characteristics of first and second person pronouns appear to make it necessary to regard them as forming a distinct category, in spite of the fact that the category consists of a very small number of lexical items.

5 Conflicting Characteristics

5.1 Introduction

The characteristics of first and second person pronouns that we have examined so far showed them to be undifferentiated from one another but differentiated from proforms and nouns. These characteristics derive from their primary function of denoting speech roles and indicating their involvement in the event or state that the predicate denotes. Since this primary function is common to both first and second person pronouns, the characteristics that derive from it occur uniformly in the case of both first and second person pronouns.

In contrast to these shared characteristics, there do occur certain other characteristics that differentiate between first and second person pronouns. These derive from certain other factors such as (i) the prominence of the speaker as compared to the addressee, as for example when the speaker is making a statement or giving an order, (ii) the prominence of the addressee as compared to the speaker, as for example when the speaker is asking a question. Languages also differ from one another in generally giving more prominence to either the first or the second person (§5.2). Differentiating characteristics that derive from these factors may appear to go against the main claim of this monograph, namely that personal pronouns and other pronouns (proforms) belong to two different word classes. This is due to the fact that these latter characteristics have the effect of showing either the first person pronoun or the second person pronoun to be similar to other pronouns and nouns on points on which the two personal pronouns are differentiated from one another.

I have been emphasizing, in the preceding sections, the characteristics of first and second person pronouns that derive from the necessity to *dissociate* them from their referents. Such an emphasis has been necessitated by the fact that this particular aspect of personal pronouns has not received sufficient attention among linguists. Personal pronouns also have an *association*, though rather distant and indirect, with their referents, and this association also induces certain characteristics in them. There are several ingenious ways that languages make use of in order to keep these two functions apart, but there do occur instances in which they co-occur among personal pronouns, and the characteristics that derive from such a co-occurrence has the effect of making personal pronouns appear Janus-like (§5.3).

5.2 Prominence of speech roles

Several languages are reported to give greater prominence to first and second person pronouns, as compared to those of the third, while representing them in different spheres of grammar. This characteristic is considered to derive from the fact that the speaker's centre of interest is primarily in the speech act participants (Mallinson and Blake 1981: 81). For example, in several languages that belong to different language families, verbal agreement for agent and patient is determined not only by the case roles themselves, but also by the fact as to whether or not the case roles are connected with speech roles. In Tangut (Tibeto-Burman), for example, verbal agreement occurs only in case roles that belong to first and second persons but not in case roles that belong to third persons. The language shows agreement for the possessor of the subject or object as well, but even here the agreement is restricted to first and second person possessors (DeLancey 1983: 104).

When personal pronouns develop bound forms or change into agreement markers, the general tendency is for the first and second person pronouns to develop such forms to begin with, and it is only at a later stage that the third person pronouns also develop such forms. In support of this generalization, Mithun (1988: 86) points out that the majority of the languages indigenous to North America contain only first and second person bound forms. Even in the case of languages that have bound forms for all persons, there is internal and/or comparative evidence that the third person markers were grammaticalized at a different time than the markers of first and second persons. For example, in the languages of the Algonquian group, first and second persons are specified by pronominal prefixes on verbs, whereas third persons of various types are specified by pronominal suffixes (Mithun 1988: 87). The paradigm (1) of Cree exemplifies this usage.

- (1) *nī-nēhiyawān* 'I speak Cree'
kī-nēhiyawān 'you speak Cree'
nēhiya-wēw 'he/she speaks Cree'

In the case of transitive verbs, on the other hand, two different agreement markers might occur together, such as the ones that denote the agent and the patient. First and second person pronouns differ from those of third in having fused forms for these pairs of agreement markers. In Caddo, for example, the agent marker for first person is *ci-* and for second person it is *si-*, but for a first person that acts on the second, the marker is *t'a-*. Similarly in Lakshota, the marker for first person agent is *wa*, second person agent *ya-*, but for first person agent acting on second person patient, it is *či-* (Mithun 1988: 87).

Several languages treat first and second persons uniformly as against third persons in such contexts, but there do occur several other languages that differentiate between the first and the second by giving prominence to one of them. The general tendency is to give prominence to first person (§5.2.1), but there do occur languages in which second person receives greater prominence as compared to first person (§5.2.2). This differentiation between first and second person pronouns, based on the speaker's centre of interest, can make it appear that first and second person pronouns form a gradation along with third person pronouns, proforms, and also other nominals (§5.2.3).

5.2.1 *Prominence of the speaker*

In the case of most languages, first person receives greater prominence than second person. Whenever there is a conflict between the two, and a language has to choose either of them for preferential treatment, it is generally the first person that wins over the second. We can regard this preferential treatment of first person as deriving from the prominence that speakers attach to the 'ego', that is, to themselves. As suggested by Kuno (1976: 433), it is easier for a speaker to 'emphathize' with himself, and to express his own point of view, than that of the addressee or someone else. Different types of examples concerning the use of first and second person pronouns can be given in support of this generalization.

For example, when an expression has to indicate a combination of first and second persons or of first and third persons, languages generally use the dual or plural forms of the first person pronoun rather than that of the second or third. The pronouns of Djaru (Australian: Tsunoda 1995: 64) given in (2) exemplify this point.

(2)	Singular	Dual	Plural
1 (Exclusive)	<i>ŋadu</i>	<i>ŋadura</i>	<i>ŋanimba</i>
1 (Inclusive)		<i>ŋali</i>	<i>ŋaliba</i>
2	<i>ŋundu</i>	<i>ŋunbula</i>	<i>ŋunra</i>
3	<i>ŋandu</i>	<i>ŋanbula</i>	<i>(ŋandu)</i>

Notice that the 'inclusive' dual form *ŋali* '1 + 2' involves a conjunction of first and second persons, whereas the inclusive plural form *ŋaliba* '1 + 2 + 3', involves a conjunction of all three persons. However, the pronominal stem that is used for deriving both these forms is that of *first* person (*ŋa-*) and not of second or third person (*ŋun-* or *ŋan-*).

Similar tendency occurs in the agreement system as well. When coordinated phrases consisting of first and second person pronouns are used in the subject position, most languages that have a verbal agreement system would use verbal forms in the dual or plural of first person, rather than that of second person. The Kannada (Dravidian) sentence (3) exemplifies this usage.

- (3) a. *na:n-u: ni:n-u: ondu sinema: no:didd-evu*
 1Sg-also 2Sg-also one cinema saw (Perf)-1Pl
 'I and you had seen a cinema'
 b. *na:n-u: avan-u: ondu sinema: no:didd-evu*
 1Sg-also 3Sg-also one cinema saw (Perf)-3Pl
 'I and he had seen a cinema'

It is only when the coordination does not involve first person that a second person plural agreement would be used in Kannada.

- (4) *ni:n-u: avan-u: ondu sinema: no:didd-iri*
 2Sg-also 3Sg-also one cinema saw (Perf)-2Pl
 'You and he had seen a cinema'

The prominence of first person as against the second is shown in several other ways. In Nocte (Tibeto-Burman: Das Gupta 1971), for example, when both first and second person pronouns occur as arguments in a transitive sentence, the verbal agreement is with first person, irrespective of whether the first person pronoun occurs as the agent or the patient. The only difference is that when first person is the agent, the agreement is with first person plural.

- (5) a. *nga-ma nang hetho-e*
 I-Erg you teach-1Pl
 'I will teach you'
 b. *nang-ma nga hethoh-ang*
 you-Erg I teach-1Sg
 'You will teach me'

There is also, apparently, an inverse marker *h* added to the verb in (5b) where the agent is second person (see §4.4.3).

Languages also manifest this prominence of first person as against the second in the relative ordering of affixes that denote them. In Tutelo (Siouan), for example, when the actor and the patient prefixes occur together in a transitive sentence, first person always precedes the second person, whatever the case role (actor, patient, or dative) that the prefixes denote (Oliverio 1997: 78).

- (6) a. *wa-yi-ki-ko:ha*
 1Sg-2Sg-Dat-call
 'I call you'
 b. *wai-yat-o-ya-ste:ka*
 1Pl-heart-Loc-2Sg-good
 'You loved us'

Helmbrecht (1999: 291) points out that the occurrence of a split between first and second/third persons in their agreement marking is more likely when the verb

expresses the intention of a participant or when it is marked for direct or sensory evidence. Among the several instances referred to by him in support of this claim is Tonkawa (Isolate: Hoijer 1933: 307). This language has two different verbal suffixes, *a:ha'a* and *a:*, that are used for expressing the strong will of the first person agent to perform the designated action. Verbal forms containing these two suffixes can be used only with first person.

- (7) a. *hewl-a:ha'a*
 catch-1.Volitional
 'I will catch him'
 b. *sa:ya yax-a:*
 1.emph eat-1.Volitional
 'I intend to eat it'

5.2.2 *Prominence of the addressee*

There are some languages, on the other hand, that are considered to be exceptional to the generalization made in the previous section. They appear to give greater prominence to second person as compared to first person. This greater prominence of the addressee manifests itself in these languages in different aspects of grammar, such as the formation of dual or plural inclusive forms, case marking, verbal agreement, coordination, etc. For example, some of the Algonquian languages form their plural inclusive forms from the pronominal base of second person rather than from that of the first person. This is true of Plains Cree, mentioned earlier (see §4.2.4), in which the inclusive form is derived by adding the plural marker of the first person to the second person stem. Wolfart explains this characteristic of Cree as resulting from the fact that, in that language, second person ranks higher than the first.

Algonquian languages also have a direct-inverse system of case marking, in which second person receives greater prominence as compared to first person. When both the core arguments of a transitive verb are speech act participants (speaker and addressee), the occurrence of the addressee as the agent and the speaker as the patient is regarded as direct and the reverse occurrence is regarded as inverse. The sentences (8a–b) of Plains Cree (Algonquian: Wolfart 1996: 400) exemplify this usage.

- (8) a. *ki-tasam-in*
 2Sg-feed (Direct)-1Sg
 'You feed me'
 b. *ki-tasamit-in*
 2Sg-feed (Inverse)-1Sg
 'I feed you'

Other Algonquian languages also give greater prominence to the second person pronoun in a similar fashion, and Goddard (1967) reconstructs for

Proto-Algonquian the direct and inverse markers **i* and **eθe(ne)* that are to be used for differentiating between (i) second person agent/first person patient (direct) and (ii) first person agent/second person patient (inverse) respectively.

Sanuma, a language belonging to the Yanomami family (Amazonian), uses its second person plural form ambiguously for denoting either 'second person plural' or 'first person inclusive plural'. Borgman (1989: 149) reports that only the context can decide whether its meaning includes first person or not. The plural suffix used, however, is the same in both first and second person pronouns.

- | | | | | |
|-----|---|---------------|--------------------|-----------------------|
| (9) | | Singular | Plural | |
| | 1 | <i>kamisa</i> | <i>kamisa-makö</i> | (Exclusive) |
| | 2 | <i>kaw</i> | <i>kamakö</i> | (Exclusive/Inclusive) |

The situation in Sanuma is rather different from the one existing in familiar languages like English. In the latter case, it is the first person plural form (*we*) that is used ambiguously for denoting either the first person exclusive meaning or the first person inclusive meaning.

Awa Pit (Barbacoan family: Curnow 1997: 187) has an interesting system of case marking in which prominence to first or second person is given depending upon the 'epistemic source' of the utterance. Verbs are marked as locutor or non-locutor depending upon whether the subject is first or second person. In statements, first person subject has the locutor marking, but in questions it has the non-locutor marking. In the latter case, second person receives locutor marking.

- | | | | | | |
|------|----|----------------------------|----------------|------------------|--|
| (10) | a. | <i>nana</i> | <i>pala</i> | <i>kumtu-s</i> | |
| | | I | plantain | eat-Locutor | |
| | | 'I am eating plantains' | | | |
| | b. | <i>nuna</i> | <i>pala</i> | <i>kumtu-y</i> | |
| | | you | plantain | eat-Nonlocutor | |
| | | 'You are eating plantains' | | | |
| (11) | a. | <i>shi = ma</i> | <i>nuna</i> | <i>kimtu-s</i> | |
| | | what = Int | you | do-Locutor | |
| | | 'What are you doing?' | | | |
| | b. | <i>mina = ma</i> | <i>na = na</i> | <i>ashaptu-y</i> | |
| | | who (Acc) = Int | I = Topic | annoy-Nonlocutor | |
| | | 'Whom am I annoying?' | | | |

5.2.3 Hierarchy of nominal categories

The occurrence of first and second person pronouns as arguments of sentences makes it relevant for them to take case markers. Since they share this characteristic with nouns and also with some of the proforms, there is a possibility of regarding personal pronouns, along with these other words, as forming a lexical

category of 'nominals'. There are, however, important differences in the way in which the various types of nominals get associated with case markers. Linguists have used these differences as the basis for establishing 'a hierarchy of nominal categories' (see Silverstein 1976, Comrie 1978, Dixon 1979, 1994, Mallinson and Blake 1981, Blake 1994).

This hierarchy is primarily based upon the case marking of core arguments, that is, the single obligatory argument (S) of intransitive sentences and the two obligatory arguments, namely the agent (A) and the patient or object (O) of transitive sentences. Personal pronouns generally show what is called the 'accusative' system of case marking, in which the two core arguments S and A are left unmarked, and the third core argument O is marked by the accusative. The remaining nominals (proforms and nouns), on the other hand, may either show (i) the above-mentioned accusative system of case marking, or (ii) what is called the 'ergative' system of case marking, in which the core arguments S and O are left unmarked and A is marked by the ergative case. (There are very few languages that are reported to show ergative case marking for both personal pronouns as well as other nominals.)

Kannada (Dravidian) shows accusative case marking for both personal pronouns as well as other nominals.

- (12) a. *appa banda*
 father came
 'Father came'
- b. *amma appan-annu karedaḷu*
 mother father-Acc called
 'Mother called father'
- (13) a. *na:nu bande*
 I came
 'I came'
- b. *ni:nu nann-annu karede*
 you me-Acc called
 'You called me'

Notice that in (12a), the noun *appa* 'father' occurring in the S function is unmarked for case. In (12b), on the other hand, the noun *amma* 'mother', occurring in the A function is unmarked but the noun *appa* 'father' occurring in the O function is marked for the accusative. A similar type of case marking is used for the pronouns *na:nu* '1Sg' and *ni:nu* '2Sg' in the sentences (13a–b).

On the other hand, Dyirbal (Australian) shows accusative case marking for first and second person pronouns, and ergative case marking for other nominals (Dixon 1994: 10, 14).

- (14) a. *ɲuma banagan^yu*
 father returned
 'Father returned'
- b. *yabu-ɲgu ɲuma buran*
 mother-Erg father saw
 'Mother saw father'
- (15) a. *ɲana banagan^yu*
 we returned
 'We returned'
- b. *n^yura ɲana buran*
 you we-Acc saw
 'You saw us'

Languages that use both accusative as well as ergative systems of case marking may include in the former (accusative) system not only personal pronouns, but also certain other words like proforms, kinship terms, and proper names. Blake (1994: 139) points out, for example, that in Gugu Yimdhirr, accusative case marking is used for all three personal pronouns, whereas in Nhanda it gets further extended to personal names and kin terms. Both these use ergative case marking for the remaining nominals. Blake points out that seven different cut-off points of this type have been attested in different languages.

In order to describe this cross-linguistic variation occurring in the extent of accusative and ergative case markings, nominals have been arranged into a hierarchy of categories, such as the one given in (16), which includes pronouns of first and second person as well (Blake 1994: 138).

- (16) *Hierarchy of nominal categories:*
 First person pronouns
 Second person pronouns
 Third person pronouns
 Proforms
 Proper names, kin terms
 Human nouns
 Animate nouns
 Inanimate nouns

It is hypothesized that the accusative marking will run from the top of this hierarchy and the ergative one from its bottom, each covering a continuous segment. The two may overlap in the middle, as for example, in Yidiny, an Australian language, in which first and second person pronouns show accusative case marking, while common nouns (and inanimate interrogatives) show ergative case marking. Third person pronouns, other proforms, proper names, and kin terms, on the other hand, show an overlap, with A being marked by the ergative case and O by the accusative case (Dixon 1977: 167).

Different reasons have been suggested for the occurrence of this hierarchy. For example, according to Silverstein (1976: 113), it expresses the semantic naturalness for lexically specified noun phrases to function as the agent or patient of a transitive sentence. Humans are more likely to be agents than patients whereas non-humans, especially inanimate objects, are more likely to be patients than agents. The occurrence of the former as patients is therefore a marked situation and the use, in these languages, of an overt case marker (accusative) for denoting human patients can be regarded as resulting from that markedness. Similarly, the occurrence of non-humans as agents rather than as patients is also a marked situation, and the use of an overt case marker (ergative) for denoting non-human agents can also be regarded as resulting from that markedness.

Mallinson and Blake (1981) point out, on the other hand, that such an explanation can account for the hierarchy of (16) only partially. It can account for the relative order of humans, animate nouns, and inanimate nouns, but not that of personal pronouns, proforms, and proper names/kin terms. Mallinson and Blake suggest an alternative explanation, namely that the hierarchy reflects a 'relative centre of interest'. The upper part of the hierarchy, according to them, derives from the fact that events tend to be seen from the point of view of speech act participants. The lower end, on the other hand, derives from the general characteristic of humans, namely that they would be most interested in other humans, and more interested in animals than inanimate objects.

5.3 Personal pronouns as Janus-like

I pointed out in the second chapter (§2.2) how the association of any information regarding the referent with personal pronouns can conflict with their primary function of denoting speech roles. In order to carry out this latter function satisfactorily, they need to remain unchanged even as the actual performers of the speech roles that they represent shift among two or more individuals. Supportive devices like using appositive noun phrases, using vocatives, using noun phrases in a matrix clause, etc. are used for denoting some of the identifying characteristics of these performers (referents of personal pronouns) in certain non-prototypical contexts (like oath-taking or telephone conversation), such that the pronouns themselves are left free of such conflicting information (§2.2.3). Further, redundant personal markers like agreement affixes are also used for specifying some of the identifying characteristics of their referents.

On the other hand, there do occur languages in which information regarding the referents is expressed through markers that are attached directly to personal pronouns, in spite of the fact that it makes the pronouns less efficient. For example, most of the Afro-Asiatic languages convey the distinction between male and female addressees either by using distinct second person pronouns or by attaching gender markers to them. The second person singular pronouns *fīy*

'masculine' and *mācə* 'feminine' of Miya (Chadic: Schuh 1998: 187) exemplify the use of distinct pronouns, whereas those of Egyptian Arabic (Semitic: Mitchell 1956: 26), namely *ʔinta* 'masculine' and *ʔinti* 'feminine', exemplify the use of gender markers.

As I have pointed out in the previous chapter (§4.3.2), the occurrence of gender distinctions in first and second person pronouns can have other explanations. They may have an entirely different function of indicating social distinctions between speech roles rather than between the persons who perform those roles. They may also have diachronic explanations that have nothing to do with their functioning in language. But in some languages at least, they appear to have the function of differentiating between different types of referents, and in such languages, we will have to concede that the pronouns have been allowed to be less efficient in their primary function of denoting speech roles.

Another situation of this type derives from the association of personal pronouns with the category of number. This association gives rise to a conflict between (i) denoting the 'number' of speech roles on the one hand and (ii) the number of referents that the pronouns represent on the other. I had suggested earlier (§4.2.2) that in the former case, the notion involved is one of conjunction whereas in the latter case it is one of plurality. The former emphasizes the differences that occur between the two conjoined entities, whereas the latter emphasizes the similarities. Differences are generally suppressed when a plural marker is used. Several languages resolve this conflict by restricting number marking for denoting the conjunction of speech roles in the case of personal pronouns. Some languages use number marking only for this purpose.

Languages may also use the distinction between independent pronouns and agreement markers for denoting this distinction between conjunction of speech roles and plurality of speech act participants. For example, in some of the languages like Finnish, the plural forms of personal pronouns, when used as polite forms for referring to single individuals, have the verb showing singular agreement (Moravcsik 1978: 359). That is, the plurality that occurs among personal pronouns represents a distinction among speech roles (see §4.3), whereas the one that occurs among agreement markers represents a distinction among the referents (speech act participants). There are also languages in which suppletive forms are used for denoting the conjunction of speech roles and number markers for denoting the plurality of referents. Alternatively, languages may use two different types of number marker for this purpose, and in some languages these may occur together with personal pronouns, with one of them representing conjunction and the other one representing plurality (§4.2.3).

The use of 'inclusory pronouns' in some languages (see §4.2.6) also appears to provide a basis for making a similar differentiation. The number distinction occurring among personal pronouns in these constructions appears to indicate the conjunction of speech roles, whereas the one occurring among person-number agreement markers appears to denote the plurality of their referents.

On the other hand, there do occur several languages in which no differentiation is made in the representation of these two notions of number. The same set of non-singular forms are allowed to ambiguously indicate both these connotations. For example, the second person plural *you* of English generally represents a conjunction of speech roles (2 + 3), but it can also represent the plurality of referents, as for example when it is used for indicating a group of addressees.

Another aspect of personal pronouns that appears to form the basis of conflicting characteristics concerns the notion of definiteness. As I have pointed out in the second chapter (§2.3), an expression is considered to be definite or indefinite depending upon the ability of the speaker (in the case of proforms) or the addressee (in the case of noun phrases) to identify its referent. Since first and second person pronouns are to be dissociated from aspects of their referents in order to carry out, satisfactorily, their primary function of denoting speech roles, they are not generally associated with any markers that have the function of denoting definiteness or indefiniteness. However, the fact that the pronouns do have easily identifiable referents makes it possible for them to manifest some of the characteristics of definite noun phrases. These appear to conflict with the other characteristics of these pronouns that derive from the irrelevance of the notion of definiteness for their functioning (§2.3.1–2).

5.4 Basis of differentiation

The gradation, mentioned earlier, that occurs between personal pronouns and proforms makes it rather difficult to apply the classical theory of categorization to them. According to this theory, properties which define a category need to be shared by all its members and hence there needs to be a sharp and clear-cut distinction between members of two different categories. Personal pronouns cannot be sharply differentiated from proforms (and nominals) because in some of its characteristics, second person pronoun appears to resemble the latter rather than first person pronoun. The Janus-like nature of these two pronouns also appears to add to these difficulties.

This problem of differentiation, however, affects other lexical categories like nouns, verbs, adjectives, etc. as well (see Bhat 1994). It has therefore been suggested by several linguists like Givón (1979), Dixon (1982), and Lakoff (1987) that an alternative theory of categorization, called 'Prototype Theory', developed by Eleanor Rosch and her associates in cognitive psychology (see Rosch 1978), be made the basis of these categories. Human categorization, according to this Prototype Theory, proceeds from central to peripheral instances, with the central instances of a given category being 'prototypical' for that category. If we apply this theory to personal pronouns, we would find the first person pronoun to be occurring as the prototypical personal pronoun.

5.5 Summary

I have pointed out in this chapter two different types of characteristics of personal pronouns that appear to conflict with the characteristics that I have described in earlier chapters. Languages generally give greater prominence to the speaker as compared to the addressee, and this has the effect of making the expressions that denote the speaker dissimilar to those that denote the addressee. The characteristics that derive from this differentiation can form the basis for establishing a hierarchy of nominals. However, there are some languages that give greater prominence to the addressee than the speaker, and further there are also languages in which the emphasis falls on the speaker or the addressee depending upon certain modal distinctions.

The second type of conflicting characteristics that I have examined in this chapter derives from the dual requirements of personal pronouns. They need to *dissociate* themselves from their referents in order to carry out, consistently, their primary function of denoting speech roles, but the fact that they do have referents of their own gives rise to an *association* with their referents. Characteristics that derive from the latter factor conflict with the ones that derive from the former factor. Languages generally avoid this conflict through the use of alternative devices for denoting the characteristics of pronominal referents, but there do occur instances in which the conflict is tolerated.

6 The Position of Third Person Pronouns

6.1 Introduction

The characterization of personal pronouns in the previous chapter assumed that only first and second person pronouns are to be regarded as belonging to the system of personal pronouns. They were contrasted with third person pronouns as well as with other types of pronouns (proforms) and nouns. Such a differentiation between first and second person pronouns on the one hand and third person pronouns on the other is in tune with the view held by several linguists concerning the position of third person pronouns. According to J. Lyons (1977: 638), for example, there is a ‘fundamental ineradicable difference’ between the two sets of pronouns, which derives primarily from the fact that only first and second person pronouns denote individuals who actually participate in the speech act. Lyons suggests that third person pronouns are only negatively defined with respect to first and second, and are obviously dispensable in favour of demonstrative pronouns. This is also the view expressed by several other linguists such as Forchheimer (1953), Benveniste (1971: 217), and Mühlhäusler and Harré (1990: 33).

There are several languages in which the system of personal pronouns consists of only first and second person pronouns. Third person pronouns are either identical with one of the sets of demonstrative pronouns (or with all of them) or are derivationally related to them. We might regard such languages as supporting the above-mentioned claim concerning the position of third person pronouns in the system of person pronouns.

For example, Lak, an East-Caucasian language (Helmbrecht 1999: 285), has personal pronouns only for first and second person. Its third person pronouns are identical with the various demonstrative pronouns.

- | | | | |
|--------|---------------------------------|-------------------------|--|
| (1) a. | Singular | Plural | |
| 1 | <i>na</i> | <i>žu</i> | |
| 2 | <i>ina</i> | <i>zu</i> | |
| b. 3 | <i>ta</i> ‘that’ | <i>wa</i> ‘this’ | |
| | <i>k’a</i> ‘that above’ | <i>ga</i> ‘that below’, | |
| | <i>mu</i> ‘that aforementioned’ | | |

Helmbrecht points out that there are several differences between first and second person pronouns given in (1a) and third person pronouns (demonstratives) given in (1b). For example, (i) the former have suppletive stems with oblique case, while the

latter require special stem extenders, and (ii) the former never receive the ergative case marker while the latter can be ergative case marked in a transitive clause.

Khasi, an Austroasiatic language (Nagaraja 1985: 10), on the other hand, has third person pronouns that are derivationally related to demonstrative pronouns. These latter pronouns are derived by adding the deictic suffixes *-ne* 'proximate', *-to* 'remote, not very far', *-tay* 'remote, visible', *-ta* 'remote, invisible', *-tey* 'up', and *-thie* 'down' to third person pronouns, namely *u* '3MSg', *ka* '3FSg' and *ki* '3Pl'. The latter (third person pronouns) also function as articles.

(2)		Masculine singular	Feminine singular	Plural
	Proximate	<i>u-ne</i>	<i>ka-ne</i>	<i>ki-ne</i>
	Remote, not far	<i>u-to</i>	<i>ka-to</i>	<i>ki-to</i>
	Remote, visible, far	<i>u-tay</i>	<i>ka-tay</i>	<i>ki-tay</i>
	Remote, invisible	<i>u-ta</i>	<i>ka-ta</i>	<i>ki-ta</i>
	Up	<i>u-tey</i>	<i>ka-tey</i>	<i>ki-tey</i>
	Down	<i>u-thie</i>	<i>ka-thie</i>	<i>ki-thie</i>

While Khasi derives its demonstrative pronouns from third person pronouns through affixation, Kharia (Munda: Biligiri 1965: 37) derives third person pronouns from demonstratives by affixing the element *kar* 'person' to them.

(3)	Demonstratives	Third person pronouns
	<i>u</i> 'this'	<i>u-kar</i>
	<i>ho</i> 'that'	<i>ho-kar</i>
	<i>han</i> 'that (remote)'	<i>han-kar</i>

We can regard these different types of languages to be supporting the assumption that third person pronouns do not belong to the system of personal pronouns. However, there are several other languages in which third person pronouns form part of the system of personal pronouns and are quite different from demonstrative pronouns. It is difficult to see how these languages can be regarded as supporting the above-mentioned claim. For example, Angami, a Tibeto-Burman language, differentiates clearly between third person pronouns and demonstrative pronouns and places the former in the system of personal pronouns (Giridhar 1980: 32).

(4)	Personal pronouns			Demonstrative determiners	Demonstrative pronouns		
	Singular	Dual	Plural		Human F	Rest	Diminutive
1 (Excl)	<i>a</i>	<i>hiê-nie</i>	<i>hiê(kô)</i>	<i>hâ</i> 'proximate to 1'	<i>hâ-pfê</i>	<i>hâ-û</i>	<i>hâ-yô</i>
(Incl)		<i>âvû</i>	<i>û(kô)</i>				
2	<i>no</i>	<i>âniê</i>	<i>niê(kô)</i>	<i>tsê</i> 'proximate to 2'	<i>tsê-pfê</i>	<i>tsê-û</i>	<i>tsê-yô</i>
3	<i>puô</i>	<i>puôniê</i>	<i>û(kô)</i>	<i>lû</i> 'remote'	<i>lû-pfê</i>	<i>lû-û</i>	<i>lû-yô</i>

Notice that the third person pronouns of this language are very similar to first and second person pronouns in their formation of dual and plural forms. On the other hand, demonstrative pronouns are quite different from them in that their formation involves the suffixing of gender markers *pfə* 'human, feminine', *u* 'rest', and *yó* 'diminutive' to demonstrative determiners. None of these demonstrative pronouns is related to third person pronouns. While the former show gender distinctions, the latter (third person pronouns, like other personal pronouns) do not.

Languages of this latter type appear to indicate that the concept of third person as a 'non-person' may not be a characteristic of all languages but only of some languages. I wish to examine this possibility in somewhat greater detail in this chapter. Would it be possible to claim that in some languages like Angami the third person pronoun is part of the system of personal pronouns, whereas in others, like English, Lak, Khasi, and most of the familiar languages, it properly belongs to the system of demonstratives? It is quite possible to hypothesize, for example, that languages may either (i) view the speech act as a *two-way affair*, taking place between a speaker and his addressee, or (ii) as a *group affair*, taking place among three or more individuals, with the speech roles denoted by the pronouns of first, second, and third person shifting among the individuals as a conversation progresses. The morphological inclusion of third person pronouns either in the system of demonstratives or in that of personal pronouns can then be regarded as the primary basis of such a typological distinction. This is only a speculation, but it is a useful speculation because it can form the basis of further inquiry regarding other possible distinctions occurring in the use of third person pronouns or other related expressions that correlate with this particular distinction, as I point out in the next section (see 6.2.2).

6.2 'Two-person' and 'three-person' languages

I am assuming that the distinction between (i) languages in which third person pronouns and demonstratives are either identical or derivationally related on the one hand, and (ii) languages in which they are quite different from one another on the other, can form the basis of a typological distinction between 'two-person' and 'three-person' languages. In the former case, the system of personal pronouns consists of only two persons, namely first and second (indicating the speaker and the addressee), whereas in the latter case it consists of three different persons, namely first, second, and third (indicating the speaker, the addressee, and a person who is different from both of them respectively).

I had an opportunity to examine a sample of 225 languages (see Appendix for a list of these languages) that contained most of the 200 languages that were recommended for the World Atlas of Language Structures (Haspelmath, Dryer, Gil, and Comrie:2005) in connection with an article that I prepared for that World

Atlas (Bhat, forthcoming). That study indicated that two-person languages slightly outnumber three-person languages: It appears to be possible to regard 126 of these languages as two-person languages, and only 99 as three-person languages.

6.2.1 Nature of the affinity

The relationship between third person pronouns and demonstratives can be either complete or partial. There are altogether 52 languages (among the 126) in which the relationship between the two can be regarded as complete. In the case of 33 of these languages, grammarians report that there is no third person pronoun as such, and any demonstrative can be used as a third person pronoun. For example, Poppe (1960: 53) points out that in Buriat (Mongolian), there is no special pronoun for third person; instead, the demonstrative pronouns *ene* 'this', *ede* 'these', *tere* 'that', and *tede* 'those' are used. On the other hand, Saltarelli (1988: 211) lists three different types of deictic pronouns, *hau* 'proximal', *hori* 'medial', and *hura* 'distal', as third person pronouns of Basque, along with other personal pronouns, but the same three pronouns are also described as having the function of demonstratives (1988: 213). Some of the additional languages of this type are Hixkaryana (Carib: Derbyshire 1979: 130), Mishmi (Tibeto-Burman: Sastry 1984: 80), Marathi (Indo-Aryan: Pandharipande 1997: 375), and Maricopa (Yuman: Gordon 1986: 58).

In the case of four additional languages, there do occur distinct third person pronouns, but demonstratives are also used as third person pronouns. In Yukulta (Tangkic; Queensland, Australia), for example, *ɲiya* is the third person pronoun, but normally the demonstratives *ɬanta* 'this one' and *ɬantinta* 'that one' are used to express third person (Keen 1983: 200).

In the remaining 15 languages, third person pronouns are derivationally related to demonstratives. The derivation involves (i) sharing a pronominal stem (6), (ii) using third person pronouns in the derivation of demonstrative pronouns (6), and (iii) using demonstrative pronouns in the derivation of third person pronouns (3). In Asheninca (Arawakan: Reed and Payne 1986: 324, 330), for example, third person pronouns share the stem *ir-* with demonstratives (in addition to the gender markers).

(5)	3rd pronouns	Demonstratives		
		Proximate	Medial	Remote
M	<i>irirori</i>	<i>irika</i>	<i>irinta</i>	<i>irintó</i>
F	<i>iroori</i>	<i>iroka</i>	<i>ironta</i>	<i>irontó</i>

The examples from Khasi (2) and Kharia (3) given earlier (§6.2) exemplify the derivation of demonstratives from third person pronouns and third person pronouns from demonstratives respectively. In Sentani (Cowan 1965: 15, 18), on the other hand, the third person pronoun *na* is reported to occur as a demonstrative in addition to proper demonstratives *bekə/dakə* 'proximate' and *bele/dike* 'remote'.

The remaining 74 languages (of the group of 126 mentioned above) can be regarded as showing a 'partial' relationship between third person pronouns and demonstratives. These relationships are of several different types. In 19 of these languages, the relationship is restricted to *remote demonstratives*. Third person pronoun is identical with the remote demonstrative in 9 of them, whereas in the remaining 10 languages the two make use of the same stem. For example, in Georgian (Hewitt 1995: 58, 76), the third person singular pronoun is *is*, which is also the remote demonstrative pronoun; it contrasts with two other demonstrative pronouns, namely *es* 'near 1' and *eg* 'near 2'. Similarly, in Kannada (Dravidian), the third person pronouns *avanu* 'he', *avaḷu* 'she', and *adu* 'it' are the same as remote demonstratives, contrasting with the proximate demonstratives *ivanu* 'this man', *ivaḷu* 'this woman', and *idu* 'this thing' respectively. In Lower Grand Valley Dani (Greater Dani family: Bromley 1981: 207), on the other hand, the third person pronoun *at* and the remote demonstrative *aty* 'that' have the same stem *at*. The latter contrasts with the proximate demonstrative *jy* 'this'.

Five of the 19 languages mentioned above show an affinity of this type only between the non-singular form of the third person pronoun and the remote demonstrative. For example, Hawaiian (Polynesian) has the third person pronoun *ia* that has the stem *lā* and *ana* in the dual and plural. It has the distal demonstrative *kēlā* (which also has the alternants *lā* and *ana*) (Elbert and Pukui 1979: 107, 110).

The affinity is restricted to *non-remote* demonstratives in the case of 14 of the remaining languages. There are 4 languages among these 14, in which the affinity is restricted to proximal demonstratives. In Asmat (southern Papua), for example, the third person pronoun is *ar*, which also functions as the proximate ('near 1') demonstrative, contrasting with the two non-proximal demonstratives *ja* 'near 2' and *er* 'distant' (Voorhoeve 1965: 155). In the case of the remaining 10 languages, the affinity is between third person pronouns and medial demonstratives. For example, Warao (Isolate: Romero-Figeroa 1997: 60) has the third person pronoun *tai*, which is identical with the medial demonstrative *tai* that contrasts with the proximate demonstrative *tama* 'this' and the remote demonstrative *ama* 'that yonder'. Mapuche (possibly Southern Andean: Smeets 1989: 104, 123) has the third person pronoun *fey* which is derivationally related to the demonstrative *tü-fey* 'that'. The latter contrasts with two other demonstratives, namely *tü-fá* 'this' and *tü-ye* 'that, over there'.

In the case of 8 of these languages, the demonstrative that is related to the third person pronoun is not specified as 'medial', but still, it does contrast with a distinct 'distal (yonder)' demonstrative. For example, Lezgian (Nakh-Daghestanian) has the demonstratives *im* 'proximal', *am* 'distal', and *at'a* 'remote (yonder)' of which *am* functions as third person pronoun (Haspelmath 1993: 184–8).

In the case of 24 of the remaining languages, third person pronouns and demonstrative pronouns share the characteristic of having the same set of *gender or noun class markers*. Alamlak (Sepik Hill: Bruce 1984: 75), Apurinã

(Arawakan: Facundes 2000: 350, 361), Trumai (Isolate: Guirardello 1999: 27), Venda (Bantu: Poulos 1990: 95), and Kisi (Niger-Congo: Childs 1995: 104) are some of the languages that belong to this group. For example, Apurinã has the third person pronouns *u-wa* (M) and *o-wa* (F) and the demonstrative pronouns *u-kira* (M), *o-kira* (F) 'distal', and *i-ye* (M), *o-ye* (F) 'proximate' (Bruce 1984: 75, 81). Kisi (Niger-Congo: Childs 1995: 104, 110) has seven different third person pronouns showing noun class (and number) distinctions (*ò*, *à*, *lè*, *là*, *ì*, *η*, and *mà*) corresponding to which it has two sets of demonstratives (proximate and remote) each showing seven different forms. The remote demonstratives, for example, are *kónj*, *kánj*, *lénj*, *lánj*, *kénj*, *mónj*, and *mánj*.

Some of the languages that were described earlier as showing identity between third person pronouns and demonstratives also show gender distinctions among those pronouns. For example, Malayalam (Dravidian: Asher and Kumari 1997: 258) has the demonstratives *avan* (M), *aval* (F), *atə* (N) 'remote', and *ivan* (M), *ival* (F), *itə* (N) 'proximal' that also function as third person pronouns. However, such languages are not included in the above mentioned group of 24 languages. Only languages that show partial affinity due to the sharing of gender or noun class markers are included in this group.

In the case of the remaining 17 languages, third person pronouns are used only for referring to humans or animate beings. In order to denote other beings or inanimate objects, these languages make use of their demonstrative pronouns. Some extend the use of the latter for humans as well. I consider these languages also as showing affinity between third person pronouns and demonstratives. In Jaqaru (Aymaran: Hardman 2000: 27), for example, the third person pronoun *upa* is used for referring to humans only; the language uses the demonstratives *aka* 'this' and *uka* 'that' for referring to non-humans. Similarly in Nivkh (Palaeo-Siberian: Gruzdeva 1998: 26), the third person pronoun *if* is used only in the case of humans; the demonstratives *hyd* 'this, that' and *tyd* 'that', on the other hand, can be used for humans as well as non-humans. In the case of some of these languages, one of the demonstratives would be specialized for this purpose. In Hungarian (Kenesei, Vago, and Fenyvesi 1998: 260), for example, the remote demonstrative *az* 'that' is used for referring to inanimate objects, while the third person pronoun *ő* is restricted to humans or animates.

It must be noted here that since the notion of affinity or 'derivational relationship' is a graded one, the division between the first group (unrelated) and the rest (related) cannot be sharp and clear-cut. It is quite possible that some of the languages that are included in the first group could also be classified as belonging to one of the latter groups. For example, Serbo-Croatian has the third person pronouns *ón* 'he', *òna* 'she', and *òno* 'it', and the demonstrative pronouns *ovaj* 'near 2' and *onaj* 'distal' (Kordić 1997: 22–5). It is possible that one or more of these pronouns (especially *onaj* 'distal') share the stem with third person pronouns. However, such languages are not regarded as showing affinity between third person pronouns and demonstratives.

Third person pronouns are considered to be related to demonstratives in this study only in the case of languages in which (i) the affinity is quite obvious or (ii) the authors have suggested that the two are historically related.

6.2.2 *Distinction between third person pronouns and demonstratives*

As I will be pointing out in detail in the ninth chapter, there are several characteristics that can be used to differentiate between third person pronouns and demonstratives (and other proforms). Most importantly, reference in the case of third person pronouns is pragmatic whereas in that of demonstratives it is semantic. The former is 'definite' or identifiable in a pragmatic (or 'linguistic') sense, whereas the latter needs more substantial 'semantic' basis in order to be regarded as 'definite' (§9.2.1). A referent becomes identifiable automatically the moment it is introduced through an expression in the former case, whereas in the latter case identity needs to be established by providing some substantial amount of information (§9.2.2). It is possible that even in the case of languages in which demonstrative pronouns are used as third person pronouns without making any changes in them, there is this differentiation between the two uses.

In spite of the occurrence of such a (possible) differentiation, the formal identity or affinity between third person pronouns and demonstrative pronouns mentioned earlier appears to be correlatable with a functional affinity, as shown by the occurrence of several co-occurring characteristics. I describe below, in detail, two of these characteristics, namely (i) and (ii).

- (i) Languages in which there is a gender or noun class distinction among third person pronouns on the one hand, and the ones in which there is no such distinction among third person pronouns on the other. The latter are more frequently three-person languages than two-person languages (§6.3).
- (ii) Languages in which demonstratives are distance-oriented on the one hand, and the ones in which demonstratives are person-oriented on the other. The latter are more frequently three-person languages than two-person languages (§6.4).

6.3 Gender distinctions

I suggested in the fourth chapter (§4.3) that the general absence of gender distinction in first and second person pronouns derives from its irrelevance for the functioning of those pronouns. Its presence in third person pronouns and nouns, on the other hand, derives from the fact that providing additional information about their referents would be useful, and sometimes essential, for identifying those referents.

- (6) a. *John asked Bill to bring his book to the class.*
 b. *John asked Mary to bring her book to the class.*

Notice that in (6a) the third person pronoun *his* can have either John or Bill as its antecedent, whereas in (6b) the pronoun *her* can have only Mary as its antecedent. (It is possible, of course, for both of them, to have an exophoric antecedent.) The gender distinction occurring in the pronoun has been able to remove the ambiguity occurring in (6a) (at least partially) by disallowing John from being the antecedent of that pronoun. Since the primary function of first and second person pronouns is to identify speech roles rather than their referents, such information would be irrelevant for their functioning. It would in fact be harmful, as it would adversely affect their primary function of denoting speech roles.

Even in the case of third person pronouns, some languages show gender distinction whereas others do not. For example, English shows a three-fold gender distinction in third person singular (*he, she, it*). There are other languages like Latvian (Nau 1998: 16) that extend the distinction (between masculine and feminine) to plural as well, but most languages that show gender distinction in third person are like English in restricting it to singular forms or at least reducing it (from M/F/N to MF/N) in their non-singular forms. On the other hand, there are several other languages in which third person pronouns show no gender distinction whatsoever. For example, Angami, mentioned earlier (§6.1(4)), does not show any gender distinction among third person pronouns even though it does show a masculine–feminine gender distinction among demonstratives.

This cross-linguistic variation concerning the occurrence of gender distinction among third person pronouns appears to correlate with the distinction between two-person and three-person languages in the sense that gender distinction in third person pronouns occurs primarily among two-person languages. In my sample of 225 languages, for example, it occurs among 49 of the 126 two-person languages, whereas among the 99 three-person languages, it occurs only in 13 languages. This latter number is even smaller as far as the contrast between third person pronouns and the other two personal pronouns is concerned, because in the case of 7 of those 13 languages gender distinction occurs in second person as well. That is, among the 51 languages in which gender distinction is restricted to third person, 45 belong to two-person languages and only 6 belong to three-person languages.

The six three-person languages in which gender distinction is restricted to third person are Tibetan (Denwood 1999: 97), Harar Oromo (Owens 1985: 98), Serbo-Croatian (Slavonic, Indo-European: Kordić 1997: 22), Jacalteco (Craig 1977: 100), Eastern Pomo (McLendon 1975: 124), Maung (Arnhem Land, Australian: Capell and Hinch 1970: 59) and Wari' (Everett and Kern 1997: 303). In Tibetan, however, the distinction occurs only in the case of ordinary third person pronouns and not in that of honorific pronouns.

Languages in which gender distinction occurs in both second and third person pronouns are altogether 10 in my sample. Four of them belong to the Chadic family (Hausa: Newman 2000, Miya: Schuh 1998, Pero: Frajzyngier 1989, and Mupun: Frajzyngier 1993), two to Semitic (Arabic: Mitchell 1956 and

Hebrew: Glinert 1989), one each to Berber (Tamazight: Penchoen 1973), Mon-Khmer (Khmu': Premsrirat 1987), and Middle Sepik (Ambulas: Wilson 1980) and one to an isolate (Tunica: Haas 1940). Nunggubuyu (Australian: Heath 1984) and Spanish (Green 1988) show gender distinction in all three persons, but not in the singular. In all these languages, however, gender is not a characteristic that differentiates third person pronouns from other personal pronouns.

There are 20 languages in my sample in which third person pronouns show a human–non-human or animate–inanimate distinction. It is possible to regard these also as showing some kind of 'gender' distinction. It is interesting that only one of them, namely Koromfe (Gur: Rennison 1997: 245), is a three-person language. Its third person pronouns *dəkɔ* 'human', *gukɔ* 'non-human' appear to be quite different from the demonstratives *hɔŋ* 'human' and *kɔŋ* 'non-human'. The remaining 19 languages are two-person languages in that all of them show some kind of affinity between third person pronouns and demonstratives. In the case of 16 of these languages, the non-human or inanimate third person is identical with demonstratives. Some of the languages of this latter type are listed in (7).

(7)	Third person pronouns	
	Human	Non-human (Demonstrative)
Dumi (Van Driem 1993: 80)	<i>im</i>	<i>tom/mom</i>
Garó (Burling 1961: 38)	<i>bia</i>	<i>ua</i>
Hungarian (Kenesei et al. 1998: 260)	<i>ő</i>	<i>az</i>
Jaqaru (Hardman 2000: 27)	<i>upa</i>	<i>aka/uka</i>
Persian (Rastorgueva 1964: 28)	<i>u/vey</i>	<i>an</i>
Quechua (Adelaar 1977: 259)	<i>pay</i>	<i>kay/çay/wak</i>
Yagua (Payne and Payne 1982: 370)	<i>níí</i>	<i>jiy-/ru-</i>

6.4 Person-oriented and distance-oriented demonstratives

As pointed out by Anderson and Keenan (1985: 282), languages manifest a distinction between two different types of deictic systems, namely 'distance-oriented' and 'person-oriented'. This distinction occurs mainly among three-term deictic systems, but there do occur some languages in which a person-oriented deictic system involves primarily a two-way opposition (see §8.2.1). In the case of three-term systems, the middle term would indicate a location that is between proximate and remote in distance-oriented systems, whereas in person-oriented systems it would indicate a location that is proximate to the addressee. This contrast between person-oriented and distance-oriented deictic systems can be exemplified with the help of the deictic systems of Plains Cree (8), which has a distance-oriented deictic system, and Maori (9), which has a person-oriented deictic system.

- (8) *Distance-oriented deictic system: Plains Cree (Algonquian: Wolfart 1973: 32)*
awa 'this'
ana 'that'
naha 'that yonder'
- (9) *Person-oriented deictic system: Maori (Polynesia: Bauer 1999: 381)*
nei 'near the speaker'
naa 'near the addressee'
raa 'distant from both the speaker and the addressee'

It appears to be possible to correlate this distinction with the distinction mentioned earlier, namely between two-person and three-person languages, in the sense that two-person languages are less frequently associated with person-oriented deictic systems as compared to three-person languages. Languages of the former type also show three-term deictic systems less frequently as compared to three-person languages. In my sample of 126 two-person languages, for example, 39 showed three-term deictic systems, of which only 14 involved person-oriented deictic systems. On the other hand, among the 99 three-person languages, 51 showed three-term deictic systems of which 31 involved person-oriented deictic systems. The following sets of languages exemplify this preferred correlation between two-person languages and distance-oriented deictic systems on the one hand (10) and three-person languages and person-oriented deictic systems on the other (11), both occurring among three-term demonstrative systems.

- (10) *Two-person languages with a distance-oriented three-way deictic distinction among demonstratives*

	Personal pronouns			Demonstrative pronouns		
	First	Second	Third	Proximate	Medial	Distal
Maricopa (Gordon 1986)	' <i>nyaash</i>	<i>mansh</i>	Dem	<i>da</i>	<i>va</i>	<i>aas</i>
Turkish (Lewis 1967)	<i>ben</i>	<i>sen</i>	<i>o</i>	<i>bu</i>	<i>şu</i>	<i>o</i>
Tiriyó (Meira 1999)	<i>wĩ</i>	<i>ëmë</i>	Dem	<i>mërë</i>	<i>mëërë</i>	<i>ohki</i>
Comanche (Charney 1993)	<i>ni̱</i>	<i>inni̱</i>	Dem	<i>i</i>	<i>ma</i>	<i>u</i>
Hopi (Hill et al. 1998)	<i>nu'</i>	<i>um</i>	<i>pam</i>	<i>ĩ</i>	<i>pam</i>	<i>mit</i>
Basque (Saltarelli 1988)	<i>ni</i>	<i>hi</i>	Dem	<i>hau</i>	<i>hori</i>	<i>hura</i>
Kharia (Billigiri 1965)	<i>iñ</i>	<i>am</i>	Dem + <i>kar</i>	<i>ho</i>	<i>u</i>	<i>han</i>
Mapuche (Smeets 1989)	<i>iñché</i>	<i>eymi</i>	<i>fey</i>	<i>tü-fá</i>	<i>tü-fey</i>	<i>tü-ye</i>

Warao (R-Figeroa 1997)	<i>ine</i>	<i>ihí</i>	<i>taí</i>	<i>tama</i>	<i>taí</i>	<i>ama</i>
Pero (Frajzyngier 1989)	<i>nè</i>	<i>kè</i>	<i>tè</i>	<i>mò</i>	<i>mù</i>	<i>téejè</i>
Kawaiisu (Zigmond 1991)	<i>níʔi</i>	<i>ʔimi</i>	Dem	<i>siʔi</i>	<i>sa,a</i>	<i>suʔu</i>
Lavukaleve (Terrill 1999)	<i>ngai</i>	<i>inu</i>	Dem	<i>fona</i>	<i>foina</i>	<i>feana</i>

(11) *Three-person languages with a person-oriented three-way deictic distinction among demonstratives*

	Personal pronouns			Demonstrative pronouns		
	First	Second	Third	Near 1	Near 2	Distant
Amele (Roberts 1987)	<i>ija</i>	<i>hina</i>	<i>uqa</i>	<i>i</i>	<i>au</i>	<i>ou</i>
Angami (Giridhar 1980)	<i>a</i>	<i>nō</i>	<i>puō</i>	<i>hâû</i>	<i>tsû</i>	<i>tûû</i>
B. Fijian (Dixon 1988)	<i>au</i>	<i>o</i>	<i>e</i>	<i>qoo</i>	<i>qori</i>	<i>yaa</i>
Chamorro (Topping 1973)	<i>hu</i>	<i>un</i>	<i>ha</i>	<i>este</i>	<i>ena</i>	<i>ayu</i>
Koasati (Kimball 1991)	<i>anó</i>	<i>isnó</i>	<i>ibisnó</i>	<i>yaʔá</i>	<i>maʔa</i>	<i>má:fa</i>
Ladakhi (Koshal 1979)	<i>nə</i>	<i>khyot</i>	<i>kho</i>	<i>i</i>	<i>ote</i>	<i>ə</i>
Lango (Noonan 1992)	<i>án</i>	<i>yìn</i>	<i>én</i>	<i>-ni</i>	<i>-nò</i>	<i>-cà</i>
Mao Naga (Giridhar 1994)	<i>ay</i>	<i>ni</i>	<i>pfo</i>	<i>hi</i>	<i>ʔi</i>	<i>lo</i>
Marshallese (Zewen 1977)	<i>ʔa</i>	<i>kwe</i>	<i>e</i>	<i>e</i>	<i>ʔe</i>	<i>eʔ</i>
Mundari (Osada 1992)	<i>añ</i>	<i>am</i>	<i>ay</i>	<i>mo</i>	<i>i</i>	<i>hi</i>
T. Quechua (Adelaar 1977)	<i>nuxa</i>	<i>xam</i>	<i>pay</i>	<i>kay</i>	<i>cay</i>	<i>wak</i>
Tagalog (Schachter et al. 1972)	<i>ako</i>	<i>ikaw</i>	<i>siya</i>	<i>ito</i>	<i>iyan</i>	<i>iyon</i>
Tinrin (Osumi 1995)	<i>nro</i>	<i>nrü</i>	<i>nrî</i>	<i>ha</i>	<i>mwâ</i>	<i>rra</i>
T. Besi (Donohue 1999)	<i>ko</i>	<i>nu</i>	<i>no</i>	<i>ana</i>	<i>atu</i>	<i>iso</i>
W.F-Aniwa (Dougherty 1983)	<i>avay</i>	<i>akoe</i>	<i>aia</i>	<i>nei</i>	<i>na</i>	<i>ra</i>

I have used the label Dem in the third column of (10) in order to indicate that any of the demonstrative pronouns that occur in the language concerned can be used as a third person pronoun. The remaining languages of (10) use only one of the demonstratives as their third person pronoun and I have indicated this restriction by placing the relevant demonstrative under the third person column as well. On

the other hand, I have not shown other distinctions like gender and number that occur among personal pronouns and demonstratives of some of these languages. There are also other distinctions, like visible–non-visible, up–down, and exophoric–anaphoric, that occur among the demonstratives of some of these languages, but these are also not shown in (10–11).

6.4.1 Exceptions to the correlation

It is perhaps possible to make a stronger claim regarding this correlation, such as for example that the occurrence a person-oriented deictic system is restricted to three-person languages. There are 14 languages in my sample that appear to go against such a claim, but at least some of them may have alternative explanations. For example, some of these exceptional cases may have resulted from an areal spread of the person-oriented deictic notion or as a recent innovation, that is, a change that has not yet penetrated into the grammatical system. Foley (1986: 75) points out that it is somewhat uncommon among Papuan languages to have a person-oriented three-term system. Rather more usual are distance-oriented demonstrative systems that choose the position of speech act (i.e. of the speaker) as the starting point and indicate the location of objects as being removed from this central point. Among the Oceanic languages, on the other hand, 'it is almost universal' to have three-term demonstrative systems of the former type (Lynch 1998: 114).

This is also true of Sinhalese (Indo-Aryan), Georgian (Kartvelian), and Zulu (Bantu). Most of the Indo-Aryan languages like Hindi, Marathi, and Punjabi, for example, have only distance-oriented two-term deictic systems. Sinhalese has apparently developed the person-oriented system only recently. In the case of Georgian also, there is a dispute as to whether the proto-language had a two-term demonstrative system or a three-term demonstrative system (Harris 1991: 23). Daghestanian languages like Lezgian (Haspelmath 1993: 187) and Bantu languages like Venda, on the other hand, do possess three or more distinct deictic terms, but the distinction is distance-oriented rather than person-oriented. It is therefore possible that the above-mentioned languages have developed this distinction only recently. Notice, however, that Godoberi, another Daghestanian language, has a person-oriented deictic system.

In the case of four of these languages, namely Hawaiian (Elbert and Pukui 1979), Maori (Bauer 1999), Samoan (Mosel and Hovdhaugen 1992), and Tuvaluan (Besnier 2000), all belonging to the Polynesian family, the affinity between third person pronouns and demonstrative pronouns derives from the fact that the non-singular forms of third person pronouns contain a *different* stem that can be identified with a demonstrative stem. It is possible that these languages are basically three-person languages and the affinity that occurs between the non-singular forms of third person pronouns and demonstratives has some other explanation.

Another important point that needs to be considered here is that in the case of languages in which demonstrative pronouns show a three-fold deictic distinction, grammarians are not generally very clear in specifying the existence or non-existence of 'proximate to the addressee' as one of the distinct notions. This is not true, of course, of all grammarians. Some provide very specific instances in order to establish the actual position of demonstratives in this regard. For example, Meira (1999: 157) clearly states that the medial demonstrative of Tiriyo', a Cariban language, cannot be regarded as denoting 'proximate to the addressee', as shown by (12).

- (12) *mërë-po* *pai* *i-pata*
 there(Medial)-Loc Tapir 3-village
 'Tapir's village is over there'

According to Meira, (12) is uttered by a person to show the way for another person who did not know where Tapir's village was. Obviously, *mërë* 'there' cannot mean 'near the addressee' in such a usage. Wolfart (1973: 33) also mentions specifically that the three-fold deictic distinction occurring in Plains Cree (Algonquian) does not correlate with the three-fold person distinction occurring among personal pronouns. It only involves a spatial distinction between proximate, medial, and distant, arranged in a linear sequence.

In the case of some languages, the notion of proximate to the addressee is mentioned as one of the *possible* meanings of the medial demonstrative. For example, Heath (1984: 269) describes the immediate demonstrative *da-* of Nunggubuyu, an Australian language, to be often denoting the addressee's location, but he suggests further that it is also used for locations conceptualized as within easy access, not necessarily closer to the addressee than to speaker. It is also used in the indefinite sense 'somewhere around here'. Notice that demonstrative pronouns can be used to differentiate between the location of the speaker and the addressee even in the case of languages like English in which there is only a two-fold distance-oriented deictic distinction.

- (13) a. John: *Is this the book that we give to Mary?*
 b. Joan: *No! That one is for Bill. I was thinking of this one.*

The remote demonstrative *that* occurring in (13b) appears to indicate an object that is near the addressee, namely John. What is relevant in the present context, therefore, is the existence of a demonstrative pronoun that is specifically meant for denoting the location of the addressee.

There are also languages in the case of which earlier grammarians had failed to understand the exact nature of the deictic system. For example, Hausa (Chadic) makes a four-fold spatial distinction among its demonstrative pronouns, involving proximate to the addressee as one of the distinct deictic notions (Newman 2000: 147).

(14)		Masculine forms	Feminine forms
	proximate to the speaker	<i>wannàn</i>	<i>wannàn</i>
	proximate to the addressee	<i>wànnan</i>	<i>wànnan</i>
	remote from both	<i>wancàn</i>	<i>waccàn</i>
	distant	<i>wàncan</i>	<i>wàccan</i>

Jaggar and Buba (1994: 388) point out that earlier grammarians had described the system as involving either (i) a two-fold spatial distinction between proximate and remote combined with a two-fold non-spatial distinction between visible and non-visible, or (ii) between several spatial distinctions, that is, as involving a distance-oriented deictic system. However, these analyses had left many of the uses of these pronouns unaccounted for. Jaggar and Buba argue that a better analysis would be to regard *nàn* as denoting 'speaker-proximity', *nan* 'hearer-proximity', *càn* 'remoteness from both' and *can* 'more distal'. This re-analysis of Hausa demonstratives appears to have been accepted by other scholars as shown in (14), which is from Newman (2000).

The point to be noted here is that this latter analysis, unlike the former one, supports the typological distinction that is being proposed in this chapter. The third person pronouns of Hausa (*shī* 'he', *ita* 'she', and *sū* 'they') are quite different from demonstrative pronouns, and hence according to the former analysis Hausa would have been an exception to that proposal. Since the correlation that is being postulated here makes use of the distinction between the occurrence vs. non-occurrence of 'proximate to the addressee' as one of the variables, grammars that wrongly specify the meaning of a medial term among demonstratives can easily appear as contradicting the correlation.

In addition to this correlation between person-oriented demonstratives and three-person languages (and between distance-oriented demonstratives and two-person languages), one can also think of a correlation between gender-marked third person pronouns and distance-oriented deictic systems (contrasting with a correlation between gender-marked third person pronouns and person-oriented deictic systems). There are 73 languages in the sample that show the former correlation whereas there are only 9 languages that show the latter correlation.

6.5 Distinctions in other areas of grammar

I have pointed out in the second chapter several different ways in which first and second person pronouns differ from other pronouns in the marking of core case relations. Languages that use the head-marking strategy, for example, may restrict subject marking to first and second persons and use 'inverse' markers when a third person referent functions as the subject (see §4.4.1). It appears that this kind of differentiation between third person subjects and non-third person subjects occurs mostly among two-person languages. The three-person languages that I have examined so far provide uniform representations to the three persons.

In the use of other strategies also, languages appear to reflect the characteristic of being a two-person or three-person language.

Awa Pit, a language belonging to the Barbacoan family of Colombia, for example, may be regarded as a two-person language. Its third person pronouns *us* (plural: *uspa*) refer to humans only. It uses the demonstrative pronouns, which make a two-fold deictic distinction between proximate (*an* 'this') and remote (*sun* 'that'), for referring to non-humans. It shows an interesting binary agreement system of person marking in the case of its verbs, which can probably be correlated with this characterization of the language as a two-person language. Verbs occur with two different person markers in this language, called locutor and non-locutor markers. The use of these markers depends upon whether the item of interest is (or is not) the epistemic source for the utterance. In the case of statements, locutor corresponds to first person, and non-locutor to second and third, whereas in the case of questions, locutor corresponds to second person, and non-locutor to first and third (Curnow 1997: 187–90).

- (15) a. *nana pala kuntu-s*
 I - plantain eat-Locutor
 'I eat plantains'
- b. *nuna pala kumyu-y*
 you plantain eat-Nonlocutor
 'You are eating plantains'
- (16) a. *m̃na = ma na = na ashaptu-y*
 who-Acc.Int I-Topic annoy-Nonlocutor
 'Who am I annoying?'
- b. *shi = ma nuna kimutu-s*
 what = Int you do-Locutor
 'What are you doing?'

Notice that the occurrence of first person pronoun (locutor) as the subject in (15a) allows the verb to occur with the locutor suffix *-s*, whereas the occurrence of second person pronoun (non-locutor) as the subject in (15b) makes it necessary to use the non-locutor suffix *-y*. The reverse is the case in (16), because it involves a question in which the second person is the locutor. Notice that in (16a) the verb contains the non-locutor suffix because its subject is first person, which is a non-locutor of a question. On the other hand, (16b) has a locutor suffix on the verb because its subject is second person, which is the locutor of a question. The point to be noted here is that in this locutor/non-locutor dichotomy, third person pronoun does not find any place.

Hausa (Newman 2000) can be regarded as a three-person language, as its third person pronouns, *shī* (Masculine), *ita* (Feminine), and *sū* (Plural), are quite different from demonstratives. The latter show a distinction between *wannān* 'this (near me)', *wànnan* 'that (near you)', *wancàn* 'that (remote from me and

you)', and *wàncan* 'that (yonder)' (see (12) given earlier in §6.4). Hausa verbs show two distinct forms when they are directly followed by a direct object: their 'B-form' is to be selected when they are directly followed by a personal pronoun direct object, and their 'C-form' is to be selected when they are directly followed by any other direct object, such as a simple noun or a non-personal pronoun. The third person pronoun functions like a 'personal pronoun' in this dichotomy. (Verbs have two other forms, called (i) A-form that occurs when the verb is not followed by an object, either because the verb is intransitive or the object is deleted or moved, and (ii) D-form that occurs when the verb is followed by an indirect object (Newman 2000: 627).)

6.6 Position of language families

The general studies and reconstructions of language families that I have been able to examine in connection with this typological distinction between two-person and three-person languages appear to support that typology on the whole. Some of the specific developments shown by daughter languages may appear to go against it, but the tendency is apparently to discard contradicting distinctions or to develop distinctions that are in tune with the position maintained by the typology.

6.6.1 Two-person languages

Proto-Indo-European is reconstructed as a two-person language, having a system of personal pronouns consisting of first (**egō/*eĝhom*) and second (**tu*) persons, and showing a two-fold proximate–remote distinction among the demonstratives. The latter are also used as third person pronouns in many of the daughter languages (Baldi 1987: 55). Most of the daughter languages have maintained this system and are therefore supportive of the typological distinction. For example, most of the Indo-Aryan languages lack a third person pronoun, and use demonstratives instead. Most of them also show a two-fold proximate–remote distinction among the demonstratives and, further, most of them also show gender distinctions that are restricted to third person pronouns (Masica 1991: 251). This is also true of other Indo-European languages.

On the other hand, some of the languages of the family have developed demonstrative systems that involve the notion of 'proximity to the addressee'. For example, the demonstrative system of Armenian is made up of stems *so* 'this', *do* 'that (close to the interlocutor)', and *no* 'that (remote)' (Ajello 1998: 216). Latin also has a similar three-term distinction among its demonstratives, namely *hic* 'proximity to the speaker', *iste* 'proximity to the interlocutor', and *ille* 'remote', according to some linguists. Vineis (1998: 293) points out, however, that in third person singular and plural, Latin usually makes use of the demonstrative pronouns for the subject forms, but in all other cases, it uses the

reflexive pronoun. Italian has generalized these oblique case forms of Latin into subject forms as well (Vincent 1988: 291).

Proto-Dravidian is also reconstructed as a 'two-person' language. Its third person pronoun is considered to be indistinguishable from demonstrative pronouns. The latter is reconstructed as making a three-fold deictic distinction between **i*- 'proximate', **u*- 'intermediate', and **a*- 'remote' in the demonstrative system (see Burrow and Emeneau 1960), but the distinction is distance-oriented. Among the modern Dravidian languages, some like Tamil, Kannada, Tulu, etc. have lost the intermediate demonstrative, whereas some like Kuvi have increased the number of distinctions. According to Israel (1979), Kuvi makes a five-fold contrast between *i:*, *e:*, *u:*, *he:*, and *hu:* among its demonstratives, but the distinction does not involve any reference to the addressee. Dravidian languages also show gender distinctions of different types that are restricted to third person pronouns. Some of the daughter languages like Tamil have inalienable possessive forms for their pronouns, but only those of first and second person are directly derivable from personal pronouns. The affix used in the case of third person is connected to the anaphoric (reflexive) pronoun.

The primary distinction among the personal pronouns of Uto-Aztecan languages is also between first and second persons (Langacker 1977: 124). For the most part, third person is represented by demonstratives or elements derived from them. The system of demonstratives may range from the relatively simple to the quite complex, but all of them are centred on a proximal–distal contrast. Langacker (1977: 99) suggests the reconstruction of **i* as proximal and **u* as distal for Proto-Uto-Aztecan.

Australian languages can also be regarded as basically two-person languages. Most of them do not have any third person pronouns as such. According to Dixon (1980: 356), they use their definite determiners, with obligatory specification of distance, in place of third person singular pronouns. The deictic distinction, shown by demonstratives, is generally between proximate and remote, but even when showing a three-fold distinction, the medial is not used to represent the notion of 'proximate to the addressee'. For example, Ngaanyatjara has the forms *ngaa*- 'this, here', *pala*- 'that, middle distant', *nyarra*- 'that, far distant', and *palunya*- 'not visible' (Dixon 1980).

The pronominal system reconstructed for Proto-Semitic by Moscati, Spitalec, Ullendorff, and von Soden (1964) also indicates that the third person pronoun is related to demonstratives, and the latter show only a two-fold proximate–remote distinction.

6.6.2 *Three-person languages*

According to Abondolo (1998: 24), Proto-Uralic can be reconstructed as having three distinct personal pronouns, namely **mV*- 'first person', **tV*- 'second person', and **sE*- 'third person', with *V* denoting any vowel and *E* denoting front

vowel. The reconstructed demonstrative pronouns are different from third person pronouns, and show a three-fold person-oriented distinction between **tE*- 'proximate', **c^jE*- 'an area closer to the addressee/anaphoric', and **tB*- 'remote' (where *B* refers to a back vowel). The function of the second demonstrative, however, is considered to be not clear. It may have denoted an area closer to the addressee, or it may have been anaphoric (or both). Some of the daughter languages have reduced this demonstrative system into a simple two-term proximal–distal dichotomy, but Abondolo notes that some of them also use one or the other of the demonstrative pronouns as third person pronouns, replacing the earlier **sE*-.

Oceanic languages are also considered to be mostly three-person languages. According to Lynch (1998: 114), it is almost universal in Oceanic languages to make a three-way directional–locational–temporal distinction among demonstratives corresponding to the three grammatical persons. The three degrees are often referred to as proximal (near the speaker), intermediate (near the addressee), and distant (away from both). Corresponding to this characteristic, we also find third person pronouns to be quite distinct from demonstrative pronouns in these languages. Very few of them show gender distinction among pronouns. Lynch (1998: 102) points out that Maringe is one of the few that have gender distinction, but its use is constrained by social factors. Female speakers use only one set of third person forms, but male speakers use two sets—one for referring to males and the other one for referring to all others (see §4.3).

6.7 Functional distinctions

The foregoing examination of the position of third person pronouns in the system of personal pronouns on the one hand and of demonstrative pronouns on the other was primarily based upon their formal characteristics. There is a need to examine the functions in which third person pronouns are used in these two types of language because it is quite possible that cross-linguistic distinctions occurring in this regard can also be correlated with the distinction between two-person and three-person languages.

As I will be pointing out in the ninth chapter (§9.2), one of the most important characteristics of proforms like demonstratives and interrogative-indefinites is that the notion of reference and identification that is involved in their use is quite different from the one that occurs in the case of definite and indefinite noun phrases. The reference is only 'linguistic' in the latter case, whereas in the former case it is more substantial, involving one of several levels of 'existence' such as that of reality, fiction, mythology, dream, etc. Similarly, the identification (or identifiability for the addressee) in the case of noun phrases is automatic, established by the mere occurrence of a noun phrase in an utterance, whereas, in the case of proforms, identification gets established only when sufficient information is made available for the addressee. It is not automatic; the

indefiniteness of a referent can persist so long as the addressee is not satisfied with the information that is made available for him (§9.4).

The position of first and second person pronouns in this regard is not very clear because identifying the referents is not their primary function. However, they can have their referents identified by ordinary noun phrases (definite or indefinite) used either in apposition or in the matrix clause (§2.3.2). Hence, we can probably regard them as having referents that need to be established only at the linguistic level. This is especially true of first and second person pronouns occurring in reported speech. Third person pronouns, on the other hand, are generally considered to be similar to definite noun phrases, which means that their reference and identifiability also needs to be merely linguistic. The question that needs to be examined here, therefore, is whether there is any cross-linguistic variation in the kind of reference and identifiability that gets associated with third person pronouns and whether such a variation can be correlated with the typological distinction between two-person and three-person languages.

6.8 Summary

I have examined in this chapter the possibility of establishing a typological distinction between languages that show affinity between third person pronouns and demonstratives ('two-person' languages) on the one hand, and the ones that do not ('three-person' languages) on the other. There are certain interesting characteristics that appear to support such a typology. For example, languages that manifest gender distinction among their third person pronouns tend to show affinity between third person pronouns and demonstratives (i.e. they tend to be 'two-person' languages), whereas languages that manifest person-oriented demonstrative systems tend to show no such affinity. The latter tend to be 'three-person' languages. There is clearly a need to make a more detailed study of this typological distinction.

PART II

Proforms

This page intentionally left blank

7 The Structure of Proforms

7.1 Introduction

Proforms are quite different from personal pronouns in having an internal dual structure. We can generally regard them as consisting of two different elements, namely a general term that denotes the scope of those proforms and a pronominal element that indicates the purpose for which they are used. Proforms occurring in most languages can be displayed in the form of a ‘table of correlatives’ in which the columns represent pronominal elements and rows represent general terms. Even in the case of languages in which the proforms do not show, synchronically, any internal structure, it is generally possible to arrange them meaningfully into such tables (§7.2).

On the other hand, there do occur languages in which proforms show more complex internal structures. These complexities derive from several sources. Languages may use a single pronominal element or general term for representing two or more concepts and use additional markers for differentiating between them. This would have the effect of making the proforms more complex in their internal structure. Complexities may also arise from the grammaticalization of proforms and the use of some of the peripheral markers like those of gender and case, and of pragmatic notions like emphasis or contrast, either in the representation of pronominal elements or of general terms (§7.3).

7.2 Paradigms of proforms

Proforms occur in the form of a set of paradigms in most languages. The general characteristic of these paradigms is that each of them represents a set of general concepts like person, thing, place, time, manner, property, etc. that is associated with one particular function such as denoting a deictic distinction (like remote vs. proximate), interrogation, indefiniteness, anaphora, etc. That is, the forms occurring in these paradigms are generally made up of (i) a term that denotes a general concept and (ii) a pronominal element that denotes a specific function. Languages differ, however, concerning the set of general concepts that are made use of in these paradigms and also concerning the set of functional distinctions that are represented through distinct paradigms in their system of proforms. This dichotomous structure of proforms, occurring in most languages, can be exemplified with the help of the following somewhat shortened paradigms of Khezha, a Tibeto-Burman language (Kapfo 1993, personal communication).

(1)	Gloss	Proximate		Remote		Interrogative	Indefinite
		Speaker	Hearer	Visible	Invisible		
Person	<i>hì</i>	<i>cə</i>		<i>whɔ́</i>	<i>śə</i>	<i>thu</i>	<i>ketə</i>
Thing	<i>hì-no</i>	<i>cə-no</i>		<i>whɔ́-no</i>	<i>śə-no</i>	<i>dá-no</i>	—
Place	<i>hì-bá</i>	<i>cə-bá</i>		<i>whɔ́-bá</i>	<i>śə-bá</i>	<i>dá-bá</i>	<i>ketə-bá</i>
Time	<i>hì-nhie</i>	<i>cə-nhie</i>		<i>whɔ́-nhie</i>	<i>śə-nhie</i>	<i>dì-chə-nhie</i>	<i>ketə-nhie</i>
Manner	<i>hì-ci</i>	<i>cə-ci</i>		<i>whɔ́-ci</i>	<i>śə-ci</i>	<i>dá-ci</i>	—
Amount	<i>hì-ze</i>	<i>cə-ze</i>		<i>whɔ́-ze</i>	<i>śə-ze</i>	<i>dá-ze</i>	—
Type	<i>hì-bì</i>	<i>cə-bì</i>		<i>whɔ́-bì</i>	<i>śə-bì</i>	<i>dì-bì</i>	<i>ketə-bì</i>

Khezha has a set of six pronominal elements that are used for differentiating between two proximate (*hi-* ‘near the speaker’ and *cə-* ‘near the hearer’) and two remote (*whɔ́-* ‘visible’ and *śə-* ‘invisible’) demonstratives and also between an interrogative and an indefinite. (It also has a distinct paradigm of relative pronouns, containing the element *zə-*, but this has only pro-nominal forms, but not any pro-adjectival or pro-adverbial forms.) These pronominal elements appear to occur by themselves (with a ‘zero’ stem) for denoting a person, whereas for denoting other general concepts, they are attached to stems like *-no* ‘thing’, *-ba* ‘place’, *-nhie* ‘time’, *-ci* ‘manner’, *-ze* ‘amount’ and *-bi* ‘type’. The interrogative has two alternative forms, namely *də* and *dì*, that occur with general terms and a third form *thu* (denoting persons) that occurs as an unstructured form. The interrogative is different from other proforms on another point, namely that it has an additional element *chə* in its temporal form *dì-chə-nhie*. The indefinite paradigm is also rather different, as it has no forms for denoting ‘something’, ‘somehow’, and ‘some amount’. The pronominal elements of all these proforms can function on their own as determiners in noun phrases.

Another example for this dual structure of proforms can be given from Kolyma Yukaghir (Palaeo-Siberian: Maslova 1998: 290). The language shows a distinction between proximate, medial, and distal demonstratives and has a distinct paradigm of interrogatives.

(2)	Proximate	Medial	Distal	Interrogative
Attributive	<i>tɪŋ</i>	<i>adɪŋ</i>	<i>taŋ</i>	<i>qadi</i>
Nominal	<i>tuõn</i>	<i>aduon</i>	<i>tawun</i>	<i>qado:n</i>
Locative	<i>ti:</i>	<i>ada:</i>	<i>ta:</i>	<i>qon</i>
Direction	<i>tɪŋide</i>	<i>adaŋide</i>	<i>ta:ŋide</i>	<i>qaŋide</i>
Source	<i>ti:t</i>	<i>ada:t</i>	<i>ta:t</i>	<i>qot</i>
Temporal	<i>te:ne</i>	<i>tuda:</i>	<i>ta:ne</i>	<i>qanin</i>
Kind	<i>ti:tmie</i>	—	<i>ta:tmie</i>	<i>qodimie</i>

In addition to these, there are a few interrogative pronouns in Kolyma Yukaghir that do not fit into these paradigms, such as *kin* ‘who, whose’, *leme/neme* ‘what’, and *noŋon* ‘what for’. In the case of these proforms and also in that of the proforms that occur in the paradigms, it is more difficult to identify the

constituent elements than it was in the case of the Khezha proforms given earlier in (1). However, the general pattern is clearly discernible, namely the attachment of a general term denoting concepts like attribution, reference, location, direction, source, time, and kind to a set of pronominal elements.

Traditional grammars of Indo-European languages often contain a table of 'correlative pronouns' that includes demonstratives, interrogatives, indefinites, and relatives. These tables are also very clearly based upon this dual structure of proforms. Columns of these tables represent pronominal elements and rows represent general concepts. The following table of Ancient Greek from Haspelmath (1997: 30) exemplifies this tradition.

(3)		Demonstratives	Interrogatives	Relatives	Indefinites
	Person	<i>hoũtos</i>	<i>tís</i>	<i>hós</i>	<i>tis</i>
	Thing	—	<i>tí</i>	—	<i>ti</i>
	Property	<i>toiósde</i>	<i>poĩos</i>	<i>hoĩos</i>	<i>poiós</i>
	Place	<i>ekeĩ</i>	<i>poũ</i>	<i>hoũ</i>	<i>pou</i>
	Time	<i>tóte</i>	<i>póte</i>	<i>hóte</i>	<i>poté</i>
	Manner	<i>hoútōs</i>	<i>pōs</i>	<i>hōs</i>	<i>pōs</i>
	Amount	<i>tosósde</i>	<i>pósos</i>	<i>hósos</i>	<i>posós</i>

Haspelmath notes that the pronominal stems in these Greek paradigms are *hou-/to-* 'demonstrative', *ti-/po-* 'interrogative-indefinite' (with the two paradigms differing in stress), and *ho-* 'relative'. The affixes that are attached to them in order to indicate ontological categories are *-ou* 'place', *-ote* 'time', *-s* 'manner', etc.

Several languages manifest compact systems of proforms of this type (as shown by paradigms given later on in this chapter), but the occurrence of such a compact system of proforms is by no means a language universal characteristic. English, for example, has a truncated system of proforms, especially in the case of its demonstrative paradigms (4) with several gaps (Quirk, Greenbaum, Leach, and Svartvik 1985: 369). The forms are not generally put into a single table as in (4), apparently because of these gaps.

(4)	Demonstratives		Interrogatives	Indefinites	
	Proximate	Remote		Specific	Non-specific
	<i>this</i>	<i>that</i>	<i>which</i>	<i>some</i>	<i>any</i>
	—	—	<i>who</i>	<i>someone</i>	<i>anyone</i>
	—	—	<i>what</i>	<i>something</i>	<i>anything</i>
	<i>here</i>	<i>there</i>	<i>where</i>	<i>(somewhere)</i>	<i>(anywhere)</i>
	—	<i>then</i>	<i>when</i>	<i>sometime</i>	<i>anytime</i>
	—	<i>thus</i>	<i>how</i>	<i>(somehow)</i>	<i>(anyhow)</i>
	—	—	<i>why</i>	—	—
	<i>hither</i>	<i>thither</i>	—	—	—

English also has certain more complex forms that contain more than two constituent elements such as *thereafter*, *therefore*, *whereat*, *whereto*, etc. Further, some of the forms included in the two indefinite paradigms, such as *somewhere/anywhere* and *somehow/anyhow* also involve more than two elements.

An examination of the grammars of several languages indicates that the majority of them list proforms that can only form truncated systems like that of English rather than compact ones like that of Khezha or Kolyma Yukaghir. There is reason to believe, however, that in the case of at least some of these languages, the system of proforms that actually exists is more like that of Khezha than that of English. One reason for this partial representation of proforms in grammars is apparently the use of English as the eliciting language. Most grammars provide a full set of forms of the interrogative paradigm that resemble the set of interrogative forms occurring in English. However, they rarely attempt to provide an equivalent set of forms for the demonstrative paradigm, apparently because English does not have such a set of forms. Hence in some of them, at least, the gaps occurring in the demonstrative paradigms may not be representing the actual state of affairs.

For example, Kapfo (1993) gives proforms of Khezha (Tibeto-Burman) that can only form a truncated system of paradigms like that of English. Most of these gaps were filled later on through elicitation (in personal communication). It is possible, of course, that these gaps represent the frequency of occurrence of the relevant forms. That is, interrogative proforms might be used more frequently in the language as compared to the various demonstrative proforms that correspond to them, and this might be the reason why the linguist was able to readily recognize the occurrence of the former but not the latter. A comparative study of the occurrence of proforms in texts would probably indicate the correctness (or incorrectness) of such a claim. The occurrence of real gaps in languages like English also apparently supports such a claim. However, the effect of using English as the eliciting language, as mentioned earlier, cannot be completely discounted.

The gaps appearing in the set of proforms given in the grammar of Iraqw, a Cushitic language (Mous 1993: 119) are of a different type. The grammar provides a full list of interrogative pronouns that are derived by attaching the interrogative suffix *-má* or *-lá* to a general noun. Iraqw also has a set of deictic suffixes, namely *í* (*ká* for neuter) 'near the speaker', *síng* 'near the addressee', *qá* 'away from both', and *dá* 'far away' that can be attached to any given noun in order to indicate the location of its referent. These deictic suffixes can also be attached to the set of general terms from which the interrogative pronouns are derived as shown by the sets of paradigms given in (5) (Mous, personal communication).

(5)		Near 1	Near 2	Near 3	Far	Interrogative
	<i>hee</i> 'man'	<i>heewí</i>	<i>heesíng</i>	<i>heeqá'</i>	<i>heedá'</i>	<i>heemá</i>
	<i>gaa</i> 'thing'	<i>gaarí</i>	<i>gaasíng</i>	<i>gaarqá'</i>	<i>gaadá</i>	<i>gaalá</i>
	<i>dii</i> 'place'	<i>diirí</i>	<i>diisíng</i>	<i>diirqá'</i>	<i>diidá'</i>	<i>diimá</i>

<i>aamo</i>	'place'	<i>aamorí</i>	<i>aamosíng</i>	<i>aamorqá'</i>	<i>aamodá'</i>	<i>aamá</i>
<i>adoo</i>	'manner'	<i>adoorí</i>	<i>adoosíng</i>	<i>adoorqá'</i>	<i>adoodá'</i>	<i>adoomá</i>
<i>daqa</i>	'moment'	<i>daqarí</i>	<i>daqasíng</i>	<i>daqarqá'</i>	<i>daqadá'</i>	<i>daqmá</i>

Mous points out, however, that none of these forms has any special characteristics that make it necessary to regard them as 'pronouns'. All are merely nominal forms that occur with deictic and interrogative affixes. It is clearly the possibility of giving pronominal equivalents in English that gave a special position for the interrogatives in Mous (1993). There are, on the other hand, a few additional interrogative pronouns that do not fit into this table, namely *xaylá* 'when', *milá* 'what', and *magá* 'how many'. There is also a set of demonstrative pronouns formed by attaching the deictic suffixes to gender markers, *ku-* 'M/N' and *ta-* 'F', with the latter functioning as general terms (*kwí*, *kwisíng*, *kuqá'*, and *kudá'* are the M/N forms).

7.2.1 Order of constituents

Most of the languages that were examined for this study have at least some proforms with a dual structure consisting of a pronominal element and a general term. In most of them, the pronominal element precedes the general term, but there are also several languages in which it *follows* the general term. Consider, for example, the paradigms of the proforms of Maybrat (a West Papuan language: Dol 1999: 107) given in (6).

(6)	Very near 1	Near 1	Far from 1	Unmarked	Interrogative
Adjectival:					
Location (specific)	<i>refo</i>	<i>reto</i>	<i>rono</i>	<i>reau</i>	<i>royo</i> 'which one'
Location (general)	<i>wefo</i>	<i>weto</i>	—	<i>wEAU</i>	—
Area	<i>tefo</i>	<i>teto</i>	—	<i>teau</i>	<i>toyo</i>
Side	<i>tifo</i>	—	<i>tino</i>	<i>tiau</i>	—
Adverbial:					
Location	<i>tofo</i>	—	<i>tono</i>	<i>toau</i>	—
Location (general)	<i>wofO</i>	—	<i>wono</i>	—	<i>woyo</i>
Area	<i>pefo</i>	<i>peto</i>	<i>peno</i>	—	—
Similar to	<i>fifo</i>	<i>fito</i>	<i>fino</i>	<i>fiau</i>	<i>fiye</i>
Presentative	<i>mefo</i>	<i>meto</i>	<i>mno</i>	<i>meau</i>	<i>miyo</i>

There are a few additional proforms in the language that do not appear to fit into this set, such as *tíya* 'how much', *títiya* 'when', and *awiya* 'who', but they too appear to involve a similar structure. Some of the forms show gender distinction and take the suffixes *i*, *ait*, or *e* in order to denote masculine gender.

Several other languages manifest similar structures in which the pronominal element follows the general term. This contrast in the ordering of constituent elements can be seen in the two sets of proforms given in (7) as well. The first two sets of languages, namely Jaqaru (Hardman 2000: 27) and Sanskrit (Speijer 1884: 204) have proforms with the Pronominal element-General term (PG) order

whereas the next two languages, namely Apurinã (Facundes 2000: 366) and Miya (Schuh 1998: 331) have proforms with the General term-Pronominal element (GP) order. All the examples given in (7) represent interrogative proforms.

(7)	Proforms with a PG order		Proforms with a GP order	
	Jaqaru	Sanskrit	Apurinã	Miya
who	<i>qa-chi</i>	<i>k-aḥ</i>	<i>ki-pa</i>	<i>wà-a</i>
where	<i>qa-si</i>	<i>k-va</i>	<i>na-pa</i>	<i>‘íkw-a</i>
how	<i>qa-misha</i>	<i>k-atham</i>	<i>kanhi-pa</i>	<i>wánkwa-a</i>
how much	—	<i>k-iyat</i>	—	<i>mən-a</i>
why	<i>qu-sit’a</i>	<i>k-utah</i>	<i>kenere-pa</i>	<i>émà-a</i>

In Miya, the initial elements of *‘íkw-a* ‘where’ and *wánkwa* ‘how’ are reported to be recognizable as *‘íy* ‘place’ and *wán* ‘like’ (Schuh 1998: 331); these occur as general terms in the proforms.

This distinction in the relative order of the constituents of proforms appears to constitute an interesting typological distinction, as it is possible to correlate it with other word-order distinctions appearing among languages. The use of GP structure for proforms appears to be one of the characteristics of verb-initial languages. I have found all the Polynesian languages that were examined for this study, such as Hawaiian (Elbert and Pukui 1979: 110), Maori (Bauer 1999: 361), Tuvaluan (Besnir 2000: 405), Samoan (Mosel and Hovdhaugen 1992: 129) and West Futuna Aniwa (Dougherty 1983: 33) to have proforms with a GP structure. Other verb-initial languages like Tagalog (Schachter and Otanes 1972: 504), Sawu (Sumba-Bima, Austronesian: Walker 1982: 40), Tinrin (Melanesian: Osumi 1995: 229), Taba (Eastern Malayo-Polynesian: Bowden 1997: 412), Apurinã (Maipure, Arawak: Facundes 2000: 366), and Miya (Chadic: Schuh 1998: 331) also show a similar order of constituents among their proforms.

The GP order occurs among the proforms of a few other Austronesian languages even though they are verb-medial rather than verb-initial. Maybrat (West Papuan: Dol 1999: 105) is a language of this type (see (6) given earlier for its proforms). Mokilese (Micronesian: Harrison 1976: 313) also shows a few interrogative proforms with GP structure, such as *amw-da* ‘why’ (*da* ‘what’). Lampung (Malayan, Western Austronesian: Walker 1976: 24) also has forms like *sapa* ‘who’, *sipa* ‘where’, *xepa* ‘how’, *ulihapi* ‘why’ that appear to show a GP structure. Another language that appears show the GP order in spite its being a verb-medial language is Vietnamese (Thompson 1991: 142).

(8)	Proximate	Distant	Indefinite
Nominal	<i>đây</i>	<i>đâ’y</i>	<i>đâu</i>
Place	<i>này</i>	<i>nọ</i>	<i>nào</i>
Extent	<i>bây</i>	<i>bâ’y</i>	<i>bao</i>
Manner	<i>vây</i>	<i>vây</i>	<i>sao</i>

Some of the Bantu and Niger-Congo languages also show GP structure in spite of the fact that they are verb-medial. For example, Diola-Fogny (Niger-Congo: Sapir 1965: 58, 71) derives its interrogatives by attaching $-\varepsilon y$ to noun class (concord) markers. It also derives its demonstratives in a similar fashion, by attaching the deictic elements $-\varepsilon$ 'here', $-u$ 'there (indefinite as to locality)', $-a$ 'there (definite locality)', $-apa$: 'over there', and $-apa:m\phi$ 'away over there' to the stem $uC(\varepsilon)-$ in which C represents the concord element. The former can also occur with certain semi-nominal forms like $ta:t-$ 'circumscribed area', $ba:b-$ 'general area', and $d\phi r-$ 'within'.

The proforms of Venda (Bantu: Poulos 1990: 105, 455) also show basically a GP structure. The language has four different sets of demonstratives, which denote spatial distinctions between (i) proximate to 1, (ii) relatively close to 1, (iii) further away from 1, and (iv) distant. They are derived by suffixing different pronominal elements to noun class markers. Similarly, its adnominal interrogatives are formed by suffixing the interrogative element $-fhi\phi$ 'which' to enumerative noun class (concord) markers. There are also a few additional interrogative proforms that involve the use of the suffix $-ni$ 'what' to general terms as seen in (9).

- (9)
- | | |
|------------------|---|
| <i>lini</i> | 'when' |
| <i>hani</i> | 'how' |
| <i>mini</i> | 'what' |
| <i>-ngafhani</i> | 'how much' (attached to noun class markers) |
| <i>-nzani</i> | 'what colour' (,,) |

It is also possible to add $-ni$ 'what' to verbs in order to ask a question about a particular object, as in *u vhona-ni* 'what do you see?' There is also a locational interrogative element $-fhi$ that can be suffixed to verbs in order to ask a question about a location (Poulos 1990: 451).

In some verb-initial languages, we find the PG order to be occurring in some proforms, and GP order to be occurring in others. In Breton (Celtic, Indo-European: Stephens 1993: 386), for example, demonstratives show a GP order, which is in tune with the fact that the language is a verb-initial one, but interrogatives show a PG order (10). This disparity probably derives from the fact that the language retains its interrogative pronouns from an earlier stage in which it was not a verb-initial language. Celtic is considered to be a branch of the Indo-European family, and the proto-language of that family is reconstructed as having a verb-final structure.

- (10) a. *Demonstratives with a GP structure*

<i>a-mañ</i>	'here'	<i>du-mañ</i>	'over here'
<i>a-ze</i>	'there'	<i>du-ze</i>	'over there'
<i>a-hont</i>	'there (remote)'	<i>du-hont</i>	'over there (remote)'

b. *Interrogatives with a PG structure*

<i>petra</i>	‘what’
<i>pelec’h</i>	‘where’
<i>pegoulz</i>	‘when’
<i>penaos</i>	‘how’
<i>pegement</i>	‘how much’

A similar type of situation occurs in Yosondúa Mixtec (Farris 1992: 35, 136). It also uses the PG order in the case of its interrogative pronouns and GP order in the case of its demonstrative pronouns. The proform for ‘place’ appears to have the word *īchī* ‘trail’ used as a general term.

(11) a. *Interrogative pronouns*

<i>ndōō</i>	‘what’
<i>nāvā</i>	‘who’
<i>náchī</i>	‘where (direction)’
<i>nándī</i>	‘from where’
<i>nāmā</i>	‘when’
<i>nāsā</i>	‘how, how much’

b. *Demonstrative pronouns*

Proximate		Remote	
<i>xáhá</i>	‘this, here’	<i>yúkán</i>	‘that, there’
<i>chāhá</i>	‘this place’	<i>chúkán</i>	‘that place’
<i>syáhán</i>	‘in this way’	<i>syúkán</i>	‘in that way’

There are other types of mixed systems in which one of the two constituent orders appears to be more prominent than the other one. For example, Kobon (Papua New Guinea: Davies 1981: 8, 161) uses two different deictic elements, *-i* ‘proximate’ and *-u* ‘remote’, for forming its demonstrative pronouns. They are suffixed to general terms in most of the instances, as shown in (12a), but there is prefixing in one instance (12b). Similarly, the interrogative is generally formed by suffixing the element *-ai* to general terms, as shown in (13a), but in one case (13b), it is prefixed.

- | | | | | | |
|---------|---------------|----------------|----|--------------|---------------|
| (12) a. | <i>rīm-ni</i> | ‘some of this’ | b. | <i>in-bō</i> | ‘in this way’ |
| | <i>rīm-nu</i> | ‘some of that’ | | <i>un-bō</i> | ‘in that way’ |
| | <i>ga-i</i> | ‘here’ | | | |
| | <i>ga-u</i> | ‘there’ | | | |
| | <i>da-ni</i> | ‘this’ | | | |
| | <i>da-nu</i> | ‘that’ | | | |

- | | | | | | |
|---------|---------------|----------------|----|----------------|------------|
| (13) a. | <i>g-ai</i> | ‘where’ | b. | <i>ai-gege</i> | ‘how much’ |
| | <i>mañ-ai</i> | ‘when’ | | | |
| | <i>m-ai</i> | ‘which, where’ | | | |
| | <i>au-ai</i> | ‘where’ | | | |

Similarly, Ambulas (Sepik: Wilson 1980: 38, 55–7) derives some of its demonstrative proforms by prefixing the deictic elements *ken* ‘proximate’, *an* ‘medial’ and *wan* ‘distal’ to general terms like *-i* ‘adjectival’, *-et* ‘goal’, *-ba* ‘location’ and *-ga* ‘similar’. It also forms interrogative proform in a similar way. However, animate demonstrative pronouns are formed by prefixing third person pronouns, *dé* ‘MSg’, *lé* ‘FSg’, *bét* ‘Dual’ and *de* ‘Pl’ to these deictic elements. That is, the structure of proforms involves PG order in all cases excepting that of animate demonstratives.

- | | | | | | |
|---------|----------------|-----------------------|----|---------------|------------------------|
| (14) a. | <i>kén-i</i> | ‘this’ | b. | <i>yan-i</i> | ‘which’ |
| | <i>kén-et</i> | ‘to here’ | | — | |
| | <i>ké-ba</i> | ‘here’ | | <i>ya-ba</i> | ‘where from, where to’ |
| | <i>ké-ga</i> | ‘like this’ | | <i>ya-ga</i> | ‘what’ |
| | — | | | <i>ya-gap</i> | ‘how many’ |
| c. | <i>dé-kén</i> | ‘this one (M)’ | | | |
| | <i>lé-kén</i> | ‘this one (F)’ | | | |
| | <i>bét-kén</i> | ‘these two (Dual)’ | | | |
| | <i>de-kén</i> | ‘these ones (Plural)’ | | | |

It is possible that these irregularities have explanations in the diachrony of these languages as in the case of languages like Breton described earlier. Another point that needs to be noted here is the possibility that some of the proforms are not lexicalized compact forms in the language but only loose structures that involve the association of a noun with a determiner that follows it as in the case of the proforms of Iraqw examined earlier (5). The occurrence of a determiner after the noun that it modifies is a characteristic of prototypical verb-object languages (Andersen 1983: 41) that may or may not be verb-initial, whereas the occurrence of GP structure among the proforms does not appear to be a characteristic of languages in which the verb is not initial.

In fact, in the case of verb-medial languages like English, the need to have a PG structure, with the pronominal element occurring first, is so strong that it appears to have forced prepositions to function as postpositions or suffixes when they are associated with proforms. There are forms like *there-in*, *where-from*, *where-as* etc. in English that appear to have resulted from this tendency.

7.2.2 Identifying the constituents

In the case of prototypical proforms, there would be no difficulty in identifying the two basic constituents, namely a pronominal element and a general term because they would be denoting prototypically quite distinct concepts. For example, pronominal elements generally denote deictic notions like being proximate, medial, or remote, visible or invisible, up or down, and other functional distinctions like being indefinite (having lack of knowledge), interrogative, or relative (anaphoric). General terms, on the other hand, denote categorical

distinctions like person, thing, place, time, manner, quality, quantity, reason, cause, etc. The sets of paradigms given above clearly manifest this semantic distinction between the two constituents of proforms.

There is a need to differentiate between these basic constituents of proforms on the one hand, and other elements that are attached to them for various syntactic and pragmatic purposes on the other. For example, some languages allow their nominal proforms to take case markers, which are clearly peripheral to the structure of those proforms. The fact that the occurrence of these markers is restricted to a part of the various paradigms like demonstratives, interrogatives, indefinites, and relatives is an indication of this point. Further, the occurrence of such markers is not a unique characteristic of proforms, since the markers occur with other nominals as well. This is also true of agreement markers that can be attached to adjectival and verbal proforms, and other similar markers. They do not form part of the structure of proforms.

For example, Kannada (Dravidian) allows all the case markers that occur with nouns to occur with nominal proforms as well. None of them is specific to proforms.

(15)		‘this person’	‘king’	‘who’	‘god’
	Nominative	<i>ivanu</i>	<i>ra:ja</i>	<i>ya:ru</i>	<i>de:varu</i>
	Accusative	<i>ivan-annu</i>	<i>ra:jan-annu</i>	<i>ya:r-annu</i>	<i>de:var-annu</i>
	Ablative	<i>ivan-inda</i>	<i>ra:jan-inda</i>	<i>ya:r-inda</i>	<i>de:var-inda</i>
	Dative	<i>ivan-ige</i>	<i>ra:jan-ige</i>	<i>ya:r-ige</i>	<i>de:var-ige</i>
	Genitive	<i>ivan-a</i>	<i>ra:jan-a</i>	<i>ya:r-a</i>	<i>de:var-a</i>
	Locative	<i>ivan-alli</i>	<i>ra:jan-alli</i>	<i>ya:r-alli</i>	<i>de:var-alli</i>

Similarly, Marathi (Indo-Aryan: Pandharipande 1997: 376) attaches the agreement markers that denote gender and number distinctions to adnominal proforms in exactly the same way in which it attaches them to adjectives. It has a set of interrogative pronouns and two sets of demonstrative pronouns (proximate and remote) that can occur in the adnominal position. All of them show distinctions of gender and number in order to agree with the head noun.

(16)	Demonstrative pronouns				Interrogatives	
	Proximate		Remote			
	Singular	Plural	Singular	Plural	Singular	Plural
M	<i>hā</i>	<i>he</i>	<i>tā</i>	<i>te</i>	<i>koṇṭā</i>	<i>koṇṭe</i>
F	<i>hī</i>	<i>hyā</i>	<i>tī</i>	<i>tyā</i>	<i>koṇṭī</i>	<i>koṇṭyā</i>
N	<i>he</i>	<i>hī</i>	<i>te</i>	<i>tī</i>	<i>koṇṭa</i>	<i>koṇṭī</i>

It is possible, however, for these forms to represent a combination of pronominal element and general term, with the gender-number markers functioning as general terms, as for example when these are used as independent pronouns rather than as modifiers (see §7.3.4).

Languages in which proforms can be used in the position of the verb allow all the relevant tense and aspect markers and also person markers to occur with them. For example, Mapuche (possibly Southern Andean: Smeets 1989: 424) has a set of deictic verbs derived from the roots *fa-* 'to become like this' and *fe-* 'to become like that' that can take the various verbal inflections like other intransitive verbs.

- (17) *ka fe-le-pa-tu-n*
 again become.like.that-Stative-hither-Iterative-Indicative (1Sg)
 'I was in the same situation as before'

Proforms may also be associated with notions like emphasis, contrast, or focus, specificity or non-specificity, precision, etc. that are meant for denoting distinctions in their pragmatic functions. It is necessary to differentiate these from the two basic constituents of proforms mentioned earlier. For example, Tümpisa Shoshone (Uto-Aztecan: Dayley 1989) has a set of deictic terms, *i* 'right here', *e* 'here, nearby', *a* 'there, visible', and *u* 'there, not visible' that can occur with certain general terms like *-nni* 'way, like', *ittün* 'kind', etc. It is possible to attach the prefix *s-* to these proforms in order to associate them with the pragmatic meaning of being the topic of discourse. Similarly, several languages allow emphatic particles to be attached to demonstrative proforms in order to provide contrastive meaning. This can be exemplified with the help of the Kannada (Dravidian) sentences given in (18).

- (18) a. *i: mudukan-annu na:nu ninne no:ḍ-idde*
 this old.man-Acc I yesterday see-was
 'I had seen this old man yesterday'
- b. *id-e: mudukan-annu na:nu ninne no:ḍ-idde*
 this-Emph old.man-Acc I yesterday see-was
 'I had seen this very old man yesterday'

It is evidently necessary to regard these different types of inflectional and pragmatic markers to be peripheral to the structure of proforms, just as it is necessary to regard them as peripheral to the structure of nouns, verbs, and adjectives that can also occur with them. As I will be pointing out later on in the next section (§7.3.4), however, the use of some of these markers as general terms in the structuring of proforms makes this differentiation between basic and peripheral constituents of proforms rather difficult to establish in some cases.

7.3 Complex structures

Several languages possess proforms that have structures that are more complex than the dual structure that we expect them to have from the point of view of the basic requirements of their functioning in language. They may show such

complexities either in the representation of the pronominal concepts or of general concepts. On the other hand, languages may also possess proforms that do not appear to have any internal structure, and this might derive from the fusion of the two constituent elements into a single constituent through grammaticalization.

Complexity in the representation of pronominal concepts or general concepts derives from the fact that in some languages a single marker is used for representing two or more concepts, and additional markers are attached to them in order to differentiate between those undifferentiated concepts. For example, a single pronominal element may be used in a language for representing the meaning of 'lack of knowledge' and additional markers may be attached to that element, or to the proform that contains the element, in order to indicate different purposes for which the lack of knowledge is expressed, such as requesting the addressee to provide the relevant information (interrogative proforms), leaving the details unspecified (specific indefinite proforms), leaving the referent itself unspecified (non-specific indefinite proforms), etc. This, however, has the effect of making the proforms more complex (§7.3.1). Similar complexity can occur in the expression of general concepts as well (§7.3.2).

7.3.1 *Complexity in the expression of functional distinctions*

In the case of some languages like Khezha (see (1) given in §7.2), proforms have a uniform structure. They occur with distinct pronominal elements for representing the various primary functions of proforms. On the other hand, there are several languages in which two or more of these primary functions are collapsed and represented through a single pronominal element. A comparison of the paradigms of Khezha with those of Yidiny (Australian: Dixon 1977), given below in (19), makes this point clear. While Khezha differentiates between interrogative and indefinite meanings through the use of two distinct sets of proforms, Yidiny uses only a single set of proforms for both these purposes.

(19)	Gloss	Definite pronouns			Interrogative/ indefinite pronouns
		Near	Far	Very far	
	Animate	<i>yipdu</i>	<i>ɲipdu</i>	<i>yupdu</i>	<i>wapdu</i>
	Inanimate	<i>yingu-</i>	<i>ɲungu-</i>	<i>yungu-</i>	<i>wapi</i>
	Adjectival	<i>yɪɲu</i>	<i>ɲuɲu</i>	<i>yɪɲu</i>	<i>wapa</i>
	Place/time	<i>yingu</i>	<i>ɲungu</i>	<i>yungu</i>	<i>wapda</i>
	Manner	<i>yingu:rup</i>	<i>ɲungu:rup</i>	—	<i>wapda:rup</i>
	Source	<i>yɪngum</i>	<i>ɲungum</i>	—	<i>wapɖum</i>
	Type	<i>yɪɲarip</i>	<i>ɲuɲarip</i>	<i>yunɲarip</i>	<i>wapda:rip</i>

Another language that has distinct sets of proforms for interrogatives and indefinites is Ainu (Tamura 2000: 232). The interrogatives generally have an initial *h-* whereas the indefinites have an initial *ne-*.

(20)	Interrogatives	Indefinites
	‘who’	<i>hunna</i> <i>nen</i>
	‘what’	<i>h̃ñta</i> <i>nep</i>
	‘where’	<i>hunak</i> <i>nei</i>
	‘when’	<i>hempara</i> <i>nei</i>
	‘how’	<i>mak</i> <i>neu</i>

Languages that have a single pronominal element for representing two or more primary functions may use additional devices for differentiating between those functions. This, however, has the effect of making the representation of pronominal concepts more complex. While representing the interrogative–indefinite distinction, for example, languages may use additional markers either in the case of the interrogative or in that of the indefinite. For example, Zuni (S. Newman 1996: 499) forms its interrogative pronouns by adding the suffix *-pi*, *-ppi*, or *-ʔpi* to indefinite pronouns (see §10.2.4 for additional languages of this type).

(21)	<i>koʔ</i>	‘something intangible’	<i>koʔ-pi</i>	‘what, how’
	<i>ko.wi</i>	‘few’	<i>ko.wi-ʔpi</i>	‘how much’

On the other hand, there are several languages that add affixes to their unmarked interrogative–indefinite proforms for specifying the indefinite rather than the interrogative meaning. The affixes may be used not only for specifying the indefinite meaning but also for differentiating between different types of indefinite meanings. For example, several Dravidian languages attach a conjunctive particle to unmarked interrogative–indefinite pronouns in order to derive non-specific indefinites, and a disjunctive particle in order to derive specific indefinites. The Kannada interrogative and indefinite pronouns given in (22), with the latter (indefinite pronouns) containing the conjunctive particle *u*: ‘and, also’ in the non-specific paradigm and the disjunctive particle *o*: ‘or’ in the specific paradigm, exemplify this usage.

(22)	Interrogatives	Indefinites	
		Non-specific indefinites	Specific indefinites
	<i>ya:-ru</i> ‘who’	<i>ya:-r-u:</i> ‘anyone’	<i>ya:-r-o:</i> ‘someone’
	<i>e:-nu</i> ‘what’	<i>e:-n-u:</i> ‘anything’	<i>e:-n-o:</i> ‘anything’
	<i>e-lli</i> ‘where’	<i>e-ll-u:</i> ‘anywhere’	<i>e-ll-o:</i> ‘somewhere’
	<i>he:-ge</i> ‘how’	<i>he:-g-u:</i> ‘anyhow’	<i>he:-g-o:</i> ‘somehow’
	<i>ya:-ke</i> ‘why’	<i>ya:-k-u:</i> ‘for any reason’	<i>ya:-k-o:</i> ‘for some reason’

Notice that the addition of these differentiating particles has made the indefinite proforms more complex in that they now possess three different elements each (see §11.2.1 for additional examples).

There are also several languages that differentiate between the interrogative and the indefinite uses of a single set of proforms by using differentiating markers elsewhere in the sentence. For example, several languages use

an interrogative particle that is attached to the verb in the case of constituent questions. When the particle is not used in the sentence, the proforms provide only indefinite meaning. Other devices used for this purpose include interrogative verb and interrogative intonation (see §10.2).

Languages may also use a single pronominal element for representing different deictic meanings and differentiate between them through the use of additional markers. For example, Wambaya (West Barkley family, Australian: Nordinger 1998: 125) has four distinct demonstrative pronouns that show a four-fold noun class distinction. The pronominal element and the general term (that denotes the noun class distinction) have fused together into compact forms. These stems can be used without any modification to denote the notion of proximity, but in order to express the notion of remoteness, they are used with the suffix *-iyaga*. All these derived and non-derived pronominal elements occur with number and case markers for showing agreement with their head noun. They can also be used as independent pronouns.

(23)	Class	Proximate	Remote
	I	<i>ini</i>	<i>iniyaga</i>
	II	<i>nana</i>	<i>naniyaga</i>
	III	<i>mama</i>	<i>mamiyaga</i>
	IV	<i>yana</i>	<i>yaniyaga</i>

Another language of this type is Kilivila, an Austronesian language of Trobriand Islands (Senft 1986: 64). Demonstrative pronouns are formed in this language by prefixing the proximate marker *m-* to one of the classificatory particles (which are reported to number about 176). The particle is followed by a word-final *-na*. In order to form remote demonstratives, the affix *-we* is placed between the classificatory particle and the word-final *-na*. That is, the remote demonstratives contain all the elements that occur in the proximate demonstratives and an additional affix *-we* as shown in (24).

(24)	Proximate	Remote	
	<i>m-to-na</i>	<i>m-to-we-na</i>	(<i>to</i> 'particle for man')
	<i>mi-na-na</i>	<i>mi-na-we-na</i>	(<i>na</i> 'particle for woman')
	<i>ma-ke-na</i>	<i>ma-ke-we-na</i>	(<i>ke</i> 'particle for wooden')

Languages that possess two or more distinct pronominal elements for denoting deictic distinctions may combine some of their pronominal elements together in order to express additional deictic distinctions. For example, Mao Naga (Tibeto-Burman: Giridhar 1994: 148) makes a three-fold deictic distinction between *hi* 'near 1', *ti* 'near 2', and *lo* 'remote from both'. In addition to this, it also indicates (i) the notion of 'relatively nearer to 2' by combining *lo* 'remote' with *ti* 'near 2' (*loti*) and (ii) the notion of 'distant (from 1)' by combining *lo* 'remote' with *hi* 'near 1' (*loohi*). These derived pronominal stems are similar to non-derived ones

in their ability to occur with general terms like *mii* 'person', *ta* 'kind', *li* 'manner', etc. to form demonstrative pronouns.

Another language of this type is Maricopa (Yuman: Gordon 1986: 55). It has four different demonstrative roots, namely *da* 'proximate', *va* 'medial', *aas* 'distant, out of sight', and *aany* 'anaphoric'. The medial demonstrative root can combine with proximate and distant demonstrative roots to give two additional demonstrative stems: *v-da* 'near, in hand' and *sva* 'distant'. The remote demonstrative *aas* has the meaning 'out of sight' in this paradigm. All these and the interrogative root *mki* can occur with different case markers to denote different nominal and adverbial connotations. (That is, case markers function as general terms in the formation of these proforms—see §7.3.4.)

Alamblak (East Sepik, Papua New Guinea: Bruce 1984: 81) has a demonstrative root to which gender and number markers can be attached to derive demonstrative pronouns that are apparently neutral to deictic distinctions. It is also possible to attach the deictic suffixes *-ar* 'near' and *-ur* 'far' to the demonstrative pronoun in order to form more complex demonstratives that specify the relevant deictic distinctions. Number and gender markers (*-r* 'M' and *-t* 'F') are attached to these derived demonstratives as well.

(25)	Neutral	Proximate	Remote
3MSg	<i>ind-r</i>	<i>ind-ar-r</i>	<i>ind-ur-r</i>
3FSg	<i>ind-t</i>	<i>ind-ar-t</i>	<i>ind-ur-t</i>
3Dual	<i>ind-f</i>	<i>ind-ar-f</i>	<i>ind-f-ur</i>
3Pl	<i>ind-m</i>	<i>ind-ar-m</i>	<i>ind-m-ur</i>

Bruce (1984: 82) notes that the demonstrative root may be left out in the case of both proximate and remote demonstratives, apparently because they do not have any specific function to perform in those expressions. That is, the language shows a tendency to reduce complex forms into simpler ones that have a dual structure.

(26)	<i>ind-ar-m</i>	'these'
	<i>ar-m</i>	'these'

The case of Mundari (Munda, Austroasiatic) is rather interesting in this context. It makes a three-fold deictic distinction between 'near 1', 'near 2', and 'distant' and under each of these it makes an additional three-fold distinction between 'nearest', 'nearer', and 'near' (Osada 1992: 68). The 'near 2' stems are derived from 'near 1' stems by reversing the order of sounds, whereas the 'distant' stems are derived from 'near 2' stems by prefixing *h-*.

(27)	Near 1	Near 2	Distant
Nearest	<i>ni</i>	<i>in</i>	<i>hin</i>
Nearer	<i>ne</i>	<i>en</i>	<i>hen</i>
Near	<i>na</i>	<i>an</i>	<i>han</i>

Demonstrative pronouns are formed from these deictic stems by attaching *y* 'animate' and *a* 'inanimate' to them. The latter apparently take the position of general terms.

(28)	Animate pronouns			Inanimate pronouns		
	Near 1	Near 2	Distant	Near 1	Near 2	Distant
Nearest	<i>niy</i>	<i>iniy</i>	<i>hiniy</i>	<i>nia</i>	<i>ina</i>	<i>hina</i>
Nearer	—	—	—	<i>nea</i>	<i>ena</i>	<i>hena</i>
Near	<i>nay</i>	<i>aniy</i>	<i>haniy</i>	<i>naya</i>	<i>ana</i>	<i>hana</i>

Notice that all the three sets of forms (i.e. altogether nine pronominal stems) are available for forming inanimate proforms, but in the case of animate proforms, only two sets, involving six pronominal stems, are available. On the other hand, when these pronominal elements are used with other general terms, the number of distinctions gets further reduced. Some of them also receive alternative connotations. For example, there are three forms that denote the notion of 'being similar', which are derived from the 'nearer' stems. There are also three additional ones that are derived from the 'near' stems, but instead of denoting the corresponding spatial distinction, they denote a pragmatic notion, namely that the similarity is contrary to what is expected by the addressee (Osada 1992: 73).

(29)	Against addressee's expectation			
	a.			
	<i>neka</i>	'like this'	<i>naka</i>	'like this'
	<i>enka</i>	'like that'	<i>anka</i>	'like that'
	<i>henka</i>	'like yonder'	<i>hanka</i>	'like yonder'

Mundari manifests a similar type of complexity among the pronominal elements that occur in the interrogative and indefinite proforms. It makes a three-fold distinction in the case of its interrogative paradigms. The stems used are *oko*, *ca*, and *ci*, of which the last two, according to Osada (1992: 74), may be related with the distinction occurring between the demonstrative bases *a* 'near' and *i* 'nearest'.

There is also a three-fold distinction in the case of the indefinite proforms. The stems used for this purpose are *oko*, *ja*, and *jeta* of which the first one, *oko*, is identical with the first interrogative stem given earlier, but it differs from the interrogative in its ability to precede the topic marker. The other two are also clearly correlatable with the remaining two interrogative stems, namely *ca* and *ci*. According to Osada (1992: 78), indefinite *jeta* is used more frequently with the negative as compared to *ja*. Both of them differ from *oko* in being translated as 'any' (non-specific), while the translation for *oko* generally involves the indefinite 'some' (specific).

We may regard all these complex structures as resulting from the use of a single pronominal element for representing two or more functions. They appear to be complex because such languages generally try to differentiate between the

various neutralized functions by attaching additional affixes to proforms in the case of one or more of those uses (or in the case of all of them). If we regard these additional markers as forming part of the pronominal element, we would be able to perceive a basic dual structure among these complex forms as well.

7.3.2 Complexity in the expression of general concepts

Languages may show a similar type of complexity in their representation of general concepts. They may use a single general term for representing two or more distinct concepts and attach additional terms in order to differentiate between them. For example, verbal proforms (demonstratives or interrogatives) may be derived (i) directly by attaching the relevant pronominal elements to general verbs like 'do', 'say', 'go' etc. or (ii) by attaching a verbalizer to fully formed proforms. In the latter case, the verbal proform would be more complex than other proforms.

For example, Maricopa (Yuman: Gordon 1986: 56) derives demonstrative verbs by directly attaching deictic prefixes to certain verbs like *duum* 'be', *wiim* 'do', *iim* 'say' and *yemk* 'go'.

- (30) a. *aany-sh* *v-wii-m*
 Dem-Subj Dem-do-Realis
 'He did it this way'
- b. *iipaa-ny-sh* *v-dik-k*
 man-Dem-Subj Dem-lie-Realis
 'The man is lying there'

On the other hand, Tarma Quechua (Adelaar 1977: 179) derives interrogative verbs from interrogative pronouns by attaching the latter to verbs like *ni* 'to say', *na* 'to do', or *ya* 'to become'. Unlike the verbal proforms of Maricopa, these involve a general term in the proform itself in addition to the verb functioning as another general term.

- (31) *ayga* 'how much' *aygani-* 'to ask how much'
 ima 'what' *imana-* 'to do what'
 imani- 'to say what'
 imaya- 'to become what'

Similarly, adjectival proforms may be derived either (i) directly by attaching pronominal elements to a general term, or (ii) indirectly from nominal proforms that already contain a dual structure, by attaching an adjectival affix to them. There are several languages that differentiate between adnominal and pronominal interrogatives by attaching different general terms to the interrogative (or indefinite) pronominal element. For example, Tibetan (Denwood 1999: 94) has the forms *ga.gi* 'which' and *ga.re* 'what' both containing the interrogative element *ga.* and a general term. On the other hand, there are also languages in

which adnominal forms are derived from pronominal ones through affixation. In Manipuri (Tibeto-Burman: Bhat and Ningomba 1997: 73), for example, *kəri* 'what' contains the pronominal element *kə-* and the general term *ri* but *kəɾəmbə* 'which' is more complex, as it contains the former as one of its constituents.

Heine, Claudi, and Hünemeyer (1991: 48, 56) propose to differentiate between basic and non-basic categories among the concepts that are represented by pronouns. The basic concepts are considered to reflect a scale or chain of 'categorical metaphors' as shown in (32).

(32) Person > Object > Activity > Space > Time > Quality

The arrangement of the categories in this chain is considered to be unidirectional, proceeding from left to right. Any category in the chain may serve to conceptualize any other category to its right. Heine, Claudi, and Hünemeyer suggest that languages would tend to have one-word proforms for referring to basic concepts, whereas for expressing non-basic categories, languages would tend to show more complex forms like two-word expressions.

In the case of interrogative pronouns, for example, Heine, Claudi, and Hünemeyer point out that the forms that involve general concepts like person (*who*), thing (*what*), and place (*where*) exhibit minimal morphological complexity, whereas the ones that involve more abstract concepts like time (*when*) and manner (*how*) are slightly more complex. Forms for other concepts like purpose and cause are even more complex, as their derivation would involve the use of one of the preceding ones (almost always the 'object' pronoun) as a constituent.

It is possible that some of the complexities that we notice among proforms reflect this basic-non-basic distinction occurring among the general concepts that are used in their derivation. However, there do occur several exceptions to this generalization. This has been noted by Heine, Claudi, and Hünemeyer in the case of two of the concepts occurring in (32), namely object and activity. These two appear to show a special cognitive relationship in that the languages surveyed by Heine, Claudi, and Hünemeyer did not morphologically differentiate between the two. On the other hand, there do occur languages in which the two are differentiated. For example, Usan (Papuan: Reesink 1987: 58) has an interrogative verb *mâi-s* 'to do what' which is derived from *mâi* 'what'. Marshallese (Micronesian: Zewen 1977: 94) also has an interrogative verb *iet* 'to do what' derived from *it* 'what'. But Lillooet (Interior Salish: Van Eijk 1997: 134, 165) has an interrogative verb *kánəm* 'to do what' that appears to be unrelated to the interrogative pronoun *stam* 'what'. Maung (Australian: Capell and Hinch 1970: 64) has *gunuga* 'what (object)' and *gigi* 'what (action)'.

Heine, Claudi, and Hünemeyer (1991: 58) find English pronouns *how* and *why* also to be exceptions to their prediction that non-basic categories would be represented by two-word expressions. They point out, however, that such cases form a small minority among the languages that they surveyed. It appears,

however, that the representation of 'manner' by a simple one-word expression is far more frequent than that of 'reason' and resembles the representation of the more basic category, 'thing', as shown by the languages given in (33).

- | | | | | | |
|------|-------------------------------|-----------------------|-------------|---------------|--------------|
| (33) | Pipil (Campbell 1985: 57) | <i>ta:</i> | 'what' | <i>ke:n</i> | 'how' |
| | Supyire (Carlson 1994: 186) | <i>jàhá</i> | 'what' | <i>dì</i> | 'how' |
| | Y. Mixtec (Farris 1992: 35) | <i>ndōō</i> | 'what' | <i>nāsā</i> | 'how' |
| | Ika (Frank 1990: 83) | <i>bema</i> | 'which one' | <i>azi</i> | 'how' |
| | Trumai (Guirardello 1999: 34) | <i>han</i> | 'what' | <i>hele</i> | 'how, what' |
| | E. Pedee (Harms 1994: 114) | <i>k^hā</i> | 'what' | <i>sā</i> | 'how, where' |
| | Japanese (Hinds 1986: 231) | <i>nan(i)</i> | 'what' | <i>doo</i> | 'how' |
| | Mbay (Keegan: 1997: 67) | <i>qí</i> | 'what' | <i>bāy</i> | 'how' |
| | Kamera (Klamer 1998: 132) | <i>nggamu</i> | 'what' | <i>nggiki</i> | 'how' |

7.3.3 Effects of grammaticalization

There is generally a tendency among languages to grammaticalize constructions that are most frequently used or the ones that involve concepts that are very closely associated with one another (Bybee 1985). We find this tendency affecting proforms such that their constituent elements lose their identity and get fused into unitary expressions. This tendency appears to affect interrogative (or indefinite) pronouns more prominently than demonstrative pronouns. There are several languages in which the constituent elements of demonstrative pronouns are easy to identify and separate from one another whereas the constituents of interrogative-indefinite pronouns are difficult or impossible to identify.

For example, Tamazight (Berber: Penchoen 1973: 21, 79) makes a two-fold distinction between proximate (*-ad/-a*) and remote (*-in*) demonstratives. The pronominal elements can be used with certain general terms to derive demonstrative pronouns as shown in (34).

- | | | | | |
|------|--------------|------------------|---------------|------------------|
| (34) | <i>wad</i> | 'this one (M)' | <i>wann</i> | 'that one (M)' |
| | <i>t ad</i> | ',, ,, (F)' | <i>tann</i> | ',, ,, (F)' |
| | <i>ayad</i> | ',, ,, (action)' | <i>ayinn</i> | ',, ,, (action)' |
| | <i>da(d)</i> | 'here' | <i>dinn</i> | 'there' |
| | <i>sya</i> | 'hereabouts' | <i>syinn</i> | 'thereabouts' |
| | <i>awra</i> | 'hither' | <i>awrinn</i> | 'thither' |

There is, on the other hand, a set of interrogative adverbials in the language whose underlying structure is rather non-transparent.

- | | | | | |
|------|--------------|----------------------|-----------------|---------|
| (35) | <i>milmi</i> | 'when' | <i>mani</i> | 'where' |
| | <i>mimš</i> | 'how' | <i>matta</i> | 'what' |
| | <i>mšta</i> | 'how much, how many' | <i>mahallig</i> | 'why' |

Penchoen (1973: 80) suggests that the question words were, very probably, originally made up of *m-* and a noun form; at least one of them, namely *mimš* 'how'

is clearly related to *imšis* ‘thus, this way’. Tamazight has an alternative way of forming constituent questions, namely to attach an initial *m-* to a noun that is to be questioned. It is also possible to attach *-mi* to prepositions as in *-s-mi* ‘with whom’, *xf-mi* ‘on what, about whom, what’, *i-mi* ‘to whom’, and *šəgg d-mi* ‘you and who’ (1973: 82).

There are also languages in which some proforms are clearly unstructured in that they contain only a pronominal element. For example, several languages use the pronominal elements of their proforms as determiners or modifiers of nouns and less frequently, of verbs. Kannada (Dravidian) is a language of this type. Its demonstrative (*i*: ‘proximate’ and *a*: ‘remote’) and interrogative (*ya:va* ‘which’) pronominal elements can be used, without being attached to any general term, as modifiers in noun phrases.

- (36) *i: pustaka* ‘this book’
 a: pustaka ‘that book’
 ya:va pustaka ‘which book’

On the other hand, there are also several languages in which determiner proforms contain a general term in addition to a pronominal element. In fact, as pointed out by Diessel (1999: 59), the majority of languages use the same form in the nominal as well as adjectival functions. In many of these languages, such a form would also contain a pronominal element as well as a general term. For example, demonstrative pronouns of Cubeo (Tucanoan: Morse and Maxwell 1999: 83) are marked for animate–inanimate distinction, with the former showing masculine–feminine distinction and the latter showing noun class distinctions. All these can be used as modifiers of nouns or as independent pronouns.

In languages of the former type, in which the pronominal element is directly associated with a noun or verb, we can regard the noun or verb as representing the general term. The occurrence of such determiners appears to depend upon other grammatical characteristics of the languages concerned rather than upon the characteristics of proforms as such (see §8.2).

7.3.4 *Use of peripheral concepts as basic constituents*

I suggested earlier (§7.2.2) the necessity to differentiate between the two basic constituents of proforms on the one hand, and peripheral constituents like inflectional markers on the other. This differentiation is rather difficult to establish in some cases because languages may use some of the affixes that generally occur as peripheral constituents for expressing basic concepts. This is, in fact, the case with some of the examples that were given in the two previous sections. Gender and number markers of Alamlak, for example (see (25)), appear to function as general terms when they are attached to the demonstrative stem *ind-*, whereas in other uses they appear to function as peripheral

constituents. Similarly in Kannada (see (22)), gender number markers occurring in proforms function as general terms.

These extended uses of peripheral markers for denoting some of the meanings that are otherwise expressed by the basic constituents of proforms need to be differentiated from their peripheral uses because there would be a stronger tendency in the former case, as compared to the latter case, to gradually lose the individuality of those markers as their function is derivational rather than inflectional. This tendency can, later on, lead to the possibility of attaching additional peripheral markers to proforms that already contain one of them. For example, case markers are frequently used as general terms in the formation of adverbial proforms. As proforms formed in this fashion get grammaticalized, it would become possible for them to take other case markers.

This tendency can be exemplified with the help of the locative proforms of Kannada (Dravidian) that were originally formed by attaching the locative marker *-lli* to the three pronominal elements, namely *i* 'proximate', *a* 'remote', and *e* 'interrogative'. This locative suffix, apparently through grammaticalization, has become part of the proforms (as a general term), and as a result, the proforms are allowed to take additional case markers like the dative *-ige* for denoting the goal and the ablative *-inda* for denoting the source.

- | | | | | | | |
|------|--------------|---------|----------------|------------------|-----------------|--------------------|
| (37) | <i>i-lli</i> | 'here' | <i>ill-ige</i> | 'to this place' | <i>ill-inda</i> | 'from this place' |
| | <i>a-lli</i> | 'there' | <i>all-ige</i> | 'to that place' | <i>all-inda</i> | 'from that place' |
| | <i>e-lli</i> | 'where' | <i>ell-ige</i> | 'to which place' | <i>ell-inda</i> | 'from which place' |

7.4 Summary

The foregoing examination of the internal structure of proforms indicates clearly, I believe, that there is a need to regard them as having a basic dual structure, made up of a pronominal element and a general term. The former indicates their function in sentences whereas the latter indicates their scope. Languages do manifest proforms with more complex internal structures and also proforms that appear to have, at least synchronically, no perceivable internal structure. We need to regard them as non-prototypical cases and relate them with prototypical cases.

In the case of proforms with more complex structures, it is generally possible to perceive an underlying dual structure, as the complexities derive from complexities occurring either in the representation of the pronominal element or the general term. On the other hand, proforms that do not show any easily recognizable internal structure do represent semantic contrasts regarding functional distinctions and distinctions in the general concepts used. It is therefore more advantageous to regard them as having a 'fused' internal structure rather than

as having no internal structure as such. I have also pointed out the need to differentiate between this dichotomous internal structure of proforms on the one hand, and peripheral elements like syntactic (inflectional) and pragmatic markers that get attached to them on the other, even though there do occur instances in which it is rather difficult to differentiate between the two.

8 Constituent Elements of Proforms

8.1 Introduction

I suggested in the previous chapter that proforms can generally be regarded as involving a dual structure, made up of a pronominal element that indicates their function (or purpose) in language and a general term that indicates their scope. We can divide the various functions in which proforms are used into three main groups, namely (i) demonstratives, (ii) interrogative-indefinites, and (iii) relative-anaphors (§8.2). The type of distinctions that are made under each of these categories differs from one language to another. There are also cross-linguistic differences in the representation of the three main functions, as any of them may be grouped with any other function. There are even languages that group all three of them into a single category and represent them with the help of a single pronominal element (§8.3).

Languages also differ in their representation of the scope of proforms, with some languages using a single general term for all purposes. We may divide the concepts that get included under these general terms into four main categories, namely nominal, adjectival, verbal, and adverbial, corresponding to the four main word classes that occur among languages. The differences that occur among languages in the representation of these four main types of general terms appear to correlate with the differences occurring in the representation of the corresponding word classes, even though on some points there do occur differences between them (§8.4). Languages may also show an overlap in their representation of functional distinctions and general concepts (§8.5).

8.2 Functional distinctions

There is a large amount of cross-linguistic variation concerning the set of functional distinctions that are shown in proforms with the help of distinct pronominal elements. On the one hand, there are languages like Lyele (Niger-Congo: Showalter 1986) in which no pronominal distinction is used for representing functional distinctions. On the other hand, there are languages like Khezha (Tibeto-Burman) in which several different pronominal elements are used in order to differentiate, clearly, between the various important functions of proforms. Other languages fall in-between these two extremes, as they differentiate between some of the functions through the use of distinct pronominal elements, but use alternative devices for differentiating between other functions.

In order to obtain a better understanding of this cross-linguistic variation, it would be useful to group the functional distinctions that are generally represented through the use of pronominal elements in proforms under three different categories, namely (i) demonstratives, (ii) interrogative-indefinites, and (iii) relative-anaphors. The primary distinction that occurs among demonstratives is a spatial one, such as for example, between proximate and remote entities. Languages may, however, make additional distinctions like proximity to the speaker vs. proximity to the addressee, visible vs. invisible, up vs. down, inside vs. outside, etc. The purpose of making these different types of distinctions among demonstrative proforms is to assist the addressee in recognizing the referent that is being identified, or the location, time, manner, cause, reason, etc. of the event that is being predicated (§8.2.1).

The primary distinction that occurs in the case of the interrogative-indefinite category is between interrogation and indefiniteness. Both involve the indication of 'lack of knowledge' on the part of the speaker, but the two differ from one another regarding the purpose for which this lack of knowledge is specified. In the case of the interrogative, the purpose is to obtain the relevant information from the addressee, whereas in the case of the indefinite, it may be any of several different purposes such as not revealing the information or being indifferent to it. The majority of world's languages use a single pronominal element for representing this unified interrogative-indefinite category. However, such languages use additional markers either in the proform or elsewhere in the sentence, like for example, with the predicate, in order to specify the purpose for which the speaker's lack of knowledge is expressed. There are also several languages that use distinct pronominal elements for representing interrogation and indefiniteness (§8.2.2).

The relative-anaphoric category has the primary function of relating a referent with another referent that is mentioned elsewhere in the discourse, or the location, time, manner, etc. of an event with that of an event mentioned elsewhere. There are several sub-functions that can be brought under this category, and languages differ in representing them either with the help of different pronominal elements like relative, correlative, anaphoric, etc. or with the help of a single pronominal element (§8.2.3).

There is apparently a characteristic that is common to all these various pronominal functions as shown by languages like Lyle (Niger-Congo) that represent all of them with the help of a single pronominal element. It is also possible to combine any two of the three different types, that is, demonstratives with interrogative-indefinites, interrogative-indefinites with relative-anaphors, and demonstratives with relative-anaphors. The characteristic that unites all these functions, I think, is simply the need to have a general term (§8.3.1).

As I suggested in the first chapter (§1.6), there is a need to differentiate between primary and extended or secondary functions of these proforms because the characteristics that they manifest in their various uses cannot otherwise be

satisfactorily accounted for. Most of these proforms have extended uses that may or may not be represented by distinctions occurring in their pronominal elements. For example, demonstrative proforms may have an extended function in which they denote temporal or anaphoric meanings. Interrogative proforms may have extended uses in which the speaker does not actually lack the relevant information. Indefinites may also have similar extended uses. However, these are clearly *extended* uses, as indicated by the fact that most of the characteristics that are shown by these proforms reflect their primary use rather than these secondary uses.

8.2.1 *Distinctions among demonstratives*

As pointed out by Diessel (1999: 36, forthcoming), the primary deictic distinction that demonstratives make in most languages is a spatial one, that is, between proximate and remote or proximate, medial, and remote. The latter distinction involves both the speaker and the addressee in some languages (with the medial being regarded as proximate to the addressee), but in some it is based only (or primarily) upon the location of the speaker. There are also languages in which other deictic distinctions like visible–invisible, above–below, in front–behind, inside–outside, etc. are represented through distinct pronominal elements, but the number of such languages is not very large.

In my sample of 225 languages, 120 show only a two-fold spatial distinction, whereas the remaining languages show spatial distinctions between three or more points, with 45 of them differentiating between speaker's and addressee's locations. For example, English makes only a two-fold distinction between proximate (*this*) and remote (*that*), whereas Basque (Isolate: Saltarelli 1988: 213) makes a three-fold distinction between proximate (*hou*), medial (*hori*), and distant (*hura*). However, the medial is perceived as merely a point midway between the proximal and distal. Tagalog (Schachter and Otnes 1972: 88), on the other hand, differentiates between *ire* 'near the speaker', *ian* 'near the addressee', and *iyon* 'distant from both', with the medial point being regarded as the location near the addressee. Languages of this latter type are described as having a 'person-oriented' system of demonstratives, in contrast to the former, which have a 'distance-oriented' system (Anderson and Keenan 1985: 282).

As I have suggested in the sixth chapter (§6.2), this perception of the medial as the location of the addressee in contrast to its perception as a point that is midway between the proximal and distal appears to form part of a typological distinction between 'three-person' and 'two-person' languages. This distinction is primarily based upon the existence of third person pronouns that form part of the system of personal pronouns on the one hand, and the ones that do not fit properly into that system, but belong to the system of demonstratives, on the other. In the latter case, the pronouns would show derivational affinity with demonstrative pronouns or they would be identical with one or more of them. That is, third person pronouns would only represent an extended anaphoric use of demonstratives.

Among the languages that have a person-oriented deictic system, some appear to use the same stem for deriving the two 'near' forms, namely 'near 1' and 'near 2', and a different stem for deriving the 'distant from both' form. The following languages appear to support such a correlation.

(1)	Near 1	Near 2	Distant
F-Aniwa (Oceanic: Dougherty 1983)	<i>nei</i>	<i>na</i>	<i>ra</i>
B. Fijian (Pacific: Dixon 1988)	<i>qoo</i>	<i>qori</i>	<i>yaa</i>
Hunzib (Tsezik: Van den Berg 1995)	<i>bəd</i>	<i>bəl</i>	<i>əg</i>
Maori (E. Polynesian: Bauer 1999)	<i>nei</i>	<i>naa</i>	<i>raa</i>
Georgian (Kartvelian: Hewitt 1995)	<i>es</i>	<i>eg</i>	<i>is</i>
Lango (Nilo-Saharan: Noonan 1992)	<i>ni</i>	<i>nò</i>	<i>cà</i>
Mokilese (Micronesian: Harrison 1976)	<i>e</i>	<i>en</i>	<i>o</i>
A. Yupik (Eskimo-Aleut: Miyaoka 1996)	<i>mat</i>	<i>tamat</i>	<i>aw</i>
Tukang Besi (Austronesian: Donohue 1999)	<i>ana</i>	<i>atu</i>	<i>iso</i>
Hausa (W. Chadic: Newman 2000)	<i>wannàn</i>	<i>wànnan</i>	<i>wancàn</i>

In some of these languages like West Futuna-Aniwa and Maori, this similarity between the first two forms apparently derives from their having the same common ancestral language. Other languages like Samoan (Mosel and Hovdhaugen 1992: 129), To'abaita (Simons 1986: 30), and Tuvaluan (Besnier 2000: 405) also show comparable sets for the three demonstrative proforms. On the other hand, there do occur some languages in which the 'near 2' form uses the same stem as the distant one. The Tagalog forms given earlier (*ire* 'near 1', *iyán* 'near 2', and *iyón* 'distant from 1 and 2') exemplify this possibility. However, my sample of languages has a higher number of languages that support the former possibility than the ones that support this latter possibility.

There are also a few languages that appear to make a primary two-way distinction between the 'near 1' and 'near 2' meanings and associate the notion of being 'distant' individually with both of them. For example, Godoberi (Daghestanian: Kibrik 1996: 42) has the proforms *ha* 'near 1' and *hu* 'near 2'. Both these can take the marker *da/do* to indicate the notion of remoteness and form two additional proforms, namely *hada* 'remote from 1' and *hudo* 'remote from 2'. That is, even in the distal form, the distinction between speaker's location

and addressee's location appears to be maintained. Koasati (Muskogean: Kimball 1991: 486) appears to be similar in that it has two distinct stems for denoting the first two meanings, namely *yaʔa* 'near 1' and *maʔa* 'near 2'. It can derive several demonstrative pronouns like *yín* 'this here', *yá* 'this', *yá:fa* 'away from the speaker', and *yólí* 'this very' from the 'near 1' stem, and several others like *má* 'that one', *má:fa* 'that over there', and *mǎ:fa* 'that way over there' from the 'near 2' stem. However, it is not clear whether these additional derived forms maintain the same meaning distinction. The language is also reported to have a third distinct stem occurring in the proform *akkō* 'very far or out of sight'.

There are only eight languages in my sample that show four or more spatial distinctions. For example, Iraqw (Cushitic: Mous 1993: 114) shows a four-way distinction between *-i'/-ká* 'near the speaker', *-sín* 'near the addressee', *-qá* 'distal', and *-dá* 'far'. Sawu (Sumba-Bima group, Austronesian: D. Walker 1982: 11) makes a five-way distinction between *oni* 'touching the speaker', *(na)(pu)ne* 'near the speaker', *(na)d'e* 'in the speaker's vicinity', *(na)(pu)nəne* 'near the addressee', and *(na)ni/nad'o* 'distant from both'.

Languages that do not show any spatial distinctions among their demonstrative determiners or demonstrative pronouns are equally rare. Further, such languages appear to make at least a two-fold proximate-remote distinction among the adverbial (locational) demonstratives, as pointed out by Diessel (1999: 38). For example, Godié (Niger-Congo, Kru: Marchese 1986) has only a single set of demonstratives that show a two-fold distinction between humans and non-humans, and a four-fold noun-class distinction among non-humans. There is a singular-plural distinction in both these cases, with the noun-class distinction being restricted to singular forms. No spatial distinctions are indicated by these demonstrative determiners and pronouns. Marchese (1986: 233) points out, for example, that even in contrastive situations like (2), the same form would be used.

- (2) a. *nàna ɲwọ kọ a zì nàna kọ*
 this heavy up it pass this up
 'This (word) is stronger than that one'
- b. *nọnọ ọ nálá nọnọ ọ wọ nálá*
 this it good this it Neg good
 'This one is good; that one is not good'

The demonstrative proforms *nàna* 'this, that' (of 2a) differs from *nọno* 'this, that' (of 2b) regarding the noun class to which its antecedent noun belongs and not regarding any deictic meaning distinction. On the other hand, Godié does make a distinction between *lɔ* 'here', *mɔ* 'there, over there', and *mɔyí* 'over there' in the case of its locative demonstratives.

Mbay (Nilo-Saharan, Central Sudanic: Keegan 1992: 113) is reported to be similar in not making any spatial distinctions in the case of its demonstrative

determiners and pronouns. It has three different demonstratives that differentiate between positions of the object: *ńtèn* 'this/that, referring to things put in a place, like books, fruits, etc.', *ńdàn* 'this/that, things left standing, like poles, trees, wall, etc.', and *ńdìn* 'this/that, living things seated, and also chairs, shoes, baskets, etc.' None of these demonstratives shows any spatial distinctions. However, in the case of locative proforms, the language does make spatial distinctions like *ngɔ̀r* 'near', *ńóó-ń* 'over there', and *mù-ũ* 'far away'.

As I have mentioned earlier, demonstratives may show other types of distinctions (i.e. in addition to spatial distinctions), like visible and invisible, above and below the speaker, in front and behind the speaker, inside and outside, etc., but such languages are not many. For example, there are only 12 languages in my sample, such as Khasi (Mon-Khmer: Nagaraja 1985: 11), Lahu (Tibeto-Burman: Matisoff 1973: 51), and Tukang Besi (Austronesian: Donohue 1999: 137) that show the up-down distinction. Visible-invisible distinction is shown by 7 languages, which include Babungo (Grassfields Bantu: Schaub 1985: 203), Khasi, and Kharia (Munda: Biligiri 1965: 97). The demonstrative pronouns of Kwakiutl (Wakashan: Boas 1911: 127) are rather interesting in this context; they show a three-fold distinction between 'near 1', 'near 2', and 'near 3' and in the case of all these three pronouns, they make a two-fold distinction between visible and invisible.

(3)		Visible	Invisible
	Near 1	- <i>k</i>	- <i>g.a</i>
	Near 2	- <i>ōx</i>	<i>ō^ε</i>
	Near 3	- <i>ēq</i>	<i>ē^ε</i>

In some languages, the forms used for representing the various spatial and non-spatial distinctions among demonstratives appear to group into fewer categories. In Khasi (Mon-Khmer: Nagaraja 1985: 11), for example, there is a dichotomous distinction between the proximate forms and the rest. The former has an initial *n*- whereas the latter have an initial *t*- as seen in (4).

(4)	- <i>ne</i>	'proximate'	- <i>to</i>	'medial'
			- <i>tay</i>	'remote'
			- <i>ta</i>	'invisible'
			- <i>tey</i>	'up'
			- <i>thie</i>	'down'

The grouping of the two proximate forms, namely 'near 1' and 'near 2', into a set in some languages, mentioned earlier (see 1), and the grouping of distant forms into speaker-related and addressee-related sets in others like Godoberi and Koasati can also be regarded as representing this tendency of grouping deictic distinctions into smaller sets. Another interesting tendency shown by several languages is a diachronic one, namely to reduce a complex system of deictic distinctions into simpler ones that consist of only two elements, such as

proximate and remote. For example, Sanskrit is originally considered to have had a four-fold distinction in the case of its demonstratives, involving (i) a distinction between proximate (*eṣaḥ* ‘this’) and remote (*asau* ‘that’) and (ii) a distinction between presence (*ayam* ‘this’) and absence (*saḥ* ‘that’). The ‘proximate’ form *eṣa* ‘this’ is considered to be more emphatic than the ‘presence’ form *ayam* ‘this’ (Speijer 1884: 204). This system has undergone several changes, and as a result Sanskrit now possesses only a two-fold proximate–remote distinction. Vestiges of the earlier system can be seen in the allomorphy of its proximate and remote demonstratives (5).

(5)	Gloss	Proximate	Remote	Relative	Interrogative
	Person (M)	<i>ayam</i>	<i>saḥ</i>	<i>yaḥ</i>	<i>kaḥ</i>
	Person (F)	<i>iyam</i>	<i>sa:</i>	<i>ya:</i>	<i>ka:</i>
	Thing	<i>idam</i>	<i>tad</i>	<i>yad</i>	<i>kad/kim</i>
	Place/Goal	<i>atra</i>	<i>tatra</i>	<i>yatra</i>	<i>kutra/kva</i>
	Source	<i>itaḥ</i>	<i>tataḥ</i>	<i>yataḥ</i>	<i>kutaḥ</i>
	Time	<i>ida:ni:m</i>	<i>tada:ni:m</i>	<i>yada:/yarhi</i>	<i>kada:/karhi</i>
	Quality	<i>i:dṛśaḥ</i>	<i>ta:dṛśaḥ</i>	<i>ya:dṛśaḥ</i>	<i>ki:dṛśaḥ</i>
	Amount	<i>iyat</i>	<i>ta:vat</i>	<i>ya:vat</i>	<i>kiyat/kiyant</i>
	Number	—	<i>tati</i>	<i>yati</i>	<i>kati</i>
	Manner	<i>ittham</i>	<i>tatha:</i>	<i>yatha:</i>	<i>katham</i>

Most of the modern Indo-Aryan languages also show only a two-fold deictic distinction between proximate and remote. In the case of Dravidian languages also, there has been a similar reduction of deictic distinctions. Proto-Dravidian is generally reconstructed as possessing a three-fold proximal–medial–distal distinction, but several modern languages like Kannada, Tamil, Tulu, etc. have lost the medial form. Similar reduction has apparently taken place among some of the Uralic languages. Abondolo (1998: 24) reconstructs a three-fold deictic distinction for Proto-Uralic, but points out that some of the daughter languages have reduced this demonstrative system into a simple two-term proximal–distal dichotomy.

C. Lyons (1999: 110) mentions the occurrence of similar reductions of three-term demonstrative systems into two-term systems or even single-term systems in some of the European languages. For example, English has reduced an earlier three-term system into a two-term system, and the relic of the earlier system can be seen among the archaic distal forms, *yon* and *yonder*. Catalan and Old French have also reduced an earlier three-term demonstrative system in a similar fashion. German and French have reduced the system further, as their demonstrative pronouns do not generally show even the proximate–remote distinction. According to Diessel (1999: 38), German has three adverbial demonstratives, *hier* ‘here’, *da* ‘there’, and *dort* ‘there’, but it employs only a single demonstrative pronoun, *dies* ‘this/that’. Colloquial German has two demonstrative pronouns, *dies* and *das*, but they do not contrast deictically.

On the other hand, there do occur some languages that have increased the number of deictic distinctions. For example, Proto-Indo-European is reconstructed as showing only a two-fold proximal–distal deictic distinction, but Armenian and Latin have developed an additional deictic distinction involving the notion of ‘proximate to the addressee’ (Ajello 1998: 216, Vineis 1998: 293). Similarly among the Indo-Aryan languages, Sinhalese has developed an additional ‘near 2’ form and also a distinct anaphoric form (Gair 1998: 113). Among the Dravidian languages also, Kuvi has developed a five-fold spatial distinction from an earlier three-fold system (Israel 1979).

8.2.2 *Distinctions among interrogative-indefinites*

As the term ‘interrogative-indefinite’ indicates, the main distinction that occurs in this functional category is between interrogation and indefiniteness. I have put them together into a single primary function because most languages use a single pronominal element for representing them among their proforms. According to Haspelmath (1997: 241), languages that differentiate between the two notions occur mainly among the languages of Africa and Oceania. Even among these languages, many use their interrogative pronouns for denoting some of the indefinite meanings (especially the non-specific ones) as well (see §10.1). For example, Kilivila (Austronesian: Senft 1986: 67) generally uses distinct stems for denoting interrogative and indefinite meanings. It has a set of interrogative pronouns that includes forms like *avela* ‘who’, *avaka* ‘what, which’, *avetula* ‘when’, and *ambeya* ‘where’; these are used in constituent questions. On the other hand, it derives indefinite pronouns by attaching one of the classificatory particles to the numeral *tala* ‘one’.

- (6) a. *avela e-ma*
 who he-come
 ‘Who is coming?’
 b. *te-tala bi-ma bi-livala avaka ku-vagi*
 man-one will-come will-tell what you-do
 ‘Someone will come and tell you what to do’

Interrogative pronouns in Kilivila can also be formed by attaching one of the classificatory particles to the stem *-vila* ‘how much’. However, this pronoun can also be used as an indefinite pronoun.

- (7) a. *na-vila vivila e-keosi-se*
 woman-how.many girl they-dance-PI
 ‘How many girls are dancing?’
 b. *te-vila e-sisu-si va simla te-vila e-sila-si*
 male-some they-stay-PI at island male-some they-sail-PI
 ‘Some of them stay at the island, some of them keep on sailing’

Languages that use the same pronominal element in the formation of interrogative and indefinite pronouns, on the other hand, generally use additional devices for differentiating between the two. The devices include (i) attaching additional affixes to these proforms, (ii) using interrogative particles elsewhere in the sentence, (iii) using interrogative verbs, and (iv) using different intonation patterns (see §10.2 for details). Jaqaru (Jaqi: Hardman 2000: 33) sentences given in (8a–b) exemplify the use of an interrogative particle as a differentiating device.

- (8) a. *qachi-s jal-sh munki*
 who-Int fall-out want
 ‘Who wants to go out?’
 b. *qusi-nh-ash uk wallmich irawi*
 what-1-maybe that girl carry
 ‘Maybe, that girls took something of mine’

As I will be pointing out later (§8.5), some languages have developed distinct stems (apparently involving distinct pronominal elements) for denoting the interrogative-indefinite meaning about human and non-human referents (or animate and inanimate referents), that is, distinct pronominal stems for ‘who’ and ‘what’. This distinction has clearly developed from the fusion of general terms that denote the human–non-human or animate–inanimate distinction with the interrogative-indefinite pronominal element. Some of these languages do not differentiate between these two stems when they are used in the formation of adverbial interrogatives, whereas some languages use only the inanimate stem for deriving such forms.

8.2.3 Distinctions among relative-anaphors

As I have pointed out in the third chapter (§3.2), there are several languages that possess anaphoric proforms whose main function is to remove the ambiguity that arises from the double use of demonstratives both as deictic words (with exophoric reference) and also as anaphoric words (with endophoric reference). These are generally included among demonstratives apparently because their main function is connected with the use of demonstratives. However, they are also related to relative pronouns in the sense that both have the same purpose of relating an entity with another entity that has been introduced elsewhere in the discourse. There are also some languages like Kusaiean (Micronesian: Lee 1975: 107) in which the same pronoun (*ma*) is used both as an anaphoric pronoun (for referring to a person or object mentioned earlier) and also as a relative pronoun.

I suggested further that there is a need to differentiate between these anaphoric pronouns, which form part of the system of proforms (demonstratives), on the one hand, and the ones that form part of the system of personal pronouns on the other. The latter are generally regarded as ‘long-distance’ reflexives. There

are several important differences between these two types of ‘anaphoric pronouns’, which derive from the fact that they belong to two entirely different systems of pronouns (personal pronouns and proforms). For example, long-distance reflexives may show personal distinctions, may be restricted to humans, and may involve the protagonist’s perspective or subjective viewpoint, whereas anaphoric proforms (demonstratives) do not show any of those characteristics. The former can only be nominal whereas the latter can also be adjectival, adverbial, or verbal (see e.g. (33) in §3.2, in which Sinhala paradigms of demonstratives, including anaphoric proforms are given). The following discussion of ‘anaphoric pronouns’ is restricted to the use of that term for denoting anaphoric proforms.

Most languages allow their demonstrative pronouns to be used as anaphoric pronouns. For example, Tinrin (Melanesian: Osumi 1995: 90) has the demonstratives *ha* ‘proximate’, *mwâ* ‘mid-distance’, and *rra* ‘distant’; all these can also be used as anaphors. On the other hand, some restrict this anaphoric use to only one (or some) of their demonstratives. Imonda (Papuan: Seiler 1985: 45), for example, has the demonstrative pronouns *ôh* ‘here’, *ed* ‘there’, of which *ed* may be used as an anaphoric pronoun. Languages may also derive their anaphoric pronouns from demonstratives. For example, Swahili (Madan 1918) has three sets of demonstrative pronouns showing noun-class distinctions. The proximate is formed by prefixing *h-* to noun class markers and the remote one by adding *-le* to the second syllable of the proximate. The third set of demonstrative pronouns denotes something mentioned before, and is derived from the proximate by changing the final vowel to *o*. Lampung (Malayan: D. Walker 1976: 16) has the demonstratives *hinji/ji* ‘this, here’, *hina/na* ‘that (near)’, *hudi/di* ‘that (far)’, and *jino* ‘previously mentioned’, with the last one apparently derived from the proximate demonstrative. Hunzib (Tsezik, Daghestanian: Van den Berg 1995: 60) has the demonstratives *bəd* ‘near 1’, *bəl* ‘near 2’ and *əg* ‘far from both’; these can be preceded by the anaphoric prefix *yo-/yə-*.

In the case of several other languages, on the other hand, anaphoric pronouns are quite different from demonstratives. Lezgian (Nakh-Daghestanian: Haspelmath 1993: 187), for example, has the demonstratives *im* ‘proximate’, *am* ‘remote’ contrasting with *ham* ‘anaphoric’. Basque (Saltarelli 1988: 207, 211) has the demonstrative pronouns *hau* ‘proximal’, *hori* ‘medial’, and *hura* ‘distal’, contrasting with the anaphoric pronouns *berau* ‘near 1’, *berori* ‘near 2’, and *bera* ‘over there’. There are also languages in which anaphoric markers can be attached to nouns like deictic determiners. For example, Guaymí (Chibchan: Quesada 1999: 220) has the anaphoric suffix *e* that can be attached to nouns when their referent is reactivated in discourse, having gone unmentioned for a while.

- (9) *mütu kāmiga näen... ngri-e...*
 pig kill go meat-Anaphoric
 ‘One kills the pig ... the meat (of the said pig) ...’

Distinctions that occur among these anaphoric pronouns that correspond to the spatial distinctions of demonstratives are generally used for denoting temporal distinctions among the antecedents. In some languages, however, anaphoric pronouns show other types of distinction. For example, Ainu (Tamura 2000: 93) makes a distinction between *ne* 'this (the one being talked about)' and *néa* 'that (the one mentioned previously)', whereas Asheninca (Campa) (Arawakan: Reed and Payne 1986: 330) makes a distinction between anaphoric and cataphoric references. The latter are derived from the former by attaching the prefix *h-* (*irika* 'M, nearby, anaphoric', *hirika* 'cataphoric').

Languages that possess distinct anaphoric pronouns may also allow their demonstrative pronouns to occur as anaphoric pronouns. Generally, grammarians do not say whether there is any difference between these two anaphoric usages. For example, according to Gair (1998: 113), Sinhala (Indo-Aryan) allows its demonstrative pronouns (*mee* 'this, near 1', *oyə* 'that, near 2', and *arə* 'over there, distal from 1, 2') as well as its anaphoric pronoun, *ee* 'that', to be used as anaphors. The only difference between the two is that the latter (anaphoric pronouns) can only have an anaphoric use. This is also true of third person pronouns in languages in which they are distinct from demonstrative pronouns. Most languages are reported to allow their third person pronouns to function as anaphoric pronouns, but the grammarians do not say whether there is any difference between the anaphoric use of such pronouns and other types of expression like demonstratives or anaphoric pronouns.

On the other hand, there are some languages in which the two usages are reported to be pragmatically different. For example, according to Terrill (2001: 67), Lavukaleve (Papuan) has a set of demonstrative (deictic) pronouns contrasting with a set of anaphoric pronouns. The former make a distinction between proximal, medial, distal, and hypothetical, with each showing a three-fold number distinction and a three-fold gender distinction (except in plural). Anaphoric pronouns also show the same type of number and gender distinctions. Both can be used anaphorically but the latter do not have any non-anaphoric (exophoric) use.

Terrill reports that in the anaphoric use, Lavukaleve makes a pragmatic distinction between activated and semi-activated referents. Deictic demonstratives (in their anaphoric use) denote activated referents, whereas anaphoric pronouns denote semi-activated referents. This is comparable to the distinction, pointed out by Comrie (1997) as occurring in Dutch between third person pronouns and demonstratives in their anaphoric use. While the former pick out a topical (expected) antecedent, the latter pick out a non-topical (non-expected) antecedent (see §9.2.3). The demonstrative pronouns of Lavukaleve resemble the third person pronouns rather than demonstratives of Dutch on this point apparently because they have the additional function of being third person pronouns. Lavukaleve does not have any third person pronouns as such that are distinct from demonstrative pronouns (Terrill 2001: 70–1).

Several Indo-Aryan languages make a distinction between relative and correlative pronouns. They use a relative pronoun for introducing a referent and a correlative pronoun for relating it with one of the participants of the main event that is being predicated. The proforms of Bengali given in (10a) and the sentence given in (10b), showing the use of these relative and correlative pronouns, exemplify this contrast (Dasgupta 1979, K. K. Rarhi, personal communication).

- (10) a.
- | | Proximate | Remote | Interrogative | Relative | Correlative |
|--------|---------------|---------------|-----------------|----------------|----------------|
| Person | <i>ey</i> | <i>oy</i> | <i>ke</i> | <i>je</i> | <i>se</i> |
| Place | <i>ekhane</i> | <i>okhane</i> | <i>konkhane</i> | <i>jekhane</i> | <i>sekhane</i> |
| Time | <i>ekhon</i> | <i>tɔkhon</i> | <i>kɔkhon</i> | <i>jɔkhon</i> | <i>tɔkhon</i> |
| Amount | <i>eto</i> | <i>oto</i> | <i>koto</i> | <i>joto</i> | <i>toto</i> |
- b. *je bajare gie chilo se ekhane ache*
 who market gone had he here is
 ‘The man who had gone to the market is here’

Notice that in (10b) the relative clause, *je bajare gie chilo* introduces a person with the help of the relative pronoun *je* and the main clause that follows it has a correlative pronoun *se* that provides the connection between the relative pronoun and the main clause.

8.3 Neutralization of functional distinctions

I suggested earlier that we might regard demonstratives, interrogative-indefinites, and relative-anaphors as forming the three main pronominal functions among proforms. There are, however, several languages that are able to combine even these main functions in different ways, apparently on the basis of the fact that they represent variations of a single function. There is at least one language, namely Lyle (see §8.3.1) that appears to combine together all the three main functions and use a single pronominal element for representing them. On the other hand, there are several other languages that combine together only two of them, such as for example demonstratives with interrogative-indefinites (§8.3.2), demonstratives with relative-anaphors (§8.3.3) or interrogative-indefinites with relative-anaphors (§8.3.4).

8.3.1 Combining all the three main functions

According to Showalter (1986: 202), Lyle (Gur, Niger-Congo) has only a single set of forms in its system of proforms. These can function as demonstratives, interrogatives, and also relatives. The paradigm involves a nine-fold noun-class distinction, with (i) human, (ii) diminutive, (iii) augmentative, and (iv) the rest functioning as four distinct classes with their plural forms forming four additional classes. The ninth class, which has no plural counterpart, denotes mass nouns and abstractions.

The proforms, as shown in (11), have three forms each, which, according to Showalter (1986: 210), are interchangeable. The first is derived with the help of the suffix *bɛ* (showing alternations due to vowel harmony) and the third with the help of the determiner *wá* or *yá*. The second is apparently a weakened form of the first.

(11)	Singular	Plural
Human	<i>mɔ́bɔ́/mɔ́ɔ́/mɔ́wá</i>	<i>byèbɛ́/byèé́/byèyá</i>
Diminutive	<i>kèbɛ́/kèé́/kèyá</i>	<i>shèbɛ́/shèé́/shèyá</i>
Augmentative	<i>kàbɔ́/kàɔ́/kàwá</i>	<i>tèbɛ́/tèé́/tèyá</i>
Rest	<i>tèbɛ́/tèé́/tèyá</i>	<i>nyèbɛ́/nyèé́/nyèyá</i>
Mass/Abstract	<i>mɔ́bɔ́/mɔ́ɔ́/mɔ́wá</i>	

Notice that the mass/abstract class is identical with the human class except for the absence of a plural form. Further, the plural forms of the augmentative class are the same as the singular forms of the Rest class.

Showalter points out that the same pronoun can be used either as a demonstrative or as an interrogative, the only difference between the two being that the final high tone of an interrogative becomes a falling tone in the demonstrative. It is possible, according to Showalter, to regard this high tone as the interrogative intonation rather than as the tone of the pronoun itself.

- (12) *jì kèé́* 'Which house?'
jì kèè 'This house'
house which/this

When these pronouns are used as demonstratives, the noun occurring with them will have a determiner suffix. On the other hand, when they are used as relatives, they are followed by a correlative pronoun.

- (13) *lò mɔ́wá ñ ní ñ títì*
person that he Cont he stumbles
'The person that stumbles ...'

8.3.2 Combining demonstratives with interrogative-indefinites

Wari', belonging to the Chapakuran family of western Brazil, makes no distinction between demonstratives and interrogatives. Demonstratives of this language show a three-fold deictic distinction between *cwa* 'proximate to speaker', *ma'* 'proximate to hearer' and *cwain* (MF), *cain* (N) 'distal'. Two of these pronouns, namely *ma'* and *cain*, can be used in a special kind of construction called complementizer construction in which the pronoun is shifted to the sentence-initial position and the inflectional morpheme is placed after it. While occurring in the adnominal position, it takes the gender markers *-on* 'MSg', *-m* 'FSg' and *-in* 'N' to form the proforms *mon*, *mam*, and *main* respectively. The complementizer

construction occurs with the interrogative intonation and functions as a constituent question (Everett and Kern 1997: 12, 49).

- (14) a. *maqui' na co ma'*
 come Realis.3Sg M that
 'He came'
- b. *ma' co mao na*
 that Realis.MF go Realis.3Sg
 'Who went?'
- (15) *main ca mao ca*
 that (N) Realis (N) go 3NSg
 'Where did he go?'
- (16) a. *xirim cain*
 house that (N)
 'that house'
- b. *cain' ca tomi' cama*
 that (N) Realis.N speak 3FSg
 'What did she say?'

Tonkawa (Hokan: Hoijer 1933: 128) is similar in using the same set of proforms as demonstratives as well as interrogative/indefinites, but it differentiates between the two by attaching an interrogative marker in the latter case. Demonstrative pronouns are formed in this language by adding a set of general terms to deictic pronominal elements, namely *de*.- 'proximate', *he*.- 'a distance away', and *we*.- 'a greater distance away'.

- | | | | | |
|------|-----------|---------------|-----------------|------------------------------------|
| (17) | | Proximate | Medial | Distal |
| | Place | <i>de.dja</i> | <i>he'e-dja</i> | — |
| | Direction | <i>de-l</i> | <i>he'e-l</i> | <i>we</i> .- 'il |
| | Manner | <i>de.-dj</i> | <i>he</i> .-dj | <i>we</i> .-dj (<i>we</i> .-dedj) |

Interrogatives are formed from one of these, namely proximate demonstratives, by prefixing the interrogative *he*- to them.

- | | | | | |
|------|----------------|---------------------|------------------|------------------------|
| (18) | <i>del</i> | 'here' | <i>he-del</i> | 'where' |
| | <i>de.wa'n</i> | 'in this direction' | <i>he-dew'an</i> | 'in which direction' |
| | <i>de.dj</i> | 'in this manner' | <i>he-dedj</i> | 'in which manner' |
| | <i>de.dja</i> | 'this place' | <i>he-dedja</i> | 'where, in what place' |

There are three additional interrogative pronouns, namely *hedju* 'what', *hedju* 'ed' 'why', and *hedwan* 'how many' which, according to Hoijer, cannot be analysed in this way. However, they do appear to have derived in a similar fashion. Indefinite proforms of this language involve further complexity, as they are derived by adding the suffix *-ax* to interrogatives.

- (19) *hedju*.-‘ax ‘something, anything, anyone’
he-dedha.-‘ax ‘somewhere, wherever’
he-dew’an.-‘ax ‘in any direction’

8.3.3 Combining demonstratives with relative-anaphors

The relative pronouns of Tuvaluan (Polynesian: Besnier 2000: 433) are formally identical with demonstrative pronouns. They make a three-fold deictic distinction: *teenei* ‘in the vicinity of 1’, *teenaa* ‘in the vicinity of 2’, and *teelaa* ‘away from both’. Their use in a relative clause, however, is optional.

- (20) *koo puke katoa nee te tagata siaamai teenaa*
 Inceptive take all Erg the man Germany that
ne vau
 Past came
 ‘The German man that came over took [it] all’

Supyire (Gur, Niger-Congo: Carlson 1994: 161) also uses the same set of proforms as demonstratives as well as relatives. However, it adds the suffix *-mù* to demonstratives in order to derive relative pronouns from them. This suffix is cognate with the suffix *-mù* ‘also’. Interrogative pronouns are differentiated from these relative pronouns through the use of a distinct set of pronominal elements for representing them. Kiribatese (Austronesian, North Hebridean: Groves, Groves, and Jacobs 1985: 61) also shows derivational relationship between demonstrative and relative pronouns. However, it differs from Supyire in showing deictic distinctions in the case of demonstratives as well as relative pronouns.

- | | | | |
|------|-----------------|----------------|-------------------|
| (21) | | Demonstratives | Relative pronouns |
| | ‘near 1’ | <i>aei</i> | <i>ae</i> |
| | ‘near 2’ | <i>anne</i> | <i>ane</i> |
| | ‘far from both’ | <i>arei</i> | <i>are</i> |

8.3.4 Combining interrogative-indefinites with relative-anaphors

There are several languages like English that combine together the function of a relative pronoun and that of an interrogative-indefinite pronoun, and use a single pronominal element (as, for example, *wh*- in English) for representing them.

- (22) a. *Who invited me?*
 b. *The man who invited me is sick.*

It appears that languages that differentiate between interrogatives and indefinites by using distinct sets of pronominal elements for representing them do not combine together the interrogative and relative functions in this fashion. That is, it is the ‘indefiniteness’ of interrogative-indefinite proforms rather than their ‘interrogativity’ that facilitates their use as relative pronouns. The general tendency,

among languages in which interrogative pronouns are unrelated to indefinite pronouns, is to use one of the demonstrative paradigms as a relative pronoun paradigm. However, it is also possible for languages in which interrogatives are *related* to indefinites to use demonstrative proforms rather than interrogative-indefinite proforms as relative proforms. German, for example, uses its demonstrative pronouns as relative pronouns, even though it has a single set of stems for denoting interrogative and indefinite meanings (Haspelmath 1997: 243).

8.4 General concepts

The general concepts that get associated with proforms indicate their scope. In the case of interrogative-indefinite proforms, for example, general concepts like person, thing, place, time, manner, etc. provide a broad description of the type of information that the speaker lacks (indefinite pronouns) and desires to obtain from the addressee (interrogative pronouns). Similarly, in the case of demonstratives and relative-anaphoric proforms, they indicate the kind of entity that is being pointed out as occurring in the speech context (demonstrative pronouns) or somewhere else in the discourse (relative-anaphoric pronouns). The kind of distinctions that are shown by these general terms would therefore depend upon the way in which languages divide human experience into categories or what Jackendoff (1992: 34) calls 'conceptual parts-of-speech'. We can therefore expect these distinctions to correlate, to a certain extent at least, with the morphosyntactic parts-of-speech of the languages under consideration, because the latter appear to correlate with the conceptual parts-of-speech (§8.4.1).

On the other hand, there are also important differences between the functions for which word-class distinctions and distinctions of general concepts are used in language. Word-class distinctions are meant for organizing the structure of sentences, whereas distinctions occurring among the general terms of proforms are meant for specifying the scope of deictic marking, interrogation, non-specification of identity, etc. The requirements of these two functions would clearly be different and these would be reflected in the occurrence of different kinds of distinctions among word classes and among the general terms of proforms (§8.4.2).

8.4.1 *Word-class distinctions*

The set of general terms that occur among proforms can be divided into four main categories, namely nominals, adjectivals, adverbials, and verbals. However, the occurrence of any of these as a distinct category in a language would generally depend upon the occurrence of a corresponding word class as a distinct part-of-speech in that language. As I have suggested elsewhere (Bhat 1994), a typological distinction can be made between languages that possess adjectives as a distinct word class on the one hand and the ones that do not on the other. In the latter case, words that correspond semantically to the adjectives of the former

type of languages would be included either in the category of nouns or of verbs. There are also languages in which no categorial distinction is made between nouns, verbs, and adjectives. This typological distinction occurring among the parts-of-speech of languages can be correlated, to a certain extent, with the distinction occurring among proforms.

For example, languages in which adjectives are included under the category of nouns would not generally show any distinction between adnominal (adjectival) and pronominal proforms. This is true of most of the Indo-Aryan languages. They do not differentiate between adjectives and nouns and, correspondingly, they also generally do not show any distinction between adnominal and pronominal proforms. This is also true of most of the Australian languages. They do not show any distinction between nouns and adjectives and, correspondingly, they use the same set of proforms either as adnominals or as pronominals. For example, Wardman (Merlan 1994: 58) allows its adjectives to be used not only as modifiers of nouns, but also as independent arguments. Its demonstrative pronouns can also serve either as modifiers within a nominal phrase or as reference-making members of a nominal phrase.

- (23) a. *ngarrugu* *dana* *yirrbag wonggo yanja*
 1 (Pl.Incl.Dat) this (Abs) back Neg 3Sg.go.Pres
 'This one doesn't move back for us'
- b. *dang-nyi* *wunggun-bu-ndi* *yibiyan-yi*
 yonder-Erg 3Sg-hit-Past man-Erg
 'That yonder man hit them'

On the other hand, languages that include adjectival words under the category of verbs have adnominal proforms that show the property of verbs, such as occurring with a relative participle. For example, Manipuri includes adjectival words under the category of verbs (Bhat and Ningomba 1997: 73) and requires them to be relativized in order to use them as modifiers of nouns. Manipuri also has proforms like the interrogative *kəɾəmbə* ‘which’ that resemble verbs in occurring with the relative particle *bə*. However, its deictic elements *-si* ‘this’ and *-du* ‘that’ can be attached directly to nouns.

In the case of languages in which any content word can be used as a predicate, we find proforms also being freely used as predicates. For example, Mundari (Austroasiatic, Munda: Hoffmann 1903) does not make any noun–verb distinction. Any of its content words, whether they translate as nominals, verbals, or adjectivals, can be used freely as predicates. One can also attach any of the inflectional markers that occur with predicates, such as tense and aspect makers, transitivity markers, and personal markers for denoting subject and object, to any of its content words. Its demonstrative and interrogative proforms are similar to these content words in that they can also be used as predicates and can occur with any of the inflectional markers that can be attached to predicates (see Bhat 1994: 217).

8.4.2 *Distinctions among different word classes*

While there is a general kind of correlation between word-class distinctions occurring among languages on the one hand and distinctions occurring among the general concepts that are used in the formation of proforms on the other, there are also certain specific types of difference that occur between the two. These differences appear to derive from some of the specific requirements of the function for which the proforms are used.

For example, nominal proforms occurring in the interrogative-indefinite paradigm generally show a distinction between human and non-human (or animate and inanimate) referents (see (33) given in §8.5 for examples). There are a very few languages, such as Tzutujil (Mayan: Dayley 1985: 227) that use the same interrogative pronoun to denote 'who' and 'what'. On the other hand, there are languages that do not show a human–non-human or animate–inanimate distinction elsewhere in the grammar, but still show this distinction in their interrogative pronouns, as has been pointed out by Haspelmath (1997: 30).

A comparable human–non-human distinction is made by several languages among their demonstrative pronouns, but this is generally associated with the gender distinction (or animate–inanimate distinction) occurring elsewhere in the language. For example, Yidiny (Australian: Dixon 1977: 181) makes a distinction between human (and higher animate) and the rest in the case of demonstratives as well as interrogatives.

(24)	Human	Inanimate
Proximate	<i>yip̣du-</i>	<i>yĩngu-</i>
Medial	<i>ɲup̣du-</i>	<i>ɲungu-</i>
Distant/invisible	<i>yup̣du-</i>	<i>yungu-</i>
Inter/Indefinite	<i>wap̣du-</i>	<i>wapi-</i>

In the case of the adjectival category, languages generally make a distinction between simple adjectivals or determiners that indicate or question the identity of an entity on the one hand, and the ones that refer to the kind or type of entity involved on the other. For example, Kannada (Dravidian) has a simple adjectival form *ya:va* 'which' and an additional complex form *entaha* 'of what type' among its adjectival interrogative forms. There are comparable forms among proximate demonstratives (*i:* and *intaha*) and remote demonstratives (*a:* and *antaha*) as well. Languages may also have a distinct proform for indicating or questioning the quantity (number or amount) of a given entity. For example, Kannada has *eṣṭu* 'how much/many', *iṣṭu* 'this much/many', and *aṣṭu* 'that much/many' as distinct adjectival proforms. Most languages have a single form for representing number and amount, but some have distinct proforms. For example, Godoberi (Daghestanian: Kibrik 1996: 39) has the interrogatives *into* for denoting 'how much' and *čamu* for denoting 'how many'.

Several languages have a distinct ordinal proform that may be derived from the cardinal proform or formed directly through a general term. In Kannada, for example, the proforms that denote amount or number take the ordinal suffix *ne*: to indicate the corresponding cardinal: *aṣṭu* 'that much/many', *aṣṭane*: 'of that cardinal number'; *iṣṭu* 'this much/many', *iṣṭane*: 'of this cardinal number'; and *eṣṭu* 'how much/many', *eṣṭane*: 'of which cardinal number'.

The adverbial category appears to show the maximum amount of differentiation among proforms. There is generally a term each for location, time, manner, and reason. The various paradigms of proforms given earlier exemplify the occurrence of these distinctions. Languages may show additional distinctions under the category of location, such as, for example, between goal and source, or between location and direction of motion. For example, Yosondúa Mixtec (Mixtecan: Farris 1992: 35) has the interrogative forms *nàchĩ* 'in which direction', *nándĩ* 'from where', and *nánũ* 'at which place'.

There are several languages that possess a distinct verbal category among their proforms. For example, Dyirbal (Australian: Dixon 1972: 49) has two distinct interrogative verbs, namely intransitive (*wiyamay*) and transitive (*wiyamal*), that can be used either by themselves as verbs, in which case they mean 'to do what', or with some other verb to mean 'to do how'. The language also has a similar set of demonstrative verbs (intransitive *yalmay* and transitive *yalmal*) that provide the meaning 'it is done like that'. Mapuche (Southern Andean: Smeets 1989: 132, 424) is another language that has interrogative and demonstrative verbs. The interrogative *chum* 'to do what, how' is used as a verb, whereas the demonstrative roots *fa-* 'to become like this' and *fe-* 'to become like that' are attached to verbs. The latter may denote the relevant proximate–distant distinction, or indicate a distinction between situationally determined (exophoric) and contextually determined (anaphoric) events.

Verbal proforms may also be derived from interrogative pronouns through the addition of verbalizing affixes to them. For example, Kayaradild (Tangkik, Australian: Evans 1995: 371) has the interrogative verb *ngaakawatha* 'do what' derived by adding the suffix *-watha* to the interrogative pronoun *ngaaka* 'who, what'. On the other hand, Venda (Bantu: Poulos 1990: 451) has the interrogative marker *-ni* occurring in the object position of verbs in order to ask the question 'what'.

Diessel (1999: 78) considers it necessary to establish a distinct category of demonstratives called 'identifiers' for some languages. They are used in copular and non-verbal clauses in order to focus the hearer's attention to entities in the surrounding situation or in the universe of discourse. For example, Ponapean (Rehg 1981: 143, 152) has a set of demonstrative identifiers that differ from demonstrative pronouns in having an initial high vowel and in not having an initial nasal.

(25)	Demonstrative identifiers		Demonstrative pronouns	
	Singular	Plural	Singular	Plural
'near 1'	<i>me(t)</i>	<i>metakan</i>	<i>ie(t)</i>	<i>ietakan</i>
'near 2'	<i>men</i>	<i>menakan</i>	<i>ien</i>	<i>ienakan</i>
'distant'	<i>mwo</i>	<i>mwohkan</i>	<i>io</i>	<i>iohkan</i>

- (26) a. *ien noumw pinselen*
 there your pencil
 'There is your pencil!'
 b. *ien* 'There it is'

Tümpisa Shoshone (Uto-Aztecan: Dayley 1989) has such identifiers occurring both as demonstratives as well as interrogative-indefinites. They are formed by adding *sün* to pronominal elements.

- (27) *i* 'right here' *isün* 'this right one here'
 e 'here, nereby' *esün* 'this is the one that...'
 hii 'what' *hiyasü* 'what kind (is it)'

In the languages mentioned above, identifiers occur as parts of copulative or non-verbal sentences. It is also possible for such forms to function as interjections or as one-word utterances. In the Havyaka dialect of Kannada (Dravidian), for example, demonstrative identifiers showing proximate-remote distinction (*ida* 'here' and *ada* 'there' which are different from demonstrative pronouns *idu* 'this one' and *adu* 'that one' and adverbs like *illi* 'here' and *alli* 'there') can occur as interjections along with fully formed copulative sentences.

- (28) a. *ida, ninna pustaka illi-ddu*
 here, your book here-is
 'Here is your book!'
 b. *ada, ninna pustaka alli-ddu*
 there, your book there-is
 'There is your book (over there)!'

Standard Kannada also has a comparable set of demonstrative identifiers, formed by adding the suffix *-go* to proximate and remote demonstrative elements (*igo* 'here', *ago* 'there'). Tulu (Dravidian) has the forms *nda* 'here' and *ava* 'there', which are also different from demonstrative pronouns and adverbs. These also have plural forms, *nde* and *ave* which indicate the plurality of the addressee, and are generally used for denoting politeness.

According to Haiman (1980: 275) Hua (Papuan) has not only interrogative verbs like *zahu-* 'what is it?', *aigahu* 'do what' or 'do how', but also an interrogative interjection *ve* used in contexts like (29).

- (29) a. *bira-a-e* 'You (unseen) over there!'
 b. *ve* 'Yes, what is it?'

There are also languages in which very few general concepts are used in the formation of proforms. Asheninca, a Pre-Andean Arawakan language appears to be of this type. Reed and Payne (1986: 328) point out that a single interrogative pronoun *t^{sh}ika* can be used in this language in most of the constituent questions that require an interrogative pronoun. It can also be used in several types of clauses that require an indefinite pronoun, that is, while occurring as complements to verbs like 'know'.

- (30) a. *t^{sh}ika* *i-t^sim-i-ka*
 Int/Indefinite 3M-exist-Nonfuture-Int
 'Who is it?'
 b. *t^{sh}ika* *pi-kant^s-i-ka*
 Int/Indefinite 2-say-Nonfuture-Int
 'What did you say?'
 c. *t^{sh}ika* *p-iyaa^s-i-ka*
 Int/Indefinite 2-go-Nonfuture-Int
 'Where did you go?'

Notice, however, that Asheninca is able to differentiate between these uses of a single interrogative-indefinite pronoun with the help of distinctions occurring elsewhere in the sentence. In (30a–c), for example, the distinction is conveyed by the content of the verb. In other cases, it might be indicated by certain inflectional elements occurring in the verb. In the case of 'when', however, a temporal marker is attached to the question word.

- (30) d. *t^{sh}ika-paite-ka* *p-iyaa^s-i*
 Int/indefinite-Temporal-Int 2-go-Nonfuture
 'When did you go?'

Read and Payne suggest that the unmarked semantic interpretation of *t^{sh}ika* is 'where'. They also point out that other closely related dialects like Perené and Ucayali of Asheninca have developed additional interrogative pronouns like *paita* 'what' and *ninka* 'who'.

8.4.3 Representation of general concepts

Several languages use the strategy of attaching some of the case markers to proforms in order to express distinctions among general concepts. For example, Lezgian (Daghestanian: Haspelmath 1993: 188) has a single set of adverbial stems for denoting the general notion of location, derived by adding the suffix *-n* to pronominal elements, namely *in-* 'proximate', *an-* 'remote', *hin-* 'interrogative'

and *san-* 'indefinite'. These stems occur with different case markers to denote different types of locations.

(31)		Proximate	Remote	Interrogative	Indefinite
	place	<i>inag</i>	<i>anag</i>	<i>hinag</i>	<i>sanag</i>
	at place	<i>ina</i>	<i>ana</i>	<i>hina</i>	<i>sana</i>
	to place	<i>iniz</i>	<i>aniz</i>	<i>hiniz</i>	<i>saniz</i>
	from place	<i>inaj</i>	<i>anaj</i>	<i>hinaj</i>	<i>sanaj</i>
	on place	<i>inal</i>	<i>anal</i>	<i>hinal</i>	—
	of place	<i>inin</i>	<i>anin</i>	<i>hinin</i>	<i>sanin</i>
	in places	<i>inra</i>	<i>anra</i>	<i>hinra</i>	<i>sanra</i>

Languages show a tendency to lexicalize such case-marked forms. This could lead to the association of such forms with additional case markers, as can be seen in the Kannada example (37) given in the previous chapter (§7.3.4). It is possible that many of the synchronically unanalysable proforms (especially the interrogative ones) have a diachronic source of this type.

Even among the proforms of a single language, the set of general terms that are used in their formation may differ from one pronominal paradigm to another. For example, as noted earlier (§7.2), most languages are described as having a fuller paradigm for interrogative-indefinites than for demonstratives. This might be partly due to the fact that languages like English that have a truncated paradigm of demonstratives have been used for eliciting information, but even after discounting for this possibility, there would still be several languages that show such a disparity.

The use of general terms may occur as a syntactic process rather than as a lexical one in some languages. However, some of these constructions would be used more frequently than others, and hence, gradually, some of them may get grammaticalized and may develop into a set of lexicalized proforms.

8.5 Overlap between pronominal functions and general concepts

Languages show some overlap between pronominal functions and general concepts in the sense that concepts that are generally represented by pronominal elements may, in some languages, be represented by general terms, or vice versa. For example, deictic distinctions like proximate and remote are generally represented by distinctions occurring among pronominal elements. However, languages may also use distinctions occurring among general terms for representing some of them. For example, Paamese (Oceanic: Crowley 1982: 85) makes a three-fold spatial distinction between *kele* 'near 1', *kaisom* 'near 2', *ekok* 'near 1, 2', and *akek* 'distal'. In addition to this, it also makes certain other deictic distinctions, but these are expressed by proforms that are derived with the help of

a set of general verbs to which the prefix *ke-* 'direction' is attached. That is, the distinctions occur as part of the set of general terms that are used in the formation of proforms rather than as part of the pronominal elements.

- (32) a. *mai* 'come' *kēmai* 'over here'
 haa 'go' *kēva* 'over there'
- b. *maa* 'come up' *kema* 'up over here'
 hinaa 'group' *kevina* 'up over there'
- c. *miitaa* 'come down' *kemīa* 'down over here'
 hiitaa 'go down' *kevīta* 'down over there'

Dyirbal (Australian: Dixon 1972: 45, 57) is another language in which the representation of some of the deictic distinctions appears to occur among general terms. It has three different deictic noun markers (determiners), namely *yala-* 'here, visible', *bala-* 'there, visible', and *ḡala-* 'not visible'. All these can occur with a set of bound forms that provide a more specific indication of the location (of the noun's referent), such as *bayḍi* 'short distance downhill', *dayi* 'short distance uphill', *bayḍa* 'medium distance downhill', *daya* 'medial distance uphill', *balbala* 'medium distance downriver', *dawala* 'medium distance upriver', etc. They can also be followed by three other forms, namely *gala* 'vertically up', *gali* 'vertically down', and *galu* 'out in front'. This same set of bound forms can also occur after a set of verb markers that provide locational qualification for the verb, such as *balu* 'towards there', *yalu* 'towards here', *baḡum* 'from there', *yaḡum* 'from here', etc.

This is also true of the person-thing distinction occurring among the interrogative pronouns of several languages (§8.4.2). In some of them, the distinction is clear in the general terms used. For example, Tunica (Haas 1940: 83) has the form *ka'ku* 'who', which contains the general term *-ku* 'masculine' attached to the interrogative element *ka*, and *ka'nahku* 'what', which contains the general term *-nahku* 'like, resembling', attached to the same interrogative element. Similarly, Kannada has the forms *ya:vanu* 'which man', *ya:vaḷu* 'which woman', and *ya:vudu* 'which thing' derived by adding the masculine, feminine, and neuter gender markers to the interrogative-indefinite element *ya:va* 'which'. On the other hand, there are several languages in which the general term has fused with the pronominal element in the case of these two forms and hence it is rather difficult to say whether the distinction occurs in the pronominal element or in the general term.

- (33)
- | | | | | |
|--------------------------|-------------|-------|-------------|--------|
| K. Nubian (Nilo-Saharan: | <i>ni</i> | 'who' | <i>me:r</i> | 'what' |
| A-Hafiz 1988) | | | | |
| T. Quechua (Isolate: | <i>pi:</i> | 'who' | <i>ima</i> | 'what' |
| Adelaar 1977) | | | | |
| Garó (Tibeto-Burman: | <i>sawa</i> | 'who' | <i>maia</i> | 'what' |
| Burling 1961) | | | | |

Awa Pit (Barbacoan: Curnow 1997)	<i>min</i>	'who'	<i>shi</i>	'what'
Nivkh (Palaeosiberian: Gruzdeva 1998)	<i>aŋ</i>	'who'	<i>sidʹ</i>	'what'
Burushaski (Isolate: Lorimer 1935)	<i>mEn</i>	'who'	<i>bESAn</i>	'what'
Mapuche (Isolate: Smeets 1989)	<i>iney</i>	'who'	<i>chem</i>	'what'
Bagrmi (Nilo-Saharan: Stevenson 1969)	<i>naŋa</i>	'who'	<i>'di</i>	'what'
Tinrin (Oceanic: Osumi 1995)	<i>jaa</i>	'who'	<i>ne</i>	'what'

In some of these languages, one may use either of the two stems while forming adverbials like 'where', 'when', 'how', 'why', etc. For example, Pipil (Uto-Aztec: Campbell 1985: 57) has interrogative forms showing distinct pronominal elements for inquiring about humans and non-humans, namely *ka*: 'who' and *ta*: 'what', but either of the two stems could be used while forming adverbials. It has *tayika* 'why' but *ka:n* 'where' and *anka* 'how much'. On the other hand, Khezha (Tibeto-Burman: Kapfo 1993) has the stem *thu* used for questioning humans and *dá* or *dí* used for questioning non-humans. All adverbial forms are derived from the latter, as can be seen from (1) in §7.2.

I described earlier the occurrence of several types of complexities in the representation of distinctions among pronominal elements and general terms (§7.3). It is possible that in some of them there has been a shift in the representation of concepts from pronominal element to general term or vice versa, such as the ones described above. Only a detailed study of the structure of the language under consideration can form the basis of a decision as to whether such forms are to be regarded as involving a complex constituent or two different constituents.

8.6 Summary

I have examined in this chapter some of the differences that occur among languages in the representation of functional distinctions (through pronominal elements) and of general concepts (through general terms). Demonstrative proforms are primarily used for locating an entity in the speech context. They have extended anaphoric uses in most languages, but many languages have special sets of proforms for carrying out this function. Interrogative and indefinite proforms, on the other hand, share the characteristic of denoting lack of knowledge on the part of the speaker. Most languages make use of this particular characteristic as the basis of the strategy that they use for representing interrogatives as well as indefinites. That is, they use a single set of proforms for carrying out the

interrogative and indefinite functions, and differentiate between them either by attaching additional markers to these proforms or by using such markers elsewhere in the sentence. I will be describing the use of this strategy in greater detail in the tenth chapter.

The representation of general concepts among proforms depends, to a certain extent, upon the type of word-class distinctions that are made by various languages. For example, languages that do not differentiate between nouns and adjectives do not generally possess a distinct adnominal proform. On the other hand, languages that allow contentive words to be used as predicates allow their proforms also to be used as predicates. However, there do occur several differences between the kind of word-class distinctions that a language possesses on the one hand, and the kind of distinctions that it shows among the general concepts that are used in the formation of its proforms on the other. These differences apparently derive from distinctions that occur in the functioning of these two syntactic devices.

9 Characteristics of Proforms

9.1 Introduction

I have pointed out in the first chapter some of the most important characteristics that can be used to differentiate between personal pronouns on the one hand and proforms on the other. In some of these characteristics proforms resemble nouns, but in several others they appear to constitute a distinct category. There are also a few characteristics that the proforms share with personal pronouns. I propose to examine these characteristics of proforms in somewhat greater detail in this chapter.

While examining the structure of proforms in the previous two chapters, we found three main types of functions, namely general, pronominal, and categorial, to be relevant for their description. (i) All proforms share a set of general concepts and hence expressing a general concept can be regarded as the first set of their main functions. (ii) The purpose for which these general concepts are used can be different for different proforms, as shown by the association of different pronominal elements with general terms in the structuring of those proforms. We can therefore regard the three main types of paradigms, namely demonstratives, interrogative-indefinites, and relative-anaphors as representing the second set of functions that need to be associated with them. (iii) There are different categorial functions in which proforms are used, and accordingly proforms occurring in each of the above-mentioned paradigms (namely demonstratives, interrogative-indefinites, and relative-anaphors) can fall into different word classes like nouns, adjectives, adverbs, and verbs. These categorial functions can therefore be regarded as their third set of main functions. We can expect all these three main functions of proforms to provide different sets of characteristics to them. I therefore propose to examine the characteristics of proforms in this chapter primarily from the point of view of this three-fold distinction occurring among their functions in language.

9.2 Denoting a general concept

The traditional definition of pronouns as words that ‘stand for nouns’ (or rather for ‘words’ in general) can be applied only partially to proforms, that is, to one of their constituent elements that function as general terms. All proforms share this characteristic of denoting a general concept rather than a specific one. However, the purpose for which such general concepts are used is different for

different paradigms of proforms. For example, they are used in the case of demonstratives in order to draw the attention of the addressee to an entity that is present in the speech context. Since the entity is identified in this case by specifying its location as being proximate or remote to the speaker, above or below him, visible or invisible to him, etc., rather than by providing a detailed description of that entity, the use of a 'general' term would be sufficient in the formation of these proforms. This is also true of anaphoric and relative proforms, as their purpose is to relate a given entity with one that is mentioned elsewhere in the sentence or discourse. They also do not need to provide a detailed description of that entity.

On the other hand, interrogative proforms are used primarily for obtaining information from the addressee regarding an unknown entity, and hence the speaker can provide only a general indication of its identity. He is therefore forced to use a general term for referring to it. Even when interrogative pronouns have some of their extended uses like testing the knowledge of the addressee, there would still be a need to use a general term. This is also true of indefinite pronouns because they are used either to indicate an unknown entity or an unimportant one about which nothing much need be specified. There is thus a common property that is shared by all these different types of proforms, and the notion of denoting a general concept correctly describes that property.

Another interesting property that all these proforms share is the notion of 'identification'. Demonstrative proforms need to identify an entity for the addressee as occurring in the speech context such that its role in an event or state can be specified. Interrogative proforms also need to identify an entity more substantially because the speaker would not otherwise be able to elicit any information about it from the addressee. Terms like 'definite' and 'indefinite' that are generally used for describing this characteristic of proforms fail to bring out the exact nature of this property and to differentiate it from the notion of identification that occurs in the case of definite and indefinite articles. The latter notion is only 'linguistic', whereas the former notion is more than linguistic. The former may involve identification in the real world, in a world of fiction, in mythology, or in a world of dreams, but in each case, it would differ from the kind of identification that needs to be established in the case of definite articles or topic markers (§9.2.1).

Several puzzling characteristics of proforms derive from this requirement, namely that they need to have a more substantial, extra-linguistic type of identification. For example, the 'indefiniteness' of interrogative and of indefinite pronouns can persist so long as the required identification is not established. It can persist even after a 'definite' noun phrase has been used for referring to an entity. On the other hand, the 'indefiniteness' of noun phrases generally ceases to exist the moment the noun phrases are used in a sentence. It is completely devoid of the property of persistence (§9.2.2). Further, a speaker can use a 'definite' demonstrative in order to introduce a new entity that is not readily identifiable for

his addressee on its own. This is because demonstratives carry with them the information that is needed for identifying an entity whereas definite articles depend upon information that the addressee already possesses. This characteristic of demonstratives allows them to be used in some contexts in which only indefinite articles can be used (§9.2.3).

These distinctions occurring in the identifiability of the referents of proforms and of noun phrases can be accounted for by regarding the definiteness (or indefiniteness) of proforms as 'semantic' and that of noun phrases as 'pragmatic'. Proforms are definite or indefinite depending upon the meaning that a speaker can convey through their use, whereas noun phrases are definite or indefinite depending upon the pragmatic context in which they are being used. The definiteness or indefiniteness of the latter does not depend upon the amount of information that the speaker is able to convey through an expression (§9.2.4).

9.2.1 *Linguistic and extra-linguistic identifications*

In most languages, there is a need to differentiate between the kind of identification that needs to be established through the use of proforms on the one hand, and the one that is established through the use of noun phrases containing an indefinite article on the other. In the latter case, a speaker would be satisfied if his addressee concedes the verbal or 'linguistic' identity of the referent concerned, whereas in the former case, he would expect the addressee to establish a more substantial 'extra-linguistic' identification of the referent. For example, a speaker who uses an interrogative pronoun in order to obtain additional information about its referent *in the real world* cannot expect his addressee to provide any information about it, so long as he does not give enough information about that referent that could help the addressee to identify it *in the real world*. On the other hand, the identity ('linguistic') that needs to be established in the case of indefinite noun phrases is only meant for allowing the speech act participants to continue their conversation.

This point is generally missed by linguists who, while describing the use of definite noun phrases, fail to notice that the *minimum requirement* for their use is simply coreference with a previously used noun phrase. It is true that in some contexts their use involves real-life identity or the so-called 'uniqueness' of referents as, for example, in contrastive situations like *the roof is leaking* vs. *a window is broken*. The use of definite article in the former case (*the roof is leaking*) derives from the fact that there would only be a single roof in a house and hence the referent of the noun *roof* is unique in the speech context. On the other hand, the use of indefinite article in the latter case (*a window is broken*) derives from the fact that there would generally be several windows in a house and hence the referent of the noun *window* is non-unique (Hawkins 1978: 101).

The occurrence of such real-life identity or uniqueness, however, does not constitute an essential requirement for the use of definite noun phrases. Hawkins points out, for example, that in order to use such a noun phrase, the speaker needs to establish only a pragmatic set inside the universe of discourse and so long as the given object is unique inside that set, he can use a definite noun phrase for referring to it. But the establishment of this pragmatic set can be achieved by merely referring to a single item and thereby including only one item in one's pragmatic set. That is, the notion of 'identity' involved in the use of definite articles is an extremely weak one. For example, while using a noun phrase like *the unicorn*, the speaker and the hearer would only be 'pretending' that the relevant referent exists (Hawkins 1978: 200). In other instances also, finding a pragmatic set or frame is all that the hearer can do. He must take everything else *on trust* from the speaker.

Heim (1988) makes use of the concept of 'file-changing semantics' for characterizing this distinction between definite and indefinite noun phrases, which clearly brings out this particular aspect of the indefiniteness of noun phrases (see also Du Bois 1980). Understanding a discourse, according to Heim, is like keeping a file in which each discourse referent is represented by a numbered file card. Whenever a new discourse referent is introduced, a new card is added to the file. On the other hand, whenever something is said about a discourse referent that had been introduced earlier in the discourse, it is entered on the card that represents that particular referent. That is, a hearer must add a new card for each indefinite noun phrase, and update an old card for each definite noun phrase. Heim points out further that a file card can fail to correspond to a referent, or two file cards can correspond to the same referent. The definite-indefinite distinction concerns only the file cards and not the actual referents.

On the other hand, the requirements for identification in the use of demonstrative pronouns and for lack of identification in the use of indefinite and interrogative pronouns are more demanding and 'strong'. Unlike the indefinite article, indefinite and interrogative pronouns allow coreference with previously mentioned referents, as for example, when the expression used earlier is not satisfactory for making the required type of 'strong' identification. They differ rather markedly from indefinite articles on this particular point. For example, after referring to an object by an indefinite noun phrase as in (1a), a speaker (or his addressee) generally has to use a definite noun phrase (or an anaphoric pronoun) in the next sentence if he is to talk about the same object, as shown in (1b). It would not be possible to use an indefinite noun phrase for the second time as shown by (2) (Hawkins 1978: 178).

- (1) a. *I had borrowed a book from the library yesterday.*
- b. *The book was very interesting and I finished reading it in a single sitting.*

- (2) *A student was standing outside the factory gate. Bill kept his eye on him. After a little while a student came up to him and asked his name.*

Hawkins points out that the second *a student* cannot refer back anaphorically to the first in (2), according to one of the most fundamental principles of reference in English.

On the other hand, a speaker can persist with his use of an interrogative or indefinite pronoun even when the verbal identification is well established and a definite noun phrase or a third person pronoun is already in use. Consider, for example, the case of (3). Supposing a speaker has uttered the sentences (1a–b) in which he has introduced a referent by an indefinite noun phrase *a book* (1a) and has referred to the same object with the help of a definite noun phrase *the book* and a pronoun *it* (1b). It is possible for his addressee to still use (3), given below, which contains an interrogative pronoun, in order to obtain additional information about that *same* referent, namely *the book* that the speaker had borrowed from the library.

- (3) *Which book did you borrow?*

This is also generally true of the use of demonstratives. They involve more than mere verbal identification. Hawkins (1978: 152) points out that while using a demonstrative like *this* and *that*, a speaker instructs his hearer to *match* the referent with some identifiable object. The identifiability may be either in the visible situation or in the discourse that occurred earlier. Further, the speaker must make it clear which of these two is being intended. No such matching is required in the case of the definite article.

- (4) a. *Pass me that bucket, will you?*
b. *Pass me the bucket, will you?*

Hawkins points out that when a demonstrative is used in sentences like (4a), the object must be visible. If this condition is not fulfilled, there would be a breakdown of communication, and the speaker could expect the addressee to ask a constituent question regarding the identity of that object. In the case of (4b), however, there is no such visibility condition.

The use of demonstrative pronouns for anaphoric use is actually an extension of their primary deictic use, as pointed out by several linguists like J. Lyons (1977: 671) and Diessel (1999: 110). There are languages like Mupun (Chadic: Frajzyngier 1993: 83) in which demonstrative pronouns have only exophoric use. Such languages generally possess distinct expressions for denoting anaphoric coreference. Even when languages allow demonstratives to be used for anaphoric purposes, they may differentiate between such demonstrative anaphors and other anaphoric expressions like third person pronouns and definite articles, as pointed out by Comrie (1997) (see §9.2.3).

I have argued elsewhere (Bhat 1981: 99) that this need to establish two distinct kinds of identification derives from a peculiar problem or dilemma that the users of language face constantly in their day-to-day use of language—a dilemma that results from the interaction of the following two distinct characteristics of language:

(i) As is well known, the use of natural language is free of stimulus. A person can use a noun phrase ‘identifyingly’ for referring to any object that he likes. It may be a non-existent one like *a lion* that he never saw (as for example when he is lying), an imaginary one like *the ghost* that he thought he saw, an improbable one like *a round square* that can never be produced, or, of course, a real one like *a lion* that he did see.

(ii) This ‘freedom of speech’ makes it impossible for the hearer to be sure as to whether or not there is actually an object or entity that he can regard as the referent of a given noun phrase. He would not also be in a position to make sure of this point in most of the situations. However, the conversation cannot proceed smoothly unless the hearer concedes the identity of the various referents on demand from the speaker because there would be a need to refer back to those various entities again and again during the course of speech.

I have put forth a ‘theory of levels’ (Bhat 1981: 162) for representing the solution that natural language users have evolved in order to get out of this dilemma. I have postulated a basic level of identification, namely the *verbal level* (called here the ‘linguistic level’) on the one hand, and a series of non-basic (extra-linguistic) levels such as those of actuality, fiction, mythology, dream, etc. on the other. The identity of a referent at the linguistic level gets conceded readily on hearing an expression that has been used referentially, whereas the identity of the same at one of the extra-linguistic levels gets established gradually on the basis of various kinds of information that become available to the hearer from time to time.

For example, when a speaker refers to an object like *a lion* (that he had seen in a forest), his addressee would readily concede its identity at the linguistic level. This is evidenced by the fact that the addressee would use a definite noun phrase or a third person pronoun while referring to it in his comments or questions about that object. However, the addressee would use his own discretion while deciding as to whether this lion is an actual one or only a fictitious one. If he suspects that the speaker might be lying, he would only place it at the level of fiction, not at the level of reality.

The existence of this duality (linguistic vs. extra-linguistic) in the identification of referents, and of several sub-levels under the extra-linguistic level, allows different individuals to concede the reference of terms like *Santa Claus* (at the linguistic level). It also allows each one of them to place its referent at different sub-levels of reference at the extra-linguistic level. Some could regard it as mythological, some (like small children) as real, and some as fictitious. Further,

all could concede that sentences like (5a) given below are correct, whereas sentences like (5b) are incorrect (see Linsky 1977: 22).

- (5) a. *Santa Claus lives in the North Pole.*
- b. *Santa Claus lives in the South Pole.*

That is, for some persons the correctness of (5a) is at the level of reality, whereas for some others it is at the level of mythology or fiction. In spite of this difference, all are able to use the same 'definite' expression while referring to that individual.

As I have pointed out (Bhat 1981: 145), there are certain speech contexts in which identification, if at all intended, must necessarily get established properly at a particular extra-linguistic level. This is especially true of imperative and interrogative sentences. For example, if a speaker is desirous of ordering someone to bring a 'real' object, he will have to establish its identification at the level of actuality for his addressee. Otherwise he should leave it unidentified, that is, he should make a non-referential (non-specific) use of that expression (see (6) below).

Similarly, if a speaker desires to inquire about a real object, either he should leave it unidentified (i.e. he should provide only a description of that object), or he should provide sufficient information for the addressee to establish its identification at the level of actuality (7).

- (6) *I borrowed this book from a student yesterday. Will you please return it to the student?*
- (7) *I was invited by a man yesterday. Do you know where he lives?*

The use of a definite noun phrase (*the student*) in (6) and a pronoun (*he*) in (7), involving definiteness, is 'grammatically' correct. This is because definiteness needs identification only at the linguistic level, and the mere use of the noun phrase *a student* in (6) or *a man* in (7) earlier in the sentence is quite sufficient for the establishment of such identification. On the other hand, the addressee needs the identification to be established at the extra-linguistic level of actuality if he is to carry out either of these two requests adequately. He needs to know the actual identity of *the student* in order to return the book to that person in (6) and the actual identity of *the man* in order to provide that person's address in (7). Neither of these two sentences contains sufficient information for establishing such extra-linguistic identification and hence neither of them can be regarded as proper requests.

In the two instances given in (6–7), the type of identification that is needed is at the level of actuality. There can also be instances in which identification would be required at other extra-linguistic levels such as that of mythology, fiction, dream, etc. For example, an addressee cannot answer an enquiry regarding a character occurring in a particular novel so long as the speaker has not properly established its identity at the level of that novel for him. What I am suggesting is that proforms are used in establishing identification in one of these extra-linguistic levels and hence their definiteness or indefiniteness belongs to those extra-linguistic levels.

9.2.2 Persistence of indefiniteness

The notion of indefiniteness that is denoted by an article differs from the one that is indicated by interrogative-indefinite pronouns regarding the notion of 'persistence'. In the case of noun phrases that contain an indefinite article, indefiniteness generally comes to an end the moment the phrases are used in a sentence or a clause. In the case of proforms, on the other hand, indefiniteness can persist even after a *definite* noun phrase that is coreferential with them has been used in a sentence. The indefiniteness of the latter would come to an end only when the addressee or someone else has provided sufficient information to satisfy the questioner.

This difference between noun phrases and proforms gets reflected in the fact that indefinite noun phrases generally do not occur in an anaphoric context. All noun phrases that occur after an indefinite (or definite) noun phrase and are coreferential with it are generally definite. Indefinite and interrogative pronouns, on the other hand, can be coreferential with previously used noun phrases in several of their uses, as shown by the sentences (8–10) from Jespersen (1943).

- (8) *There is somebody who is responsible for it, and that somebody is he.*
- (9) *You have got something in your pocket. What is it?*
- (10) *Somebody drove me to the Fallowfield . . . Who was the somebody?*

Notice that in (8), the second *somebody* (with a demonstrative preceding it) is anaphoric to the first *somebody*. Similarly, in (9), an interrogative pronoun is in the anaphoric position and in (10), both an interrogative pronoun as well as an indefinite pronoun occurs in an anaphoric position.

Similar uses of indefinite and interrogative pronouns in contexts of coreference can be noticed in other languages as well. In Kannada (Dravidian), for example, indefinite as well as interrogative pronouns can be anaphoric to expressions (definite or indefinite) that have been used earlier in the discourse.

- (11) a. *alli obba huḍuga kuḷit-idda:ne*
 there one boy sitting-is
 'There is a boy sitting there'
- b. *avanu ya:ru*
 he who
 'Who is he?'
- c. *ya:ro: nana-ge paricayav-illad-avanu*
 someone me-Dat acquaintance-not-they
 'Someone that I do not know'

Notice that the interrogative pronoun *ya:ru* 'who' occurring in (11b) is anaphoric (coreferential) with the indefinite noun phrase *obba huḍuga* 'a boy' occurring in (11a), and also with the pronoun *avanu* 'he' occurring in (11b). Similarly, the

indefinite pronoun *ya:ro*: 'someone' occurring in (11c) can be interpreted as being anaphoric to all the above-mentioned words, that is, the indefinite noun phrase *obba huḍuga* 'a boy', *avanu* 'he', and *ya:ru* 'who'.

Harris (1984: 73) finds the use of demonstratives in apposition to interrogatives in sentences like (12) of Georgian to be rather puzzling.

- (12) *es ra gkonia, papav, amaši?*
 this what you.have.it, grandfather this.in
 'What is this you have in here, grandfather?'

The demonstrative *es* 'this' and the interrogative *ra* 'what' are in apposition in (12). The former is definite (its referent is 'known') whereas the latter is indefinite (its referent is 'unknown'), but both have the same referent. This appears to make the construction, according to Harris, logically inconsistent. How can a referent identified by a demonstrative pronoun as 'known' later on be indicated by an indefinite pronoun as 'unknown'? My claim is that this puzzle derives from the use of the term 'definite' in two entirely different meanings.

Philosophers of language also find this persistence of the indefiniteness of interrogative pronouns to be puzzling. For example, Boer and Lycan (1975: 301) point out the possibility of someone asking a constituent question of the following type in which the referent that is being questioned is very clearly 'definite'.

- (13) *Who is this Erving Smedly, the office-boy who lives at
 851 N Kumquat Lane?*

The interrogative pronoun *who*, occurring in (13), refers to an individual whose name, profession, and even the address are already known to the speaker. How can we claim that an expression that refers to such an individual is 'indefinite' for the speaker? It is evident that the 'indefiniteness' that occurs here is quite different from that of indefinite noun phrases like *an office boy*. The speaker is trying, in this case, to establish a 'real-life' identity. There is apparently some additional information about *the office boy* that he considers unknown to him, and it is that lack of knowledge that prompts him to use the interrogative pronoun *who*.

Hintikka (1974: 140) tries to account for this persistency of constituent questions through the postulation of an additional condition on answers to such questions. He suggests that the referents of noun phrases that are used as substitutions to interrogative pronouns in answers to constituent questions *must be known to the questioner*. Responses to constituent questions, according to him, may have two different functions. Over and above the function of providing the questioner with a substitution value, say 'A', a response may have to serve the additional purpose of supplying enough background information to enable the questioner to know who is being referred to by 'A'; when the questioner does not know who 'A' is, he would come up with another constituent question, 'But who is he?'

9.2.3 Introducing new entities

Proforms differ from definite articles in their ability to introduce new entities. They make such entities identifiable to the addressee with the help of information that is contained in them. For example, demonstratives generally contain spatial information like proximate or remote from the speaker, which help the addressee to locate an entity. Definite articles on the other hand, merely *inform* the addressee that a given participant is identifiable for him. They do not generally provide any specific information for doing so (see Davis and Saunders 1997: 93). This particular characteristic of demonstratives makes it possible for them to occur in some of the contexts in which only *indefinite* noun phrases can occur. For example, Hawkins (1978: 150) points out that there are certain contexts in English that are suitable to the indefinite article rather than the definite article, but still demonstratives can be used in those contexts.

- (14) a. *There was an Englishman in a pub...*
b. *There was this Englishman in a pub...*
- (15) a. *What is wrong with Bill? Oh, he went out with a woman last night who was nasty to him.*
b. *What is wrong with Bill? Oh, he went out with this woman last night who was nasty to him.*

According to Comrie (1997), when demonstratives are used as anaphors in Dutch, they pick out a *non-topical* antecedent, whereas when third person pronouns are used as anaphors, they pick out a topical antecedent. Comrie derives this distinction from the fact that demonstratives are marked expressions as compared to ordinary third person pronouns, which are unmarked expressions. A topic is generally expected to continue and hence unmarked third person pronouns (or definite articles) are used in order to represent the continued occurrence of a topic. Non-continuation, on the other hand, is unexpected, and hence marked demonstratives are used in order to indicate that, instead of the expected topic, a non-topic is being introduced or reintroduced.

C. Lyons (1999: 40, 151) finds the English demonstrative *such* to be problematic because even though a demonstrative, the meaning that it conveys, namely 'of that type' appears to denote indefinite reference. Should we therefore call it a 'demonstrative'? However, several languages like Turkish, Japanese, Lithuanian, Albanian, etc. have pronominal forms that formally belong to the paradigms of demonstratives for denoting this meaning. For example, Turkish (Lewis 1967: 71) has a set of demonstrative paradigms that are derived by using the pronominal elements *bu* 'close to the speaker', *şu* 'a little further away', and *o* 'distant'. One of these paradigms, namely the proximate one, contains the form *büyle* that denotes 'such' (16).

(16)		Proximate	Medial	Distal
	Place	<i>burada</i>	<i>şurada</i>	<i>orada</i>
	Direction	<i>buraya</i>	<i>şuraya</i>	<i>oraya</i>
	Source	<i>buradan</i>	<i>şuradan</i>	<i>oradan</i>
	Type/kind	<i>büyle</i>	<i>şüyle</i>	<i>öyle</i>

C. Lyons examines two different options in this context, namely (i) to regard the expressions that provide the meaning ‘such’ in various languages as ‘indefinite demonstratives’ or (ii) to assume that they are non-demonstratives. He opts for the latter alternative, but this leaves the formal similarity between the words that have the meaning ‘such’ and other types of demonstratives in several languages an unresolved puzzle. On the other hand, if we regard the definiteness of demonstratives as deriving from the meaning that they contain, it is quite possible to regard *such* as ‘definite’ and therefore as rightly belonging to the paradigm of demonstratives (see, however, §9.4 regarding the notion of ‘reference’).

We can account for the above-mentioned difference between proforms and noun phrases by assuming that the definiteness of proforms is ‘semantic’ whereas that of noun phrases is ‘pragmatic’. The former depends upon the amount of information that a speaker possesses regarding the entity that the proforms denote, whereas the latter depends upon the speaker’s assumptions regarding the identifiability of the referent in a given speech context. It does not depend upon the amount of information that the addressee possesses about that referent. The addressee may have to concede the identifiability of a referent purely on the basis of its mere mention in a previous utterance.

9.3 Using proforms for different purposes

The second set of characteristics that proforms manifest derives from their pronominal functions. These are different for different paradigms of proforms. For example, demonstratives are generally used for pointing out the location of an entity to the addressee whereas interrogatives are used, prototypically, in order to obtain information about an entity from the addressee. In the former case, the speaker has the relevant information for locating the entity, whereas in the latter case he does not. In fact, the speaker expects the addressee to supply the necessary information for him in the latter case. (There are, of course, exceptional cases, like a person asking a question even when he has the answer, or using a question for getting something done, but these need to be regarded as extended uses of questions.)

This distinction between demonstratives and interrogatives is generally represented by grammarians with the help of the notion of a distinction between ‘definiteness’ and ‘indefiniteness’. Demonstratives are considered to be definite, whereas interrogatives are considered to be indefinite. Notice,

however, that this definiteness or indefiniteness of proforms depends upon the *speaker's* possession of information whereas that of definite and indefinite articles (or of noun phrases that contain them) depends upon the speaker's assumption regarding the *addressee's* possession of information. It is only the addressee who is not expected to possess the information when an indefinite article is used. Further, in the case of proforms, both the 'knowledge' as well as 'lack of knowledge' regarding the identity of a given referent can co-occur in the same individual as shown by the co-occurrence of demonstratives with indefinites (*that someone*). This would be impossible in the case of definite and indefinite articles.

Interrogative pronouns manifest two other characteristics, namely (i) occurring in the sentence-initial or focus position and (ii) having a corresponding indefinite pronoun in a presupposition. Both these characteristics can become problematic if we do not differentiate between the pragmatic and semantic notions of indefiniteness. We need to assign only the latter to interrogative pronouns (§9.3.1). There are also interesting differences between interrogative and indefinite pronouns, and between interrogative-indefinite and relative pronouns. Some of these will be discussed in the following chapter (§10.2). Another important point on which the definiteness and indefiniteness of proforms differ from those of noun phrases is that the latter is restricted to nominal expressions whereas the former need to be extended to adjectival, adverbial, and even verbal expressions (§9.3.2).

9.3.1 Indefiniteness of interrogatives

The occurrence of interrogative pronouns in the sentence-initial position, which is regarded as the position of the topic, has been rather problematic for linguists who claim that they are 'indefinite'. For example, English does not generally allow indefinite noun phrases to occur in the sentence-initial position, whereas interrogative pronouns are obligatorily moved to the sentence-initial position. That is, one of the most prominent characteristics of the interrogative pronouns of English, namely the so-called 'wh-movement', appears to go against the contention that they are indefinite.

Halliday (1973) makes a distinction between theme–rheme contrast on the one hand and given–new contrast on the other. The former, according to him, concerns the clause structure, whereas the latter concerns the structure of information unit. Theme is what the clause is about. A speaker is free to use any item that he likes as the theme of a clause. The hearer might or might not be familiar with it. It can even be a 'new' item. On the other hand, 'given' is the recoverable part of an information unit. It depends upon what the hearer can recover from the speech context. Halliday considers the occurrence of interrogative pronouns in the sentence-initial position in English as resulting from the fact that they are the themes of constituent questions. The speaker is trying to obtain information

about that particular element in such questions. Since the position of the theme is the sentence-initial one, interrogatives are moved to that position.

Another characteristic of interrogative pronouns, which also appears to go against the claim that they are pragmatically indefinite like the indefinite article, concerns the presupposition of sentences in which the interrogative pronouns are used. Constituent questions are generally considered to involve presuppositions that contain an indefinite pronoun in place of the interrogative pronoun, as shown in the pairs of sentences (17–19).

- (17) a. *Who left the door open?*
b. *Someone left the door open.*

- (18) a. *What did he eat?*
b. *He ate something.*

- (19) a. *Where did he go?*
b. *He went somewhere.*

This relationship between interrogative and indefinite pronouns is generally considered to support the claim that interrogative pronouns are indefinite. Notice, however, that the (b) sentences of (17–19), which contain an indefinite pronoun, are the *presuppositions* of the corresponding (a) sentences. We can associate a presupposition that contains a corresponding indefinite noun phrase only with definite noun phrases but not with indefinite noun phrases. The possibility of associating such a presupposition with interrogative pronouns is another indication of the fact that a different kind of indefiniteness (namely a semantic one) is involved in the case of interrogatives.

9.3.2 *Range of occurrence*

I suggested in the previous chapter that the primary function of proforms is to provide a general concept that can be used in different functions like deictic location, request for information, anaphoric coreference, etc. This general concept need not necessarily be a nominal one. It can be an adjectival or an adverbial one or even a verbal one. Several scholars have noticed this association of demonstrative, interrogative, and indefinite pronouns with different word classes, but the problem that it could cause for the claim that the *same* definite–indefinite distinction occurs in the case of both noun phrases as well as proforms does not appear to have been recognized.

When we examine the use of demonstratives in the nominal function as in (20), or in a nominal modifier or determiner function as in (21), there would not be any difficulty in regarding them as ‘definite’.

- (20) *This will hold two litres of milk.*

- (21) *I bought this jug in Mumbai.*

On the other hand, when we examine their use in certain more complex adnominal contexts, as for example, in constructions like (i) 'this kind of' or 'this type of' and (ii) 'this number of' or 'this amount of', the assignment of definiteness to them appears to be rather problematic. Several languages have distinct demonstrative words for denoting the meanings of these complex constructions, and these, in spite of their being 'demonstrative' appear to be 'indefinite'. I have mentioned earlier (§9.2.3) the problem of the demonstrative *such* in English and its equivalents in other languages. It is an isolated word in the system of English demonstratives and hence it can easily be regarded as not forming part of the demonstrative paradigm (C. Lyons 1999: 157). But comparable words occurring in other languages are not isolated instances of that type. They form part of a *paradigm* of demonstratives and hence the question as to whether they are to be regarded as definite or indefinite (if demonstratives are characterized as definite in general) cannot be brushed aside so easily.

Similarly, there are several languages in which the paradigms of demonstrative proforms include verbal proforms (see §8.4.2). For example, Sawu (Sumbabima, Austronesian; Walker 1982: 25) has three demonstrative verbs, namely *nee* 'be near the speaker', *nei* 'be distant from the speaker', and *nəne* 'be near the addressee' that occur as intransitive verbs.

- (22) *nee* *noo* *pa* *d'əmu*
 be.here 3Sg Loc loft
 'He is here in the loft (where I am)'

The question that needs to be examined here is whether the notion of definiteness-indefiniteness that is considered to be one of the important differences between demonstratives and interrogative-indefinites is to be regarded as extending to the whole paradigm or only to a portion of it (namely the nominal one). C. Lyons (1999: 118) appears to suggest that we need to restrict this distinction to nominal proforms. However, we would then have to find some other notion for differentiating between these non-nominal demonstratives and non-nominal interrogative-indefinites.

For example, the adjectival and adverbial demonstrative forms (and also nominal ones) of Kannada also have correlatable interrogative-indefinite forms.

- | | | |
|------|------------------------------|------------------------------|
| (23) | Remote demonstratives | Interrogative-indefinites |
| | <i>antaha</i> 'of that type' | <i>entaha</i> 'of what type' |
| | <i>aṣṭu</i> 'that much' | <i>eṣṭu</i> 'how much' |
| | <i>alli</i> 'there' | <i>elli</i> 'where' |
| | <i>a:ga</i> 'then' | <i>ya:va:ga</i> 'when' |
| | <i>ha:ge</i> 'like that' | <i>he:ge</i> 'how' |
| | <i>avaru</i> 'those persons' | <i>ya:ru</i> 'which persons' |
| | <i>adu</i> 'that thing' | <i>ya:vudu</i> 'which one' |

Similarly, Tuvaluan (Polynesian: Besnier 2000: 411, 29) has demonstrative verbs derived by prefixing the stem *pee-* to deictic elements *-nei* 'near 1', *-naa* 'near 2', and *-laa* 'distant', and corresponding to this, it also has an interrogative verb derived by prefixing the stem *pee-* to the interrogative element *-fea*.

- (24) a. *ttoeaina teenaa e peenaa*
 the.old.man that Non.past thus
 'The old man is like that'
- b. *e peefea ana fooliga*
 Non.past how her feature
 'What does she look like?'

There is evidently a need to establish a single characteristic that can consistently differentiate between demonstrative proforms and interrogative-indefinite proforms that occur in these two paradigms. If we assume that the distinction between demonstratives and interrogatives-indefinites involves a semantic notion of definiteness whereas the one between definite and indefinite noun phrases involves only a pragmatic notion of definiteness, we can assume further that semantic definiteness has a greater range of occurrence as compared to pragmatic definiteness. The latter is restricted to noun phrases whereas the former can extend to adjectives, adverbs, and even verbs.

9.4 Associating with different categories

The paradigms of proforms generally contain forms that can be used for different sentential functions like referring to a participant in an event, predicating an event, and modifying an argument or a predicate. We can expect them to show different characteristics depending upon the sentential function in which they are primarily used. These different sentential functions are indicated in the proforms by the set of general terms that are used as one of their constituent elements. Languages may also use inflectional markers for differentiating between these different categorial uses.

Proforms that occur in the referential function resemble nouns in taking number, gender, and case inflections. They are, however, different from nouns in some respects, and these differences derive from some of the general characteristics of proforms that have been described in earlier sections. For example, proforms differ from nouns in the kind of referentiality that is associated with them. While noun phrases can be referential or non-referential primarily from a pragmatic point of view, proforms show this contrast primarily from a semantic point of view (§9.4.1). This difference in the referentiality of noun phrases and proforms is related to the distinction, described earlier (§9.2.1) that occurs in the kind of identifiability that is associated with them.

Proforms used in the function of nominal modifiers or determiners also resemble other modifiers (adjectives) and determiners in showing characteristics of agreement with the head noun. Similarly, proforms used in the verbal function resemble verbs or auxiliaries in showing tense, aspect, and mood distinctions and also other distinctions like transitivity, person-number-gender agreement, etc. Even here, differences occurring between these proforms and other members of the categories can be seen as deriving from the general characteristics of proforms.

What is interesting about these categorial characteristics of proforms is that they cut across the various paradigms of proforms like demonstratives, interrogative-indefinites and relative-anaphors. For example, comparable sets of characteristics are shown by the adjectival forms occurring in the demonstrative, interrogative, indefinite and relative proforms. Cross-linguistic differences in this regard would reflect either constraints that occur elsewhere in the grammar or they would derive from the different pronominal purposes for which the proforms are used (§9.4.2).

9.4.1 *Notion of pragmatic reference*

We can differentiate between nominal proforms and other types of proforms on the basis of the fact that the notion of reference can be associated primarily with nominal proforms. However, the fact that the distinction between nominal and adverbial categories is not sharp and clear-cut makes this differentiation among proforms somewhat difficult to maintain. This is also true of the distinction between nominal and adjectival proforms. There are several languages in which there is no clear-cut formal distinction between these two categories and this fact gets reflected among the characteristics that are manifested by the proforms of those languages.

The notion of reference itself is rather confusing in spite of the fact (or, maybe, because of the fact) that a lot has been written on this topic. The most important point that needs to be noted here, I think, is that in the case of noun phrases, reference is basically a pragmatic concept, whereas in that of proforms it is more prominently a semantic concept. As in the case of the notion of definiteness described earlier (§9.2.1), using the same set of terms for referring to both these notions of reference is one of the reasons why there are so many puzzling characteristics that are associated with the notion of reference.

In the case of noun phrases, reference depends crucially upon the *intention* of the speaker. If a speaker intends a given noun phrase to be referential, it will have a referent whether a corresponding object actually 'exists' or not. On the other hand, if the speaker intends it to be non-referential, it would have no referent whatsoever, even if there is a corresponding object that is very familiar to the speaker as well as his addressee. Factors such as the amount of information that the speaker has about the referent, the possibility or the impossibility of the

addressee 'recognizing' the referent, the reality or unreality of the referent, and so on are completely irrelevant in this regard. The nature of our use of language is such that an addressee *has to concede* the establishment of a referent immediately on hearing a noun phrase that has been intended to be referential by the speaker (see §9.2.1).

Further, the question as to whether a speaker would use a given noun phrase referentially or non-referentially in a given context would generally depend upon the *purpose* for which he is using that expression. For example, if he is making a general statement or a habitual one about a 'kind of' object or entity, he would have no reason to refer to a specific object. Similarly, if his purpose is to allow his addressee to choose an object, as for example when he tells the latter '*Bring a detective novel from the library*' there would be no possibility of his using the expression *a detective novel* referentially, that is, with the intention of referring to a particular book.

Another important point is that it is only the speaker who has the freedom to impart a referential or non-referential function to a given noun phrase. The addressee has no choice whatsoever in this respect. If he fails, in a given context, to interpret properly the intention that underlies the use of a given noun phrase, he would only be considered to have misunderstood or misinterpreted the sentence. While writing about the referentiality of noun phrases, linguists generally refer to such noun phrases as having two different *interpretations* or *readings* (see Jackendoff 1972: 279, Givón 1973: 96, Hawkins 1978: 205). This is clearly a mistake because, in a given context, a noun phrase can have only a single interpretation or reading (referential or non-referential) depending upon whether the speaker has intended it to occur in one way or the other. Consider, for example, the following sentence:

(25) *Bill didn't eat a large cake.*

As pointed out by Hawkins (1978: 205), while uttering this sentence, a speaker might have intended to deny either (i) the existence of a large cake that Bill ate (non-referential use), or (ii) Bill's eating of a particular large cake (referential use). An addressee would be expected to correctly guess the intention of the speaker in a particular use of that sentence and interpret the sentence accordingly. Hence, Hawkins's following comment regarding the 'interpretation' of this sentence cannot be correct. He writes: 'Since the hearer lacks the knowledge that there definitely is or definitely is not some particular large cake about which a negative claim is being made, both specific [referential] and nonspecific [non-referential] interpretations are possible'. It is evident, however, that the addressee does not have any such option here. He has to choose the right meaning that the speaker has intended to express through that noun phrase.

Another point that needs to be noted here is that reference is a sub-function of speech act, and hence, it can be established only by the speaker of an utterance and not by the believer or thinker of a proposition. Most of the

examples that linguists provide while discussing the referentiality of noun phrases involve 'non-referring' believers or thinkers as subjects. It is difficult to see how such sentences can be referentially ambiguous from the point of view of those subjects. Consider, for example, the following often-quoted sentence:

(26) *John wants to catch a fish.*

The noun phrase *a fish* occurring in this sentence is considered to be ambiguous depending upon whether *John* has a specific fish in mind or not. However, any ambiguity that occurs in (26) can only depend upon the intention of its speaker—that is, whether he intends it to refer to a specific fish (referential) or to no fish in particular (non-referential). He might be describing the desire of John correctly or incorrectly through such a use, but the point is that the 'referentiality' of the noun phrase *a fish* can only depend upon the speaker's intention and not upon the type of desire that occurred in the mind of John.

Several linguists have tried to associate the notion of referentiality with that of an 'opaque' context (see Quine 1960, Jackendoff 1972, Givón 1973, Lumsden 1988). Noun phrases are considered by these scholars to show an ambiguity between referential and non-referential meanings only when sentences containing them are embedded beneath a verb of 'propositional attitude'. For example, Jackendoff (1972: 280) lists verbs such as *want*, *intend*, *look for*, *wish to*, *plan*, *attend*, *expect*, etc. as providing the environment for such an ambiguity. He considers these verbs to contain a semantic marker called 'unrealized' which, according to him, provides the basis for the non-referential use. Givón (1973: 110) considers the non-opaque environment to be occurring under the scope of a 'non-factive' sentence modality such as negation, future, interrogation and conditional. Habitual and generic sentences are also considered by him to be behaving like opaque environments.

The postulation of this 'opaque' environment is apparently an effort to establish a 'syntactic' basis for the occurrence of the pragmatic referential–non-referential distinction in the meaning of noun phrases. Cole (1978) rightly points out that the non-referential meaning is not due to any syntactic property of the sentences but rather due to the attributive 'use' of the relevant noun phrases. However, he too fails to notice that this distinction (or 'ambiguity') is not restricted to sentences that are embedded beneath a verb of propositional attitude.

Actually, Donnellan shows this claim to be clearly mistaken in several of his articles (see, for example, Donnellan 1978). He points out that a 'definite description' (like *Smith's murderer*) can be used by a speaker even in an assertion like *Smith's murderer is insane* either in a referential or in an attributive (non-referential) sense. In the former case, the intention would be to enable his audience to pick out whom or what he is talking about, whereas in the latter case it would be in order to state something about whoever or whatever is so and so. That is, there is no need for the environment to be 'opaque' in order to assign a

non-referential meaning to a given noun phrase. What is needed is only the intention of the speaker to have no specific referent as such.

This point can also be exemplified with the help of an indefinite noun phrase occurring in a sentence like *I have seen an elephant*. It is possible for the noun phrase *an elephant* of this sentence to have been used non-referentially in spite of the fact that the sentence, being in perfect tense, can only be regarded as non-opaque. For example, if the speaker had in fact seen not just a single elephant, but many elephants, he would not be using the noun phrase *an elephant* in the above sentence for referring to any particular elephant. His intention in using it would only be a non-referential one (see Donnellan 1978 and also Bhat 1981: 35).

Languages do possess specific devices for indicating this pragmatic distinction between referential and non-referential noun phrases. For example, there are several languages in which articles are to be attached to noun phrases in order to indicate that they denote a specific entity. Bemba (Bantu: Givón 1984: 412) is a language of this type. The major distinction among the noun phrases of this language is between referential and non-referential uses. Nouns take a prefix of the form VCV to mark referentiality and a prefix of the form VC to mark non-referentiality.

- (27) a. *a-a-fwaaya* *ici-tabo*
 he-Past-want Ref-book
 (i) 'He wanted the book'
 (ii) 'He wanted a (specific) book'
- b. *a-a-fwaaya* *ci-tabo*
 he-Past-want NonRef-book
 'He wanted some (non-specific) book'

Another language that appears to use articles for denoting a distinction between referential and non-referential uses of noun phrases is Wintu (belonging to the Wintun family of northern California). There are two thematic aspect suffixes in this language, namely the generic *s* (with several alternants like *w*, *m*, *↓*, etc.) and the particular *t~h* (Pitkin 1984: 210). These are used to derive nouns from non-substantives and also to prepare nouns to receive case suffixes. The former denotes a mass in general, a continuum, a group, or simply an unspecified, non-particular individual, class, or genus. The latter, on the other hand, denotes a live, animate, personified, or whole individual, a group considered as a unit, or a punctual action.

Historically, according to Pitkin (1984: 203), the generic form was unmarked. The particular seems to have originated from a topicalizing/foregrounding suffix **t* with something like the force of a definite article (like the disjunctive post-clitic *to* which marks new information) and several other suffixes with the shape *t*. The various allomorphs of the generic suffix, on the other hand, appear to have originated from derivational verb suffixes.

Even in the case of English, the use of articles is primarily meant for differentiating between pragmatic reference and non-reference. According to Chesterman (1991: 26), the basic distinction in English is between zero article and the rest, with the former being of unlimited extensivity and the latter of limited extensivity. The former (zero article) signifies the most abstract sense of a noun, whereas the latter (definite or indefinite article), when added to a noun, has the effect of restricting or limiting its extensivity in such a way that the noun comes to represent a distinct entity or a set of entities.

Chesterman points out that this basic idea about the limiting nature of articles was originally described by Guillaume (1919). Hjelmslev (1928) also referred to it when he claimed that both definite and indefinite articles are morphemes of concretization. When a common noun is used with an article (definite or indefinite), it generally denotes a specific referent (limited extensivity), whereas when it is used without any article, it provides the abstract or 'generic' meaning of unlimited extensivity.

- (28) a. *He has been digging up gold here for years.*
 b. *He has hidden the gold inside the pillow.*
- (29) a. *He is studying life.*
 b. *He saved a life.*

Notice that the nouns *gold* in (28a) and *life* in (29a) have unlimited extensivity as they are used without any article. Their use with the article *the* in (28b) or with *a* in (29b) has the effect of limiting this extensivity.

While the distinction between the occurrence vs. non-occurrence of an article in a noun phrase indicates, in English, primarily the distinction between referential and non-referential uses, the distinction between definite (*the*) and indefinite (*a/an* or *some*) articles indicates a distinction between identifiable and non-identifiable referents. The latter also show number distinction as one can use *a/an* for singular referents and *some* for plural (or mass) referents.

- (30) a. *John has bought a car.*
 b. *He has left the car with his girl friend.*
 c. *He also gave her some money for buying petrol.*

Unlike languages like Bemba, mentioned earlier, however, English is not very consistent in maintaining this primary distinction between referential and non-referential meanings as it allows its articles to be used even in non-referential contexts. For example, both definite and indefinite articles can be used in generic contexts.

- (31) a. *The tiger will become extinct soon.*
 b. *A tiger needs a larger enclosure.*

We need to regard this generic use of articles as an *extended* one because it actually conflicts with their primary function of differentiating between extensive

(non-referential) and non-extensive (referential) connotations of nouns. The ambiguity between referential and non-referential (or attributive) uses of definite and indefinite noun phrases that has been described in detail by Donnellan (1978) and several other scholars is actually a result of this extended use of articles for denoting non-referential meanings. As I will be pointing out in the next section, English shows a similar ambiguity in its use of specific indefinite pronouns, though at a different level of identification.

9.4.2 *Notion of semantic reference*

While the notion of reference that is connected with noun phrases is pragmatic, the one that is connected with proforms is clearly semantic. When a speaker uses referential (nominal) proforms, he expects them to indicate entities in one of the extra-linguistic levels like the real world, fiction, mythology, dream, etc. (see §9.2.1). Their referentiality is not something that the addressee has to concede, but rather one that he has to grasp properly. Being a semantic notion, the referentiality of proforms is also constrained by various types of semantic factors. For example, as I have pointed out earlier (§9.4.1), the notion of 'opaque environment' is irrelevant for a description of the distinction between referential and non-referential uses of noun phrases; on the other hand, the notion is quite relevant for a description of a similar distinction occurring in the use of proforms.

For example, Kannada makes a distinction between referential (specific) and non-referential (non-specific) indefinite pronouns by using a disjunctive particle in the former case and a conjunctive particle in the latter case. The former can be used in indicative sentences whereas the latter cannot (Bhat 1997: 375).

- (32) a. *avanu ellig-o: ho:g-idda:ne*
 he where-or gone-is
 'He has gone somewhere'
- b. **avanu ellig-u: ho:g-dda:ne*
 he where-and gone-is
 '*He has gone anywhere'

On the other hand, both specific and non-specific indefinites can be used in negative, interrogative, and subjunctive sentences.

- (33) a. *avanu ellig-o: ho:gal-illa*
 he where-or gone-not
 'He did not go somewhere'
- b. *avanu ellig-u: ho:gal-illa*
 he where-and go-not
 'He did not go anywhere'

Similar constraints are shown by other languages in which proforms make a distinction between referential and non-referential forms. For example, English

has a set of non-specific indefinite pronouns derived by using the pronominal element *any* that can be used only in certain restricted contexts called 'negative polarity contexts'. Direct or indirect negations, questions, conditionals or standards of comparison are considered to be contexts of this nature (Haspelmath 1997: 36).

- (34) a. *I did not see anything.*
 b. **I saw anything.*
- (35) a. *Did anyone see you?*
 b. **Anyone saw me.*
- (36) *If anyone comes, show him in.*

English differs from Kannada, however, in not restricting the use of its specific indefinite pronouns, derived by using the pronominal element *some*, to specific use. It allows them to be used in non-specific contexts as well.

- (37) a. *I saw something.*
 b. *Did someone see you?*

While the pronoun *something* in (37a) is generally regarded as referential, the same in (37b) can apparently be referential or non-referential. This double use of the semantically referential indefinite pronoun is similar to the double use of pragmatically referential articles (whose primary function is to limit the extensivity of nouns, and thereby to make them referential) in generic and other non-specific contexts, mentioned in the previous section (§9.4.1).

Salishan languages appear to have a system of determiners that are sensitive to semantic rather than pragmatic distinctions. According to Matthewson (1998: 112), these languages differentiate between entities that are asserted to exist and the ones that are not. In St'át'imcets, for example, nouns can be used either (i) with the enclitic *a* (along with one of the deictic markers *ti* 'present', *ni* 'absent' and *k^wu* 'remote') to indicate an entity that the speaker asserts as existing, or (ii) with the determiner *k^wu* when no such assertion is made. The latter occurs within the scope of non-factual operators such as the negative marker, the polar question marker, or the modal *kelh* 'might'. It cannot be used in contexts that induce existential force, such as an ordinary declarative sentence.

- (38) a. *təx^wp-mín-łkan ti púk^w-a łkúnša*
 buy-App1-1Sg Det book-Det today
 'I bought a book today'
- b. *təx^wp-mín-łkan kł k^wu púk^w natx^w*
 buy-App1-1Sg might Det book tomorrow
 'I might buy a book tomorrow'

Matthewson points out further that the determiners of these languages do not encode the pragmatic notions of definiteness and specificity. Their use is based

only on the speaker's knowledge. They ignore the hearer's knowledge. The following pairs of sentences of Sechelt, quoted from Beaumont (1985: 188) exemplify the fact that they do not encode definiteness (Matthewson 1998: 33).

- (39) a. *t'i s'uxwt-as lhe řulhka? slánay*
 Fact saw-he Det snake woman (novel)
 'He saw a snake-woman ...'
 b. *t'i tlúm s-kwal-s lhe slhánay*
 Fact then Nom-speak-her Det woman (familiar)
 'Then the woman said ...'

The same determiner *lhe* occurs in both (39a) and (39b) even though in (39b) the reference is to a previously specified person ('snake woman') and hence identifiable (definite).

9.4.3 *Notion of scope*

Opaque sentences are generally considered to be amenable to 'scope' analysis. That is, the ambiguity of noun phrases that occur in such sentences is considered to result from the possibility that the noun phrases may either be (i) outside the scope of the relevant modality element (and thereby giving rise to referential meaning), or (ii) inside its scope (and thereby giving rise to non-referential meaning). Since the notion of 'opaque context' is irrelevant as far as the pragmatic distinction between referential and non-referential noun phrases is concerned (see §9.4.1), it is evident that the notion of scope cannot also be relevant as far as this pragmatic distinction is concerned.

On the other hand, there is a possibility that this notion of scope is relevant for the description of proforms because the notion of opaque context is relevant for them (see §9.4.2). However, as pointed out by Fodor (1970: 37), it has not been possible to provide any 'general account of just what it is about a semantic contrast which makes variation in the scope of a symbol, rather than the different choice of symbols, an appropriate representational device'.

Dummett (1973: 10) contends that Frege did establish such an explanation for scope analysis. Frege is actually said to have 'extended' this analysis from its application to certain mathematical expressions like $4 \times 5 + 2$ (in which $+$ may be either inside or outside the scope of $\times$) to natural language sentences. Frege's explanation, according to Dummett, is that there are different steps that need to be used for arriving at the result of these expressions, and the scope difference represents a difference in the *ordering* of these steps. In the mathematical expression $4 \times 5 + 2$, for example, if the multiplication forms the first step ($4 \times 5 = 20$), and the addition the second step ($20 + 2 = 22$) the final result would be 22. On the other hand, if the addition forms the first step ($5 + 2 = 7$), and multiplication the second step ($4 \times 7 = 28$), the final result would be 28.

As I have pointed out (Bhat 1981: 84), Frege was quite correct in extending this notion of scope analysis to sentences like *John will buy a house and a cycle or a car* (in which *or* may be either inside or outside the scope of *and*). This is because the two meanings that such sentences provide can ultimately be regarded as derivable from the relative ordering of the two steps that one needs to use while interpreting the two operators *and* and *or*. The sentence can, in fact, be bracketed in exactly the same way in which the earlier mathematical expression can be bracketed.

(40) a. $(4 \times 5) + 2 = 22$

b. $4 \times (5 + 2) = 28$

(41) a. *John will buy (a house and a cycle) or a car.*

b. *John will buy a house and (a cycle or a car).*

According to Dummet (1973) Frege extended this explanation further to sentences that contain quantifiers, such as (42), and thereby provided an explanation for the scope distinction that occurs in them.

(42) *Everybody envies somebody.*

a. *some x (everybody envies x).*

b. *every y (y envies somebody).*

The problem with this explanation, however, is that sentences that contain only a single quantifier, such as (43), also show the same type of ambiguity between referential and non-referential uses of a quantifier as the one shown by the earlier sentence.

(43) *John envies somebody.*

On the other hand, there is no possibility of using the notion of 'relative order of steps' for explaining this meaning distinction in the case of (43). This is because (43) can allow only a single derivational step. There is no possibility of applying the notion of scope to (43) either, because it does not involve any opaque contexts.

I believe that the meaning distinction occurring in sentences like (42) and (43) derives from the existence of two different *derivations*, rather than from a difference in the relative order of steps as claimed by Frege. In (43), for example, one can differentiate between the two meanings by postulating two distinct derivations, namely *John envies x* to *John envies somebody* in the case of the referential meaning, and *John envies x, y, or z* to *John envies somebody* in the case of the non-referential meaning. In the case of (42) also, a similar distinction in the nature of derivation used can form the basis of the meaning distinction. That is, the distinction results from the use of *somebody* either (i) as a semantically referential expression or (ii) as a semantically non-referential expression.

In languages like Kannada in which specific and non-specific indefinites are clearly differentiated from one another, distinct indefinite pronouns are used in order to express the two meanings of sentences like *Everybody envies somebody*. This, I believe, supports my claim that the distinction (or ambiguity) of sentences of this type occurs in the pronoun itself and not in the ‘scope’ of some other entity.

- (44) a. *ellar-u: ya:rann-o: nambutta:re*
 all-also who-or believe
 ‘Everybody believes somebody (specific)’
 b. *ellar-u: ya:rann-a:dar-u: nambutta:re*
 all-also who-become.Cond-also believe
 ‘Everybody believes somebody (non-specific)’

Kannada attaches the disjunctive particle *o*: ‘or’ to an interrogative (unmarked indefinite) pronoun in order to derive specific indefinites and the conjunctive particle *u*: ‘also’ to the same for deriving non-specific indefinites (see §11.2.1). In the latter case, one can further attach the conditional form of the verb *a:gu* ‘to become’ in order to limit the extensivity of the pronoun as in (44b).

9.5 Summary

I have argued in this chapter that there is a need to differentiate between two different types of reference, and correspondingly two different types of identifiability. I have proposed to differentiate between these two notions of reference (and identifiability) by regarding one of them as ‘pragmatic’ and the other one as ‘semantic’. In most languages, the pragmatic notion occurs in the use of definite and indefinite noun phrases, and also third person pronouns, whereas the semantic notion occurs in the use of proforms. When a noun phrase is used with an article, its referentiality or non-referentiality depends upon the *intention* of the speaker. Even in a most non-opaque context, a noun phrase can have a referential or a non-referential connotation depending upon what the speaker intends.

The identifiability of a referent in the case of a noun phrase that has been used referentially is automatic. The moment an indefinite noun phrase is used in a sentence, its referent becomes ‘identifiable’. There is generally no option for the addressee or someone else but to concede its identity. Any reference to that referent in the ensuing part of the discourse is therefore made with the help of a definite noun phrase or a ‘definite’ third person pronoun. Hence, the identifiability can only be regarded as ‘linguistic’ and pragmatic.

On the other hand, the referentiality or non-referentiality of proforms depends upon a more substantial basis. Correspondingly, the ‘definiteness’ of a referent that is introduced by proforms also depends upon the availability of a more substantial basis for identification. It is quite possible for the addressee to feel

unsatisfied with the available basis for identification in the case of a proform. He may demand additional information before the notion of indefiniteness is given up. Further, the notion of 'opaque context' is relevant in the case of proforms but not in that of noun phrases. In an opaque context, the speaker may not be in a position to provide the necessary information for establishing the necessary identification and hence he would be forced to use proforms that do not involve any referentiality. In the case of noun phrases, on the other hand, the opaqueness of contexts has no relevance whatsoever because the kind of identification that needs to be established in their case is only linguistic (or pragmatic) and can be established in any type of context.

10 Interrogative–Indefinite Puzzle

10.1 Introduction

Interrogative and indefinite pronouns are considered to be either identical in form, or derivationally related, in the majority of the world's languages. For example, in Lakhota (a Siouan language spoken in North and South Dakota), interrogative and indefinite pronouns are identical in form. They are differentiated from one another by the fact that in constituent questions the pronoun occurs with the question particle *he* whereas in declarative sentences it occurs without such a particle (Van Valin 1993: 98).

- (1) a. *šúka ki táku yaxtáka he*
dog the what bite Int
'What did the dog bite?'
b. *šúka ki táku yaxtáke*
dog the something bite
'The dog bit something'

In Kannada (Dravidian), on the other hand, indefinite and interrogative pronouns are considered to be derivationally related. There is a distinction between specific and non-specific indefinite pronouns, with the former being derived by adding the disjunctive particle *o*: 'or' to interrogative pronouns, and the latter by adding the conjunctive particle *u*: 'also' to the same.

- (2) a. *ra:ju ellige ho:da*
Raju where went
'Where did Raju go?'
b. *ra:ju ellig-o: ho:da*
Raju where-or went
'Raju went somewhere'
c. *ra:ju ellig-u: ho:d-a:nu*
Raju where-also go-may
'Raju may go anywhere'

According to Ultan (1978: 230), this relatedness between interrogative and indefinite pronouns is a near-universal characteristic of natural languages. He finds all the languages excepting two of his sample of 79 to be showing such an identity or derivational relationship. Haspelmath (1997: 26) suggests, on the other hand, that this affinity between indefinite and interrogative pronouns may not be as ubiquitous as Ultan's data suggest, because he finds only 63 of his

sample of 100 languages to be showing that affinity. He finds the indefinite pronouns of the remaining languages to be made up of an indefinite article (or the numeral ‘one’) and a generic noun like person, thing, place, time, etc. They may also involve the use of only a generic word. However, many of these latter type of languages may also allow their interrogative pronouns to be used in the meaning of indefinite pronouns in some restricted contexts.

For example, Tzutujil (Mayan: Dayley 1985: 70) has an indefinite pronoun derived from *juun* ‘one’ that is distinct from the interrogative pronoun *naq* ‘who, what, which’ (oblique form *choq*). However, it also has an indefinite pronoun *ya-naq-ta* ‘whoever’ (involving *ya* ‘only’, *naq* ‘who’ and *ta* ‘irreal’), that is, an indefinite pronoun that shares its stem with the interrogative. Further, several additional languages are similar to the former type in allowing the interrogative pronoun to occur in the so-called indirect questions which, even though called ‘questions’, do not actually involve the notion of interrogativity (see §11.3). This leaves the number of languages that really distinguish between interrogative and indefinite pronouns to be rather small as compared to the ones that use the same pronominal element for representing both of them.

10.1.1 Puzzle about the indefinite–interrogative relationship

One interesting question that has been raised in connection with this relationship between indefinite and interrogative pronouns concerns the nature of this affinity. Is this affinity one of monosemy or polysemy? That is, are we to regard interrogatives and indefinites as representing (i) two different uses of one and the same set of pronouns or (ii) as the uses of two different sets of pronouns that have different meanings but only happen to be similar or identical in form?

By taking the former alternative, one can claim that both these types of sentence involve ignorance or an ‘information gap’. In the case of an interrogative usage, the gap represents the speaker’s inability to provide the necessary information, whereas in that of an indefinite usage, it represents either his inability or his unwillingness to do so. Consider, for example, the following two sentences:

- (3) a. *Someone stole my bike.*
- b. *Who stole my bike?*

In (3a), the speaker has used an indefinite pronoun, namely *someone*, either because he does not know the exact identity of the person who stole his bike or because he does not desire to specify that identity to his addressee. In (3b), on the other hand, he does not have the necessary information for identifying that person, and is in fact trying to obtain it from his addressee by uttering (3b).

According to Haspelmath (1997: 175), Karcevski (1969) uses the term ‘ignorative pronoun’ for referring to indefinite as well as interrogative pronouns. I point out (Bhat 1989) that the same set of pronouns is used not only as indefinite and interrogative pronouns, but also as the pronouns of relative clauses and

exclamatory sentences in several languages. I argue that all these usages have the common characteristic of indicating an information gap.

An alternative analysis that has been suggested for this interrogative–indefinite affinity is to regard the two sets of pronouns to be distinct in their meaning and explain the affinity on the basis of the fact that interrogative pronouns are semantically more complex structures than the corresponding indefinite pronouns. While they share the property of denoting ignorance or information gap, they differ from one another in that the interrogative pronouns denote, in addition, the notion of a ‘request for information’. The fact that questions containing an interrogative pronoun *presuppose* statements that contain an indefinite pronoun in place of the interrogative pronoun (see Katz and Postal 1964: 116, Quirk et al. 1972: 396) appears to support such an explanation. Consider, for example, the following pairs of sentences in which (a) presupposes the corresponding (b).

- (4) a. *Who left the door open?*
b. *Someone left the door open.*
- (5) a. *What did he eat?*
b. *He ate something.*

Klima (1964: 252) makes use of this relationship between indefinite and interrogative pronouns as the basis of his claim that the latter ‘incorporate’ the former. That is, interrogative pronouns like *what* can be considered, according to him, to be ‘transformationally’ derived from an underlying question operator WH plus an indefinite pronoun like *something*.

This explanation for the relationship between interrogative and indefinite pronouns is quite attractive as it correctly describes the semantic relationship that exists between the two types of pronouns. But the problem, as pointed out by Haspelmath (1997: 176), is that the explanation would require the interrogative pronouns to be derived from indefinite pronouns. In actuality, however, the derivation appears to proceed in the *opposite* direction. There are many languages in which indefinite pronouns are reported to derive from interrogative pronouns through the addition of an affix or a particle (§11.2), whereas there are very few languages in which interrogative pronouns are reported to derive from indefinite pronouns in a similar fashion (§10.2.4). Further, interrogative pronouns, as pointed out by Haspelmath, are resistant to etymological analysis in most languages. They also appear to be the slowest changing elements in any language.

Assuming that the derivation occurs in the opposite direction, that is, from interrogatives to indefinites is also equally problematic because there can be no functional or semantic explanation that can form the basis of such a derivation. Take, for example, the ‘derivation’ of indefinite pronouns from interrogative pronouns through the addition of conjunctive or disjunctive particles, that has been reported to occur in several languages (see §11.2.1). The Kannada examples

given earlier (2a–b) exemplify this derivation. What could be the semantic or functional explanation for this derivation? How is it that a question about an unknown entity changes into a statement about it when the notion of conjunction or disjunction is added to it? In view of the difficulty of providing a satisfactory answer to this question, Haspelmath (1997: 163) suggests that the derivation may be regarded as proceeding from *indefinites* to indefinites rather than from *interrogatives* to indefinites. That is, he considers the two sets of pronouns to be polysemous, with the conjunctive and disjunctive particles being added to pronominal bases that are distinct in meaning from interrogative bases (apparently to provide non-specific and specific indefinite meanings—see §11.2). This, however, leaves the puzzle about the affinity between interrogative and indefinite pronouns unresolved.

10.1.2 *As part of a system of proforms*

I would like to suggest in this connection that the affinity between interrogative and indefinite pronouns appears to be puzzling mainly because linguists have been merely comparing the two types of pronoun with one another. As we have seen in the seventh chapter, the two form part of a larger system of proforms that includes, in addition to the paradigms of indefinite and interrogative pronouns, one or more paradigms of demonstratives and also of relatives and correlatives. We have to examine this whole system of proforms and see how interrogative and indefinite pronouns fit into that system. The affinity between the two will become clear only when we examine the way in which the various paradigms that form part of a system of proforms interact with one another.

As we have seen in the seventh chapter, a cross-linguistic study of the paradigms of proforms appears to suggest that there is a primary function (or requirement) that is common to all of them, namely to indicate a general concept such as a person, thing, place, time, manner, etc. The pronouns are grouped into different paradigms on the basis of the *purpose* for which these general concepts are utilized. In the case of demonstratives, for example, there is generally no need to provide details regarding those concepts because the speaker would be using deictic factors like being near the speaker or away from the speaker for identification. In the case of relative pronouns also, details regarding the participant need not be provided because their purpose is merely to identify the participant as being the same as the one that is identified elsewhere in the matrix clause of the sentence. In the case of indefinite pronouns on the other hand, details are not provided regarding a concept either because the speaker considers them to be irrelevant or because he does not possess them. The latter is generally true of interrogative pronouns as well (see §10.2).

Some languages use different paradigms for these different purposes, whereas some combine together two or more purposes and use the pronouns of a single paradigm for representing them. Languages of this latter type use other aspects of

sentence structure for differentiating between the individual purposes that are combined together and represented by a single paradigm.

Consider, for example, two of the purposes that a language can have for using a general concept in a sentence, namely (i) obtaining additional information about that concept (interrogative use) and (ii) merely leaving the concept unspecified (indefinite use). Supyire (Gur family, Niger-Congo) differentiates between these two purposes by using a set of interrogative pronouns in the former case, and a different set of indefinite pronouns in the latter case, as shown in (6a–b) (Carlson 1994: 533).

- (6) a. *wà na wá mɛŋi i*
 someone Prog be.there there.Def at
 ‘There is someone over there’
 b. *jǒ u a kù bò yɛ*
 who 3Sg Perf it kill Int
 ‘Who killed it?’

On the other hand, Lakhota combines the two purposes and represents them with the help of a single paradigm of pronouns. However, it differentiates between the two purposes by specially marking the interrogative use; in this latter use of those pronouns, (i) it employs an additional question particle *he* and (ii) it has the relevant pronoun in focus (see 1a–b given earlier).

Similarly, Bengali (Indo-Aryan) differentiates between (i) the purpose of obtaining additional information about a concept (interrogative use) and (ii) the purpose of indicating coreference with a concept that has been expressed elsewhere in the sentence (relative use) by using two distinct sets of pronouns, called *k*-words and *j*-words, as shown in 7a–b (Dasgupta 1979).

- (7) a. *ke bajare gie chilo*
 Int market gone had
 ‘Who had gone to the market?’
 b. *je bajare gie chilo se ekhane ache*
 Rel market gone had he here is
 ‘The man who had gone to the market is here’

On the other hand, English combines the two purposes and uses a single set of pronouns, called *wh*-words, for both of them, as can be seen in the translations of sentences (7a–b). However, it differentiates between the two purposes by (i) using an interrogative intonation and (ii) shifting the pronoun to the initial position in the former case, and by placing the coreferential noun phrase immediately before the pronoun in the latter case.

Thus the solution that I wish to suggest for the indefinite–interrogative puzzle is that the affinity between the two types of pronouns derives from the fact that they share the notion of ‘lack of information’. This is true not only from the

semantic or logical point of view as pointed out above, but also from the formal or grammatical point of view. Linguists like Klima (1964) and Haspelmath (1997) have been trying to find a solution to the interrogative–indefinite puzzle in the pronoun itself, whereas my claim here is that the differentiation does not occur in the pronoun itself in these languages, but elsewhere in the sentence. The languages use the same unmarked indefinite pronoun in both constituent questions as well as statements involving the indefinite notion. However, they differentiate between the two by using one or more *non-pronominal* devices along with the pronominal device and express through them additional meanings such as interrogation and focus in the case of constituent questions and specificity or non-specificity in the case of indefinite pronouns.

10.1.3 Semantic differences

There are two important semantic differences between sentences that contain indefinite pronouns on the one hand, and constituent questions that are generally reported to contain interrogative pronouns on the other. They are (i) the notion of interrogativity and (ii) the notion of focus. In order to claim that the pronouns occurring in constituent questions are only ‘indefinite’ pronouns in most of the languages, it is therefore necessary to show that those languages possess alternative devices for expressing these two meanings. That is, we need to show that with the help of such additional devices, they are generally able to differentiate between constituent questions on the one hand and sentences that contain indefinite pronouns on the other (see Bhat 2000).

The fact that constituent questions differ from sentences that contain indefinite pronouns by the occurrence of interrogativity as an additional concept can be seen in (8a–b).

- (8) a. *Mary went somewhere.*
- b. *Where did Mary go?*

Notice that (8b) differs from (8a) by the fact that (8b) expresses interrogativity whereas (8a) does not. In (8a), the speaker merely expresses his lack of knowledge regarding the location of Mary’s travel whereas in (8b) he also asks the addressee to provide information about it. However, it is not necessary for a sentence to contain an ‘interrogative’ pronoun in order to express the notion of interrogativity. For example, (8c), called a polar question, also expresses the notion of interrogativity even though it does not contain any interrogative pronoun. In fact, a sentence need not even contain a pronoun in order to express interrogativity, as shown by (8d).

- (8) c. *Did Mary go somewhere?*
- d. *Did Mary go to the market?*

A constituent question like (8b) is similar to a polar question like (8c–d) in that both of them denote the notion of interrogativity. However, it differs from the

latter in having the scope of interrogation restricted to a single constituent, namely the one represented by the pronoun. Sentence (8b) asks for information only about the *location* of Mary's travel whereas sentences (8c–d) ask for information regarding the whole action. Interrogative sentences that do not contain any 'interrogative' pronouns (or any pronouns) can also have the interrogativity restricted to a single constituent, as for example, through the use of a focus construction as in (8e).

(8) e. *Is it to the market that Mary went?*

It is evident from the preceding discussion, I believe, that there is no need to use an 'interrogative pronoun' in English, in order to express either of the two meanings that differentiate constituent questions from sentences that contain indefinite pronouns. This appears to be the case in the majority of the world's languages, that is, languages that are reported to show identity or derivational relationship between interrogative and indefinite pronouns. Hence, there is no reason why the notion of interrogativity must be attached to 'interrogative' words in such languages, especially since it gives rise to unnecessary puzzles and paradoxes.

10.1.4 *Solution to the puzzle*

The solution that I wish to suggest here for the interrogative–indefinite puzzle is based upon this absence of any necessity to regard the pronouns occurring in constituent questions of most languages as 'interrogative'. As mentioned earlier, the distinction between constituent questions on the one hand, and statements with an unknown constituent on the other (i.e. between sentences like *who arrived?* and *somebody arrived*), is represented in some languages by a distinction in the set of pronouns used, whereas in others it is expressed elsewhere in the sentence, using the same set of pronouns (or derivationally related ones) in both the contexts. Pokot, a Nilo-Hamitic language (Baroja 1989), for example, is of the former type. In that language, a set of interrogative pronouns (like *ngo* 'who', *nyonö* 'which one', and *onö* 'where') contrasts with a completely different set of indefinite pronouns (like *anim* 'someone', *kitikon* 'anyone', and *kareran* 'anywhere'). Ainu (Tamura 2000: 232) is another language in which interrogative and indefinite pronouns are quite different from one another (9).

(9)	Interrogatives		Indefinites
	<i>hunna</i>	'who'	<i>nen</i>
	<i>húta</i>	'what'	<i>nep</i>
	<i>hunak</i>	'where'	<i>nei</i>
	<i>hempara</i>	'when'	<i>nei</i>
	<i>mak</i>	'how'	<i>neu</i>

English belongs only partly to this group, as only some of its indefinite pronouns (like *someone* and *anyone*) are unrelated to interrogatives. In addition to

these, it has indefinite pronouns like *somewhere* and *anywhere* or *somehow* and *anyhow* that are related to interrogative pronouns.

Maricopa (Yuman: Gordon 1986: 62), on the other hand, is a language of the latter type, as it uses one and the same set of pronouns in both these types of sentence. The two are differentiated, however, by the use of an additional interrogative particle in constituent questions (10a–b).

- (10) a. *mki-sh* *hav-ii*
 someone-Subj enter-Int
 ‘Who came in?’
- b. *mki-sh* *hav-sh*
 someone-Subj enter-Perf
 ‘Someone entered’

The interrogative–indefinite puzzle does not affect languages of the first type because they do not show any affinity between indefinite pronouns that are used in declarative sentences and pronouns that are used in the corresponding constituent questions. On the other hand, it does affect languages of the second type mainly because grammarians have been generally assuming that any pronoun that occurs in a constituent question and stands for the constituent that is being questioned is an ‘interrogative’ pronoun, regardless of its form and regardless of whether or not the interrogative meaning is otherwise expressed in the sentence. If it is possible to discard this assumption, and to assume instead that the pronouns that occur in the constituent questions of such languages are *indefinite* pronouns, the puzzle will automatically disappear from these latter languages.

This alternative analysis is clearly the most desirable one from the semantic point of view as well. Constituent questions presuppose declarative sentences that contain the corresponding indefinite pronouns. They differ from the latter only in the additional fact (i) that they request for information (interrogation) and (ii) that they restrict that request to a particular constituent, namely the one that is indicated to be unknown (focus), as I have pointed out earlier (§10.1.3). If these additional meanings of constituent questions can be assigned to other aspects of those sentences, there should be no problem in regarding the pronouns that occur in them as identical with the pronouns that occur in the corresponding declarative sentences.

I will show in the following sections (§10.2–4) how languages that use the same pronouns, or derivationally related ones, in constituent questions as well as statements can be regarded as having only indefinite pronouns in their constituent questions. This is made possible by the fact that the languages express the above two additional meanings of interrogation and focus through alternative devices in their constituent questions. Such an analysis is also helpful in providing an explanation for the derivation of non-specific and specific indefinite pronouns through the use of conjunctive and disjunctive particles, as I point out in the next chapter (§11.2). It is also helpful in resolving the interrogative–relative puzzle (§11.4).

10.2 Request for information

Most of the languages that use a single set of pronouns in both declarative sentences as well as constituent questions have been reported to use one or more additional devices in their constituent questions. The meaning of interrogation ('request for information') can easily be assigned to one of these additional devices instead of assigning them to the pronoun itself. The devices used include interrogative intonation (§10.2.1), interrogative particle (§10.2.2), and interrogative mood (§10.2.3). The languages also use these devices in other types of sentences that involve interrogativity, such as polar questions. That is, the devices appear to express the meaning of interrogativity in these uses as well and hence they support the claim that the meaning of interrogativity may be assigned to these devices rather than to the interrogative pronoun in the case of constituent questions in these languages. There are also a few languages in which these interrogative particles are directly attached to indefinite pronouns. This has the effect of making interrogative pronouns to be marked structures as compared to indefinite pronouns (§10.2.4).

10.2.1 Use of interrogative intonation

The device that is used most commonly by languages in order to indicate a request for information is interrogative intonation. Even languages that make use of other devices like a question particle or a distinct interrogative pronoun may utilize intonation as a supporting device. In English, for example, a statement is changed into a polar question by moving the verb to the initial position, but in addition to this, the intonation of the sentence is also generally changed from a statement intonation to a question intonation. One can even change a statement into a question by merely changing its intonation in many languages like Damascus Arabic (Ambros 1977: 84), Margi (Afroasiatic, Chadic: C. Hoffmann 1963: 98), and Tagalog (Schachter and Otnes 1972: 501).

On the other hand, there are languages like Chinese in which the use of interrogative intonation is the only way in which a statement can be changed into a question. Xu (1990: 357) considers the pronouns of Chinese to be functioning 'ambiguously' as indefinites and interrogatives, with the sentences being disambiguated (as statements or questions) on the basis of the type of intonation that occurs with them.

- (11) *zheli que-le shenme*
 here miss-Aspect what
 (i) 'There is something missing here'
 (ii) 'What is missing here?'

According to Xu (1990), the nuclear head of intonation is on *shenme* 'what' when it is a question and on *que* 'miss' when it is a statement. The alternative analysis

that is being suggested here is to regard the pronoun *shenme* as indefinite in both the meanings (11i, ii), and to assign the meaning of interrogation to the intonation. In Vietnamese (Thompson 1991: 143) also, sentences that contain indefinite pronouns are to be interpreted as constituent questions if they occur with sustaining intonation, whereas if they occur with fading intonation they are to be interpreted as statements. Thompson considers the term ‘interrogative’ to be inappropriate for these pronouns.

Kamaiurá, a Tupí language of Brazil, also allows sentences that contain question words to be interpreted as questions or statements depending upon the intonation that occurs with them (Brandon and Seki 1984: 96).

- (12) *awa* *o-‘ut*
 who 3-come
 (i) ‘Who is coming?’
 (ii) ‘People are coming/someone is coming’

Similar uses of intonation for differentiating between statements and questions that share the same set of pronouns have been reported to occur in Penutian languages like Klamath (Barker 1964) and Siuslaw (Frachtenberg 1922: 584). In Yaqui (Uto-Aztec), both polar questions and constituent questions carry a special intonation pattern that is characteristic of questions. They are also usually marked by the particle *haisa* occurring at the beginning of the sentence (Lindenfeld 1974: 33). In Hixkaryana (Carib), constituent questions can be used as rhetorical questions where the intent is not to ask for information, but to give it, or to express surprise, frustration, annoyance, or the like. Derbyshire (1979: 19) points out that in these rhetorical usages constituent questions do not occur with the normal question intonation.

10.2.2 Use of interrogative particles

An alternative device used by several languages for denoting interrogation is the use of an interrogative particle (or affix), along with the interrogative pronoun, in their constituent questions. The same particle occurs in polar questions as well. The case of Lakhota (Siouan: Van Valin 1993: 98), described earlier (§10.1) exemplifies this device.

- (13) a. *šúka ki táku yaxtáke*
 dog the something bite
 ‘The dog bit something’
 b. *šúka ki táku yaxtáka he*
 dog the what bit Int
 ‘What did the dog bite?’
 c. *šúka ki igmú wə yaxtáka he*
 dog the cat a bit Int
 ‘Did the dog bite a cat?’

Notice that the interrogative/indefinite pronoun *táku* 'what/something' occurs with the interrogative particle *he* in (13b) to indicate the interrogative meaning, whereas in (13a) it occurs without that particle and has no interrogative meaning. In (13c), which is a polar question, on the other hand, the particle occurs by itself and provides interrogative meaning.

Grammarians generally regard the particles as well as the pronouns that occur in such constituent questions to be 'interrogative'. Van Valin (1993), for example, considers the word *táku* of (13b) to be an interrogative pronoun, even though in a sentence like (13a) it occurs with an indefinite meaning. Such an analysis is actually indefensible because in all sentences in which the notion of interrogation is to be specified, the element that needs to be used obligatorily is the interrogative particle and not the interrogative pronoun, as shown by (13c). On the other hand, the meaning that needs to be obligatorily assigned to all sentences that contain an 'interrogative-indefinite' pronoun is the notion of indefiniteness and not the notion of interrogativity, as shown by (13a-b). There appears to be no basis, therefore, for regarding the pronoun as providing the interrogative meaning in languages of this type.

There are several other languages, belonging to different parts of the world, that may be analysed in a similar fashion. For example, Wintu, a Penutian language (Pitkin 1984: 124, 269) is reported to be using an interrogative suffix (which is *wi*, after vowels and *i*, elsewhere) in both polar questions and constituent questions. Pronouns that occur in constituent questions can also occur in declarative sentences, but in such a usage they are not accompanied by the interrogative suffix. In Ika, a Chibchan language of Colombia (Frank 1990: 79), all interrogative sentences are marked by one of the verb phrase-final suffixes or particles *-e*, *-o*, or *no*; *-e* occurs in questions regarding past time and the others occur in questions referring to present, future or distant past. Sentences that do not involve interrogation may contain an interrogative pronoun but not the interrogative suffix.

- (14) a. *ini-sin* *na-nas-e*
 who-with 2Sg-came-Int
 'With whom did you come?'
 b. *beru-sin* *na?ku-in*
 burro-with came-Witness
 'I came with the donkey'

Assuriní, a Tupí language of Brazil (Brandon and Seki 1984: 94) uses the question particle *pa* along with an interrogative pronoun in constituent questions. In the absence of this particle, the interrogative pronoun denotes an indefinite meaning. Brandon and Seki (1984: 100) consider it possible to regard the interrogative pronouns of this language as basically indefinite pronouns or as words that have semantically general meanings. The question

particle occurs in polar questions as well. In Carib (Cariban family), constituent questions require an interrogative particle *ko* in addition to a pronoun that indicates the questioned constituent. Hoff (1968: 271) describes the pronoun as an ‘indefinite demonstrative’. It provides only indefinite meaning in the absence of the interrogative particle. In Sanuma (Amazonian, Yanomami family: Borgman 1989: 66), the distinctive feature common to all questions is the lack of a glottal stop at the end of the sentence. Borgman reports that the non-interrogative sentences almost always have a sentence-final glottal stop, and when they do not, the context would clearly prohibit the interpretation of the utterance as a question.

According to Newman (2000: 493), sentences containing a question word in Hausa (Chadic) add a Q-morpheme at the end. This morpheme consists of length and low tone. It attaches itself to the immediately preceding syllable. It seems to lengthen short final vowels and add a low tone to words ending in high tone and thereby producing a fall. Newman suggests that historically the morpheme had a segmental shape, which was most probably **a*, and is still preserved by some Chadic languages. Regarding the occurrence of interrogative pronouns in Grebo, a Kru language, Innes (1966: 109) states clearly that such words, occurring initially in constituent questions, are not the ones that signify question. It is rather the question marker *ε* or *a*, that occurs after the verbal stem or after the tense affix, that signifies question. Innes points out further that the former (interrogative pronouns) also occur in statements. Other languages that are reported to have such question particles or affixes in constituent questions include Mupun (Chadic: Frajzyngier 1993: 366), Wararo (a language isolate of Venezuela: Romero-Figeroa 1997: 68), Mapuche (possibly Southern Andean: Smeets 1989: 132), Cantonese (Sino-Tibetan: Matthews and Yip 1994: 324), Godoberi (Daghestanian: Kibrik 1996: 38), Jaqaru (Jaqi family: Hardman 2000: 33), Epena Pedee (Choco family of Colombia: Harms 1994: 122), Bagirmi (Central Sudanic: Stevenson 1969: 41), and several others.

The case of Nivkh (Palaeosiberian: Gruzdeva 1998: 46)) is rather interesting in this context. It uses *either* an interrogative intonation *or* an interrogative particle to denote interrogativity in its constituent questions. It has two interrogative particles, *-na* and *-ata/-at*, of which one is usually used in constituent questions. However, the particle may be omitted when the speaker uses a special question intonation.

- (15) a. *ap* *p'ry-d'-at*
 who come-Finite-Int
 ‘Who came?’
- b. *či* *sid'-na* *j-yrsu-d'* (uttered with question intonation)
 you who-Int Obj-pursue-Finite
 ‘Who do you pursue?’

10.2.3 *Use of interrogative mood*

Languages may use special interrogative verbs along with interrogative pronouns in their constituent questions. We may regard these verbs as developing from earlier verb + interrogative particle constructions through grammaticalization. In Koasati (Muscogean: Kimball 1991: 301, 424), for example, interrogative–indefinite pronouns give interrogative meaning only when they occur in a sentence that contains an interrogative verb. These interrogative verbs are formed by inserting a glottal stop between the penultimate and ultimate syllables of a verbal form and by accenting each of the flanking syllables. If there is a consonant cluster between the syllables, the glottal stop is deleted and only the accentuation indicates interrogation.

- (16) a. *ná:s-on cíbáʔná*
 what-Obj want (Int)
 ‘What do you want?’
 b. *ná:s-on ca-sobáy-kó-hco-k*
 something-Obj 1Sg-know-3Neg-Habitual
 ‘I know nothing’

According to Miyaoka (1996: 340), Central Alaskan Yupik (Eskimo-Aleut) uses a verb in the interrogative mood, along with interrogative pronouns, in its constituent questions. Some of the interrogative pronouns function like indefinite pronouns or adverbs when used with a non-interrogative verb. The verbs of Tauya (Papuan: MacDonald 1990: 208) also have a distinct interrogative mood that is used in both polar questions (suffix *-nae*) and constituent questions (suffix *-ne*).

- (17) a. *ne-ni ʔati na-fe-a-nae*
 3Sg-Erg say 2Sg-Transitive-3Sg-Int
 ‘Did he tell you?’
 b. *we fofe-ʔe-ne*
 who come-3Sg-Fut-Int
 ‘Who came?’

Cubeo (Tucanoan: Morse and Maxwell 1999: 20, 144) is another language that shows a modal distinction between indicatives, interrogatives, and imperatives in the case of its verbs. The interrogative mood is used in both content questions as well as polar questions.

10.2.4 *Derived interrogatives*

Instead of allowing the verb to attract the interrogative particle and later on grammaticalizing the verb + interrogative particle construction into an interrogative verb, languages may allow their interrogative pronouns to attract the

interrogative particle, and thereby have 'derived' interrogative pronouns. For example, Southeastern Pomo (Pomoan, Hokan: Moshinsky 1974: 105) has a set of stems that can occur with the interrogative suffix *-ʔe* to form question words, and with a different suffix, namely *-do*, to form a 'quotative'.

(18)	Interrogative	Quotative
'what'	<i>ʔawi-ʔe</i>	<i>ʔawi-do</i>
'how many'	<i>bsin-ʔe</i>	<i>bsin-do</i>
'when'	<i>btey-ʔe</i>	<i>btey-do</i>
'who'	<i>ca-ʔe</i>	<i>ca-do</i>
'which one (inanimate)'	<i>hel-ʔe</i>	<i>hel-do</i>
'which one (animate)'	<i>hiy-ʔe</i>	<i>hiy-do</i>
'where'	<i>hey-ʔe</i>	<i>hey-do</i>

Yaqui (Uto-Aztecan: Lindenfeld 1974: 33) attaches the interrogative particle *sa* to interrogative pronouns when they occur in constituent questions.

- (19) *hita-sa kari weče-ka*
 which-Int house fall-realized
 'Which house fell down?'

Dedrick and Casad (1999: 90) describe the occurrence of similar structures in Sonora Yaqui. Another language of this type is Assuriní (a Tupí-Garaní language), which, according to Brandon and Seki (1984: 89), has an interrogative particle *pa* that accompanies the interrogative word in constituent questions. Lele, a Chadic language (Burquest 1986: 100), also has interrogative pronouns like *kóngó* 'which (M)' and *tóngó* 'which (F)' that appear to derive from the corresponding indefinite pronouns *kóng* (MSg) and *tóng* (FSg) in a similar fashion. In Warao, a language isolate of Venezuela (Romero-Figeroa 1997: 68), interrogative pronouns always co-occur with the verbal suffix *-ra*, which is said to be a general question marker. In copulative sentences, this particle is attached directly to the interrogative pronoun.

- (20) a. *kasikaha ihi ribuae-ra*
 what you say-Int
 'What did you say?'
 b. *kasikaha-ra tatuma-mo*
 What-Int these-Abl
 'From these, what is it?'

According to Gair and Paolillo (1997: 21), interrogative pronouns in Sinhalese (Indo-Aryan) virtually always co-occur with a question marker *də*. It may occur either adjacent to the interrogative pronoun or following the verb or non-verbal predicate. The marker occurs immediately after the interrogative pronoun when the questions occur as focus sentences. Gair and Paolillo point out that the association of interrogative words with focusing is so strong that in some

dictionaries interrogative + *də* have been listed as units: *koheedə* ‘where’, *mokakdə* ‘what’, etc. Nivkh (Palaeosiberian) is also reported to attach the question particle either to the finite verb or to some other word that is emphasized in the question. Gruzdeva (1998: 46) points out that, in most cases, it is the question word that is emphasized and therefore the question particle gets attached to it (see (15a–b) given earlier). Ngiyambaa (Australian: Donaldson 1980: 148, 262), on the other hand, has two different ‘knowledge’ clitics, one of which must occur with question words (called ‘indeterminates’). The clitic *-wa:* is used with them in constituent questions, whereas the clitic *-ga:* is used with them when lack of knowledge is to be specified.

- (21) a. *minja-wa:* = *ndu dhayi*
 what-Int = you ate
 ‘What did you eat?’
 b. *minja-ga:* = *ndu dhayi*
 what-Ignorance = you ate
 ‘I don’t know what you ate’

10.3 Restricting the request to a constituent

Constituent questions are different from polar questions in that they not only make a request for information, but also restrict that request to a particular constituent, namely the one that is being indicated as involving lack of information (see §10.1.3). Languages that differentiate between indefinite and interrogative pronouns can apparently use the latter for denoting that the question is only about that particular constituent. However, languages that do not differentiate between the two may use additional devices for indicating that the request for information concerns a single constituent. Shifting the pronoun to the focus position is one such device that is used by several languages. This focus position may be the sentence-initial position (§10.3.1) or some other one like the preverbal position (§10.3.2). Languages may also use focus particles, or focus constructions for a similar purpose (§10.3.3). Several languages that use the same set of pronouns in both constituent questions as well as other types of sentences make it obligatory to use focused constructions only in the former case but not in the latter case.

10.3.1 *Shifting the pronoun to the sentence-initial position*

The device that is used most commonly by languages for restricting a given request for information to a particular constituent is moving that constituent to the sentence-initial position. In English, for example, the pronouns of constituent questions (*wh*-words) are generally shifted to the sentence-initial position. This

apparently has the effect of placing the pronoun in focus and thereby restricting the request for information (interrogation) to that constituent.

Haspelmath (1997: 170) points out that the interrogative pronouns of Classical Greek (22a–b) and Modern German (23a–b) can also be used as indefinite pronouns but they are usually clause-initial in their interrogative use whereas in their indefinite use they must cliticize to the preceding word and hence cannot be clause-initial.

- (22) a. *tís ēlthen?*
 who came
 ‘Who came?’
 b. *ēlthén tis*
 came who
 ‘Someone came’
- (23) a. *Wer kommt da?*
 who comes there
 ‘Who is coming?’
 b. *Da kommt wer*
 there comes who
 ‘Someone is coming’

In several Australian languages also, the same set of pronouns is used as interrogatives and indefinites. The two uses may be differentiated, however, by the fact that the pronouns shift to the sentence-initial position in the interrogative use. According to Austin (1981: 151), for example, Diyari has a single set of pronouns functioning as interrogatives and indefinites. As interrogatives, the pronouns must occur in the sentence-initial position. If they occur in the non-initial position, they are to be interpreted as indefinites. This is also true of Bagandji, another Australian language (Hercus 1982: 171), in which interrogatives always occur in the sentence-initial position. They may convey an indefinite meaning when used as part of a sentential complement, i.e. when they have a non-initial usage. The interrogatives of Yindjibarndi (Australian: Wordrick 1982: 164) are also reported to be always appearing in the initial position of the sentence.

Some of the Uto-Aztecan languages also show this characteristic of shifting the interrogative pronoun to the sentence-initial position. According to Miller (1996: 699), for example, constituent questions are formed in Shoshone by placing the interrogative-indefinite words in the sentence-initial position. The same words are used in indefinite-word sentences, but the word then comes in its usual position in the sentence.

- (24) a. *hakke in puikka*
 who you see
 ‘Who did you see?’

- b. *ni kian hakke puikka*
 I perhaps someone saw
 'I saw someone'

According to Langacker (1977: 51), Uto-Aztec languages typically require the interrogative words to be preposed to initial position. There are some exceptions, however, like Luiseño, where the interrogative word can apparently occur in any position (though it shows a tendency to be preposed). On the other hand, Luiseño differs from Shoshone in using the polar question marker in its constituent questions as well.

Another language that requires the interrogative pronouns to be shifted to the sentence-initial position is Epena Pedee (Choco family: Harms 1994: 122).

10.3.2 *Shifting to other focus positions*

Shifting the pronoun to the sentence-initial position actually has the effect of placing it in focus. Languages may use other positions in the sentence as focus positions and in such languages the pronouns of constituent questions would be shifted to that position rather than to the sentence-initial position. For example, according to Masica (1991: 386), most of the Indo-Aryan languages have the interrogative pronoun normally occurring in the position right before the verb. This position is also considered to be the focus position in these languages. However, when the interrogative pronoun occurs as an adjective, it would occupy the preverbal position along with the noun that it modifies.

In Dravidian languages also, the focus position in a sentence is preverbal, and interrogative pronouns are generally placed in that position. However, the languages also use focus constructions for this purpose, as I point out in the next section. Harrar Oromo of north-eastern Ethiopia (Owens 1985: 207) is another language that shows this tendency of placing the question word close to the verb. According to Saltarelli (1988: 5, 7), question words occur in Basque (a language isolate) obligatorily in the focus position, which is immediately before the synthetic verb or before the main verb of a periphrastic form. While any element of a statement can be focused, in the case of a question, the focus must be on the question word itself.

10.3.3 *Use of focus particles*

Another device that is used for restricting the request for information to a particular constituent in constituent questions is the addition of a focus particle. In Mangarayi, an Australian language (Merlan 1989: 6), for example, interrogative pronouns, which tend to be the first element in constituent questions, generally occur with the focus clitic *bayi*, or its reduced form *ba*, when clause-initial.

- (25) *janaŋgari-ba ja-yag*
 where-Focus 3Sg-go
 ‘Where is she going?’

These interrogative pronouns can also be used for denoting indefinite meaning, as for example with the prohibitive particle *ŋiñjag*. They do not take the clitic *ba* in such indefinite uses (Merlan 1989: 119).

- (26) *ŋiñjag ŋiñja ŋiŋa-m*
 Prohibitive who arrive(Past)-Neg
 ‘Nobody arrived’

In Taba (eastern Polynesian: Bowden 1997: 412) also, the focus marker *e* is almost invariably attached to the interrogative pronoun in constituent questions. In Trumai (language isolate: Guirardello 1999: 34), on the other hand, interrogatives always occur in the sentence-initial position, but many times they are followed by the particle *in* ‘focus’.

10.3.4 Use of focus constructions

Dravidian languages generally show a tendency to use the focus construction in their constituent questions. This construction has the effect of bringing one of the constituents of the sentence into focus, and in the case of constituent questions, it would be the questioned element that is brought into focus. The languages use this construction in their polar questions as well, when the question has to be restricted to a single constituent. In such cases, the polar question particle would be shifted to that particular constituent. The following sentences of Havyaka Kannada (Dravidian) exemplify these two uses.

- (27) a. *ni: ellige ho:d-du*
 you where went-it
 ‘Where is it that you went?’ (‘Where did you go?’)
 b. *ni: mane-g-o: ho:d-du*
 you home-Dat-Int went-it
 ‘Is it to the house that you went?’
 c. *ni: manege ho:yidey-o:*
 you home-Dat went-Int
 ‘Did you go home?’

In both (27a) and (27b), the sentences involve a focus construction in which the verb is changed into its nominalized form. This focus construction has the effect of restricting the question to a particular constituent. In (27a) the question gets restricted to the constituent that is indicated as unknown. In (27b), on the other hand, there is no unknown element as such, but the constituent to which the question is restricted is specified by attaching the polar question particle *o:*,

which otherwise occurs in the sentence-final position as in (27c), to the focused element.

Another device that is used by several languages in order to place the questioned constituent of constituent questions in focus is to employ a clefted construction or a relative clause construction. According to Hewitt (1979: 10), for example, constituent questions occurring in Abkhaz (north-west Caucasian) require the verb to be in the non-finite (relativized, pseudo-clefted) form. The verb takes the suffix *da* if the question is about humans and *y* (or *zə-y* or *ze-y*) otherwise.

- (28) a. *yɔ́-q'a-da*
 who-be-Int
 'Who is it?'
 b. *də-z-ca-x'è-y*
 he-why-go-Perfect-Int
 'Why has he already gone?'

Constituent questions formed on noun phrases are also pseudo-clefted either explicitly (i.e. the construction involving the interrogative pronoun is used with a relativized verb), or implicitly (i.e. the relativized verb incorporates the interrogative particle *da* 'who') in Abkhaz (Hewitt 1979: 21).

In Tamazight (Afro-Asiatic, Berber: Penchoen 1973: 79), all constituent questions are clearly cleft constructions. They are basically of the same structure as non-verbal identity sentences. Masai (Nilo-Saharan, Nilotic: Hollis 1905) is another language that requires its verbs to be relativized, if the interrogative is the subject. This is also true of Kanuri (Nilo-Saharan, Saharan: Lucas 1937: 32), which requires its constituent questions that have the verb denoting a positive past happening to be put into relative past form. If the predicate of such sentences is not a verb, but any other predicative expression, the particle *go* (which is what is left of an obsolete verb 'to be') is generally added to it. According to Aikhenvald (1998: 261), Warekena (northern subgroup of Maipuran) marks the predicate of constituent questions with the relative marker *li* when the questioned constituent is the subject or object.

10.4 Combining the two notions

Languages may combine together the two notions described above, namely interrogation and focus, and use a single device for denoting both of them. In such languages, the device used in constituent questions, namely question intonation or question particle, would be *different* from the one that is used in polar questions.

10.4.1 Use of distinct intonations

According to Schachter and Otnes (1972: 505), constituent questions of Tagalog have their own characteristic intonation patterns, which are different from those of polar questions. In the former case, there is a rise from the mid- position, whereas in the latter case, there is a full fall from high pitch. Similarly in Maori (eastern Polynesian: Bauer 1999: 5), constituent questions are marked by a rise on the question word and a fall sentence-finally. Polar questions, on the other hand, have raised pitch throughout.

In both these cases, we may regard the two types of questions to be indicating interrogativity (or request for information) by their intonations, with the difference that in the case of constituent questions, the intonation would also be specifying the constituent to which the request is restricted. In Maori, for example, this is accomplished by situating the rise of intonation on the questioned constituent.

10.4.2 Use of distinct question particles

There are several languages that differentiate between polar questions and constituent questions through the use of distinct question particles. Many of the Tibeto-Burman languages are reported to be of this type. For example, Angami, belonging to the Naga subgroup of Tibeto-Burman, uses the question particle *gā* in constituent questions, and *mē* in polar questions (Giridhar 1980: 81). As in the previous case, we may regard both these question particles as indicating a request for information, with the difference that the particle *gā* has the additional meaning of restricting that request to a single constituent, namely the one denoted by the interrogative-indefinite pronoun.

- (29) a. *nō kî ra vó gā*
 you where go Int
 ‘Where did you go?’
 b. *nō mhâ tsəliêté mē*
 you eaten Int
 ‘Have you eaten?’

Lahu, another Tibeto-Burman language, is similar to Angami in its use of distinct question particles. It uses the particle *lā* in polar questions and *le* in constituent questions (Matisoff 1973: 371). In Grierson (1903, 1904 and 1908), several other Tibeto-Burman languages like Lhoke (Central Tibetan group), Sunwal (Himalayan), Limbu, Miri (Assam group), Rangkhoh (Old Kuki group), etc. are reported as using distinct question particles in constituent questions and polar questions. For example, Limbu uses the particle *gō* in constituent questions and *bē* in polar questions (1908: 288). Kachin uses *mā* in constituent questions and *ī* or *khā* in polar questions. However, in some languages

like Lepcha (particle *a*) and Rangkhol (particle *mō*) the same particle occurs in both polar questions and constituent questions (Grierson 1904: 185, 1908: 238).

Some of the non-Tibeto-Burman languages that use distinct question markers in their polar and constituent questions are Jaqaru (Aymaran: Hardman 1966: 80), Huallaga Quechua (Weber 1986), Epina Pedee (Choco family: Harms 1994: 114), and Margi (Afro-Asiatic, Chadic: C. Hoffmann 1963: 98). In Margi, for example, the usual interrogative particle for polar questions is *yá*, whereas for constituent questions, it is *rá* (or *rìí* for greater emphasis). There are other particles like *réè* used if the question is spoken louder than usual, especially when talking over a distance, *ḡárí* used in polar questions that express doubt, and *wò* used at the end of negative questions. In the case of Tauya, a Papuan language (MacDonald 1990: 208), on the other hand, a comparable distinction occurs in the interrogative modal form of the verb. Verbs of this language take either the suffix *nae* or *ne* to form the interrogative mood. The former occurs in polar questions and the latter in constituent questions (see sentences (17a–b) given earlier).

10.5 Non-marking and double-marking

The foregoing examination of languages in which the same pronouns or derivationally related ones are used as interrogative and indefinite pronouns reveals that the languages generally possess additional devices in their constituent questions for indicating the two additional meanings that need to be expressed in them. They are (i) interrogation (request for information) and (ii) focus (restriction of that request to a particular constituent). It is not necessary, therefore, to regard these languages as using an ‘interrogative’ pronoun in their constituent questions. It would be sufficient if we regard them as using only indefinite pronouns in those sentences.

There do occur some languages, however, in which such additional devices are lacking, and as a result the use of the same set of pronouns in both statements and constituent questions gives rise to ambiguity. For example, according to J. Hoffmann (1903), interrogative pronouns of Mundari, an Austro-Asiatic language (Munda family), can be used either as interrogatives or as indefinites. The word *cikan* can mean either ‘to do what to oneself’ or ‘to do something to oneself’. Similarly, *oko* can mean either ‘what, which’ or ‘some, any’. Cook (1966: 339) gives sentences of the following type to illustrate this double usage.

- (30) *oko kami menai*
 what work is
 (i) ‘What work is there?’
 (ii) ‘There is some work’

- (31) *oko horólóm jagarkena*
 what man speak
 (i) 'With whom did you speak?'
 (ii) 'You spoke with someone'

There is no clue, according to Cook (1966: 234), even in the intonation pattern of these sentences, to signal a question.

Similar ambiguity is reported to occur in some of the Australian languages as well. According to Dixon (1980: 372), for example, it is possible to regard interrogative-indefinite pronouns in the Australian languages in general to be indicating, simultaneously, an indefinite specification and a request for further information. Notice, however, that in several Australian languages interrogative use is differentiated from indefinite use by moving the pronoun to the sentence-initial position (see §10.3.1). Some also possess special indefinite markers that can be used to differentiate between the two uses. Tiwi (Osborne 1974: 56), for example, has *kuwani* 'who' contrasting with *aramu-kuwani* 'someone' and *kamini* 'what' contrasting with *aramu-kamini* 'something'.

There is no need, on the other hand, for languages that differentiate between interrogative and indefinite pronouns with the help of distinct sets of pronouns, to use any additional devices in their constituent questions in order to express the meaning of interrogation and focus. Their use of a distinct set of interrogative pronouns would be sufficient for indicating these additional meanings as well. Thus, in Manam, an Austronesian language of the Oceanic subgroup, interrogative pronouns show no affinity with indefinite pronouns. Polar questions of this language are marked by a rise in intonation, but constituent questions are not; the intonation of the latter follows the same pattern as that of declarative sentences. Further, interrogative pronouns occupy the same sentence position as their corresponding counterparts in declarative sentences (Lichtenberk 1983: 391, 398).

Another language of this type is Lango, a Western Nilotic (Nilo-Saharan) language. According to Noonan (1992: 166), indefinite pronouns do not, *per se*, occur in this language. In their place, generic nouns like *ḡat* 'person' (for 'someone') and *ḡinn-ḡrḡ* 'thing' (for 'something') may be used. Polar questions of this language are distinguished from their corresponding declarative sentences by intonation, but constituent questions are marked only by interrogative words. They otherwise possess no syntactic properties that distinguish them from their corresponding declaratives. Interrogative words also occur in whatever slot that is appropriate to their grammatical use (Noonan 1992: 242).

There are, however, some languages that use some of the above-mentioned devices in their constituent questions in spite of their having distinct interrogative pronouns. For example, Trumai, a language isolate spoken in Brazil, does not appear to show any affinity between interrogative and indefinite words. Its indefinite words are derived from generic words meaning 'one' or 'people'. However, its interrogative words always occur in the first position of constituent

questions and are frequently followed by the focus marker *in* (or by the morpheme *iyi* followed by *in*) (Guirardello 1999: 35). Supyire (Niger-Congo, Gur family: Carlson 1994: 533) also differentiates between interrogative pronouns and indefinite pronouns, but still it uses sentence-initial question particles like *ye* along with interrogative pronouns in constituent questions. Its locative questions may take the sentence-final particle *k*. Carlson points out further that the majority of constituent questions in Supyire are in the form of cleft-focus constructions. The question word is fronted to the focus position at the head of the clause. We may regard these languages as double marking the relevant meanings.

I believe that the occurrence of a few 'non-marking' languages that leave the notion of interrogation or focus unspecified in constituent questions, in spite of their not having a distinct set of interrogative pronouns, does not discredit the claim of the present chapter. As pointed out in §10.1.3, it is necessary to regard such languages as using a single set of pronouns in both these types of sentences, and since the meaning that is common to both these uses is 'lack of information', we may regard them as indefinites rather than as interrogatives. They apparently allow the context to indicate that the expression of one's lack of knowledge is to be taken as a request (or demand) for information in some of their uses. Dixon (1977: 372) suggests this to be the possibility in the case of Australian languages.

On the other hand, there is no need to regard languages like Trumai and Supyire, which double-mark the notion of interrogation, that is, both with the help of a distinction among their pronouns and also through the use of distinct intonation or question particle, as discrediting the present claim because double-marking is a frequently employed device among languages. What I am suggesting here is that regarding languages of the former type also as double-marking the notion of interrogation unnecessarily leads to puzzles and paradoxes and is therefore to be avoided.

10.6 Summary

I have argued in this chapter that there is no *need* to regard languages that show affinity (identity or derivational relationship) between interrogative and indefinite pronouns as using any interrogative pronouns in their constituent questions. Instead, they can be regarded as having only unmarked indefinite pronouns in those sentences. The purpose of using these pronouns in such sentences is merely to indicate that the speaker lacks knowledge regarding a particular constituent. There are two other meanings that need to be expressed in constituent questions, namely (i) a request for information (interrogation) and (ii) restriction of that request to a particular constituent (namely the indefinite pronoun); these meanings are generally expressed, in these languages, with the help of additional devices; for example, devices like the use of question particles or question intonation are used for denoting interrogation, whereas devices like the use of

focus particles or focus constructions are used for denoting that the interrogation is restricted to a particular constituent.

I have suggested further that such an analysis would help us to resolve the puzzle regarding the affinity between interrogative and indefinite pronouns. According to this analysis, the affinity rightly represents the meaning that is common to both interrogatives and indefinites, namely the denotation of lack of knowledge regarding a particular constituent. On the other hand, the assumption that constituent questions contain an ‘interrogative’ pronoun in these languages is incapable of resolving this puzzle.

11 Other Related Puzzles

11.1 Introduction

I suggested in the previous chapter that the puzzle concerning the affinity between interrogative and indefinite pronouns, occurring in the majority of world's languages, can be resolved by regarding both constituent questions as well as other sentences that contain such pronouns as having only *indefinite* pronouns and not any interrogative pronouns as such. The meaning of interrogativity as well as that of focus is expressed in such languages not with the help of any pronouns but rather with the help of alternative devices such as the use of interrogative particles, interrogative intonation, interrogative verb, focus particles, focus constructions, etc. I wish to point out in this chapter that such an analysis has certain additional advantages, namely that it allows us to resolve certain other puzzles that are related to the above one: (i) the puzzle about the derivation of indefinite pronouns (§11.2), (ii) the puzzle about indirect questions (§11.3), and (iii) the puzzle about the affinity between interrogative and relative pronouns (§11.4).

11.2 Derivation of indefinite pronouns

As I mentioned in the previous chapter, grammarians generally describe the affinity between interrogative and indefinite pronouns as involving the derivation of indefinite pronouns from interrogative ones. My claim that constituent questions occurring in languages that show such affinity contain only *indefinite* pronouns and not any interrogative pronouns in their constituent questions makes it necessary to re-examine the nature of derivations that are involved among the indefinite pronouns of these languages.

There are two main types of derivation of indefinite pronouns that have been reported to occur among these languages. They involve (i) the use of a disjunctive particle to provide specific indefinite meaning and (ii) the use of a conjunctive particle to provide non-specific indefinite meaning. Havyaka Kannada (Dravidian) exemplifies both these types of derivations (1a–c).

- (1) a. *allige a:ru ho:-ku*
there who go-Subjunctive
'Who would go there?'

- b. *allige a:r-o: ho:-ku*
 there who-or go-Subjunctive
 'Somebody would go there'
- c. *allige a:ru-de ho:-ku*
 there who-also go-Subjunctive
 'Anybody would go there'

The pronoun *a:ru* 'who' occurs by itself in (1a), and the sentence is a constituent question. In (1b), the same pronoun occurs with the disjunctive particle *o:* 'or', but it has the specific indefinite meaning. In (1c) also, it occurs with the conjunctive particle *de* and has the non-specific indefinite meaning. Both (1b) and (1c), unlike (1a), are statements and not questions.

The term 'derivation' is rather misleading in both these instances because the particles are attached to the noun phrase that contains the indefinite pronoun and not directly to the pronoun. For example, they follow the plural and case markers and also the nouns that are modified by the pronominal words.

- (2) a. *avā ya:va mane-ge ho:-ku*
 he which house-Dat go-Subjunctive
 'Which house might he go?'
- b. *avā ya:va mane-g-o: ho:-ku*
 he which house-Dat-or go-Subjunctive
 'He might go to some house'
- c. *avā ya:va mane-g-ude ho:-ku*
 he which house-Dat-also go-Subjunctive
 'He might go to any house'

In view of our analysis of constituent questions like (2a) as containing indefinite pronouns rather than interrogative pronouns, non-interrogative sentences like (2b–c) can be regarded as *continuing* the use of indefinite pronouns in certain additional constructions in which they are associated with disjunctive and conjunctive particles. There is no need to regard them as involving the 'derivation' of indefinite pronouns. This is also true of other Dravidian languages and also several other languages in which such derivations have been reported to occur. The 'indefinite' marker occurs as a clitic or even as an independent word rather than as an affix in several of them (see §11.2.1)

Grammarians generally derive indefinite pronouns from interrogative pronouns through the addition of disjunctive or conjunctive particles. As I have pointed out in the previous chapter (§10.1.1), Haspelmath (1997: 163) questions the correctness of this derivation on the basis of the fact that there cannot be any explanation for a derivational process that changes interrogatives into indefinites through the addition of disjunctive or conjunctive particles. How can such an addition change 'a request for information' into 'a statement about lack of knowledge'? His suggestion is, therefore, to regard indefinite pronouns as

derivationally *unrelated* to interrogatives (even though the two are formally similar or identical) and to derive affixed indefinites from un-affixed indefinites.

This derivation of specific and non-specific indefinite pronouns from underlying (unmarked) indefinite bases is easier to explain because, as pointed out by Haspelmath (1997: 165), there is a close relationship between universal quantification and logical conjunction on the one hand, and between existential quantification and logical disjunction on the other. This relationship has long been recognized by logicians. For a finite number of entities x , a universal statement is equivalent to conjunction, and an existential statement is equivalent to disjunction:

- (3) a. $(\forall x) f(x) \equiv f(x_1) \ \& \ f(x_2) \ \dots \ \& \ f(x_n)$
 b. $\exists (x) f(x) \equiv f(x_1) \ \vee \ f(x_2) \ \dots \ \vee \ f(x_n)$

The derivation that is being suggested here is to regard indefinite pronouns occurring in constituent questions (i.e. pronouns that are generally regarded as ‘interrogative pronouns’) as unmarked indefinite pronouns. The addition of a conjunctive particle to such pronouns has the effect of making them non-specific indefinite pronouns or universal pronouns, having meanings like ‘anyone’ or ‘everyone’. On the other hand, the addition of a disjunctive particle to the same has the effect of making them specific indefinite pronouns, having meanings like ‘someone’. Some languages use unmarked indefinite pronouns in all these meanings, whereas some restrict the use of such unmarked pronouns to constituent questions and attach these and other affixes or particles for using them in other contexts. There are also languages that attach interrogative particles to such pronouns in order to use them in constituent questions (see §10.2.4).

This possibility of regarding specific and non-specific indefinite meanings as resulting from the interaction of indefinite pronouns with disjunctive and conjunctive meanings respectively is rejected by Haspelmath (1997: 166) mainly because he finds too many natural language instances to be going against that relationship. It appears, however, that these apparently contradictory instances occur primarily in the case of languages in which *no distinction* is made between specific and non-specific indefinites. Languages that do make a distinction between the two appear to be generally consistent in deriving non-specific indefinites with the help of conjunctive particles and specific indefinites with the help of disjunctive particles.

11.2.1 *Differentiating languages*

Most of the Dravidian languages differentiate between specific and non-specific indefinites. They are also consistent in associating the specific indefinite meaning with disjunctive particles and non-specific indefinite meaning with conjunctive particles. For example, interrogative pronouns of Kannada are consistently used (i) with the particle *u*: ‘also’ in order to denote non-specific indefiniteness and

(ii) with the particle *o*: 'or' in order to denote specific indefiniteness. The following pairs of sentences exemplify this contrast.

- (4) a. *ninne illige ya:r-u: bar-al-illa*
 yesterday here who-also come-Inf-Neg
 'Nobody came here yesterday'
- b. *ninne illige ya:r-o: bar-al-illa*
 yesterday here who-or come-Inf-Neg
 'Somebody did not come here yesterday'
- (5) a. *avanu e:n-u: ma:ḍa-bahudu*
 he what-also do-may
 'He may do anything'
- b. *avanu e:n-o: ma:ḍa-bahudu*
 he what-or do-may
 'He may do something'

Similar correlations occur in other Dravidian languages also, as shown in (6):

(6)	Conjunctive particle	Disjunctive particle
Tamil	<i>enkey-um</i> 'anywhere'	<i>enkey-o</i> : 'somewhere'
Malayalam	<i>ent-um</i> 'anything'	<i>ent-o</i> : 'something'
Telugu	<i>ev-ar-u</i> : 'anybody'	<i>ev-ar-o</i> : 'somebody'
Tulu	<i>e:r-la</i> : 'anybody'	<i>e:r-o</i> : 'somebody'
Toda	<i>o:y-um</i> 'anybody'	<i>or-istry</i> 'somebody'
Malto	<i>nere-goṭe</i> 'anyone'	<i>nere-beḍi</i> 'someone'

In Tulu, *la*: denotes 'also' and *o*: denotes 'and'. Droeze (1884) reports that, in Malto, the word *goṭe* has the meaning 'even, also'. It is possible that there is a comparable correlation in the remaining Dravidian languages as well, but unfortunately, the published grammars provide very little information regarding the formation of indefinite pronouns.

Similar correlations are reported to occur in some of the Indo-Aryan languages as well. For example, according to Kakati (1972), Assamese (Indo-Aryan, spoken in Assam) adds the suffix *o* 'also' (derived from Sanskrit *api*) and *ba* 'or' (derived from Sanskrit *va*:) to form non-specific and specific indefinites, respectively. However, the disjunctive suffix *ba* 'or' follows the conjunctive suffix *o* 'also'.

- (7) *kon* 'who' *kon-o* 'anyone'
kon-o-bo 'someone'

Haspelmath (1997: 167) mentions the occurrence of a similar derivation in Nanay, a Tungusic language (based on Onenko 1980) and Yakut, a Turkic language (based on Ubrjatova 1982). In Nanay, for example, interrogative pronouns take the marker *daa* 'also' to provide non-specific indefinite meaning and the

marker *nuu* 'or' to provide specific indefinite meaning. Japanese (Martin 1975: 1073, Hinds 1986) is another language in which a correlation exists between the conjunctive-disjunctive distinction on the one hand, and the non-specific-specific indefiniteness on the other. Interrogative pronouns like *dare* 'who' can take the particle *mo* 'and' to denote 'anyone' or 'everyone' and the particle *ka* 'or' to denote 'someone'.

Interestingly, the Japanese particle *ka* 'or' also occurs (optionally) as a question marker in both sentence questions and constituent questions. Similar usages of disjunctive particles occur, especially in polar questions, in certain other languages like Kannada (Dravidian), Karok (Hokan: Bright 1957), Usan (Papuan: Reesink 1987: 293), and Khmer (Austroasiatic, Mon-Khmer: Gorgoniyev 1966) as well. On the other hand, Manipuri (Tibeto-Burman) uses its sentence question particle *no* (which, however, is *not* used as a disjunctive suffix) for deriving specific indefinite pronouns from interrogative pronouns. Manipuri also uses the particle *su* 'also' for deriving non-specific indefinite pronouns from interrogative pronouns (Bhat and Ningomba 1997: 79).

- (8) *kəna* 'who' *kəna-no* 'someone'
 kəna-su 'anyone'
kəri 'what' *kəri-no* 'something'
 kəri-su 'anything'

There are some additional languages, which, even though not making a distinction between specific and non-specific indefinites through derivation, do make use of the conjunctive particle for denoting non-specific indefinite meaning or the disjunctive particle for denoting specific indefinite meaning. For example, Tarma Quechua, spoken in the Andes of central Peru, uses its bare interrogatives for denoting specific indefiniteness, whereas for denoting non-specific indefiniteness, it adds the conjunctive suffix *si* or *bi* meaning 'even', 'also', or 'too' to them (Adelaar 1977: 254). In another Peruvian language, Jaqaru, the addition of *psa* 'also, and' is reported to change interrogative words into non-specific indefinite words (Hardman 1966).

- (9) *kawi* 'where' *kaw.psa* 'anywhere, nowhere'
 qači 'who' *qač.psa* 'anybody, nobody'

To'abaita (Oceanic: Simons 1986) is another language that is reported to use *bana* 'just, only' with interrogative pronouns in order to derive non-specific indefinite pronouns: *taa* 'what': *taa bana* 'whatever, anything'; *tei* 'who': *nitei bana* 'whoever, anyone'. Tagalog (of Philippines) is also reported to change most of its interrogative words into indefinite expressions, usually equivalent to English expressions that involve *any*, *no*, or *ever* (i.e. non-specific indefinites), by prefixing *kahit(na)* 'even' (Schachter and Otnes 1972: 505). Similarly, Sonora Yaqui (Uto-Aztecan) changes its positive (specific) indefinite pronouns (which

also occur in constituent questions) into expressions that denote ‘any-’ or ‘-ever’ by attaching *húni* ‘even’ to them (Dedrick and Casad 1999: 249).

- (10) *hábe* ‘someone’ *hábe húni* ‘anyone, whoever’
híta ‘a thing’ *híta húni* ‘however’

On the other hand, Maori (Austronesian, Oceanic: Bauer 1999: 372) is reported to use a disjunctive word, namely *rannei* ‘or’, with its interrogative pronouns in order to indicate specific indefinite meaning, as seen in (11).

- (11) *hea* ‘where’ *hea rannei* ‘somewhere’
aha ‘what’ *aha rannei* ‘something’

The claim that these derivations of specific and non-specific indefinites involve the addition of disjunctive and conjunctive particles respectively to an unmarked indefinite pronoun is supported by the occurrence of languages in which such derivations are reported to involve indefinite pronouns that are not related to interrogative pronouns. In Garo (Tibeto-Burman: Burling 1961: 40), for example, the suffix *-ba* ‘also, too’ can be attached to indefinite pronouns like *mamun* ‘something’, which are not related to interrogative pronouns, to provide forms like *mamunba* ‘anything’.

11.2.2 Non-differentiating languages

There are several languages, on the other hand, that do not represent the distinction between specific and non-specific indefinites through a distinction in the indefinite paradigms occurring in them. In such languages, either the disjunctive or the conjunctive particle (or both) may be used for deriving pronouns that have specific or non-specific indefinite connotation. For example, Margi, a Chadic language, derives indefinite pronouns from interrogative pronouns by adding either the conjunctive particle *yé* or the disjunctive particle *kó* ‘or’ or both (C. Hoffmann 1963). All of them appear to have only a non-specific indefinite meaning.

- (12) *màrì* ‘which’ *màryè* ‘any, every’
mì ‘what’ *kómì* ‘anything’
kómìyé ‘anything’

Kanuri (Nilo-Saharan) is another language in which indefinites can be derived from interrogatives either by using the collective suffix *si* or the disjunctive particle *yaye* ‘or’ (Lucas 1937: 32). These derived forms appear to have non-specific indefinite meaning in both the cases.

- (13) *aví* ‘what’ *aví-so* ‘whatever’
ndú ‘who’ *ndú-so* ‘everyone’
ndú-yàye ‘whosoever’

In the older stages of Sanskrit, interrogative pronouns provided the indefinite sense by themselves, but in the later stages, they required the addition of particles

like *ca* 'and' or *api* 'also' for this purpose. However, there is no differentiation between specific and non-specific indefinite meanings; forms like *kim-api* or *kim-cid* can denote 'something' as well as 'anything'. Most of the modern Indo-Aryan languages follow Sanskrit on this point. They generally possess only a single set of indefinite pronouns, which can, however, be used both in specific and non-specific meanings. In Bengali, for example, indefinite pronouns are derived by adding the emphatic marker *o* (which is diachronically derivable from Sanskrit *api* 'also') to interrogative pronouns and the derived form has both specific as well as non-specific meaning (K. K. Rarhi, personal communication).

- (14)
- | | | | |
|---------------|---------|----------------|-----------------------|
| <i>ke</i> | ‘who’ | <i>keu</i> | ‘someone, anyone’ |
| <i>kəkhon</i> | ‘when’ | <i>kəkhono</i> | ‘sometime, anytime’ |
| <i>kothay</i> | ‘where’ | <i>kothaw</i> | ‘somewhere, anywhere’ |
- (15) a. *ram kotha-w gie che*
Ram where-Emph gone has
‘Ram has gone somewhere’
- b. *ram kotha-w jay ni*
Ram where-Emph gone not
‘Ram has not gone anywhere’
- (16) *se kich-u pay-ni*
he what-Emph got-not
(i) ‘He did not get anything’
(ii) ‘He did not get something’

There are, however, other devices in Bengali that can be used for differentiating between specific and non-specific meanings. For example, one can place the word *kono* 'which' before *kichu* 'what' occurring in a sentence like (16) given above in order to provide, unambiguously, the non-specific meaning as in (17).

- (17) *se kon-o kich-u pay-ni*
 he which-Emph what-Emph get-not
 'He did not get anything'

Similar ambiguity between specific and non-specific meanings is reported to occur in the use of indefinite pronouns in several other Indo-Aryan languages, as shown in (18).

- | | | | | | |
|------|----------|-------------|-------|-------------|-------------------|
| (18) | Gujarati | <i>kəṇ</i> | ‘who’ | <i>koi</i> | ‘some, any’ |
| | Kashmiri | <i>kus</i> | ‘who’ | <i>kāh</i> | ‘someone, anyone’ |
| | Sindhi | <i>kēru</i> | ‘who’ | <i>ko</i> | ‘someone, anyone’ |
| | Maithili | <i>ke</i> | ‘who’ | <i>keo</i> | ‘someone, anyone’ |
| | Konkani | <i>kōṇ</i> | ‘who’ | <i>kōṇi</i> | ‘someone, anyone’ |
| | Gojri | <i>kun</i> | ‘who’ | <i>koe</i> | ‘some, any (M)’ |
| | | | | <i>kae</i> | ‘some, any (F)’ |

Some of the Tibeto-Burman languages are also reported to derive indefinite pronouns that do not differentiate between specific and non-specific meanings from interrogatives in a similar fashion, but such languages generally occur adjacent to Indo-Aryan languages like Assamese and Bengali or are otherwise influenced by them. For example, Garo, spoken in Assam (Tibeto-Burman), adds the suffix *-ba* ‘also, too’ to interrogative pronouns in order to derive indefinite pronouns that may have specific or non-specific usage (Burling 1961).

- (19) *sawa* ‘who’ *sao-ba* ‘somebody’
 mai ‘what’ *mai-ba* ‘something’
 je ‘which’ *je-ba* ‘something, anything’

Grierson (1908: 185, 367) refers to the use of the conjunctive particle *sō* ‘and, also’ for deriving indefinite pronouns from interrogatives in Khambu (*āsē* ‘who’, *ā-sō* ‘anybody’), and the particle *yāng* ‘even, ever’ in Gurung for the same purpose (*sū* ‘who’, *sū-yāng* ‘anybody’). Tibeto-Burman languages that do not show this Indo-Aryan influence, on the other hand, generally use ‘interrogative pronouns’ by themselves as indefinites (or have a different set of indefinite pronouns).

11.2.3 Possible counter-examples

Haspelmath (1997: 166) refers to the occurrence of certain languages that appear to contradict the above-mentioned generalization, namely that the addition of conjunctive particle provides non-specific indefinite meaning, and the addition of a disjunctive particle provides specific indefinite meaning. He points out, for example, that in Ossetic (Indo-European, Iranian) the addition of *-dær* ‘also, even’ to the interrogative pronoun provides specific indefinite meaning and that of *-fændy* ‘or’ provides non-specific (free choice) indefinite meaning. This appears to be the opposite of what we expect from (3). However, Ossetic also has another expression, *dæriddær*, in which it appears to attach the former element to interrogative pronouns in order to indicate the same meaning that is denoted by the latter construction (Haspelmath 1997: 281). That is, the language needs to be included in the category of languages that do not differentiate between specific and non-specific indefinites.

On the other hand, the occurrence of *na* ‘or’ or *tunci* ‘or’ with interrogative pronouns in Korean for denoting the notion of ‘free choice’ indefiniteness need not be regarded as conflicting with the above claim because, as Haspelmath pointed out (1997: 313), both these suffixes originally had the meaning ‘whoever it may be’. Their use in the meaning of free-choice indefinites is therefore to be expected. Further, Korean has a different suffix, namely *-nka* that appears to contain the question particle *ka*, which is used with interrogative pronouns to derive indefinite pronouns. Its use is comparable to the use of Japanese *-ka* ‘or, question particle’ mentioned earlier, but the fact that the indefinite pronouns

derived from the use of this suffix may be specific or non-specific makes it necessary to regard Korean as belonging to the non-differentiating set.

There are also a few languages like Portuguese and Romanian that appear partially to contradict the claim made in (3). They do not differentiate between specific and non-specific indefinites and hence they belong to the second set of languages mentioned above. However, they do make use of a disjunctive expression for deriving comparative and free-choice indefinites, that is, indefinites that are non-specific in their meaning. For example, the Romanian expression *ori-ce* 'or-what' provides the non-specific meaning 'anything'. However, as pointed out by Haspelmath (1997: 168), the disjunctive expressions of these languages derive from earlier expressions that had 'want' or 'it may be' as their connotation. For example, Romanian *ori* derives from Latin *velis* 'you want (subjunctive)'. Hence, their use for denoting free-choice indefinites can be regarded as deriving from that particular connotation rather than directly from their disjunctive meaning.

11.2.4 *Distinction between 'even' and 'at least'*

There are some languages that differentiate between the use of an indefinite pronoun with the notion of 'and', 'even', or 'too' on the one hand, and with the notion of 'at least' on the other. The former association provides a universal non-specific indefinite meaning, as described earlier in §11.2.1, whereas the latter association (with 'at least') provides a non-universal, non-specific indefinite meaning. One important difference between these two notions of indefiniteness is that the former can occur in a negative context, whereas the latter cannot. The latter can only occur in irrealis contexts like subjunctive, interrogative, or future. Neither of them, however, can occur in realis contexts.

For example, Kannada (Dravidian) attaches the conjunctive particle directly to interrogative pronouns in order to derive universal non-specific indefinites (see (4a) given earlier in §11.2.1). On the other hand, it attaches the conditional form of the verb *a:gu* 'to become' (namely *a:dare*) followed by the conjunctive particle *u:* 'and' to interrogative pronouns in order to derive non-universal, non-specific indefinites. This latter structure can occur only in irrealis contexts like subjunctive and interrogative but not in negative contexts (see (21a) given below), whereas the former can occur in negative contexts as well (20a). The structure that is derived by using the conditional form of *a:gu* 'become' with the conjunctive particle *u:* 'and' has the meaning 'at least' in Kannada as, for example, in the phrase *maneg-a:dar-u:* 'to the house at least' (*manege* 'to the house').

- (20) a. *avanu ellig-u: ho:gal-illa*
 he where-and go-not
 'He did not go anywhere'

- b. *avanu ellig-u: ho:d-a:nu*
 he where-and go-may
 'He may go anywhere'
- (21) a. **avanu ellig-a:dar-u: ho:gal-illa*
 he where-become(Cond)-and go-not
- b. *avanu ellig-a:dar-u: ho:d-a:nu*
 he where-become(Cond)-and go-may
 'He may go somewhere'

This derivation of non-universal, non-specific indefinites through the association of the notion of 'at least' with that of indefiniteness, I think, is easy to understand. The notion of 'at least', unlike the notion of 'and', 'even', or 'too', assigns a given characteristic to only a *portion* of the possible referents of the expression to which it is attached, as shown in (22a–b).

- (22) a. *Even John will come to the meeting.*
 b. *At least John will come to the meeting.*

In (22a) the use of *even* provides the implication that others would anyway come to the meeting, whereas in (22b) the use of *at least* does not provide any such implication. It only indicates the certainty of John's coming. Similarly in (20a–b) given earlier, the association of the notion of 'and, also' with an unmarked indefinite location allows the characterization to affect all possible locations, whereas in (21b) the association of the notion of 'at least' with the same restricts it to affect only a part of the possible locations.

Both these types of indefinites contrast with specific indefinite in Kannada in that the latter can occur in realis contexts (23a) whereas the former cannot (23b–c).

- (23) a. *avanu ellig-o: ho:da*
 he where-or went
 'He went somewhere'
- b. **avanu ellig-u: ho:da*
 he where-and went
 '*He went anywhere'
- c. **avanu ellig-a:dar-u: ho:da*
 he where-become(Cond)-and went

English uses the pronoun *somewhere* for denoting specific indefinites as well as non-universal non-specific indefinites and hence it is rather difficult to indicate the ungrammaticality of (23c) through English glosses.

Haspelmath (1997: 159) refers to the use of expressions that denote the notion of 'at least' in the derivation of marked indefinites in several languages. There is apparently a distinction among these languages, similar to the one noticed earlier in the use of the two notions 'and, even' and 'or', with unmarked indefinites,

namely between languages that differentiate between non-universal and universal non-specific indefinites on the one hand, and the ones that do not make any such distinction on the other. Kannada clearly belongs to the former set as shown by the examples given above. This is also true of several other Dravidian languages like Tamil, Tulu, Telugu, and Malayalam (see Bhat 1989). On the other hand, most of the languages mentioned by Haspelmath, such as Latvian, Modern Greek, Hungarian, and Lezgian, make no such distinction between the two meanings.

For example, Lezgian has the form *wuž xajit' ani* 'anybody' that contains the expression *xajit' ani* 'at least', and is derivationally similar to the expression *ya:ra:daru:* 'someone' occurring in Kannada, but it has 'free choice' as one of its meanings. This is not true of Kannada. When used in a comparative construction, as seen in (24a), the construction occurring in Lezgian apparently provides the universal non-specific indefinite meaning (Haspelmath 1997: 296), whereas the one occurring in Kannada can only provide a non-universal non-specific indefinite meaning. That is, (24a) of Lezgian appears to provide a superlative meaning but (24b) of Kannada can only provide a comparative meaning. In Lezgian, *xajit'ani* is made up of *xun* 'become, be', *t'a* 'conditional', and the focus marker *ni* 'also, even', and in Kannada, *a:daru:* is made up of *a:gu* 'become', *re* 'conditional', and *u:* 'also'.

- (24) a. *jusufa ne-laj xajit'ani q^hsan-diz mani-jar luhu-zwa*
 Jusuf who-Super at.least good-Adv song-PI say-Impf
 'Jusuf sings better than anyone'
- b. *ra:ju ya:r-indal-a:daru: heccu canna:gi ha:da-bahudu*
 Raju who-from-at.least much well sing-may
 'Raju may sing better than somebody'

11.2.5 Additional distinctions

There are several other distinctions that are associated with indefinite pronouns, occurring in different languages. All these involve the derivation of marked indefinite pronouns from unmarked indefinite pronouns through their association with some specific notion. For example, according to Haspelmath (1997: 45), Russian makes a distinction among its indefinite pronouns concerning the knowledge of the speaker. It has a *-to* series of indefinites (*kto-to* 'someone', *gde-to* 'somewhere', etc.) that can be used only if the speaker cannot identify the referent. Another distinction that is made by some languages is between simple non-specific indefinite meaning and a special kind of non-specific indefinite meaning in which the choice of a referent is left to the addressee. Kannada uses the emphatic particle *e:* for denoting the latter meaning.

- (25) *ya:r-e: bar-ali na:nu hedaruv-ud-illa*
 who-Emph come-Subjunctive I fear-it-not
 'Whoever may come, I do not fear'

There are several languages that differentiate between affirmative and negative indefinites by attaching a negative marker to the indefinite pronoun. In English, for example, *no* functions both as a negative marker as well as a non-specific marker.

- (26) a. *He went somewhere.*
 b. *He did not go anywhere.*
 c. *He went nowhere.*

Wichita (North Caddoan; Rood 1976: 10) has a set of indefinite noun roots, namely *kiya?* 'person', *kiri?* 'thing', *ka:?* 'place', *si:?* 'event', *ckinc* 'amount', and *sis* 'extent', that can occur with the interrogative prefix *e:-*, indefinite prefix *ka:-* or the negative prefix *ká:* to form three different paradigms of proforms.

- | | | | |
|------|-----------------------|-----------------------------|---------------------------|
| (27) | Interrogative | Indefinite | Negative |
| | <i>e:kiri?</i> 'what' | <i>ka:kiri?</i> 'something' | <i>ká:kiri?</i> 'nothing' |
| | <i>e:kiya?</i> 'who' | <i>ka:kiya?</i> 'somebody' | <i>ká:kiya?</i> 'nobody' |

There are also languages in which the negative marker is attached to an indefinite pronoun rather than directly to the general term. For example, Sonora Yaqui (Dedrick and Casad 1999: 252) derives its negative indefinite pronouns by prefixing the negative particle *kaá-* to the respective positive forms of indefinite pronouns.

- | | | | | |
|------|---------------|-------------|----------------|-------------|
| (28) | <i>hábe</i> | 'someone' | <i>kaábe</i> | 'no one' |
| | <i>híta</i> | 'something' | <i>kaíta</i> | 'nothing' |
| | <i>hak</i> | 'somewhere' | <i>kaák</i> | 'nowhere' |
| | <i>haibu</i> | 'now' | <i>kaibu</i> | 'never' |
| | <i>háčini</i> | 'somehow' | <i>kaáčini</i> | 'in no way' |

In these latter types of language, we may perhaps regard negation as a peripheral constituent (§7.3.4) rather than as part of the pronoun itself. Haspelmath (1997: 192–234) provides a detailed description of negative indefinites.

11.2.6 Grammaticalization of marked indefinites

There is an interesting difference between languages that maintain the relevance of the various structures mentioned above for indicating different types of marked indefinite meanings on the one hand, and the ones that do not maintain this relevance and allow the structures to be used for denoting related concepts on the other. In the former case the internal structure of the constructions appear to remain intact, whereas in the latter case there is a tendency to grammaticalize the constructions and reduce them into affixed forms or compact lexical items. For example, Dravidian languages maintain the distinction between the use of disjunctive particles (with unmarked indefinites) for denoting specific indefinites and of conjunctive particles for denoting

non-specific indefinites. The constructions are also maintained as having transparent structures (see §11.2.1).

On the other hand, Indo-Aryan languages like Hindi, Bengali, etc. do not maintain the distinction between specific and non-specific indefinites in this derivation (see §11.2.2) and correspondingly they also tend to lexicalize the structures into compact items. In Bengali, for example, it is rather difficult to separate the element that differentiates between marked and unmarked indefinites as shown in (29) (Dasgupta 1979).

(29) Unmarked indefinites		Marked indefinites	
(interrogative pronouns)			
<i>ke</i>	'who'	<i>keu</i>	'someone, anyone'
<i>kake</i>	'to whom'	<i>kauke</i>	'to someone, to anyone'
<i>kokhon</i>	'when'	<i>kōkhono</i>	'sometime, anytime'
<i>kothay</i>	'where'	<i>kothaw</i>	'somewhere, anywhere'

Diachronically, the marked indefinites are derived from the unmarked ones by the addition of the suffix *-o* that goes back to Sanskrit *api* 'or', as mentioned earlier.

In several European languages, certain compact expressions that denote indefinite meaning have developed from longer structures involving concepts like '(I) don't know who', 'who knows what', 'whoever it may be', 'no matter what', etc. Some of these, like the ones derived from the phrase 'who knows what', appear to involve the notion of interrogation in their derivation and not merely indefiniteness. Haspelmath (1997: 129–56) provides a detailed description of these derivations.

(30)	Middle High German	<i>ne veiz wer</i>	'I don't know who'	<i>neizwer</i>	'somebody'
	Lithuanian	<i>kas žinokas</i>	'who knows who'	<i>kažkas</i>	'somebody'
	Latin	<i>qui-vis</i>	'who-want'	<i>qui-vis</i>	'anybody'

11.3 Puzzle about indirect 'questions'

I have pointed out in the foregoing sections how the analysis of constituent questions occurring in most languages as involving only unmarked indefinite pronouns and not any interrogative pronouns as such allows us to resolve the puzzle about the derivational relationship that is reported to occur in those languages between interrogative and indefinite pronouns. I have been able to show that the puzzle derives from an erroneous analysis of those languages as possessing 'interrogative' pronouns. Another puzzle that can be resolved with the help of this reinterpretation or reanalysis of constituent questions concerns the occurrence of the so-called 'indirect questions' that contain embedded constituent questions in several languages. Consider, for example, the sentences given in (31a–b).

- (31) a. *Who has gone home?*
 b. *I want to find out who has gone home.*

We can regard (31b) as incorporating (31a). Further, since the desire to obtain information has been expressed only indirectly through a matrix clause in (31b), we can also regard it as involving an ‘indirect question’. However, the occurrence of other complex sentences of this type (32a–c) makes this concept of an ‘indirect question’ rather difficult to maintain.

- (32) a. *I don’t want to know who has gone home.*
 b. *I know who has gone home.*
 c. *Don’t tell me who has gone home.*

A speaker would be using the sentences (32a–c) in contexts in which he very specifically *does not* want his addressee to provide any information regarding the constituent *who* occurring in the embedded clause *who has gone home* of those sentences. It is therefore difficult to see how this embedded clause can be described as involving interrogativity, even of the ‘indirect’ type. If it involves the notion of ‘request to provide information’ it would clearly be contradicting the matrix clauses.

On the other hand, if we regard the pronoun *who* occurring in (31a) as an unmarked indefinite pronoun and assign the notion of interrogativity to the interrogative intonation that co-occurs with it, there would be no need to regard either (31b) or (32a–c) as involving the notion of interrogation. That is, they need not be regarded as ‘questions’ simply because their embedded clauses are not uttered with an interrogative intonation.

A cross-linguistic study of constructions that correspond to these indirect ‘questions’ supports the latter analysis. In Dravidian languages like Kannada, for example, sentences containing embedded constituent questions can be uttered either with the intonation of a declarative sentence or with that of an interrogative sentence. Depending upon the intonation used, the sentences can be either statements or questions.

- (33) *ellige ho:guvud-endu ra:ju he:lidda:ne*
 where going-that Raju said
 (i) ‘Raju has told (me) where (we) are going’
 (ii) ‘Where did Raju say that (we) are going?’

In the case of sentences like (34a–b), on the other hand, Kannada allows only declarative intonation to be used, apparently because the use of interrogative intonation would make the sentences contradictory.

- (34) a. *ellige ho:guvud-endu nanage gottide*
 where going-that me know
 ‘I know where (we) are going’

- b. *ellige ho:guvud-endu he:la-be:ḍa*
 where going-that say-not
 'Don't tell (me) where (we) are going'

It is difficult to account for this constraint if we are to assume that the notion of interrogativity is expressed by interrogative pronouns in Kannada. In (33), for example, if interrogativity is expressed by the pronoun *ellige* 'where' itself, how is it that the use or non-use of intonation determines the interrogativity of that sentence? Similarly, if the use of *ellige* 'where' with its 'interrogative' meaning is not disallowed in (34a–b), why is the use of an interrogative intonation disallowed?

Bengali (Indo-Aryan) also allows sentences that contain embedded constituent questions to be interpreted either as questions or as statements, depending upon the intonation used (K. K. Rarhi, personal communication).

- (35) *boyṭa kothay ache se tamake bollo*
 book where is he you (Dat) said
 (i) 'He told you where the book is'
 (ii) 'Where did he tell you the book is?'

According to Rarhi, (35) would be interpreted as (i) a statement if it is spoken with a falling intonation, whereas it would be interpreted as (ii) a question if it is spoken with a rising intonation.

On the other hand, several languages that use an interrogative particle/affix or an interrogative mood for denoting the notion of interrogativity (§10.2.2–3) retain the 'interrogative' pronoun in indirect questions but not the interrogative particle/affix or interrogative mood. Rather strangely, however, these constructions are still described by some of the grammarians as indirect 'questions'. For example, Matthews and Yip (1994: 334) point out that in Cantonese (Chinese), constituent questions contain a question particle in addition to an interrogative pronoun. In indirect 'questions', the question particle may be omitted.

- (36) a. *léih jyuh hái bīndouh a*
 you live at where Int
 'Where do you live?'
 b. *yàhndeih mahn ngóh jyuh hái bīndoh*
 people ask me live at where
 'People ask me where I live'

This is also true of several other languages like Godoberi (Daghestanian: Kibrik 1996: 38), Ngiyambaa (Australian: Donaldson 1980: 148, 262), Hausa (Chadic: Newman 2000), and Ika (Chibchan: Frank 1990). In Hausa, for example, constituent questions contain a question word and also a 'q-morpheme' that occurs at the end of the sentence. This morpheme has the effect of lengthening the final short vowels and adding a low tone to words ending in high tone

and thereby producing a fall (see §10.2.2). Indirect questions are formed by attaching the marker *kō* ‘or’ to a clause that contains a question word. This clause, however, does not have the essential phonological feature associated with direct questions, namely the final vowel lengthening and final low tone (q-morpheme) (Newman 2000: 501). Similarly, Ika uses sentence-final particles like *e*, *o*, or *no*, along with the question word, to form its constituent questions, but in indirect questions, declarative particles like *ni* ‘certainty’ or *in* ‘witness’ are used instead. However, the ‘question’ words continue to be used in these indirect questions (Frank 1990: 86).

In southeastern Pomo (Hokan-Siouan: Moshinsky 1974: 105), all interrogative pronouns occur in two forms: (i) the basic form, ending in the interrogative suffix *?e*, occurs in constituent questions, whereas (ii) the ‘quotative’ form, in which the interrogative suffix *?e* is replaced by the quotative suffix *-do*, occurs in the narrative.

(37)	Question words	Quotative words
	<i>?awi-?e</i> ‘what’	<i>?awi-do</i>
	<i>?sin-?e</i> ‘how many’	<i>?sin-do</i>
	<i>?tey-?e</i> ‘when’	<i>?tey-do</i>
	<i>?a-?e</i> ‘who’	<i>?a-do</i>
	<i>?ey-?e</i> ‘where’	<i>?ey-do</i>
	<i>?el-?e</i> ‘which one’	<i>?el-do</i>
	<i>?iy-?e</i> ‘ „ (animate)’	<i>?iy-do</i>

On the other hand, languages that differentiate between interrogative and indefinite pronouns and use distinct sets of expressions for representing them *replace* their interrogative pronouns by indefinite ones while forming indirect ‘questions’. For example, Ainu (Tamura 2000: 232) has interrogative pronouns like *hunna* ‘who’, *hñta* ‘what’, and *hunak* ‘where’ which are quite different from the corresponding indefinite pronouns, namely *nen* ‘who’, *nep* ‘what’, and *nee* ‘where, when’. The former are used in constituent questions whereas the latter are used in indirect ‘questions’. The translation of the latter as ‘who’, ‘what’, etc. is apparently caused by their use in indirect questions, but Tamura notes that they also occur in declarative sentences in which they indicate notions like ‘somebody’.

Heath (1998: 183) points out that in Koyra Chiini (Songhay family), when constituent questions are embedded under verbs of uncertainty like ‘wonder’ or ‘enquire’ and also ‘know’ or ‘see’, the question word is replaced by a non-interrogative noun phrase that is normally indefinite in form. The remainder of the constituent question surfaces as a relative clause.

- (38) a. *mey* *ŋga* *sii* *nee*
 who Focus not.be here
 ‘Who is not here?’

- b. *ay si bey bor kaa koy*
 1Sg Neg know person Rel go
 ‘I don’t know who has gone’

11.4 Interrogative–relative puzzle

Another pronominal affinity that appears to be as puzzling as the one between interrogative and indefinite pronouns is the affinity between interrogative and relative pronouns. Several languages, like English, use the same set of pronouns in both constituent questions as well as relative clauses.

- (39) a. *Whom did she marry?*
 b. *The man whom she married is a lawyer.*
- (40) a. *Where does she live?*
 b. *The village where she lives is near the seashore.*

On the other hand, there are several other languages that have two distinct sets of pronouns, of which one is used in constituent questions and the other one in relative clauses. Most of the Indo-Aryan languages are of this type. For example, Bengali uses distinct sets of pronouns in its constituent questions and relative clauses. The former have an initial *k*- whereas the latter have an initial *j*-. In addition to these, Bengali also has a set of correlative pronouns that occur in the matrix clause and identify the antecedent of the relative pronoun that occurs in the relative clause (Dasgupta 1979, K. K. Rarhi, personal communication).

(41)		Interrogative	Relative	Correlative
	Person	<i>ke</i>	<i>je</i>	<i>se</i>
	Person-to	<i>kake</i>	<i>jake</i>	<i>take</i>
	Time	<i>kɔkhon</i>	<i>jɔkhon</i>	<i>tɔkhon</i>
	Place	<i>konkhane</i>	<i>jekthane</i>	<i>sekhane</i>
	Amount	<i>kɔto</i>	<i>jɔto</i>	<i>tɔto</i>

- (42) a. *ke bajare gie chilo*
 who market gone had
 ‘Who had gone to the market?’
- b. *je bajare gie chilo se ekthane ache*
 who market gone had he here is
 ‘The man who had gone to the market is here’

As I mentioned earlier (§8.3.4), languages that use the same pronoun as relatives and interrogatives also show affinity between interrogatives and indefinites. On the other hand, languages that have distinct sets of interrogative and indefinite pronouns do not appear to use their interrogative pronouns as relative pronouns. Instead, they use one of their demonstrative pronouns for this purpose.

If the former type of languages are assumed to have only unmarked indefinite pronouns and not any interrogative pronouns in their constituent questions, the above-mentioned puzzle about relative pronouns would become easier to resolve. What we have to explain now is only the way in which indefinite pronouns can function as relative pronouns.

This is clearly less puzzling than an affinity between interrogative and relative pronouns because we only need to account for the use of 'indefinite' pronouns as 'definite' relative pronouns. However, as I have pointed out in the ninth chapter (§9.2), the notion of 'indefiniteness' that is connected with indefinite pronouns is not the same as the notion of indefiniteness that is connected with noun phrases. The latter is linguistic (anaphoric) whereas the former is extra-linguistic. Since relative pronouns are similar to noun phrases in their 'definiteness' (involving an anaphoric relationship), the use of indefinite pronouns in their formation would not involve any contradiction. Indefinite pronouns can also be used in contexts in which there is an antecedent, as for example, in phrases like *that somebody* (see §9.2.2).

When we examine the position of relative pronouns in the whole system of proforms, their affinity with other proforms like demonstrative and indefinite pronouns becomes even more transparent. As I have pointed out in the seventh chapter, the characteristic that is common to all proforms is the denotation of a set of general concepts like person, thing, place, time, manner, kind, number, purpose, etc. Proforms are differentiated from one another by the purpose for which these general concepts are used. Some languages fail to differentiate between two or more of these purposes like relatives and indefinites, but they use alternative devices either in the pronouns themselves or elsewhere in the clause for expressing the differences.

For example, English uses the same set of proforms as relative and interrogative (unmarked indefinite) pronouns, but it differentiates between the two by placing the former (relative pronoun) immediately after its antecedent and shifting the latter (interrogative pronoun) to the sentence-initial position or placing it before the noun that it modifies.

- (43) a. *I have seen the man who wrote this book.*
 b. *Who wrote this book?*
- (44) a. *I have read the book which sold two million copies.*
 b. *Which book sold two million copies?*

There is a somewhat different kind of relative-interrogative puzzle that has affected the study of relative and interrogative pronouns of Indo-Aryan languages. It does not derive from the *affinity* between these two types of pronouns because the two are distinct in most of these languages. Instead, it derives from the application of traditional assumptions to these proforms. Traditionally, the

relative pronouns of Indo-Aryan languages are contrasted with interrogative pronouns. Following this tradition, and also the tradition of Generative Grammars, Dasgupta (1979) describes the former as [−INT] words and the latter (interrogatives) as [+INT] words. As a result, some of the constraints on the occurrence of these two pronouns in languages like Bengali have become problematic for him. For example, the interrogative pronouns of Bengali can take the marker *-o* in order to indicate specific (45a) or non-specific (45b) indefinite meanings (see §11.2.2). However, the corresponding relative pronouns do not allow such a derivation. It is rather difficult to account for this constraint with the help of the feature [±INT].

- (45) a. *ram kothaw gie che*
 Ram where-M gone has
 ‘Ram has gone somewhere’
 b. *ram kothaw jay ni*
 Ram where-M gone not
 ‘Ram hasn’t gone anywhere’

On the other hand, both interrogative as well as relative pronouns can be used in exclamatory sentences (Dasgupta 1979: 262). There is in fact an interesting difference between the uses of these two sets of words in such sentences, as shown below:

- (46) (a) *ki sundor lag che tomake*
 what beautiful look is you-Obj
 ‘How beautiful you are!’
 (b) *je sundor lag che tomake karon tumi*
 what beautiful look is you-Obj because you
 holud roner sari pore cho
 yellow color sari wear have
 ‘How beautiful you look, as you are wearing a yellow sari!’

K. K. Rarhi (personal communication) points out that an exclamatory sentence containing a *j*-word (relative pronoun) (like 46b) has to be followed by a clause that provides the context for exclamation, whereas one containing a *k*-word (interrogative pronoun) can occur on its own. This is an additional constraint that cannot also be explained with the help of the feature [±INT].

On the other hand, the assumption that interrogative pronouns of Bengali are unmarked indefinites, and its relative pronouns are basically anaphoric in their usage, requiring a correlative pronoun to be associated with them, can account for both these constraints. It is the indefiniteness of the former that allows it to denote specific and non-specific varieties of indefiniteness (§11.2), and it is the anaphoric nature of the latter that makes it necessary for it to have an antecedent specified in the following clause even in an exclamatory use. Such an assumption

would also be in tune with the suggestion made by Gonda (1954, 1955), namely that the Indo-Aryan distinction between interrogatives and relatives derives from a Proto-Indo-European distinction between the notion of ‘unspecified individuality’ and of ‘focusing’ or ‘introducing’ an element of the clause.

This view is supported by another constraint that affects the use of relative pronouns in Bengali. According to Klaiman (1977), Bengali uses *je* as a clause-initial complementizer and the word *bole* (which is the participial form of the verb *bol* ‘say’) as a clause-final complementizer.

- (47) a. *ami sunechi je apni dhakay jachen*
 I heard.have that you Dacca(Loc) going.are
 b. *apni dhakay jachen bole ami sunechi*
 you Dacca(Loc) going.are saying I heard.have
 ‘I have heard that you are going to Dacca’

Klaiman also notes that similar clause-final complementizers, derived from the verb ‘say’, occur in Marathi (*mhanu:n*) and Sinhalese (*kiala:*) as well. This constraint apparently derives from the focusing or introducing characteristic of relative pronouns.

Dravidian languages use both the participial strategy as well as the correlative strategy for deriving their relative clauses. They are generally considered to have borrowed the correlative strategy from Indo-Aryan languages, even though they do not make use of any distinct set of relative pronouns for this purpose. Instead, they use their interrogative–indefinite pronouns as relative pronouns. The Kannada sentences (48a–b) exemplify the use of participial and correlative relative constructions.

- (48) a. *avanu bared-iruv-a pustakavannu na:nu o:d-idde:ne*
 he wrote-has-Rel book I read-have
 ‘I have read the book that he has written’
 b. *avanu y:ava pustakavannu bared-idda:n-o: adannu*
 he which book wrote-has-or it
 na:nu o:d-idde:ne
 I read-have
 ‘I have read the book that he has written’

On the other hand, Laxmi Bai (1983) argues that both these strategies are native to Dravidian languages. For example, she points out that the correlative strategy is used even in Old Tamil texts like *Tolkappiam* and *Tirukkural* (belonging to the first century AD) for denoting notions like the indefinite ‘whoever’ and ‘whatever’. She points out further that even in modern Dravidian languages the two strategies (participial and correlative) are consistently employed for denoting *different* functions. For example, English sentences with an indefinite head noun, or with a definite but non-specific head noun (like

whenever or *whatever*) generally get translated by the correlative relative clause and not by the participial relative clause.

Another interesting point that supports such a claim is that the correlative relative clauses of these languages appear to form part of a larger set of constructions in which the proforms employed are very clearly indefinite pronouns. I pointed out earlier the occurrence of two distinct sets of marked indefinite pronouns in Kannada, namely specific and non-specific, derived by adding the disjunctive and conjunctive particles respectively to unmarked indefinite (interrogative) pronouns (§11.2.1). These particles need not be attached directly to indefinite pronouns. There can be other expressions occurring between the two, such as the head noun to which the indefinite pronoun is attached as a modifier, or even a nominalized verb as shown in (48) and (49) of Havyaka Kannada.

- (48) a. *a:nu enta ma:ḍidd-akk-o: i: kaṣa bayndu*
 I what did-it-Dat-or this difficulty come.has
 'This difficulty has come to me for doing something' (specific indefinite)
- b. *ni:nu enta ke:ḷuttar-ude ad-akke enna-tre uttara iddu*
 you what ask-if-also it-Dat me-with answer is
 'Whatever you may ask, there is an answer to it with me'
- (49) a. *ni: enta ke:ḷuttay-o: ad-akke enna-tre uttara iddu*
 you what ask-or it-Dat me-with answer is
 'I have an answer to whatever you may ask'
- b. *ni: enta ke:ḷuttar-ude ad-akke enna-tre uttara iddu*
 you what ask-even it-Dat me-with answer is
 'I have an answer to whatever you may ask'

Notice that the two sentences in (48) contain only specific (48a) and non-specific (48b) indefinite pronouns but not any relative clauses as such, whereas the two sentences of (49) contain relative clauses in which indefinite pronouns have been used as correlative pronouns. The relative clause ends in a disjunctive particle in (49a) whereas in (49b) it ends in a conjunctive particle. Grammarians generally describe only the former as a relative clause, but the latter does not appear to be very different from it. We can actually regard the two as involving the use of specific (49a) and non-specific (49b) indefinite pronouns if we regard the disjunctive and conjunctive particles as part of the indefinite pronouns rather than that of the verb.

We can expect these correlative clauses of Havyaka to manifest characteristics that derive from their having specific and non-specific indefinite pronouns as relative pronouns. For example, unlike the correlative relative clauses of Hindi, those of Havyaka cannot follow the matrix clause.

Comrie (1998: 62) points out that the formation of relative clauses using the Relative Pronoun Strategy is quite exceptional outside Europe. In this strategy,

the position relativized is indicated inside the relative clause by means of a pronominal element, which is case-marked in order to indicate its syntactic-semantic role within the relative clause. The English sentences (43a) and (44a) given earlier exemplify this strategy.

There has been a tendency to describe any language that uses a relative pronoun as manifesting a variation of this strategy. Even languages that do not have a relative pronoun, but a relative participle instead, are described as manifesting the same strategy with the difference that a 'gap' occurs in place of the relative pronoun. Comrie (1998: 83) points out that much typology of relative clauses, like much of the formal grammatical study of relative clauses, has been concerned with such concepts as extraction or accessibility, which characterize the Relative Pronoun Strategy. These notions may not be applicable to languages like Japanese that make use of a participial strategy. He considers it necessary to differentiate between languages that are subject to extraction constraints and the ones that are not.

The comparison of Indo-Aryan and Dravidian correlative relative clauses appears to indicate that they represent another area of relativization in which rethinking is needed.

11.5 Summary

I have pointed out in this chapter some of the additional puzzles that can be resolved on the basis of the assumption that in the case of languages in which there is affinity between interrogative and indefinite pronouns, constituent questions do not contain any interrogative pronouns as such, but only indefinite pronouns. Such an assumption removes the puzzle from the 'derivation' of specific and non-specific indefinites through the use of disjunctive and conjunctive particles respectively that occurs in several languages: The derivation can be shown to be a predictable result of the interaction of the particles with indefinite pronouns.

It also removes the puzzle from the occurrence of the so-called 'indirect questions' which do not involve any question (request for information) as such in most instances. The affinity shown by some languages like English between interrogative pronouns and relative pronouns also loses its puzzling characteristic in view of the fact that we now have to relate relative pronouns with unmarked indefinites and not with interrogative pronouns. Similarly, some of the puzzles that are connected with the assumption that the relative pronouns of Indo-Aryan languages contrast with the pronouns of constituent questions in being 'non-interrogative' are also resolved by the assumption that the two contrast concerning the notion of indefiniteness and not that of interrogativity.

12 Concluding Remarks

12.1 Definitions for pronouns

One of the purposes of this monograph has been to differentiate between personal pronouns on the one hand and the rest of the pronouns, called ‘proforms’, on the other, so that a firm basis for formulating a satisfactory set of definitions for the notion of ‘pronouns’ can be established. The traditional definition of pronouns as words that ‘stand for’ nouns (or other expressions) is generally considered to be unsatisfactory because it leaves several questions unanswered and further it is not properly applicable to most pronouns. My starting point therefore was an enquiry concerning the reason for this dissatisfaction regarding the traditional definition of pronouns. I argued that the dissatisfaction derives primarily from the fact that pronouns do not together form a single category. They do not therefore share any substantial set of properties that can form the basis of a single definition.

The most important difference between personal pronouns and proforms concerns their primary function. Personal pronouns are used primarily for denoting speech roles like ‘being the speaker’ and ‘being the addressee’ of the sentence in which they occur. Proforms, on the other hand, are used for employing a set of general concepts in different functions like locating an entity, denoting one’s lack of knowledge about it, obtaining information about it from the addressee, or relating it with some other entity. This important functional distinction between personal pronouns and proforms gives rise to an interesting formal distinction between the two, namely that personal pronouns involve single-element stems whereas proforms involve two-element stems, consisting of an element that denotes a general concept and another one that indicates the function or purpose for which the concept is used.

There are several other differences that occur between personal pronouns and proforms, which also derive from this primary functional distinction occurring between the two. For example, personal pronouns cannot be regarded as either definite or non-definite. They have an obligatory anaphoric relationship with all pronouns that belong to the same person (singular or plural). It is only when they are used for denoting speech roles that belong to different speech contexts that ambiguities arise, and languages have developed special ‘logophoric’ pronouns for removing these ambiguities. Personal pronouns also have an obligatory ‘referential’ relationship with noun phrases that indicate the identity of their referents. They cannot take modifiers or complements on their own because that would conflict with their primary function of denoting speech roles. Their association with grammatical categories like gender and number is generally

meant for indicating distinctions in speech roles rather than among the individuals who perform those roles. These unique characteristics of first and second person pronouns make it necessary to regard them as constituting a distinct category of their own.

Proforms, on the other hand, are distinct from personal pronouns as well as nouns in their having the function of indicating general concepts that may belong to any of the various lexical categories, such as nouns, verbs, adjectives, or adverbs. They may also occur as determiners that can be associated with specific lexical items that belong to nominal, verbal, adjectival, or adverbial categories. The notion of definiteness or indefiniteness that is associated with them is quite different from the one that is associated with nouns. It is semantic in the case of proforms whereas in that of noun phrases, it is only pragmatic. Proforms also differ from nouns in the kind of reference that they establish. These unique characteristics of proforms make it necessary to regard them as belonging to a distinct category of their own.

I believe that these characteristics, functional and formal, can together form the basis for a satisfactory set of definitions for pronouns:

Personal pronouns Single-element expressions that have the denotation of speech roles as their primary function.

Proforms Two-element expressions that indicate a general concept and a function such as (i) locating an entity, (ii) denoting one's lack of knowledge about an entity, (iii) obtaining information about an entity, (iv) identifying an entity as the same as the one denoted by some other expression, (v) relating an entity with some other entity in which that general concept is employed.

It is evident that these two definitions can cover only the prototypical personal pronouns and proforms. There would be exceptional instances in which both personal pronouns as well as proforms may need to perform additional functions, and in such usages they may manifest additional characteristics. These need to be regarded as instances that deviate from prototypical pronouns. In the case of personal pronouns, on the other hand, there is also the possibility of their manifesting conflicting characteristics, deriving from the fact that their functional need to dissociate from their referents conflicts with the fact that they do have referents of their own.

Another important point that needs to be noted here is that we may have to differentiate between (i) languages in which the primary function of denoting speech roles is assigned to independent expressions that occur as arguments of sentences on the one hand, and (ii) languages in which this function is assigned to bound elements (clitics or affixes) that occur as parts of the predicate on the other. The characterization of personal pronouns given in the first part of this monograph applies only to independent expressions (free pronouns) occurring in the former type of languages. It is possible that some of these characteristics are

also shown by the bound pronouns of the latter type of languages, as they also have the function of denoting speech roles, but this possibility is yet to be studied in detail.

This differentiation between the pronouns of two different types of language is necessary because in languages of the former type (called 'free-pronoun' languages), bound agreement markers occurring in the predicate may have entirely different functions to perform, such as denoting the gender and number of the *referents* of personal pronouns, and hence they need not show any of the characteristics that derive from the function of denoting speech roles. In the case of the latter type of languages (called 'bound-pronoun' languages), on the other hand, independent pronouns are not used consistently for denoting speech roles. They are used only when there is a need to emphasise or contrast the participants with others. We cannot therefore expect them to show any of the characteristics that derive from the function of denoting speech roles.

The present study of pronouns is primarily an examination of the characteristics of independent personal pronouns. It can therefore be regarded as primarily the characterization of independent pronouns occurring in free-pronoun languages. As mentioned earlier, there is a need to make a detailed study of the bound forms of pronouns (clitics or agreement markers) that occur in bound-pronoun languages in order to find out the relevance of any of these characteristics to them. The data available to me on these bound-pronoun languages is not detailed (or extensive) enough to draw any conclusions in this regard.

12.2 Plurality and conjunction

This in-depth study of personal pronouns and proforms has brought out certain other characteristics of these expressions that have been the source of puzzles and paradoxes. The association of personal pronouns with the category of number, for example, has been rather puzzling because unlike nouns and other pronouns, personal pronouns manifest several idiosyncratic characteristics in this association. While nouns generally have a single plural form, personal pronouns tend to have either (i) two or more non-singular forms or (ii) two of them may share a single non-singular form. The occurrence of the exclusive-inclusive distinction that cuts across the number distinction in the case of first and second person pronouns has also been rather puzzling.

I have suggested that this puzzle can be resolved by regarding the association of personal pronouns with the category of number as involving, primarily, the notion of conjunction rather than plurality. The two notions differ in that conjunction emphasizes the *differences* that occur among the entities that are conjoined, whereas plurality emphasizes the *similarity* that occurs among them. When personal pronouns are associated with a marker for conjunction, they indicate different combinations of speech roles, such as 1 + 2, 1 + 3, 2 + 3, etc.

which leads to the disparity in the number of singular and non-singular forms, mentioned above. Languages apparently differentiate between conjunction and plurality by using suppletive forms or distinct number markers in the former case.

12.3 Pragmatic and semantic identity and reference

The characterization of proforms has brought out another important point that, I think, has been the source of confusion regarding notions like reference, specificity, definiteness, and opacity. The notion of identification that is involved in the case of third person pronouns and noun phrases that contain articles is only pragmatic. It is only 'linguistic' and gets established automatically the moment a referring expression is used in a sentence. The addressee has no choice but to concede the identity (or 'definiteness') in this case. On the other hand, the identification that is involved in the use of proforms is more substantial and can be regarded as semantic. It is not merely linguistic. It needs to be established in one of the 'extra-linguistic' levels such as that of reality, mythology, fiction, dream, etc. There is scope for the addressee to be unsatisfied with the information that has been provided by the speaker in the case of this latter type of identification. The 'indefiniteness' of an entity can therefore persist so long as the addressee (or the speaker himself) is not satisfied by the information that is available for identification.

There is a similar kind of distinction in the notion of 'referentiality' that is associated with these two types of expressions. The referentiality of third person pronouns and noun phrases that contain articles depends crucially upon the *intention* of the speaker. Notions like opacity, familiarity, or unfamiliarity, etc. are completely irrelevant in this regard. Any given expression of this type can be referential or non-referential irrespective of the amount and type of information that it carries. Its referentiality depends only upon the intention of the speaker, which in its turn would, of course, depend upon the purpose for which the relevant expression is being used.

Proforms, on the other hand, are referential or non-referential in a more substantial fashion and hence we can regard the notion of referentiality that is associated with them as semantic rather than pragmatic. Demonstrative pronouns, for example, locate an object or entity in a specific level of reference. Interrogatives seek information regarding the identity of an entity in one of these specific levels. They cannot perform their function properly unless the entities are properly grounded in their proper levels. For example, a question about a character in a novel (such as *who is he?* *where did he go?* *why did he go?* etc.) cannot be answered unless the character has a proper existence in that novel, and the nature of that existence is made clear for the addressee. Because of this difference, the specific–non-specific (or referential–non-referential) distinction

in the case of proforms gets constrained by notions like negation and mood or opacity. Languages generally possess distinct proforms (like English *some* and *any*) that get constrained by these semantic distinctions.

12.4 Interrogation and indefiniteness

The affinity between different sets of proforms, such as interrogatives and indefinites or interrogatives and relatives is another aspect of proforms that has been the source of puzzles and paradoxes. For example, in most of the languages that have been studied so far, interrogative and indefinite pronouns are either identical or derivationally related. In several of these languages, the derivation (of indefinite pronouns) involves the addition of a conjunctive or disjunctive particle to interrogative pronouns. This is a puzzling derivation because how can the notion of interrogation change into one of indefiniteness when it is associated with the notion of conjunction or disjunction?

The solution that I suggest for this puzzle is that in these languages (which show affinity between interrogative and indefinite pronouns) there is actually no interrogative pronoun as such. They only have unmarked indefinite pronouns that can be marked for specificity by attaching disjunctive particles and for non-specificity by attaching conjunctive particles. The notions of interrogativity and focus (restricting the former to a specific constituent) are expressed in these languages not by a pronoun but rather by one or more of several additional devices like using a question particle or question intonation, using the verb in the interrogative mood, using a focus particle or focus construction, and so on. That is, the basis of the puzzle about the affinity between interrogative and indefinite pronouns is an erroneous analysis of constituent questions in these languages.

This solution is also helpful in removing certain additional puzzles that are connected with the use of interrogative pronouns in these languages. For example, interrogative pronouns occur in the so-called 'indirect questions' in which the notion of interrogation is completely absent. The assumption that the languages do not have any interrogative pronouns as such, but only indefinite pronouns in their constituent questions makes it easy to explain such usages. Similarly, the use of interrogative pronouns as relative pronouns in languages in which interrogatives are related to indefinites can be explained more easily if the pronouns are regarded as indefinites rather than as interrogatives.

I hope that the present study of pronouns has brought a semblance of order into the chaotic world of pronouns.

Appendix

List of 225 Languages (with Family Affiliations) Used as a Sample (and References for the same)

The list of languages given here includes most of the 200 languages that were recommended for the World Atlas of Language Structures (Haspelmath, Dryer, Gil, and Comrie: 2005). I have used these languages as a sample primarily for determining the relationship between third person pronouns and demonstratives, as described in the sixth chapter (see Bhat, forthcoming). The numbers that follow the language names in this list indicate the kind of relationship that exists between third person pronouns and demonstratives (1 no relationship, 2 third person pronoun related to all demonstratives, 3 related to remote demonstratives only, 4 related to non-remote demonstratives only, 5 related by gender markers, and 6 related for non-human reference only—see §6.2.1). This distinction forms the basis of a typological distinction between ‘two-person’ and ‘three-person’ languages.

Secondarily, I have used these languages as a sample for establishing a correlation between the above-mentioned distinction on the one hand, and (i) a distinction between languages that show gender distinction in third person and the ones that do not (§6.3), and (ii) a distinction between languages that show a ‘distance-oriented’ deictic system and the ones that show a ‘person-oriented’ deictic system (§6.4), on the other. I have also used the same set of languages as a sample while describing the type of spatial distinctions that occur among the demonstratives of different languages (§8.2.1). However, I have not used any language sample as such while describing other aspects of personal pronouns and proforms.

1. Ainu 1

TAMURA, SUZUKO 2000. *The Ainu Language*. Tokyo: Sanseido Co.

2. Alambalak (Sepik Hill) 5

BRUCE, LES 1984. *The Alambalak Language of Papua New Guinea (East Sepik)*. Canberra: Australian National University.

3. Aleut (Eskimo-Aleut) 2

BERGSLAND, KNUT 1997. *Aleut Grammar*. Fairbanks: Alaska Native Language Center.

4. Ambulas (Middle Sepik) 2

WILSON, PATRICIA R. 1980. *Ambulas Grammar*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

5. Amele (Gum, Trans-New Guinea) 1

ROBERTS, JOHN R. 1987. *Amele*. London: Croom Helm.

6. Angami (Naga, Tibeto-Burman) 1

GIRIDHAR, P. P. 1980. *Angami Grammar*. Mysore: Central Institute of Indian Languages.

7. Apurinã (Arawakan) 5

FACUNDES, SIDNEY DA SILVA 2000. ‘The Language of the Apurina People of Brazil (Maipure/Arawak)’. SUNY at Buffalo dissertation.

8. Arabic (Egyptian) (Semitic) 5

MITCHELL, T. F. 1956. *An Introduction to Egyptian Colloquial Arabic*. Oxford: Oxford University Press.

9. Armenian (Eastern) (Indo-European) 3

KOZINTSEVA, NATALIA 1995. *Modern Eastern Armenian*. Munich: Lincom Europa.

10. Asheninca (Campa) (Pre-Andean Arawakan) 2

REED, JUDY, and PAYNE, DAVID L. 1986. 'Asheninca (Campa) Pronominals'. In Ursula Wiesemann (ed.), *Pronominal Systems*, 323–31. Tübingen: Günter Narr Verlag.

11. Asmat (Papuan) 4

VOORHOEVE, CLEMENS L. 1965. *The Flamingo Bay Dialect of the Asmat Language*. The Hague: M. Nijhoff.

12. Athapare (Kiranti, Tibeto-Burman) 3

EBERT, KAREN H. 1997. *Athpare Grammar*. Munich: Lincom Europa.

13. Awa Pit (Barbacoan) 6

CURNOW, TIMOTHY J. 1997. 'A Grammar of Awa Pit (Cuaiquer): An Indigenous Language of Western Colombia'. Australian National University dissertation.

14. Babungo (Grassfields Bantu) 5

SCHAUB, WILLI 1985. *Babungo*. London: Croom Helm.

15. Bagirmi (Nilo-Saharan) 1

STEVENSON, R. C. 1969. *Bagirmi Grammar*. Khartoum: University of Khartoum.

16. Barasano (Tucanoan) 3

JONES, WENDELL, and JONES, PAULA 1991. *Barasano Syntax*. Arlington: University of Texas.

17. Basque (Isolate) 2

SALTARELLI, MARIO 1988. *Basque*. London: Croom Helm.

18. Bawm (Tibeto-Burman) 1

REICHLE, VERENA 1981. *Bawm Language and Lore*. Bern: Peter Lang.

19. Brahui (Dravidian) 4

BRAY, DANYS DE S. 1986. *The Brahui Language*, part I. Calcutta: Govt. Printing. Reprinted in 1986 by Gian Publishing House, Delhi.

20. Bukiyip (Arapesh) 4

CONRAD, ROBERT J. 1991. *An Outline of Bukiyip Grammar*. Canberra: Australian National University.

21. Buriat (Mongolian) 2

POPPE, NICHOLAS N. 1960. *Buriat Grammar*. Bloomington: Indiana University.

22. Burmese (Tibeto-Burman) 1

OKELL, JOHN 1969. *A Reference Grammar of Colloquial Burmese*. London: Oxford University Press.

23. Burushaski (Isolate) 3

LORIMER, D. L. R. 1935. *The Burushaski Language*, i: *Introduction and Grammar*. Oslo: H. Ascheoug & Co.

24. Canela-Krahô (Je, Amazonian) 1
POPJES, JACK, and POPJES, JO 1986. 'Canela-Krahô'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, i. 128–99. Berlin: Mouton de Gruyter.
25. Cantonese (Chinese, Sino-Tibetan) 1
MATTHEWS, STEPHEN, and YIP, VIRGINIA 1994. *Cantonese: A Comprehensive Grammar*. New York: Routledge.
26. Cayuvava (Isolate) 2
KEY, HAROLD H. 1967. *Morphology of Cayuvava*. The Hague: Mouton.
27. Chaldean (Modern) (Semitic) 2
SARA, SOLOMON I. 1974. *A Description of Modern Chaldean*. The Hague: Mouton.
28. Chamorro (Austronesian) 1
TOPPING, DONALD M. (with the assistance of Bernadita C. Dungca). 1973. *Chamorro Reference Grammar*. Honolulu: University Press of Hawaii.
29. Chemehuevi (Numic, Uto-Aztecan) 2
PRESS, MARGARET L. 1979. *Chemehuevi: A Grammar and Lexicon*. Berkeley: University of California Press.
30. Chinantec (Lealao) (Oto-Manguean) 4
RUPP, JAMES E. 1989. *Lealao Chinantec Syntax*. Arlington: University of Texas.
31. Chinese (Mandarin) (Sino-Tibetan) 6
LI, CHARLES N., and THOMPSON, SANDRA A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
32. Chukchi (Chukotko-Kamchatkan) 1
BOGORAS, WALDEMAR 1922. 'Chukchee'. In Franz Boas (ed.), *Handbook of American Indian Languages*, ii. 631–903. Washington: Govt. Printing Office.
33. Chuvash (Turkic) 4
KRUEGER, JOHN R. 1961. *Chuvash Manual: Introduction, Grammar, Reader, and Vocabulary*. Bloomington: Indiana University.
34. Cree (Plains) (Algonquian) 1
WOLFART, H. CHRISTOPH 1973. *Plains Cree: A Grammatical Study*. Philadelphia: American Philosophical Society.
35. Cubeo (Tucanoan) 2
MORSE, NANCY L., and MAXWELL, MICHAEL B. 1999. *Cubeo Grammar*. Arlington: University of Texas.
36. Dani (Lower Grand Valley) (Dani-Kwerba) 3
BROMLEY, H. MYRON 1981. *A Grammar of Lower Grand Valley Dani*. Canberra: Australian National University.
37. Diegueno (Yuman, Hokan) 2
MILLER, AMY W. 1990. 'A Grammar of Jamul Diegueño'. University of California, San Diego, dissertation.
38. Diola-Fogny (Niger-Congo) 5
SAPIR, J. DAVID 1965. *A Grammar of Diola-Fogny*. Ibadan: Cambridge University Press.

39. Djaru (Pama-Nyungan) 1
TSUNODA, TASAKU 1995. *The Djaru Language of Kimberley, Western Australia*. Canberra: Australian National University.
40. Dong (Kam-Tai) 1
LONG, YAOHONG, and ZHENG, GUOQIAO 1998. *The Dong Language in Guizhou Province, China*. Translated from Chinese by D. NORMAN GEARY. Arlington: University of Texas.
41. Dumi (Kiranti, Tibeto-Burman) 6
VAN DRIEM, GEORGE 1993. *A Grammar of Dumi*. Berlin: Mouton de Gruyter.
42. English (Germanic, Indo-European) 2
QUIRK, RANDOLPH, GREENBAUM, SIDNEY, LEECH, GEOFFREY, and SVARTVIK, JAN 1985. *A Comprehensive Grammar of the English Language*. Harlow: Longman.
43. Epena Pedee (Chocó) 1
HARMS, PHILIP LEE 1994. *Epena Pedee Syntax*. Arlington: University of Texas.
44. Erromangan (Oceanic, Austronesian) 4
CROWLEY, TERRY 1998. *An Erromangan (Sye) Grammar*. Honolulu: University of Hawaii Press.
45. Evenki (Tungus, Altaic) 1
NEDJALKOV, IGOR 1997. *Evenki*. London: Routledge.
46. Fijian (Boumaa) (Pacific, Austronesian) 1
DIXON, ROBERT M. W. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
47. Finnish (Uralic) 1
SULKALA, HELENA, and KARJALAINEN, MERJA 1992. *Finnish*. London: Routledge.
48. French (Italic, Indo-European) 5
HARRIS, MARTIN 1997. 'French'. In Martin Harris and Nigel Vincent (eds.), *The Romance Languages*, 209–45. London: Routledge.
49. Futuna-Aniwa (West) (Oceanic, Austronesian) 1
DOUGHERTY, JANET W. D. 1983. *West Futuna-Aniwa: An Introduction to a Polynesian Outlier Language*. Berkeley: University of California Press.
50. Garo (Tibeto-Burman) 6
BURLING, ROBBINS 1961. *A Garo Grammar*. Poona: Deccan College.
51. Georgian (Kartvelian) 3
HEWITT, B. GEORGE 1995. *Georgian: A Structural Reference Grammar*. Amsterdam: John Benjamins.
52. German (Germanic, Indo-European) 5
CURME, GEORGE O. 1922. *A Grammar of the German Language*. Revised second edition. New York: Fredrick.
53. Godié (Kru) 6
MARCHESE, LYNELL 1986. 'The Pronominal System of Godié'. In Ursula Wiesemann (ed.) 1986. *Pronominal Systems*, 217–56. Tübingen: Günter Narr Verlag.
54. Godoberi (Daghestanian) 5
KIBRIK, ALEXANDR (ed.) 1996. *Godoberi*. Munich: Lincom Europa.

55. Gooniyandi (Bunaban, North Australian) 1
MCGREGOR, WILLIAM 1990. *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins.
56. Grebo (Kru, Niger-Congo) 5
INNES, GORDON 1966. *An Introduction to Grebo*. London: University of London.
57. Greenlandic (West) (Eskimo-Aleut) 4
FORTESCUE, M. 1984. *West Greenlandic*. London: Croom Helm.
58. Guaraní (Tupí) 1
GREGORES, EMMA, and SUÁREZ, JORGE A. 1967. *A Description of Colloquial Guaraní*. The Hague: Mouton.
59. Guugu Yimidhirr (Australian) 6
HAVILAND, JOHN 1979. 'Guugu Yimidhirr'. In R. M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian Languages*, i. 27–180. Amsterdam: John Benjamins.
60. Hausa (West Chadic) 1
NEWMAN, PAUL 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
61. Hawaiian (Oceanic, Austronesian) 3
ELBERT, SAMUEL H., and PUKUI, MARY KAWENA 1979. *Hawaiian Grammar*. Honolulu: University of Hawaii Press.
62. Hebrew (Modern) (Semitic) 2
GLINERT, LEWIS 1989. *The Grammar of Modern Hebrew*. New York: Cambridge University Press.
63. Hindi (Indo-Aryan) 2
MCGREGOR, R. S. 1995. *Outline of Hindi Grammar: With Exercises*. Oxford: Oxford University Press [third edition revised and enlarged].
64. Hixkaryana (Carib) 2
DERBYSHIRE, DESMOND C. 1979. *Hixkaryana*. Amsterdam: North-Holland.
65. Hungarian (Uralic) 6
KENESSEI, ISTVÁN, VAGO, ROBERT M., and FENYVESI, ANNA 1998. *Hungarian*. London: Routledge.
66. Hunzib (Tsezik, Daghestanian) 2
VAN DEN BERG, HELMA 1995. *A Grammar of Hunzib*. Munich: Lincom Europa.
67. Ika (Chibchan) 1
FRANK, PAUL 1990. *Ika Syntax*. Arlington: University of Texas.
68. Ila (Bantu) 5
SMITH, EDWIN 1907. *A Handbook of the Ila Language*. London: Oxford University Press.
69. Imonda (Papuan) 1
SEILER, WALTER 1985. *Imonda, a Papuan Language*. Canberra: Australian National University.
70. Indonesian (Sundic, Austronesian) 1
SNEDDON, JAMES N. 1996. *Indonesian: A Comprehensive Grammar*. New York, London: Routledge.

71. Ingush (Nax, Nax-Dagestanian) 1

NICHOLS, J. 1994. 'Ingush'. In RIEKS SMEETS (ed.), *The Indigenous Languages of the Caucasus*, iv. 79–145. Delmar, New York: Caravan Books.

72. Iraqw (Southern Cushitic) 1

MOUS, MAARTEN 1993. *A Grammar of Iraqw*. Hamburg: Buske.

73. Irish (Celtic, Indo-European) 3

EOIN, GEARÓID M. 1993. 'Irish'. In MARTIN J. BALL (ed.), *The Celtic Language*, 101–44. London: Routledge.

74. Jacalteek (Mayan) 1

GRINEVALD CRAIG, COLETTE 1977. *The Structure of Jacalteek*. Austin: University of Texas Press.

75. Japanese (Isolate) 1

HINDS, JOHN 1986. *Japanese*. London: Croom Helm.

76. Jaqaru (Aymaran) 6

HARDMAN, MARTHA J. 2000. *Jaqaru*. Munich: Lincom Europa.

77. Ju'hoan (!xu, Khoisan) 5

DICKENS, PATRICK 1992. *Ju|'hoan Grammar*. Windhoek: Nyae Nyae Development Foundation.

78. Juang (Munda) 1

MATSON, D. M. 1964. 'A Grammatical Sketch of Juang, a Munda Language'. University of Wisconsin dissertation.

79. Kambara (Malayo-Polynesian) 2

KLAMER, MARIAN 1998. *A Grammar of Kambara*. Berlin: Mouton de Gruyter.

80. Kannada (Dravidian) 3

BHAT, D. N. S. 1978. *Pronominalization*. Pune: Deccan College.

81. Kanuri (Saharan, Nilo-Saharan) 1

CYFFER, NORBERT 1998. *A Sketch of Kanuri*. Cologne: Köppe.

82. Kawaiisu (Numic, Uto-Aztecan) 2

ZIGMOND, MAURICE L., BOOTH, CURTIS G., and MUNRO, PAMELA 1991. *Kawaiisu: A Grammar and Dictionary with Texts*. Berkeley: University of California Press.

83. Kayah Li (Eastern) (Karen) 1

SOLNIT, DAVID B. 1997. *Eastern Kayah Li: Grammar, Texts, Glossary*. Honolulu: University of Hawaii Press.

84. Kayardild (Tangkic, Australian) 6

EVANS, NICHOLAS D. 1995. *A Grammar of Kayardild, with Historical-Comparative Notes on Tangkic*. Berlin: Mouton de Gruyter.

85. Kewa (West-Central, Trans-New Guinea) 1

FRANKLIN, KARL JAMES 1971. *A Grammar of Kewa, New Guinea*. Canberra: Australian National University.

86. Khalkha (Mongolian) 2

STREET, JOHN C. 1963. *Khalkha Structure*. Bloomington: Indiana University Press.

87. Kharia (Munda) 2
BILIGIRI, H. S. 1965. *Kharia: Phonology, Grammar and Vocabulary*. Pune: Deccan College.
88. Khasi (Mon-Khmer) 2
NAGARAJA, K. S. 1985. *Khasi: A Descriptive Analysis*. Pune: Deccan College.
89. Khmu (Mon-Khmer) 1
PREMSRIRAT, SUWILAI 1987. *A Khmu Grammar*. Canberra: Australian National University.
90. Kilivila (Oceanic, Austronesian) 4
SENFT, GÜNTER 1986. *Kilivila: The Language of the Trobriand Islanders*. Berlin: Mouton de Gruyter.
91. Kiribatese (North Hebridean, Austronesian) 1
GROVES, TERAB'ATA R., GROVES, GORDON W., and JACOBS, RODERICK 1985. *Kiribatese: An Outline Description*. Canberra: Australian National University.
92. Kisi (Atlantic, Niger-Congo) 5
CHILDS, G. TUCKER 1995. *A Grammar of Kisi, A Southern Atlantic Language*. Berlin: Mouton de Gruyter.
93. Koasati (Muskogean) 1
KIMBALL, GEOFFREY 1991. *Koasati Grammar*. Lincoln, Nebraska: University of Nebraska Press.
94. Kobon (Karam, Trans-New Guinea) 1
DAVIES, JOHN 1981. *Kobon*. Amsterdam: North-Holland.
95. Korean (Isolate) 1
SOHN, HO-MIN 1999. *The Korean Language*. Cambridge: Cambridge University Press.
96. Koromfe (Gur, Niger-Congo) 1
RENNISON, JOHN R. 1997. *Koromfe*. London: Routledge.
97. Koyraboro Senni (Songhay, Nilo-Saharan) 1
HEATH, JEFFREY 1999. *A Grammar of Koyraboro (Koroboro) Senni: The Songhay of Gao, Mali*. Cologne: Rüdiger Köppe Verlag.
98. Koyra Chiini (Songhay) 1
HEATH, JEFFREY 1999. *A Grammar of Koyra Chiini: The Songhay of Timbuktu*. Berlin: Mouton de Gruyter.
99. Kusaiean (Micronesian) 1
LEE, KEE-DONG 1975. *Kusaiean Reference Grammar*. Honolulu: University of Hawaii Press.
100. Kwakiutl (Wakashan) 1
BOAS, FRANZ 1911. 'Kwakiutl'. In Franz Boas (ed.), *Handbook of American Indian Languages*, 423–558. Washington: Govt. Printing Press.
101. Lahu (Tibeto-Burman) 1
MATISOFF, JAMES A. 1973. *A Grammar of Lahu*. Berkeley: University of California Press.
102. Lakhota (Siouan) 1
VAN VALIN, ROBERT D. 1977. 'Aspects of Lakhota Syntax'. University of California, Berkeley, Ph.D. dissertation.

103. Lampung (Samba-Bima, Austronesian) 6
WALKER, DALE F. 1976. *A Grammar of the Lampung Language*. Jakarta: Badan Penyelenggara Seri Nusa Universitas Katolik Atma Jaya Jakarta.
104. Lango (Nilo-Saharan) 1
NOONAN, MICHAEL 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.
105. Latvian (Baltic, Indo-European) 5
NAU, NICOLE 1998. *Latvian*. Munich: Lincom Europa.
106. Lavukaleve (East Papuan) 2
TERRILL, ANGELA 1999. 'A Grammar of Lavukaleve: A Papuan Language of the Solomon Islands'. Australian National University dissertation.
107. Lepcha (Tibeto-Burman) 1
MAINWARING, COLONEL G. B. 1876. *A Grammar of the Róng (Lepcha) Language*. Calcutta: Baptist Mission Press.
108. Lezgian (Nax-Dagestanian) 4
HASPELMATH, MARTIN 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.
109. Lillooet (Interior Salish) 1
VAN EIJK, JAN 1997. *The Lillooet Language: Phonology, Morphology, Syntax*. Vancouver: University of British Columbia Press.
110. Limbu (Tibeto-Burman) 3
VAN DRIEM, GEORGE 1987. *A Grammar of Limbu*. Berlin: Mouton de Gruyter.
111. Luvale (Bantu) 5
DOKE, C. M. 1949. *A Grammar of Luvale*. Johannesburg: Witwatersrand University Press.
112. Macushi (Carib) 2
ABBOTT, MIRIAM 1991. 'Macushi'. In Desmond Derbyshire and Geoffrey Pullum (eds.), *Handbook of Amazonian Languages*, iii. 23–160. Berlin: Mouton de Gruyter.
113. Malayalam (Dravidian) 2
ASHER, R.E., and KUMARI, T. C. 1997. *Malayalam*. London: Routledge.
114. Maltese (Semitic) 5
BORG, ALBERT, and AZZOPARDI-ALEXANDER, MARIE 1997. *Maltese*. London: Routledge.
115. Mangarayi (North Australian) 2
MERLAN, FRANCESCA 1989. *Mangarayi*. London: Routledge.
116. Manipuri (Meitei) (Tibeto-Burman) 2
BHAT, D. N. S., and NINGOMBA, M. S. 1997. *Manipuri Grammar*. Munich: Lincom Europa.
117. Maori (Eastern Polynesian) 3
BAUER, WINFRED 1999. *Maori*. London: Routledge.
118. Mapuche (Isolate, possibly Southern Andean) 4
SMEETS, CATHARINA, J. M. 1989. 'A Mapuche Grammar'. Rijksuniversiteit Leiden dissertation.
119. Marathi (Indo-Aryan) 2
PANDHARIPANDE, RAJESHWARI V. 1997. *Marathi*. London: Routledge.

120. Margi (Chadic) 1
HOFFMAN, C. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.
121. Maricopa (Yuman, Hokan) 2
GORDON, LYNN 1986. *Maricopa Morphology and Syntax*. Berkeley: University of California Press.
122. Marshallese (Melanesian, Austronesian) 4
ZEWEN, FRANÇOIS-XAVIER NICOLAS 1977. *The Marshallese Language: A Study of its Phonology, Morphology, and Syntax*. Hamburg: Reimer.
123. Martuthunira (Pama-Nyungan) 1
DENCH, ALAN CHARLES 1995. *Martuthunira: A Language of the Pilbara Region of Western Australia*. Canberra: Australian National University.
124. Maung (Yiwaiddjan, North Australian) 5
CAPELL, ARTHUR, and HINCH, H. E. 1970. *Maung Grammar: Text and Vocabulary*. The Hague: Mouton.
125. Maybrat (West Papuan) 5
DOL, PHILOMENA 1999. *A Grammar of Maybrat: A Language of the Bird's Head, Irian Jaya*. Leiden: Grafisch Bedrijf.
126. Mbay (Central Sudanic, Nilo-Saharan) 1
KEEGAN, JOHN 1997. *A Reference Grammar of Mbay*. Munich: Lincom Europa.
127. Micmac (Algonquian) 1
HEWSON, JOHN, and FRANCIS, BERNARD 1990. *The Micmac Grammar of Father Pacifique* (translated and retranscribed). Winnipeg, Manitoba: Algonquian and Iroquoian Linguistics.
128. Mishmi (Digaro) (Tibeto-Burman) 2
SASTRY, G. D. P. 1984. *Mishmi Grammar*. Mysore: Central Institute of Indian Languages.
129. Miwok (Southern Sierra) (Penutian) 2
BROADBENT, SYLVIA M. 1964. *The Southern Sierra Miwok Language*. Berkeley: University of California Press.
130. Mixtec (Chalcatongo) (Oto-Manguean) 1
MACAULAY, MONICA 1996. *A Grammar of Chalcatongo Mixtec*. Berkeley: University of California Press.
131. Mixtec (Ocotepéc) (Oto-Manguean) 1
ALEXANDER, RUTH MARY 1988. 'A Syntactic Sketch of Ocotepéc Mixtec'. In C. Henry Bradley and Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages*, i, 151–304. Arlington: University of Texas.
132. Mixtec (Yosondúa) (Oto-Manguean) 1
FARRIS, EDWIN R. 1992. 'A Syntactic Sketch of Yosondúa Mixtec'. In C. Henry Bradley and Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages*, iv, 1–171. Arlington: University of Texas.
133. Miya (West Chadic) 1
SCHUH, RUSSELL G. 1998. *A Grammar of Miya*. Berkeley: University of California Press.

134. Mokilese (Micronesian) 1
HARRISON, SHELDON R. 1976. *Mokilese Reference Grammar*. Honolulu: University of Hawaii Press.
135. Mparntwe Arrente (Pama-Nyungan) 1
WILKINS, DAVID P. 1989. 'Mparntwe Arrente (Aranda): Studies in the Structure and Semantics of Grammar'. Australian National University dissertation.
136. Mundani (Bantu) 1
PARKER, ELIZABETH 1986. 'Mundani Pronouns'. In Ursula Wiesemann (ed.) 1986. *Pronominal Systems*, 131–66. Tübingen: Günter Narr Verlag.
137. Mundari (Munda) 1
OSADA, TOSHIKI 1992. *A Reference Grammar of Mundari*. Tokyo: Tokyo University of Foreign Studies.
138. Mupun (West Chadic) 6
FRAJZYNGIER, ZYGMUNT 1993. *A Grammar of Mupun*. Berlin: Dietrich Reimer Verlag.
139. Naga (Mao) (Tibeto-Burman) 1
GIRIDHAR, P. P. 1994. *Mao Naga Grammar*. Mysore: Central Institute of Indian Languages.
140. Nahuatl (Tetelcingo) (Aztecan) 2
TUGGY, DAVID H. 1979. 'Tetelcingo Nahuatl'. In Ronald W. Langacker (ed.), *Studies in Uto-Aztecan Grammar*, ii: *Modern Aztec Grammatical Sketches*, 1–140. Arlington: University of Texas.
141. Nama (Khoekhoe) (Khoisan) 1
HAGMAN, ROY S. 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University.
142. Ndyuka (Creole) 1
HUTTAR, GEORGE L., and HUTTAR, MARY 1994. *Ndyuka*. London: Routledge.
143. Nez Perce (Sahaptian, Penutian) 1
AOKI, HARUO 1970. *Nez Perce Grammar*. Berkeley: University of California Press.
144. Ngiti (Central Sudanic) 1
KUTSCH LOJENGA, CONSTANCE 1994. *Ngiti: A Central-Sudanic Language of Zaire*. Cologne: Rüdiger Köppe Verlag.
145. Ngiyambaa (Pama-Nyungan) 2
DONALDSON, TAMSIN 1980. *Ngiyambaa: The Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
146. Nivkh (Palaeosiberian or Palaeo-Asiatic) 6
GRUZDEVA, EKATERINA 1998. *Nivkh*. Munich: Lincom Europa.
147. Nkore-Kiga (Bantu) 5
TAYLOR, CHARLES 1985. *Nkore-Kiga*. London: Croom Helm.
148. Nubian (Kunuz) (Nilo-Saharan) 1
ABDEL-HAFIZ, AHMED SOKARNO 1988. 'A Reference Grammar of Kunuz Nubian'. State University of New York, Buffalo, dissertation.
149. Nunggubuyu (Norlth Australian) 1
HEATH, JEFFREY 1984. *Functional Grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.

150. Oneida (Iroquoian) 1
ABBOTT, CLIFFORD 2000. *Oneida*. Munich: Lincom Europa.
151. Oromo (Harar) (East Cushitic) 1
OWENS, JONATHAN 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Buske.
152. Paamese (Oceanic, Austronesian) 2
CROWLEY, TERRY 1982. *The Paamese Language of Vanuatu*. Canberra: Australian National University.
153. Palauan (Philippine Austronesian) 2
JOSEPHS, LEWIS S. 1975. *Palauan Reference Grammar*. Honolulu: University of Hawaii Press.
154. Passamaquoddy (Algonquian) 1
LEAVITT, ROBERT M. 1996. *Passamaquoddy-Maliseet*. Munich: Lincom Europa.
155. Paumarí (Arauan, Arawakan) 2
CHAPMAN, SHIRLEY, and DERBYSHIRE, DESMOND C. 1991. 'Paumarí'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, iii. 161–352. Berlin: Mouton de Gruyter.
156. Pero (West Chadic) 5
FRAJZYNGIER, ZYGMUNT 1989. *A Grammar of Pero*. Berlin: Dietrich Reimer Verlag.
157. Persian (Iranian) 6
RASTORGUEVA, V. S. 1964. *A Short Sketch of the Grammar of Persian*. Bloomington: Indiana University.
158. Pipil (Aztecán) 3
CAMPBELL, LYLE 1985. *The Pipil Language of El Salvador*. Berlin: Mouton de Gruyter.
159. Pirahã (Mura) 3
EVERETT, DANIEL L. 1986. 'Pirahã'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, i. 200–325. Berlin: Mouton de Gruyter.
160. Pitta-Pitta (Pama-Nyungan) 2
BLAKE, BARRY J., and BREEN, J. GAVAN 1971. *The Pitta-Pitta Dialects*. Melbourne: Monash University.
161. Ponapean (Micronesian, Austronesian) 1
REHG, KENNETH L. 1981. *Ponapean Reference Grammar*. Honolulu: University of Hawaii Press.
162. Pomo (Eastern) (Hokan) 1
MCLENDON, SALLY 1975. *A Grammar of Eastern Pomo*. Berkeley: University of California Press.
163. Pomo (Southeastern) (Hokan) 2
MOSHINSKY, JULIUS 1974. *A Grammar of Southeastern Pomo*. Berkeley: University of California Press.
164. Punjabi (Indo-Aryan) 2
BHATIA, TEJ K. 1993. *Punjabi: A Cognitive-Descriptive Grammar*. London: Routledge.

165. Quechua (Imbabura) (Isolate) 6
 COLE, PETER 1982. *Imbabura Quechua*. Amsterdam: North-Holland.
166. Quechua (Tarma) (Isolate) 6
 ADELAAR, W. F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary*. Lisse: Peter Ridder Press.
167. Rama (Macro-Chibchan) 2
 GRINEVALD, COLETTE G. 1988. *A Grammar of Rama*. Report of National Science Foundation BNS 8511156.
168. Rapanui (Polynesian) 3
 DU FEU, VERONICA 1996. *Rapanui*. London: Routledge.
169. Rejang (Sundic, Austronesian) 1
 MCGINN, RICHARD 1982. *Outline of Rejang Syntax*. Jakarta: Badan Penyelenggara Seri Nusa Universitas Katolik Atma Jaya Jakarta.
170. Remo (Munda) 1
 FERNANDIS, FRANK 1969. 'A Grammar of Remo'. University of North Carolina dissertation.
171. Returā (Tucanoan) 2
 STROM, CLAY 1992. *Retuarā Syntax*. Arlington: University of Texas.
172. Sahaptin (Northern) (Penutian) 1
 JACOBS, MELVILLE 1931. *A Sketch of Northern Sahaptin Grammar*. Seattle: University of Washington Press.
173. Samoan (Polynesian) 3
 MOSEL, ULRIKE, and HOVDHAUGEN, EVEN 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
174. Sango (Creole) 1
 SAMARIN, WILLIAM J. 1967. *A Grammar of Sango*. The Hague: Mouton.
175. Sawu (Sumba-Bima, Austronesian) 1
 WALKER, ALAN T. 1982. *Grammar of Sawu*. Jakarta: Badan Penyelenggara Seri Nusa Universitas Katolik Atma Jaya Jakarta.
176. Semelai (Aslian, Austroasiatic) 1
 KRUSPE, NICOLE 1999. 'A Grammar of Semelai'. University of Melbourne dissertation.
177. Sentani (Papuan) 2
 COWAN, H. K. J. 1965. *Grammar of the Sentani Language: With Specimen Texts and Vocabulary*. The Hague: Martinus Nijhoff.
178. Serbo-Croatian (Slavic) 1
 KORDIC, SANJEZANA 1997. *Serbo-Croatian*. Munich: Lincom Europa.
179. Shona (Bantu) 5
 BRAUNER, SIEGMUND 1995. *A Grammatical Sketch of Shona Including Historical Notes*. Cologne: Rüdiger Köppe.
180. Sinhalese (Indo-Aryan) 2
 GAIR, JAMES W., and PAOLILLO, JOHN C. 1997. *Sinhala*. Munich: Lincom Europa.

181. Slave (Athapaskan) 6

RICE, KEREN 1989. *A Grammar of Slave*. Berlin: Mouton de Gruyter.

182. Spanish (Italic) 1

GREEN, JOHN N. 1988. 'Spanish'. In Martin Harris and Nigel Vincent (eds.), *The Romance Languages*, 79–130. London: Croom Helm.

183. Supyire (Gur, Niger-Congo) 2

CARLSON, ROBERT 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

184. Taba (Austronesian) 2

BOWDEN, JOHN 1997. 'Taba (Makian Dalam): Description of an Austronesian Language of Eastern Indonesia'. University of Melbourne dissertation.

185. Tagalog (Philippine Austronesian) 1

SCHACHTER, PAUL, and OTANES, FE T. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press. [reprinted in 1983].

186. Tamazight (Ayt Ndhir) (Berber) 1

PENCHOEN, THOMAS G. 1973. *Tamazight of the Ayt Ndhir*. Los Angeles: Undena Publishing.

187. Tauya (Papuan) 1

MACDONALD, LORNA 1990. *A Grammar of Tauya*. Berlin: Mouton de Gruyter.

188. Thai (Kam-tai) 1

CAMPBELL, STUART, and SHAWEEVONGSE, CHAUN 1957. *The Fundamentals of the Thai Language*. New York: Paragon Book Gallery.

189. Tibetan (Tibeto-Burman) 1

DENWOOD, PHILIP 1999. *Tibetan*. Amsterdam: John Benjamins.

190. Tidore (Papuan) 2

VAN STADEN, MIRIAM 2000. 'Tidore: A Linguistic Description of a Language of the North Moluccas'. Universiteits Leiden dissertation.

191. Tonkawa (Hokan) 1

HOIJER, HARRY 1933. *Tonkawa: An Indian Language of Texas*. New York: Columbia University Press.

192. Toqabaqita (Oceanic, Austronesian) 1

SIMONS, LINDA 1986. 'The Pronouns of To'abaita (Solomon Islands)'. In Ursula Weismann (ed.), *Pronominal Systems*, 21–35. Tübingen: Günter Narr Verlag.

193. Totonac (Misantla) (Totonac-Tepehua) 1

MACKAY, CAROLYN JOYCE 1991. 'A Grammar of Misantla Totanac'. University of Texas, Austin, dissertation.

194. Trumai (Isolate) 5

GUIARDELLO, RAQUEL 1999. 'A Grammar of Trumai'. Rice University, Houston, Texas, dissertation.

195. Tukang Besi (Austronesian) 1

DONOHUE, MARK 1999. *A Grammar of Tukang Besi*. Berlin, New York: Mouton de Gruyter.

196. Tümpisa Shoshone (Numic, Uto-Aztecan) 2
DAYLEY, JON P. 1989. *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.
197. Tunica (Tunica, Gulf) 1
HAAS, MARY R. 1940. *Tunica*. New York: J. J. Augustin.
198. Turkana (Nilotic) 1
DIMMENDAAL, GERRIT 1982. 'The Turkana Language'. University of Leiden dissertation.
199. Turkish (Turkic, Altaic) 3
LEWIS, G. L. 1967. *Turkish Grammar*. Oxford: Oxford University Press.
200. Tutelo (Siouan) 1
OLIVERIO, GIULIA R. M. 1997. 'A Grammar and Dictionary of Tutelo'. University of Kansas dissertation.
201. Tuvaluan (Polynesian) 3
BESNIER, NIKO 2000. *Tuvaluan: A Polynesian Language of the Central Pacific*. London: Routledge.
202. Tzutujil (Mayan) 2
DAYLEY, JON P. 1985. *Tzutujil Grammar*. Berkeley: University of California Press.
203. Urubu-Kaapor (Tupi-Guarani) 1
KAKUMASU, JAMES 1986. 'Urubu-Kaapor'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, i. 326–403. Berlin: Mouton de Gruyter.
204. Usan (Papuan) 1
REESINK, GER P. 1987. *Structures and their Functions in Usan: A Papuan Language of Papua New Guinea*. Amsterdam: John Benjamins.
205. Venda (Bantu) 5
POULOS, G. 1990. *A Linguistic Analysis of Venda*. Pretoria: Via Afrika.
206. Vietnamese (Viet-Muong, Mon-Khmer) 3
THOMPSON, LAURENCE C. 1991. *A Vietnamese Reference Grammar*. Honolulu: University of Hawaii Press.
207. Wambaya (West Barkly, North Australian) 2
NORDLINGER, RACHEL 1998. *A Grammar of Wambaya, Northern Territory (Australia)*. Canberra: Australian National University.
208. Wangkawiru (Western Desert, Australian) 4
ECKERT, PAUL, and HUDSON, JOYCE 1994. *Wangkawiru: A Handbook for the Pitjantjatjara Language Learner*. Underdale, SA: University of South Australia.
209. Warao (Isolate) 4
ROMERO-FIGUEROA, ANDRES 1997. *A Reference Grammar of Warao*. Munich: Lincom Europa.
210. Wardman (Yangmanic, Australian) 2
MERLAN, FRANCESCA C. 1994. *A Grammar of Wardaman: A Language of the Northern Territory of Australia*. Berlin: Mouton de Gruyter.

211. Wari' (Chapacura-Wanhan) 1
EVERETT, DANIEL L., and KERN, BARBARA 1997. *Wari': The Pacaas Novos Language of Western Brazil*. London: Routledge.
212. Wichita (Caddoan) 1
ROOD, DAVID S. 1976. *Wichita Grammar*. New York: Garland.
213. Wikchamni (Yokuts, Penutian) 1
GAMBLE, GEOFFREY 1978. *Wikchamni Grammar*. Berkeley: University of California Press.
214. Wintu (Wintun, Penutian) 4
PITKIN, HARVEY 1984. *Wintu Grammar*. Berkeley: University of California Press.
215. Yagua (Peba-Yaguan) 6
PAYNE, DORIS L., and PAYNE, THOMAS E. 1990. 'Yagua'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, ii. 249–474. Berlin: Mouton de Gruyter.
216. Yaqui (Sonora) (Southern Uto-Aztecan) 1
DEDRICK, JOHN M., and CASAD, EUGENE H. 1999. *Sonora Yaqui Language Structures*. Tucson: University of Arizona Press.
217. Yidiny (Pama-Nyungan) 2
DIXON, ROBERT M. W. 1977. *A Grammar of Yidiny*. Cambridge: Cambridge University Press.
218. Yimas (Nor-Pondo, Sepik-Ramu) 2
FOLEY, WILLIAM A. 1991. *The Yimas Language of New Guinea*. Stanford, Calif.: Stanford University Press.
219. Yoruba (Defoid, Niger-Congo) 1
AWOBULUYI, A. OLADELE 1978. *Essentials of Yoruba Grammar*. Ibadan: University Press.
220. Yuchi (Macro-Siouan) 5
WAGNER, GÜNTER 1934. *Yuchi*. New York: Columbia University Press.
221. Yukaghir (Kolyma) (Isolate) 1
MASLOVA, ELENA 1999. 'A Grammar of Kolyma Yukaghir'. University of Bielefeld dissertation.
222. Yukulta (Tangkic, North Australian) 2
KEEN, SANDRA 1983. 'Yukulta'. In R. M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian languages*, iii. Amsterdam: John Benjamins.
223. Yupik (Central) (Eskimo-Aleut) 1
MIYAOKA, OSAHITO 1996. 'Sketch of Central Alaskan Yupik, an Eskimoan Language'. In Ives Goddard (ed.), *Handbook of American Indians*, xvii: *Languages*, 325–63. Washington: Smithsonian Institution.
224. Yurok (probably Algonquian) 2
ROBINS, R. H. 1958. *The Yurok Language: Grammar, Texts, Lexicon*. Berkeley: University of California Press.
225. Zulu (Bantu) 5
POULOS, G., and BOSCH, S. E. 1997. *Zulu*. Munich: Lincom Europa.

References

- ABBOTT, CLIFFORD 2000. *Oneida*. Munich: Lincom Europa.
- ABDEL-HAFIZ, AHMED S. 1988. 'A Reference Grammar of Kunuz Nubian'. State University of New York, Buffalo, dissertation.
- ABONDOLO, D. 1998. 'Introduction'. In D. Abondolo (ed.), *The Uralic Languages*, 1–42. London: Routledge.
- ADELAAR, W. F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary*. Lisse: Peter Ridder Press.
- AIKHENVALD, ALEXANDRA Y. 1998. 'Warekena'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, iv. 225–439. Berlin: Mouton de Gruyter.
- AJELLO, R. 1998. 'Armenian'. In A. G. Ramat and P. Ramat (eds.), *The Indo-European Languages*, 197–227. London: Routledge.
- AMBROS, A. 1977. *Damascus Arabic*. Malibu: Undena Publications.
- ANDERSEN, PAUL K. 1983. *Word Order Typology and Comparative Constructions*. Amsterdam: John Benjamins.
- ANDERSON, STEPHEN R., and KEENAN, EDWARD L. 1985. 'Deixis'. In Timothy Shopen (ed.), *Language Typology and Syntactic Descriptions*, iii: *Grammatical Categories and the Lexicon*, 259–302. Cambridge: Cambridge University Press.
- ANSCOMBE, G. E. M. 1975. 'The First Person'. In Samuel Guttenplan (ed.), *Mind and Language*, 45–65. Oxford: Clarendon Press.
- ASHER, R. E., and KUMARI, T. C. 1997. *Malayalam*. London: Routledge.
- AUSTIN, PETER 1981. *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- BAKER, MARK C. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- BALDI, PHILIP 1981. 'Indo-European'. In Bernard Comrie (ed.), *The World's Major Languages*, 31–67. London: Croom Helm.
- BARKER, M. A. R. 1964. *Klamath Grammar*. Berkeley: University of California Press.
- BAROJA, TOMÁS H. 1989. *Analytical Grammar of the Pokot Language*. Trieste: Università de Trieste.
- BAUER, WINFRED 1999. *Maori*. London: Routledge.
- BEAUMONT, RONALD 1985. *The Sechelt Language*. Penticton: Theytus Books.
- BENVENISTE, ÉMILE 1971. *Problems in General Linguistics*. Translated by M. E. Meek. Coral Gables: University of Miami Press.
- BERNTSEN, MAXINE, and NIMBKAR, J. 1982. *Marathi Structural Patterns*, Book 1. Pune: Granthali and Marathi Abhyas Parishad.
- BESNIR, NIKO 2000. *Tuvaluan*. London: Routledge.
- BHAT, D. N. S. 1967. *A Descriptive Analysis of Tulu*. Pune: Deccan College.
- 1978. *Pronominalization*. Pune: Deccan College.
- 1981. *Identification*. Trivandrum: Dravidian Linguistics Association.

- 1989. *Wh-words*. A report submitted to the University Grants Commission, New Delhi.
- 1994. *The Adjectival Category*. Amsterdam: John Benjamins.
- 1997. 'Indefinite Pronouns in Kannada'. In Jane H. Hill, P. J. Mistry, and Lyle Campbell (eds.), *The Life of Language: Papers in Linguistics in Honor of William Bright*, 369–83. Berlin: Mouton de Gruyter.
- 1999. *The Prominence of Tense, Aspect and Mood*. Amsterdam: John Benjamins.
- 2000. 'The Indefinite-Interrogative Puzzle'. *Linguistic Typology*, 4: 365–400.
- 2005. 'Third person pronouns and demonstratives'. In Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- and NINGOMBA, MANGI SINGH 1997. *Manipuri Grammar*. Munich: Lincom Europa.
- BILIGIRI, HEMMIGE, S. 1965. *Kharia*. Pune: Deccan College.
- BLAKE, BARRY J. 1994. *Case*. Cambridge: Cambridge University Press.
- BOAS, FRANZ 1911. 'Introduction'. In Franz Boas (ed.), *Handbook of American Indian Languages*, part 1, 1–83. Oosterhout, NB: Anthropological Publications.
- BOER, S. E., and LYCAN, W. G. 1975. 'Knowing Who'. *Philosophical Studies*, 28: 299–344.
- BOGORAS, WALDEMAR 1922. 'Chukchee'. In Franz Boas (ed.), *Handbook of American Indian Languages*, 631–903. Washington: Govt. Printing Press.
- BOHNHOFF, L. E. 1986. 'Yag Dii Pronouns'. In Ursula Wiesemann (ed.), 103–29.
- BORGMAN, D. M. 1989. 'Sanuma'. In D. C. Derbyshire and G. K. Pullum (eds.), *Handbook of Amazonian Languages*, ii. 15–248. Berlin: Mouton de Gruyter.
- BOWDEN, JOHN 1997. 'Taba (Makian Dalam): Description of an Austronesian Language from Eastern Indonesia'. University of Melbourne dissertation.
- BRADLEY, DAVID 1993. 'Pronouns in Burmese-Lolo'. *Linguistics of the Tibeto-Burman Area*, 16/1: 157–215.
- BRAINE, JEAN C. 1970. 'Nicobarese Grammar (Car Dialect)'. University of California, Berkeley, dissertation.
- BRANDON, FRANK R., and SEKI, LUCY F. 1984. 'Moving Interrogatives without an Initial+WH Node in Tupi'. In E. D. Cook and D. B. Gerdtz (eds.), *The Syntax of Native American Languages (Syntax and Semantics, xvi)*, 77–103. New York: Academic Press.
- BRESNAN, JOAN, and MCHOMBO, SAM A. 1987. 'Topic, Pronoun and Agreement in Chichewa'. *Language*, 63: 741–82.
- BRIGHT, WILLIAM 1957. *The Karok Language*. Berkeley: University of California Press.
- BROMLEY, H. MYRON 1981. *A Grammar of Lower Grand Valley Dani*. Canberra: Australian National University.
- BRUCE, LES 1984. *The Alamlak Language of Papua New Guinea (East Sepik)*. Canberra: Australian National University.
- BURCHULADZE, G. T. 1979. 'Voprosy stanovlenija ličnogo sprjaženija v laskom jazyke', *Annual of Ibero-Caucasian Linguistics*, 8: 176–246.
- BURLING, ROBBINS 1961. *A Garo Grammar*. Pune: Deccan College.
- BURQUEST, DONALD A. 1986. 'The Pronoun System of some Chadic Languages'. In Ursula Wiesemann (ed.), 71–101.
- BURROW, T., and EMENEAU, M. B. 1960. *A Dravidian Etymological Dictionary*. London: Oxford University Press.

- BYBEE, JOAN 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- CAMPBELL, LYLE 1985. *The Pipil Language of El Salvador*. Berlin: Mouton de Gruyter.
- CAPELL, ARTHUR, and HINCH, H. E. 1970. *Maung Grammar*. Berlin: Mouton de Gruyter.
- CARLSON, ROBERT 1994. *A Grammar of Supyre*. Berlin: Mouton de Gruyter.
- CHAFE, W. 1985. 'The Realis-Irrealis distinction in Caddo, the North Iroquoian Languages and English'. In J. L. Bybee and S. Fleischman (eds.), *Modality in Grammar and Discourse*, 349–365. Amsterdam: John Benjamins.
- CHAPMAN, SHIRLEY, and DERBYSHIRE, DESMOND S. 1991. 'Paumari'. In D. S. Derbyshire and G. K. Pullum (eds.), *Handbook of Amazonian Languages*, v. 161–352. Berlin: Mouton de Gruyter.
- CHARNEY, JEAN O. 1993. *A Grammar of Comanche*. Lincoln: University of Nebraska Press.
- CHESTERMAN, ANDREW 1991. *On Definiteness: A Study with Special Reference to English and Finnish*. Cambridge: Cambridge University Press.
- CHILDS, G. TUCKER 1995. *A Grammar of Kisi*. Berlin: Mouton de Gruyter.
- CHOMSKY, NOAM 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- CLEMENTS, G. N. 1975. 'The Logophoric Pronoun in Ewe: Its Role in Discourse'. *Journal of West African languages*, 2: 141–77.
- COLE, PETER 1978. 'On the Origins of Referential Opacity'. In P. Cole (ed.), *Syntax and Semantics*, ix, 1–22. New York: Academic Press.
- COMRIE, BERNARD 1978. 'Ergativity'. In Winfred P. Lehmann (ed.), *Syntactic Typology: Studies in the Phenomenology of Language*, 329–74. Austin: University of Texas.
- 1997. 'Pragmatic Binding: Demonstratives as Anaphors in Dutch'. *Berkeley Linguistic Society*, 23: 50–61.
- 1998. 'Rethinking the Typology of Relative Clauses'. *Language Design*, 1: 87–97.
- CONRAD, ROBERT J. 1991. *An Outline of Bukiyip Grammar*. Canberra: Australian National University.
- COOK, WALTER A. 1966. 'A Descriptive Analysis of Mundari'. Georgetown University dissertation.
- COOKE, JOSEPH R. 1968. *Pronominal Reference in Thai, Burmese and Vietnamese*. Berkeley: University of California Press.
- CORBETT, GREVILLE 2000. *Number*. Cambridge: Cambridge University Press.
- CORNISH, FRANCIS 1999. *Anaphora, Discourse, and Understanding*. Oxford: Oxford University Press.
- COWAN, H. K. J. 1965. *Grammar of the Sentani Language*. The Hague: Martinus Nijhoff.
- CRAIG, COLETTE, G. 1977. *The Structure of Jacalteco*. Austin: University of Texas Press.
- CROFT, W. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.
- CROWLEY, JERRY 1982. *The Paamese Language of Vanuatu*. Canberra: Australian National University.
- 1998. *An Erromangan (Syé) Grammar*. Honolulu: University of Hawaii Press.
- CULY, CHRISTOPHER 1994a. 'Aspects of Logophoric Marking'. *Linguistics*, 32: 1055–94.
- 1994b. 'A note on logophoricity in Dogon'. *Journal of African Languages and Linguistics*, 15: 113–25.
- 1997. 'Logophoric Pronouns and Point of View'. *Linguistics*, 35: 845–59.

- CURNOW, TIMOTHY J. 1997. 'A Grammar of Awa Pit (Cuaiquer): An Indigenous Language of South-Western Colombia'. Australian National University dissertation.
- CYSOUW, MICHAEL 2001. 'The paradigmatic structure of person marking'. Katholieke Universiteit, Nijmegen, dissertation.
- DAS GUPTA, K. 1971. *An Introduction to the Nocte Language*. Shillong: North East Frontier Agency.
- DASGUPTA, PROBAL 1979. 'Questions and Relative and Complement Clauses in a Bangla Grammar'. New York University dissertation.
- DAVIES, JOHN 1981. *Kobon*. Amsterdam: North-Holland.
- DAVIS, PHILIP W., and SAUNDERS, ROSS 1997. *A Grammar of Bella Coola*. Missoula, MT: University of Montana.
- DAYLEY, JON P. 1985. *Tzutujil Grammar*. Berkeley: University of California Press.
- 1989. *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.
- DEDRICK, JOHN M., and CASAD, EUGENE H. 1999. *Sonora Yaqui Language Structure*. Tucson: University of Arizona Press.
- DELANCEY, SCOTT 1983. 'Tangut and Tibeto-Burman Morphology'. *Linguistics of the Tibeto-Burman Area*, 7(2): 100–8.
- DENWOOD, PHILLIP 1999. *Tibetan*. Amsterdam: John Benjamins.
- DERBYSHIRE, DESMOND C. 1979. *Hixkaryana*. Amsterdam: North-Holland.
- DIESSEL, HOLGER 1999. *Demonstratives: Form, Function and Grammaticalization*. Amsterdam: John Benjamins.
- forthcoming. 'Spatial Deixis: A Cross-Linguistic Study'.
- DILLER, ANTHONY 1994. 'Thai'. In Cliff Goddard and Anna Wierzbicka (eds.), *Semantic and Lexical Universals: Theory and Empirical Findings*, 149–70. Amsterdam: John Benjamins.
- DIXON, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- 1977. *A Grammar of Yidiny*. Cambridge: Cambridge University Press.
- 1979. 'Ergativity'. *Language*, 55: 59–138.
- 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.
- 1982. *Where Have all the Adjectives Gone?* Berlin: Mouton de Gruyter.
- 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- 1994. *Ergativity*. Cambridge: Cambridge University Press.
- DOKE, C. M. 1949. *A Grammar of Luvale*. Johannesburg: Witwatersrand University Press.
- DOL, PHILOMENA H. 1999. 'A grammar of Mybrat'. University of Leiden dissertation.
- DONALDSON, TAMSIN 1980. *Ngiyambaa: The Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- DONNELLAN, KEITH S. 1978. 'Speaker Reference, Descriptions, and Anaphora'. In Peter Cole (ed.), *Syntax and Semantics*, ix: *Pragmatics*, 47–68. New York: Academic Press.
- DONOHUE, MARK 1999. *A Grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
- DOUGHERTY, JENET W. D. 1983. *West Futuna-Aniwa: An Introduction to a Polynesian Outlier Language*. Berkeley: University of California Press.
- DROESE, ERNEST 1884. *Introduction to the Malto Language*. Agra: Secundra Orphanage Press.

- DU BOIS, JOHN W. 1980. 'Beyond Definiteness: The Trace of Identity in Discourse'. In Wallace L. Chafe (ed.), *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*, 203–74. Norwood, NJ: Ablex Publishing Company.
- DUMMETT, M. 1973. *Frege: Philosophy of Language*. London: Duckworth.
- DURIE, MARK 1985. *A Grammar of Acehnese: On the Basis of a Dialect of North Aceh*. Dordrecht: Foris.
- ELBERT, SAMUEL H., and PUKUI, MARYK 1979. *Hawaiian Grammar*. Honolulu: University of Hawaii Press.
- ERTESCHIK-SHIR, NOMI 1997. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- EVANS, NICHOLAS D. 1995. *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- EVERETT, DANIEL 1986. 'Pirahã'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, i. 200–48. Berlin: Mouton de Gruyter.
- and KERN, B. 1997. *Wari': The Pacaas Novos Language of Western Brazil*. London: Routledge.
- FACUNDES, SIDNEY DA S. 2000. 'The Language of the Apurinã People of Brazil (Maipure/ Arawak)'. State University of New York, Buffalo, dissertation.
- FALTZ, LEONARD M. 1977. 'Reflexivization: A Study in Universal Syntax'. University of California, Berkeley, datorial dissertation.
- FARRIS, EDWIN R. 1992. 'A Syntactic Sketch of Yosondúa Mixtec'. In C. Henry Bradley and Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages*, 11–172. Arlington: University of Texas and Summer Institute of Linguistics.
- FODOR, JANET D. 1970. *The Linguistic Description of Opaque Contexts*. New York: Garland.
- FOLEY, WILLIAM A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- 1991. *The Yimas Language of New Guinea*. Stanford, Calif.: Stanford University Press.
- FORCHHEIMER, P. 1953. *The Category of Person in Language*. Berlin: Walter de Gruyter.
- FORTESCUE, MICHAEL 1984. *West Greenlandic*. London: Croom Helm.
- FORTUNE, G. 1955. *An Analytical Grammar of Shona*. London: Longmans, Green and Co.
- FRACHTENBERG, LEO J. 1922. 'Siuslawan (Lower Umpqua)'. In F. Boas (ed.), *Handbook of American Indian languages*, part 2, 431–629. Oosterhout, NB: Anthropological Publications.
- FRAJZYNGIER, ZYGMUNT 1989. *A Grammar of Pero*. Berlin: Dietrich Reimer Verlag.
- 1993. *A Grammar of Mupun*. Berlin: Dietrich Reimer Verlag.
- FRANK, PAUL 1990. *Ika Syntax*. Arlington: University of Texas and Summer Institute of Linguistics.
- GAIR, JAMES W. 1970. *Colloquial Sinhalese Clause Structures*. The Hague: Mouton.
- 1998. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages*. Oxford: Oxford University Press.
- and PAOLILLO, JOHN C. 1997. *Sinhala*. Munich: Lincom Europa.
- GENSLER, ORIN D. MS. 'On reconstructing 2fsg *m in Afroasiatic'.
- GIRIDHAR, P. P. 1980. *Angami Grammar*. Mysore: Central Institute of Indian Languages.
- 1994. *Mao Naga Grammar*. Mysore: Central Institute of Indian Languages.

- GRVÓN, TALMY 1973. 'Opacity and Reference in Language: An Inquiry into the Role of Modalities'. In John P. Kimball (ed.), *Syntax and Semantics*, 2: 95–121. New York: Academic Press.
- 1979. *On Understanding Grammar*. New York: Academic Press.
- 1984. *Syntax: A Functional-Typological Introduction*, vol. i. Amsterdam: John Benjamins.
- 2001. *Syntax: An Introduction*, vol. i. Amsterdam: John Benjamins.
- GLINERT, L. 1989. *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.
- GODDARD, CLIFF, and WIERZBICKA, ANNA 1994. *Semantic Lexical Universals: Theory and Empirical Findings*. Amsterdam: John Benjamins.
- GODDARD, EVES (ed.) 1996. *Handbook of American Indians*, xvii: 17: *Language*. Washington: Smithsonian Institution.
- GODDARD, P. E. 1967. 'Athapaskan (Hupa)'. In F. Boas (ed.), *Handbook of American Indian Languages*, part 2, 85–204. Oosterhout, NB: Anthropological Publications.
- GONDA, J. 1954. 'The Original Character of the Indo-European Relative Pronoun *io*'. *Lingua*, 4: 1–41.
- 1955. 'Notes on the Indo-European *k^ui-* and *k^uo-* pronouns'. *Lingua*, 4: 241–85.
- GORDON, LYNN 1986. *Maricopa Morphology and Syntax*. Berkeley: University of California Press.
- GORGONYEV, Y. A. 1966. *The Khmer Language*. Moscow: Nauka Publishing House.
- GREEN, JOHN N. 1988. 'Spanish'. In Martin Harris and Nigel Vincent (eds.), *The Romance Languages*, 79–130. London: Routledge.
- GREGORES, EMMA, and SUÁREZ, JORGE A. 1967. *A Description of Colloquial Guaraní*. The Hague: Mouton.
- GRIERSON, GEORGE A. 1903. *Linguistic Survey of India*, iii, part 2: *Tibeto-Burman Family*. Reprinted in 1967 by Motilal Banarsidass, Delhi.
- 1904. *Linguistic Survey of India*, iii, part 3: *Kuki-chin Group*. Reprinted in 1967 by Motilal Banarsidass, Delhi.
- 1908. *Linguistic Survey of India*, iii, part 1: *Himalan Dialects, North Assam Group*. Reprinted in 1967 by Motilal Banarsidass, Delhi.
- GROVES, TERAB'ATA R., GROVES, GORDON W., and JACOBS, RODERICK 1985. *Kiribatese: An Outline Description*. Canberra: Australian National University.
- GRUZDEVA, EKATERINA 1998. *Nivkh*. Munich: Lincom Europa.
- GUILLAUME, GUSTAVE 1919. *Le Problème de l'article et sa solution dans la langue française*. Reissued in 1975. Quebec: Presses de l'Université Laval.
- GUIRADELLO, RAQUEL 1999. 'A Reference Grammar of Trumai'. Rice University dissertation.
- HAAS, MARY 1940. *Tunica Grammar*. New York: J. J. Augustin.
- HAGMAN, ROY 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University.
- HAHN, FERD 1908. *Grammar of Kurukh Language*. Reprinted by Mittal Publications, Delhi.
- HAIMAN, JOHN 1980. *Hua: A Papuan Language of Eastern Highlands*. Amsterdam: John Benjamins.
- HALLIDAY, M. A. K. 1973. *Explorations in the Functions of Language*. London: Arnold.
- HARDMAN, M. J. 1966. *Jaqaru: Outline of Phonological and Morphological Structure*. The Hague: Mouton.

- HARDMAN, M. J. 2000. *Jaqaru*. Munich: Lincom Europa.
- HARMS, PHILLIP L. 1994. *Epena Pedee Syntax*. Arlington: University of Texas Press and summer Institute of Linguistics.
- HARRIS, ALICE C. 1984. 'Georgian'. In William S. Chisholm, Jr, Louis T. Milic, and John A. C. Greppin (eds.), *Interrogativity*, 63–112. Amsterdam: John Benjamins.
- 1991. 'Overview on the History of the Kartvelian Languages'. In Alice C. Harris (ed.), *The Kartvelian Languages*, 9–83. Delmare, New York: Caravan Books.
- HARRISON, SHELDON R. 1976. *Mokilese Reference Grammar*. Honolulu: University Press of Hawaii.
- HASPELMATH, MARTIN 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.
- 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.
- MS. 'Explaining the Ditransitive Person-Role Constraint'.
- forthcoming. 'Coordination'. To appear in Timothy Shopen (ed.), *Language Typology and Linguistic Description*, 2nd edition. Cambridge: Cambridge University Press.
- HASPELMATH, MARTIN, DRYER, MATTHEW, GIL, DAVID, COMRIE, BERNARD (eds.) 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- HAWKINS, JOHN A. 1978. *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.
- HEATH, JEFFREY 1984. *Functional Grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- 1998. *A Grammar of Koyra Chiini*. Berlin: Mouton de Gruyter.
- HEGÈGE, CLAUDE 1982. *La Structure de langues*. Paris.
- HEIM, IRENE R. 1988. *The Semantics of Definite and Indefinite Noun Phrases*. New York: Garland.
- HEINE, BERND, CLAUDI, ULRIKE, and HÜNNEMEYER, FRIEDERIKE 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- HELMBRECHT, JOHANNES 1996. 'The Syntax of Personal Agreement in East Caucasian Languages'. *Sprachtypol. Univ. Frsch. (STUF)* 49: 127–48.
- 1999. 'The Typology of 1st Person Marking and its Cognitive Background'. In Masako K. Hiraga, Chris Sinha, and Sherman Wilcox (eds.), *Cultural, Psychological and Typological Issues in Cognitive Linguistics*, 285–97. Amsterdam: John Benjamins.
- forthcoming. 'Grammar and Function of *we*'. To appear in Anna Duszak (ed.), *US and THEM across Languages and Cultures*. Amsterdam: John Benjamins.
- HERCUS, L. A. 1982. *The Bagandji Language*. Canberra: Australian National University.
- HEWITT, BRIAN G. 1979. *Abkhaz*. Amsterdam: North-Holland.
- 1995. *Georgian: A Structural Reference Grammar*. Amsterdam: John Benjamins.
- HILL, HARRIET 1995. 'Pronouns and Reported Speech in Adioukrou'. *Journal of West African Languages*, 25: 87–106.
- HILL, KEN C., SEKAQUAPTEWA, E., and BLACK, M. E. (eds.) 1998. *Hopi Dictionary*. Tucson: University of Arizona Press.
- HINDS, JOHN 1986. *Japanese*. London: Croom Helm.
- HINTIKKA, J. 1974. 'Questions about Questions'. In M. K. Munitz and P. K. Unger (eds.), *Semantics and Philosophy*, 103–58. New York: New York University Press.
- HJELMSLEV, LOUIS 1928. *Principes de grammaire Générale*. Copenhagen: Hält and Søn.
- HOFF, B. J. 1968. *The Carib Language*. The Hague: Martinus Nijhoff.
- HOFFMANN, C. 1963. *A Grammar of Margi Language*. London: Oxford University Press.
- HOFFMANN, JOHN 1903. *Mundari Grammar*. Calcutta: Bengal Secretariat Press.

- HOIJER, HARRY 1933. *Tonkawa: An Indian Language of Texas*. New York: Columbia University Press.
- HOLLIS, ALFRED C. 1905. *The Masai: Their Language and Folklore*. Oxford: Clarendon Press.
- HOPPER, P. L., and THOMPSON, S. A. 1984. 'The Discourse Basis for Lexical Categories in Universal Grammar'. *Language*, 60: 703–52.
- HOUIS, M. 1977. 'Plan et description systématique des langues négro-africaines'. *Afrique et Langage*, 7: 5–65.
- HUANG, YAN 2000. *Anaphora: A Cross-Linguistic Approach*. Oxford: Oxford University Press.
- HUDDLESTONE, RODNEY, and PULLUM, GEOFFREY K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- HUMBOLDT, ALEXANDER VON 1830. 'Über den Dualis'. Reproduced in *Wilhelm von Humboldts Gesammelte Schriften*. ed. by Königlich Preussische Akademie der Wissenschaften, vi. 4–30, 1907. Berlin: B. Behr.
- HYMAN, LARRY M., and COMRIE, BERNARD 1981. 'Logophoric Reference in Gokana'. *Journal of African Languages and Linguistics* 3: 19–37.
- INNES, GORDON 1966. *An Introduction to Grebo*. London: University of London.
- ISRAEL, M. 1979. *A Grammar of Kuvī Language*. Trivandrum: Dravidian Linguistics Association.
- JACKENDOFF, RAY S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- 1992. *Languages of the Mind*. Cambridge, Mass.: MIT Press.
- JACOBSON, SVEN A. 1995. *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language* (with Yup'ik readings written by Anna W. Jacobson). Fairbanks: Alaska Native Language Center, University of Alaska.
- JACQUESSON, FRANÇOIS 2001. 'Person-Marking in TB Languages of North-Eastern India'. *Linguistics of the Tibeto-Burman Area*, 24: 113–44.
- JAGGAR, PHILIP J., and BUBA, MALAMI 1994. 'The Space and Time Adverbials *nan/can* in Hausa: Cracking the Deictic Code'. *Language Sciences*, 16: 387–421.
- JAKOBSON, ROMAN 1971. *Selected Writings*, vol. ii. The Hague: Mouton.
- JELINEK, ELOISE 1984. 'Empty Categories, Case and Configurationality'. *Natural Language and Linguistic Theory* 2: 39–76.
- and DEMERS, RICHARD A. 1994. 'Predicates and Pronominal Arguments in Straits Salish'. *Language*, 70: 697–736.
- JESPERSEN, OTTO 1922. *Language: Its Nature, Development and Origin*. London: George Allen and Unwin.
- 1924. *The Philosophy of Grammar*. London: George Allen and Unwin.
- 1943. *A Modern English Grammar on Historical Principles*, part 7, *Syntax*. Copenhagen: Munksgaard.
- KAKATI, BANIKANTA 1972. *Assamese: Its Formation and Development*. Gauhati: Lawyer's Book Stall.
- KAKUMASU, J. 1986. 'Urubu-Kaapor'. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, i. 326–403. Berlin: Mouton de Gruyter.
- KAPFO, K. 1993. 'A Descriptive Analysis of Khezha'. University of Mysore dissertation.
- KARCEVSKI, SERGE 1969. 'Sur la parataxe et la syntaxe en russe'. In R. Godel (ed.), *A Geneva School Reader*. Bloomington: Indiana University Press.

- KATZ, JERROLD J., and POSTAL, PAUL M. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press.
- KEEGAN, JOHN M. 1997. *A Reference Grammar of Mbay*. Munich: Lincom Europa.
- KEEN, S. 1983. 'Yukulta'. In R. W. D. Dixon and Barry J. Blake (eds.), *Handbook of Australian Languages*, iii. 191–304. Amsterdam: John Benjamins.
- KENESEI, ISTVÁN, VAGO, ROBERT M., and FENYVESI, ANNA. 1998. *Hungarian*. London: Routledge.
- KEY, HAROLD H. 1967. *Morphology of Cayuvaya*. The Hague: Mouton.
- KIBRIK, ALEXANDER E. 1996. *Godoberi*. Munich: Lincom Europa.
- KIMBALL, G. D. 1991. *Koasati Grammar*. Lincoln: University of Nebraska Press.
- KLAIMAN, M. H. 1977. 'Bengali Syntax: Possible Dravidian Influence'. *International Journal of Dravidian Linguistics*, 6: 303–7.
- KLAMER, M. 1998. *A Grammar of Kambara*. Berlin: Mouton de Gruyter.
- KLIMA, EDWARD S. 1964. 'Negation in English'. In J. A. Fodor and J. J. Katz (eds.), *The Structure of Language*, 246–323. New York: Prentice-Hall.
- KORDIĆ, SANJEŽANA 1997. *Serbo-Croatian*. Munich: Lincom Europa.
- KOSHAL, SANYUKTA 1979. *Ladakhi Grammar*. Delhi: Motilal Banarsidass.
- KRISHNAMURTI, BHADRIRAJU 1968. 'Dravidian Personal Pronouns'. In B. Krishnamurti (ed.), *Studies in Indian Linguistics*, 189–205. Annamalainagar: Annamalai University.
- KUNO, SUSAMU 1976. 'Subject, Theme and the Speaker's Empathy: A Re-examination of Relativization Phenomena'. In C. Li (ed.), *Subject and Topic*, 417–44. New York: Academic Press.
- LAKOFF, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- LANGACKER, RONALD L. 1976. *Non-distinct Arguments in Uto-Aztecan*. Berkeley: University of California Press.
- 1977. *An Overview of Uto-Aztecan Grammar*. Arlington: University of Texas and Summer Institute of Linguistics.
- LAXMI BAI, B. 1983. 'Some Notes on Correlative Constructions in Dravidian'. In V. Z. Acson and R. L. Leed (eds.), *For Gordon Fairbanks*. Honolulu: University of Hawaii Press.
- LEE, KEE-DONG 1975. *Kusaiean Reference Grammar*. Honolulu: University of Hawaii Press.
- LEVINSON, STEPHEN C. 2000. *Presumptive Meaning: The Theory of Generalized Conversational Implications*. Cambridge, Mass.: MIT Press.
- LEWIS, G. L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
- LI, CHARLES, and THOMPSON, SANDRA A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- LICHTENBERK, FRANTISEK 1983. *A Grammar of Manam*. Honolulu: University of Hawaii Press.
- 2000a. 'Inclusory pronominals'. *Oceanic Linguistics*, 39: 1–32.
- 2000b. 'Reciprocals without Reflexives'. In Zygmunt Frajzyngier and Tracy S. Curl (eds.), *Reciprocals: Forms and Functions*, 31–62. Amsterdam: John Benjamins.
- LINDENFELD, JACQUILINE 1974. *Yaqi Syntax*. Berkeley: University of California Press.
- LINSKY, L. 1977. *Names and Descriptions*. Chicago: University of Chicago Press.

- LIPINŠKI, EDWARD 1997. *Semitic Languages: Outline of a Comparative Grammar*. Louvain: Vitgeverij Peeters en Departement Oosterse Studies.
- LOJENGA, CONSTANCE K. 1994. *Ngiti: A Central Sudanic Language of Zaïre*. Cologne: Rüdiger Köppe Verlag.
- LORIMER, D. L. R. 1935. *The Burushaski Language*. Oslo: H. Ascheong and Co.
- LOUST, ÉMILE 1918. *Étude sur le dialecte berbère des Ntifa. Grammaire, textes*. Paris: Leroux.
- LUCAS, JOHANNES 1937. *A Study of the Kanuri Language*. Oxford: Oxford University Press.
- LUMSDEN, M. 1988. *Existential Sentences*. London: Croom Helm.
- LYNCH, JOHN 1998. *Pacific Languages: An Introduction*. Honolulu: University of Hawaii Press.
- LYONS, CHRISTOPHER 1999. *Definiteness*. Cambridge: Cambridge University Press.
- LYONS, JOHN 1968. *Introduction to Theoretical Linguistics*. London: Cambridge University Press.
- 1977. *Semantics*, 1 and 2. London: Cambridge University Press.
- MACAULAY, M. 1996. *A Grammar of Chalcatongo Mixtec*. Berkeley: University of California Press.
- MCCLOSKEY, JAMES, and HALE, KENNETH 1984. 'On the Syntax of Person-Number Inflection in Modern Irish'. *Natural Language and Linguistic Theory*, 1: 487–533.
- MACDONALD, LORNA 1990. *A Grammar of Tauya*. Berlin: Mouton de Gruyter.
- McKAUGHAN, HOWARD P. 1972. 'Semantic Components of Pronoun Systems: Maranao'. *Word*, 15: 101–2.
- McKAY, GRAHAM R. 1978. 'Pronominal Person and Number Categories in Rembarrnga and Djeebbana'. *Oceanic Linguistics*, 17: 27–37.
- McLENDON, S. 1975. *A Grammar of Eastern Pomo*. Berkeley: University of California Press.
- MADAN, A. C. 1918. *A Handbook of Swahili Language as Spoken at Zanzibar*. London: Society for Promoting Christian Knowledge.
- MAHAPATRA, B. P. 1979. *Malto: An Ethnosemantic Study*. Mysore: Central Institute of Indian Languages.
- MALLINSON, GRAHAM, and BLAKE, BARRY J. 1981. *Language Typology: Cross-Linguistic Studies in Syntax*. Amsterdam: North-Holland.
- MARCHESE, LYNELL 1986. 'The Pronominal System of Godie'. In Ursula Weisemann (ed.), 217–56.
- MARTIN, SAMUEL E. 1975. *A Reference Grammar of Japanese*. New Haven: Yale University Press.
- MASICA, COLIN P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- MASLOVA, ELENA 1998. 'A Grammar of Kolyma Yukaghir'. MS, University of Bielefeld.
- MATISOFF, JAMES A. 1973. *The Grammar of Lahu*. Berkeley: University of California Press.
- MATTHEWS, STEPHEN, and YIP, VIRGINIA 1994. *Cantonese*. London: Routledge.
- MATTHEWSON, LISA 1998. *Determiner Systems and Quantificational Strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.
- MEIRA, S. 1999. 'A Grammar of Tiriyó'. Rice University dissertation.

- MERLAN, FRANCESCA 1994. *A Grammar of Wardman*. Berlin: Mouton de Gruyter.
- 1989. *Mangarayi*. London: Routledge.
- MILLER, WICK R. 1996. 'Sketch of Shoshone, a Uto-Aztecan Language'. In Eves Goddard (ed.), 325–63.
- MITCHELL, T. F. 1956. *An Introduction to Egyptian Colloquial Arabic*. Oxford: Clarendon Press.
- MITHUN, MARIANNE 1988. 'The Development of Bound Pronominal Paradigms'. In Winfred P. Lehmann and Helen-Jo J. Hewitt (eds.) *Language Typology 1988: Typological Models in Reconstruction*, 85–104. Amsterdam: John Benjamins.
- 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.
- MİYAKA, OSAHITO 1996. 'Sketch of Central Alaskan Yupik, an Eskimoan Language'. In Eves Goddard (ed.), 325–63.
- MORAVCSIK, EDITH 1978. 'Agreement'. In Joseph H. Greenberg (ed.), *Universals of Human Language*, iv: Syntax, 331–74. Stanford, Calif.: Stanford University Press.
- MORSE, NANCEY L., and MAXWELL, MICHAEL B. 1999. *Cubeo Grammar*. Arlington: University of Texas and Summer Institute of Linguistics.
- MOSCATI, SABOTINO, SPITALEC, ANTON, ULLENDORFF, EDWARD, and VON SODEN, WOLFRAM 1964. *An Introduction to the Comparative Grammar of Semitic Languages: Phonology and Morphology*. Wiesbaden: Otto Harrassowitz.
- MOSEL, ULRIKE, and HOVDHAUGEN, EVEN 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- MOSHINSKY, JULIUS 1974. *A Grammar of Southeastern Pomo*. Berkeley: University of California Press.
- MOUS, MARTEN 1993. *A Grammar of Iraqw*. Hamburg: Helmut Buske Verlag.
- MÜHLHÄUSLER, PETER, and HARRÉ, ROM 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford: Basil Blackwell.
- NAGARAJA, K. S. 1985. *Khasi: A Descriptive Analysis*. Pune: Deccan College.
- NATER, H. F. 1984. *The Bella Coola Language*. Ottawa: National Museums of Canada.
- NAU, NICOLE 1998. *Latvian*. Munich: Lincom Europa.
- NEWMAN, PAUL 2000. *The Hausa Language*. New Haven: Yale University Press.
- NEWMAN, STANLEY 1996. 'Sketch of the Zuni Language'. In Eves Goddard (ed.), 483–506.
- NGUYEN, DINH-HOA 1997. *Vietnamese*. Amsterdam: John Benjamins.
- NICHOLS, JOHANNA 1986. 'Head-Marking and Dependent-Marking Grammars'. *Language*, 62: 56–119.
- NOONAN, MICHAEL 1992. *Lango*. Berlin: Mouton de Gruyter.
- NORDINGER, RACHEL 1998. *A Grammar of Wambaya, Northern Territory*. Canberra: Australian National University.
- OLIVERIO, GIULIA R. M. 1997. 'A Grammar and Dictionary of Tutelo'. University of Kansas dissertation.
- ONENKO, SULUNGU N. 1980. *Nanajsko-Russkij slovar*. Moscow: Russkij jazyk.
- OSADA, TOSHIKI 1992. *A Reference Grammar of Mundari*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- OSBORNE, C. R. 1974. *The Tiwi Language*. Canberra: Australian Institute of Aboriginal Studies.

- OSUMI, MIDORI 1995. *Tinrin Grammar*. Honolulu: University of Hawaii Press.
- OWENS, JONATHAN 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske Verlag.
- PANDHARIPANDE, RAJESHWARI V. 1997. *Marathi*. London: Routledge.
- PARKER, ELIZABETH 1986. 'Mundani Pronouns'. In Ursula Weisemann (ed.), 131–65.
- PARRET, HERMAN 1980. 'Demonstratives and the I-sayer'. In Johan Van Der Auwera (ed.), *The Semantics of Determiners*, 96–111. London: Croom Helm.
- PAYNE, D. L. and PAYNE, T. E. 1982. 'Yagua'. In D. C. Derbyshire and G. K. Pullum (eds.), *Handbook of Amazonian Languages*, ii. 249–474. Berlin: Mouton de Gruyter.
- PENCHEON, THOMAS G. 1973. *Tamazight of the Ayt Nahir*. Los Angeles: Undena Publications.
- PITKIN, HARVEY 1984. *Wintu Grammar*. Berkeley: University of California Press.
- PLANK, FRANS 1989. 'On Humboldt on the Dual'. In R. Corrigan, F. Eckman, and M. Noonan (eds.), *Linguistic Categorization*, 293–333. Amsterdam: John Benjamins.
- 1996. 'Domains of the Dual, in Maltese and in General'. *Vista di Linguistica*, 8(1): 123–40.
- and SCHELLINGER, WOLFGANG 1997. 'The Uneven Distribution of Gender over Number: Greenberg Nos. 37 and 45'. *Linguistic Typology*, 1: 53–101.
- POPPE, NICHOLAS 1960. *Buriat Grammar*. Bloomington: Indiana University.
- POSTAL, PAUL M. 1970. 'On So-called Pronouns in English'. In Rodrick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, 56–82. Waltham, Mass.: Ginn and Company.
- 1979. *Some Syntactic Rules in Mohawk*. New York: Garland.
- POULOS, GEORGE 1990. *A Linguistic Analysis of Venda*. Pretoria: Via Afrika Limited.
- PREMSRIRAT, SUWILAI 1987. *Khmu, a Minority Language of Thailand*. Canberra: Australian National University.
- QUESADA, J. D. 1999. 'Chibchan, with Special Reference to Participant-Highlighting'. *Linguistic Typology* 3: 209–58.
- QUINE, WILLARD V. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.
- QUIRK, RANDOLF, GREENBAUM, SIDNEY, LEECH, GEOFFREY, and SVARTVIK, JAN 1972. *A Grammar of Contemporary English*. London: Longman.
- 1985. *A Comprehensive Grammar of English*. London: Longman.
- RADFORD, ANDREW 1993. 'Head-Hunting: On the Trail of the Nominal Janus'. In Greville G. Corbett, Norman M. Fraser, and Scott McGlashan (eds.), *Heads in Grammatical Theory*, 73–113. Cambridge: Cambridge University Press.
- RASTORGUEVA, V. S. 1964. *A Short Sketch of the Grammar of Persian*. Bloomington: Indiana University.
- REED, JUDY, and PAYNE, DAVID L. 1986. 'Asheninca (Campa) Pronominals'. In Ursula Wiesemann (ed.), 323–31.
- REESINK, GER P. 1987. *Structures and their Functions in Usan*. Amsterdam: John Benjamins.
- REHG, KENNETH L. 1981. *Ponapean Reference Grammar*. Honolulu: University of Hawaii Press.
- REINHART, TANYA 1981. 'Pragmatics and Linguistics: An Analysis of Semantic Topics'. *Philosophica*, 27: 53–94.
- and REULAND, ERIC 1993. 'Reflexivity'. *Linguistic Inquiry*, 24: 657–720.
- RENNISON, JOHN R. 1997. *Koromfe*. London: Routledge.

- REULAND, ERIC 2000. 'The Five Structures of Grammar: Anaphoric Relations'. In Zygmunt Frajzyngier and Traci S. Curl (eds.), *Reflexives: Forms and Functions*, 1–40. Amsterdam: John Benjamins.
- RIGTER, B. 1980. 'Determiners and the Syntax of Pronominals, Relativization and Modifier shift'. In Johan Van Der Auwera (ed.), *The Semantics of Determiners*, 135–55. London: Croom-Helm.
- ROBERTS, JOHN R. 1987. *Amele*. London: Croom Helm.
- ROMERO-FIGEROA, ANDRÉS 1997. *A Reference Grammar of Warao*. Munich: Lincom Europa.
- ROOD, DAVID S. 1976. *Wichita Grammar*. New York: Garland.
- ROSCH, E. 1978. 'Principles of Categorization'. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum Associates.
- SALTARELLI, MARIO 1988. *Basque*. London: Croom Helm.
- SAPIR, J. DAVID 1965. *A Grammar of Diola-Fogny*. Cambridge: Cambridge University Press.
- SASTRY, G. D. P. 1984. *Mishmi Grammar*. Mysore: Central Institute of Indian Languages.
- SCHACHTER, PAUL, and OTANES, FE T. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.
- SCHAUB, WILLI 1985. *Babungo*. London: Croom Helm.
- SCHUH, RUSSELL G. 1998. *A Grammar of Miya*. Berkeley: University of California Press.
- SCHWARTZ, LINDA 1988. 'Asymmetrical Feature Distribution in Pronominal "coordinations"'. In Michael Barlow and Charles A. Ferguson (eds.), *Agreement in Natural Language*, 237–49. Stanford, Calif.: Centre for the Study of Language and Information.
- SEILER, WALTER 1985. *Imonda, A Papuan Language*. Canberra: Australian National University.
- SELLS, PETER 1987. 'Aspects of Logophoricity'. *Linguistic Inquiry*, 18: 445–79.
- SENF, GÜNTER 1986. *Kilivila: The Language of Trobriand Islands*. Berlin: Mouton de Gruyter.
- SHOWALTER, CATHERINE 1986. 'Pronouns in Lyle'. In Ursula Weisemann (ed.), 205–16.
- SILVERSTEIN, MICHAEL 1976. 'Hierarchy of Features and Ergativity'. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*, 112–71. Canberra: Australian Institute of Aboriginal Studies.
- SIMONS, LINDA 1986. 'The Pronouns of To'abaita (Solomon Islands)'. In Ursula Weisemann (ed.), 21–35.
- SMEETS, CATHARINA J. M. 1989. 'A Mapuche Grammar'. Rijksuniversiteit Leiden, dissertation.
- SMITH-STARK, T. CEDRIC 1974. 'The Plurality Split'. In Michael W. La Galy, Robert A. Fox, and Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting, Chicago Linguistic Society*, 657–71. Chicago: Chicago Linguistic Society.
- SPEIJER, J. S. 1884. *Sanskrit Syntax*. Reprinted in 1973. Delhi: Motilal Banarsidass.
- STEPHENS, JANIG 1993. 'Breton'. In Martin J. Ball (ed.), *The Celtic Languages*, 349–409. London: Routledge.
- STERLING, L. 1993. *Switch-Reference and Discourse Representation*. Cambridge: Cambridge University Press.
- STEVENSON, R. C. 1969. *Bagirmi Grammar*. Khartoum: University of Khartoum.

- SUGAMOTO, N. 1989. 'Pronominality: A Noun–Pronoun Continuum'. In R. Corregan, F. Eckman, and M. Noonan (eds.), *Linguistic Categorization*, 267–91. Amsterdam: John Benjamins.
- TAMURA, SUZUKO 2000. *The Ainu Language*. Tokyo: Sanseido.
- TERRILL, ANGELA 1999. 'A grammar of Lavukaleve: A Papuan Language of the Solomon Islands'. Australian National University dissertation.
- 2001. 'Activation Levels in Lavukaleve Demonstratives: *oia* versus *foia*'. *Linguistic Typology*, 5: 67–90.
- THOMPSON, LAURENCE C. 1991. *A Vietnamese Reference Grammar*. Honolulu: University of Hawaii Press.
- TOPPING, DONALD M. 1973. *Chamorro Reference Grammar*. Honolulu: University of Hawaii Press.
- TSUNODA, TASAKA 1995. *The Djaru Language of Kimberley, Western Australia*. Australian National University.
- UBRIATOVA, E. I. (ed.) 1982. *Grammatika sovremennogo jakutskogo literaturnogo jazyka*, i. Moscow: Nauka.
- ULTAN, RUSSELL 1978. 'Some General Characteristics of Interrogative Systems'. In Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik (eds.), *Universals of Human language*, iv: *Syntax*, 211–48. Stanford, Calif.: Stanford University Press.
- VAN DER BERG, HELMA 1995. *A Grammar of Hunzib*. Munich: Lincom Europa.
- VAN DRIEM, G. 1993. *A Grammar of Dumi*. Berlin: Mouton de Gruyter.
- VAN EIJK, JAN 1997. *The Lillooet Language*. Vancouver: UBC Press.
- VAN VALIN, ROBERT D. JR. 1993. 'A Synopsis of Role and Reference Grammar'. In Robert D. Van Valin, Jr. (ed.), *Advances in Role and Reference Grammar*. Amsterdam: John Benjamins.
- VEERMAN-LEICHSENRING, ANNETTE 2000. 'Popolocan Independent Personal Pronouns: Comparison and Reconstruction'. *International Journal of American Languages*, 66: 318–59.
- VINCENT, N. 1988. 'Italian'. In M. Harris and N. Vincent (eds.), *The Romance Languages*. Beckenham: Croom Helm.
- VINEIS, EDOARDO 1998. 'Latin'. In Anna G. Ramat and Paolo Ramat (eds.), *The Indo-European Languages*, 261–321. London: Routledge.
- VOORHOEVE, C.-L. 1965. *The Flamingo Bay Dialect of the Asmat Language*. The Hague: Martinus Nijhoff.
- WALES, KATIE 1996. *Personal Pronouns in Present-Day English*. Cambridge: Cambridge University Press.
- WALI, KASHI, and KOUL, ONKAR N. 1997. *Kashmiri*. London: Routledge.
- WALKER, ALAN T. 1982. *A Grammar of Sawu*. Jakarta: University of Alma Jaya.
- WALKER, DALE F. 1976. *A Grammar of Lampung Language : The Pesisir Dialect of Way Lima*. Jakarta: Badan Penyelenggara Seri NUSA.
- WEBER, D. J. 1986. *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. Berkeley: University of California Press.
- WIESEMANN, URSULA 1986a. 'Grammaticalized Coreference'. In Wieseemann (ed.), 437–63.
- (ed.) 1986b. *Pronominal Systems*. Tübingen: Günter Narr Verlag.
- WILSON, PATRISIA R. 1980. *Ambulas Grammar*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

- WOLFART, H. CHRISTOPH 1973. *Plains Cree: A Grammatical Study*. Philadelphia: American Philosophical Society.
- 1996. 'Sketch of Cree, an Algonquian language'. In Eves Goddard (ed.), 390–439.
- WORDRICK, F. J. F. 1982. *The Yindjibarndi Language*. Canberra: Australian National University.
- XU, LIEJIONG 1990. 'Remarks on LF Movement in Chinese Questions'. *Linguistics*, 28: 355–82.
- ZAHARLICK, ANNA M. 1982. *Tanoan Studies: Passive Sentences in Picuris*. Ohio State University Working Papers in Linguistics, 26: 34–48.
- ZEWEN, FRANÇOIS-XAVIER N. 1977. *The Marshallese Language: A Study of its Phonology, Morphology and Syntax*. Berlin: Verlag von Dietrich Reimer.
- ZHU, SHENSHENG 1997. 'A Dynamic Semantics Theory of Chinese Anaphora'. University of Arizona dissertation.
- ZIGMOND, MAURICE L., BOOTH, CURTIS G., and MUNRO, PAMELA 1991. *Kawaiisu: A Grammar and Dictionary with Texts*. Berkeley: University of California Press.
- ŽIRKOV, L. I. 1955. *Lakskij jazyk*. Moscow: Akademija Nauk SSSR.

Index of Authors

- Abbott C. 114, 287
Abbott M. 284
Abdel-Hafiz 197, 286
Abondolo 148–9, 181
Adelaar 81, 103, 140, 142, 169, 197, 254, 288
Aikhenvald 25, 244
Ajello 147, 182
Alexander 288
Ambros 234
Andersen 161
Anderson 14, 16, 140, 177
Anscombe 40–1
Aoki 286
Asher 91, 137, 284
Austin 84, 241
Awobuluyi 291
Azzopardi-Alexander 284
- Baker 17, 24
Baldi 147
Barker 235
Baroja 232
Bauer 98, 141, 143, 158, 178, 245, 255, 284
Beaumont 222
Benveniste 40, 92, 132
Bergsland 277
Bernsten 18
Besnier 143, 178, 189, 214, 290
Bhat 2, 34, 79, 81, 83, 85, 87, 100, 115, 130,
135, 170, 190–1, 205–6, 218, 220, 223, 227,
231, 254, 260, 282, 284
Bhatia 287
Biligiri 133, 141, 180, 283
Blake 46, 121, 126–8, 287
Boas 97, 105, 180, 283
Boer 208
Bogoras 98, 279
Bohnhoff 62, 64–5
Booth 282
Borg 284
Borgman 125, 237
Bosch 291
Bowden 98, 158, 243, 289
Bradley 30, 101, 112
- Braine 89
Brandon 235–6, 239
Brauner 288
Bray 278
Breen 287
Bresnan 17, 28
Bright 254
Broadbent 285
Bromley 136, 279
Bruce 136–7, 167, 277
Buba 145
Burchuladze 26
Burling 140, 197, 255, 257, 280
Burquest 63, 239
Burrow 148
Bybee 171
- Campbell, L. 171, 198, 287
Campbell, S. 289
Capell 139, 170, 285
Carlson 88, 289
Casad 29, 239, 255, 261, 291
Chafe 88
Chapman 287
Charney 141
Chesterman 41, 219
Childs 137, 283
Chomsky 75
Claudi 170
Clements 71
Cole 217, 288
Comrie 12, 62–3, 65, 73, 126, 134, 185, 204,
209, 270–1, 277
Conrad 98, 278
Cook 246–7
Cooke 112
Corbett 97, 102–7
Cornish 40
Cowan 135, 288
Craig 25, 29, 139, 282
Crowley 45, 103, 196, 287
Culy 59, 61, 65, 72–3
Curme 280
Curnow 56, 125, 146, 198, 278

- Cyffer 282
 Cysouw 16, 93

 Das Gupta 123
 Dasgupta 186, 230, 262, 266, 268
 Davies 50, 98, 101, 160, 283
 Davis 209
 Dayley 25, 163, 192, 194, 227, 290
 Dedrick 29, 239, 255, 261, 291
 DeLancey 121
 Demers 17
 Dench 285
 Denwood 139, 169, 289
 Derbyshire 135, 235, 281, 287
 Dickens 282
 Diessel 50, 172, 177, 179, 181, 193, 204
 Diller 30
 Dimmendaal 290
 Dixon 4–5, 7, 116, 126–7, 130, 142, 148, 164,
 178, 192–3, 197, 247–8, 280, 291
 Doke 98, 284
 Dol 98, 157–8, 285
 Donaldson 240, 264, 286
 Donnellan 41, 217–18, 220
 Donohue 20, 142, 178, 180, 289
 Dougherty 142, 158, 178, 280
 Droese 253
 Dryer 134, 277
 Du Bois 42, 203
 Du Feu 288
 Dummett 222
 Durie 97, 105

 Ebert 278
 Eckert 290
 Elbert 136, 143, 158, 290
 Emeneau 148
 Eoin 282
 Erteschik-Shir 42
 Evans 193, 282
 Everett 82, 87–8, 139, 188, 287, 291

 Facundes 117, 137, 158, 277
 Faltz 76
 Farris 112, 160, 171, 193, 285
 Fenyvesi 81, 137
 Fernandis 288
 Fodor 222
 Foley 17, 143, 291
 Forchheimer 132
 Fortescue 75, 281

 Fortune 5
 Frachtenberg 235
 Frajzyngier 52, 55, 59–60, 65, 73, 139, 142,
 204, 237, 286–7
 Francis 285
 Frank 171, 236, 264–5, 281
 Franklin 282
 Frege 222–3

 Gair 3, 76–7, 182, 185, 239, 288
 Gamble 291
 Geary 280
 Gil 277
 Gensler 110
 Gil 134
 Giridhar 14, 93–4, 133, 142, 166, 245,
 277, 286
 Givón 16, 28, 41, 130, 216–18
 Glinert 140, 281
 Goddard 25, 117, 124
 Gonda 269
 Gordon 75, 114, 135, 141, 167, 169, 233, 285
 Gorgoniyev 254
 Green 110, 140, 289
 Greenbaum 2, 155, 280
 Gregores 106, 281
 Grierson 245–6, 257
 Groves, G. W. 189, 283
 Groves, T. R. 189, 283
 Gruzdeva 76, 107, 137, 198, 237, 240, 286
 Guillaume 219
 Guirardello 137, 171, 243, 248, 289

 Haas 109, 140, 197, 290
 Hagman 19, 50, 101, 286
 Hahn 100
 Haiman 45, 194
 Hale 28
 Halliday 211
 Hardman 137, 140, 157, 183, 237, 246,
 254, 282
 Harms 171, 237, 242, 246, 280
 Harré 132
 Harris 143, 208, 280
 Harrison 158, 178, 286
 Haspelmath 7, 76, 98, 104, 117, 134, 136, 143,
 155, 182, 184, 190, 192, 195, 221, 226–9,
 231, 241, 251–3, 257–62, 277, 284
 Haviland 281
 Hawkins 202–4, 209, 216
 Heath 77, 110, 140, 144, 265, 283, 286

- Hegège 115
 Heim 42, 203
 Heine 170
 Helmbrecht 26, 97, 123, 132
 Hercus 241
 Hewitt 22, 136, 178, 244, 280
 Hewson 285
 Hill, H. 65
 Hill, K. C. 141
 Hinch 139, 170, 285
 Hinds 31, 171, 254, 282
 Hintikka 208
 Hjemslev 219
 Hoff 237
 Hoffmann, C. 104, 234, 255, 285
 Hoffmann, J. 191, 246
 Hoijer 124, 188, 289
 Hollis 244
 Hopper 2
 Houis 5
 Hovdhaugen 143, 158, 178
 Huang 61–2, 69–70, 72, 75–6, 81, 83
 Huddleston 93
 Hudson 290
 Humboldt 107
 Hünemeyer 170
 Huttar 286
 Hyman 62–3, 65

 Innes 237, 281
 Israel 148, 182

 Jackendoff 190, 216–17
 Jacobs, M. 288
 Jacobs, R. 189, 283
 Jacobson 107
 Jacquesson 27, 115
 Jaggar 145
 Jakobson 40
 Jelinek 17
 Jespersen 2, 16, 39, 48, 71, 92–3, 207
 Jones, P. 278
 Jones, W. 278
 Josephs 287

 Kakati 253
 Kakumasa 76, 290
 Kapfo 153, 156, 218
 Karcevski 227
 Karjalainen 280
 Katz 228

 Keegan 26, 171, 179, 285
 Keen 135, 291
 Keenan 14, 16, 140, 177
 Kenesei 81, 137, 140, 281
 Kern 87, 139, 188, 291
 Key 25, 279
 Kibrik 76, 178, 192, 237, 264, 280
 Kimball 142, 179, 238, 283
 Klaiman 269
 Klamer 22, 114, 171, 282
 Klima 228, 231
 Kordic 137, 139, 288
 Koshal 142
 Koul 74
 Kozintseva 278
 Krishnamurti 100
 Krueger 279
 Kruspe 288
 Kumari 91, 137, 284
 Kuno 122

 Lakoff 2, 130
 Langacker 85, 148, 242
 Laxmi Bai 269
 Leavitt 287
 Lee 183, 283
 Leech 2, 280
 Levinson 76–7
 Lewis 98, 141, 209, 290
 Li 54, 82, 279
 Lichtenberk 86, 88, 103–4, 247
 Lindenfeld 235, 239
 Linsky 206
 Lipiński 110
 Lojenga 61, 286
 Lorimer 198, 278
 Loust 110
 Lucas 244, 255
 Lumsden 217
 Lycan 208
 Lynch 143, 149
 Lyons, C. 47–51, 71, 181, 209, 210, 213
 Lyons, J. 2–3, 13, 16, 92, 132, 204

 Macaulay 96, 285
 McCloskey 28
 MacDonald 238, 246, 289
 McGinn 288
 McGregor, R. S. 282
 McGregor, W. 281
 McKaughan 102

- Mackay 289
 McKay 102
 McLendon 78, 139
 Madan 184
 Mahapatra 100
 Mainwaring 284
 Mallinson 121, 126, 128
 Marchese 5, 179, 280
 Martin 254
 Masica 23, 147, 242
 Maslova 154, 291
 Matisoff 180, 245, 283
 Matson 282
 Matthews 237, 264, 279
 Matthewson 221–2
 Maxwell 81, 172, 238
 Mchombo, 17, 28
 Meira 141, 144
 Merlan 191, 242–3, 284, 290
 Miller 241, 279
 Mitchell 129, 139, 278
 Mithun 84, 103, 117, 121
 Miyaoka 178, 238, 291
 Moravcsik 106, 129
 Morse 81, 172, 238, 279
 Moscati 148
 Mosel 143, 158, 178, 288
 Moshinsky 239, 265, 287
 Mous 54, 109, 156–7, 179, 282
 Mühlhäusler 132
 Munro 282
- Nagaraja 133, 180, 283
 Nater 26
 Nau 139, 284
 Nedjalkov 280
 Newman, P. 52, 139, 144–6, 178, 237, 264–5, 291
 Newman, S. 165
 Nguyen 111
 Nichols 22, 29, 282
 Nimbkar 18, 23–4
 Ningomba 79, 85, 115, 170, 191, 254, 284
 Noonan 142, 178, 247, 284
 Nordinger 166, 290
- Okell 278
 Oliverio 24, 123, 290
 Onenko 253
 Osada 14, 142, 167–8, 286
 Osborne 247
- Osumi 20, 105, 142, 158, 184, 198
 Otanes 158, 177, 234, 245, 254, 289
 Owens 76, 81, 139, 242, 287
- Pandharipande 135, 162, 284
 Paolillo 239, 288
 Parker 5, 59, 286
 Parret 40
 Payne, D. L. 135, 140, 185, 195, 278, 291
 Payne, T. E. 140, 291
 Penchoen 21, 140, 171, 244, 289
 Pitkin 5, 218, 236, 289
 Plank 107, 110–11
 Popjes, J. 279
 Popjes, Jo 279
 Poppe 135, 278
 Postal 48, 50, 114, 228
 Poulos 137, 159, 193, 290–1
 Premrsirat 109, 140, 283
 Press 279
 Pukui 136, 143, 158, 281
 Pullum 93
- Quesada 184
 Quine 217
 Quirk 2, 155, 228, 280
- Radford 50
 Rarhi 186, 256, 264, 266, 268
 Rastorgueva 140
 Reed 135, 185, 195, 278
 Reesink 170, 254, 290
 Rehg 193, 287
 Reichle 278
 Reinhart 42, 76, 81
 Rennison 140, 283
 Reuland 76, 81–2
 Rice 289
 Rigter 50
 Roberts 45, 88, 142, 277
 Robins 291
 Romero-Figeroa 98, 136, 142, 237, 239, 290
 Rood 261, 291
 Rosch 2, 130
 Rupp 279
- Saltarelli 51, 76, 135, 141, 177, 184, 242, 278
 Samarin 288
 Sapir 159, 279
 Sara 279

- Sastry 27, 135, 285
Saunders 209
Schachter 142, 158, 234, 245, 254, 289
Schaub 69, 73, 180, 278
Schellinger 110–11
Schuh 129, 139, 158, 285
Schwartz 104
Seiler 107, 184, 281
Seki 235
Sells 71
Senft 166, 182, 283
Showalter 5, 8, 175, 186, 187
Shaweevongse 289
Silverstein 126, 128
Simons 178, 254, 289
Smeets 136, 141, 163, 193, 198, 237, 284
Smith 281
Smith-Stark 105
Sneddon 281
Sohn 283
Solnit 282
Speijer 157, 181
Stephens 159
Sterling 43, 65, 72–3
Stevenson 81, 98, 198, 237, 278
Street 282
Strom 288
Suárez 106, 281
Sugamoto 31
Sulkala 280
Svartvik 2, 155, 280

Tamura 164, 185, 232, 265, 277
Taylor 286
Terrill 142, 185, 284
Thompson, L. 158, 235, 280
Thompson, S. A. 2, 54, 82, 279
Topping 142, 279
Tsunoda 122, 280
Tuggy 286

Ubrjatova 253
Ullendorff 148
Ultan 226

Vago 81, 137
Van Der Berg 75, 86, 281
Van Driem 140, 280, 284
Van Eijk 170, 284
Van Staden 289
Van Valin 226, 235–6, 283
Veerman-Leichsenring 103
Vincent 93, 148
Vineis 147, 182
Von Humboldt 17
Von Soden 148
Voorhoeve 106, 136, 278

Wales 3, 31
Wali 19, 74
Walker, A. T. 158, 179, 184, 288
Walker, D. F. 158, 213, 284
Weber 246
Wierzbicka 25
Wiesemann 62, 64–5, 71
Wilkins 286
Wilson 109, 140, 161, 277
Wolfart 101, 116, 124, 141, 144, 279
Wordrick 241

Xu 234

Yip 237, 264, 279

Zaharlick 117
Zewen 99, 142, 170, 285
Zheng 280
Zhu 82–3
Zigmond 142, 282
Žirkov 26

Index of Languages

- Abkhaz 22, 244
Acehnese 97, 99, 100, 105
Adioukrou 65
Ainu 164, 185, 232, 265, 277
Alamblak 136, 167, 172, 277
Aleut 277
Ambulas 109, 140, 161, 277
Amele 45, 88, 142, 277
Ancient Greek 155
Angami 14, 133–4, 139, 142, 245, 277
Angas 63
Apurinã 117, 136–7, 158, 277
Arabic 139, 234, 278
Armenian 147, 182, 278
Asheninca 135, 185, 195, 278
Asmat 106, 136, 278
Assamese 253, 257
Assuriní 236, 239
Athapare 278
Awa Pit 56, 125, 146, 198, 278
Aztec 17
- Babungo 69, 73, 180, 278
Bagandji 241
Bagirmi 81, 98, 237, 278
Barasano 278
Basque 51, 76, 135, 141, 177, 184, 242, 278
Bawm 278
Bella Coola 26
Bemba 41, 218, 219
Bengali 186, 230, 256–7, 262, 264, 266, 268–9
Boumaa Fijian 4, 142, 178, 280
Brahui 278
Breton 159, 161
Bukiyip 98, 278
Buriat 135, 278
Burmese 30, 111–12, 278
Burushaski 198, 278
- Caddo 88, 121
Canela-Khraho 279
Cantonese 237, 264, 279
Carib 135, 235, 237
- Catalan 181
Cayuvava 25, 279
Central Alaskan Yupik 107, 178, 238, 291
Chaldean 279
Chalcatongo Mixtec 96, 99, 285
Chamorro 142, 279
Camehuevi 279
Chicheŵa 17, 28
Chinantec 279
Chinese 54, 69, 76, 82, 234, 264, 279
Chukchee 98, 279
Chuvash 279
Colloquial German 181
Comanche 141
Cree 101, 116, 121, 124, 140–1, 144, 279
Cubeo 81, 172, 238, 279
- Diegueno 279
Diola-Fogny 159, 279
Diyari 84, 241
Djaru 122, 280
Dong 280
Donno so 59
Dumi 140, 280
Dutch 76, 82, 185, 209, 280
Dyirbal 126, 193, 197
- Eastern Pomo 77, 139, 239, 265, 287
Efik 61
Egyptian Arabic 129
English 8, 11–12, 24, 31, 41, 48–9, 50–1, 56, 59,
60–1, 63, 68, 71, 79, 81, 83, 86–7, 95, 98,
104, 125, 130, 134, 139, 144, 155–7, 161,
170, 177, 181, 189, 196, 204, 209, 211, 213,
219–21, 230, 232, 234, 240, 254, 259, 261,
266–7, 269, 271, 276, 280
Epena Pedee 171, 237, 242, 280
Erromangan 45, 51, 103, 280
Evenki 280
Ewe 71–2
- Finnish 129, 280
French 181, 280

- Garo 140, 197, 255, 257, 280
 Gbaya 73
 Georgian 136, 143, 178, 208, 280
 German 50–1, 181, 190, 241, 280
 Godie 5, 179, 280
 Godoberi 76, 143, 178, 180, 192, 237, 264, 280
 Gojri 256
 Gokana 62–4, 66, 74
 Gooniyandi 281
 Grebo 237, 281
 Greenlandic 281
 Guaraní 76, 106, 281
 Guaymí 184
 Guugu Yimidhirr 127, 281
 Gujarati 256
 Gurung 257
- Harar Oromo 76, 81, 139, 287
 Hausa 52, 139, 144–6, 178, 237, 264, 281
 Havyaka 27, 45–6, 49, 51, 58–61, 194, 243, 250, 270
 Hawaiian 136, 143, 158, 281
 Hebrew 140, 281
 Hindi 143, 262, 270, 281
 Hixkaryana 135, 235, 281
 Hopi 107, 141
 Hua 45, 194
 Hungarian 81, 106, 137, 140, 260, 281
 Hunzib 75, 86, 178, 184, 281
- Ide 27
 Igbo 72
 Ika 171, 236, 264–5, 281
 Ila 281
 Imonda 107, 184, 281
 Indonesian 281
 Ingush 282
 Iraqw 54, 109, 156, 161, 179, 282
 Irish 28, 107, 282
 Italian 76, 148
- Jacaltepec 25, 29, 139, 282
 Japanese 30–1, 111–12, 171, 209, 254, 257, 271, 282
 Jaqaru 137, 140, 157–8, 183, 237, 246, 254, 282
 Ju/hoan 282
 Juang 282
- Kachin 245
 Kamaiurá 235
- Kambara 22, 114, 171, 282
 Kannada 11, 13–14, 21, 27, 44–6, 49, 51, 55, 58–9, 67–8, 73, 76, 80–1, 83, 86, 98, 100, 109, 115, 122–3, 126, 136, 148, 162–3, 165, 172–3, 181, 192–4, 196–7, 207, 213, 220–1, 224, 226, 228, 243, 250, 252, 254, 258–60, 263–4, 269–70, 282
 Kanuri 244, 255, 282
 Karok 254
 Kashmiri 74, 256
 Kawaiisu 142, 282
 Kayah Li 282
 Kayaradild 193
 Kewa 282
 Khalkha 282
 Khambu 257
 Kharia 133, 135, 141, 180, 283
 Khasi 133–5, 180, 283
 Khezha 153–6, 164, 175, 198
 Khmer 109, 140, 180, 254
 Khmu 109, 140, 283
 Kilivila 166, 182, 283
 Kiribatese 189, 283
 Kisi 137, 283
 Kimball 142, 179, 238
 Klamath 235
 Koasati 142, 179, 180, 238, 283
 Kobon 50, 98, 101, 160, 283
 Kolyma Yukaghir 154, 156
 Konkani 256
 Korean 257, 283
 Koromfe 140, 283
 Koyraboro Senni 283
 Koyra Chiini 283
 Kunuz Nubian 197, 286
 Kurukh 100
 Kusaiean 183, 283
 Kuvi 148, 182
 Kwakiutl 97, 99–100, 105, 180, 283
- Ladakhi 142
 Lahu 101, 180, 245, 283
 Lak 26–7, 132, 134
 Lakhota 121, 226, 230, 235, 283
 Lampung 158, 184, 284
 Lango 142, 178, 247, 284
 Latin 115, 147–8, 182, 258, 262
 Latvian 139, 260, 284
 Lavukaleve 142, 185, 284
 Lele 62, 64, 66, 74, 239
 Lepcha 246, 284

- Lezgian 7–8, 76, 98, 136, 143, 184, 195,
260, 284
Lhoke 245
Lillooet 170, 284
Limbu 245, 284
Lithuanian 111, 209, 262
Lower Grand Valley Dani 136, 279
Luiseño 242
Luvale 98, 284
Lyle 5, 8, 175, 186
- Macushi 284
Maithili 256
Malayalam 91, 96, 137, 253, 260, 284
Maltese 284
Malto 100, 253
Manam 88, 247
Mandani 5
Mangarayi 242, 284
Manipuri 79, 85, 115, 170, 191, 254, 284
Mao Naga 93–4, 142, 166, 286
Maori 98, 140–1, 143, 158, 178, 245, 255, 284
Mapuche 136, 141, 163, 193, 198, 237, 284
Maranao 102
Marathi 18–19, 23, 67, 86, 135, 143, 162,
269, 284
Margi 104, 234, 246, 255, 285
Maricopa 75, 114, 135, 141, 167, 169, 233, 285
Maringe 149
Marshallese 99, 142, 170, 285
Marthuthunira 285
Masai 244
Maung 139, 170, 285
Maybrat 98, 157, 158, 285
Mbaya 26, 171, 179, 285
Micmac 285
Middle High German 262
Miri 245
Mishmi 27, 135, 285
Miwok 285
Miya 129, 139, 158, 285
Mohawk 114
Mokilese 158, 178, 286
Mparntwe Arrente 286
Mundani 286
Mundari 14, 142, 167, 168, 191, 246, 286
Mupun 52, 55, 59, 60–4, 66, 73, 139, 204,
237, 286
- Nahuatl 286
Nama Hottentot 19–20, 50, 101, 286
- Nanay 253
Ndyuka 286
Ngaanyatjara 148
Ngiti 61, 286
Ngiyambaa 240, 264, 286
Nhanda 127
Nicobarese 89
Nivkh 76, 107, 137, 198, 237, 240, 286
Niz Perce 286
Nkore-Kiga 286
Nocte 123
Norwegian 76
Ntifa 110
Nungubuyu 77, 110, 140, 144, 286
- Ocotepec Mixtec 285
Old French 181
Oneida 114, 287
Ossetic 257
- Paamese 196, 287
Palauan 287
Passamaquoddy 287
Paumari 287
Perené 195
Pero 65, 110, 139, 142, 287
Persian 140, 287
Picurís 117
Pipil 171, 198, 287
Pirahã 82, 88, 287
Pitta-Pitta 287
Pokot 232
Polish 107
Ponapean 193, 287
Portuguese 258
Punjabi 143, 287
- Quechua 140, 142, 197, 246, 288
- Rama 288
Ranghol 245–6
Rapanui 288
Rejang 288
Rembarrnga 102
Remo 288
Returã 288
Romanian 258
Russian 260
- Sahaptin 288
Samoan 143, 158, 178, 288

- Sango 61, 288
 Sanskrit 8, 92, 157–8, 181, 253, 255, 262
 Sanuma 125, 237
 Sawu 158, 179, 213, 288
 Sechelt 222
 Semelai 288
 Sentani 135, 288
 Serbo-Croatian 137, 139, 288
 Shona 5, 288
 Shoshone 241–2
 Sindhi 256
 Sinhalese 3, 76–7, 143, 182, 239, 269, 288
 Siuslaw 235
 Slave 289
 Sonora Yaqui 29, 239, 254, 261, 291
 Southeastern Pomo 287
 Spanish 51, 110–11, 117–18, 140, 289
 St'at'imcets 221
 Straits Salish 17
 Sunwal 245
 Supyire 88, 98, 171, 189, 230, 248, 289
 Swahili 184
- Taba 98, 158, 243, 289
 Taburi 72
 Tagalog 104, 142, 158, 177–8, 234, 245, 254, 289
 Tamazight 21, 140, 171–2, 244, 289
 Tamil 76, 148, 181, 253, 260, 269
 Tangut 121
 Tarma Quechua 81, 103, 169, 254, 287
 Tauya 238, 246, 289
 Telugu 76, 253, 260
 Thai 30, 111–12, 289
 Tibetan 139, 169, 245, 289
 Tidore 289
 Tinrin 20, 105, 142, 158, 184, 198
 Tiriyo 141, 144
 Tiwi 247
 To'abaita (Toqabaqita) 103, 178, 254, 289
 Toda 253
 Tonkawa 124, 188, 289
 Totonac 289
 Trumai 137, 171, 243, 247–8, 289
 Tuburi 73
 Tukang Besi 20, 142, 178, 180, 289
 Tulu 100–1, 148, 181, 194, 253, 260
- Tümpisa Shoshone 163, 194, 290
 Tunica 109, 140, 197, 290
 Turkana 290
 Turkish 98, 141, 209, 290
 Tutelo 24, 123, 290
 Tuvaluan 143, 158, 178, 189, 214, 290
 Tzutujil 25, 192, 227, 290
- Urubu Kaapor 76, 290
 Ucayali 195
 Usan 170, 254, 290
- Venda 137, 143, 159, 193, 290
 Vietnamese 111, 158, 235, 290
- Wambaya 166, 290
 Wangkawiru 290
 Warao 98, 136, 142, 239, 290
 Wardman 191, 290
 Warekena 25, 244
 Wari' 87, 139, 187, 291
 Warlpiri 17
 Welsh 28
 West Futuna Aniwa 142, 158, 178, 280
 West Greenlandic 75
 Wichita 261, 291
 Wikchamni 291
 Wintu 5, 218, 236, 291
- Yag Dii 62, 64–5
 Yagua 140, 291
 Yakut 253
 Yapese 46
 Yaqui 235, 239
 Yidiny 4, 7, 116, 127, 164, 192, 291
 Yimas 291
 Yindjibarndi 241
 Yoruba 291
 Yosondúa Mixtec 112, 160, 171, 193, 285
 Yuchi 291
 Yukaghir 291
 Yukulta 135, 291
 Yurok 291
- Zulu 143, 291
 Zuni 165

Index of Subjects

- addressee's prominence
 - inclusive form 124–5
 - statement vs. question 125
- African languages 182
- Afro-Asiatic languages 109, 128
- agreement 15, 16–18
 - consistency 23–4
 - first inclusive 27
 - grammatical-anaphoric 17–18
 - locutor vs. non-locutor 125, 146
 - obligatoriness 24
 - personal vs. nominal 23
 - singular for plural nouns 116–17
 - source 26–7
- Algonquian languages 116–17, 124
- American linguistics 17
- anaphoric coreference 62, 69, 74, 79, 80, 204, 212
- anaphoric pronoun
 - constraints regarding antecedents 67, 71
 - denoting coreference 71
 - endophoric third person 67
 - extended use 69–70
 - long-distance reflexive 69, 74, 183
 - non-obligatoriness 72–3
 - personal pronoun vs. demonstratives 74–5
 - pragmatic differences 185
 - subject constraint 69
 - vs. cataphoric pronoun 185
 - vs. demonstrative pronoun 184–5
 - vs. logophoric pronoun 67–74
- anaphoric proform
 - category distinctions 77
 - local vs. non-local 77
 - notion of definiteness 78
 - point of view 77
 - restriction to humans 77
 - single element vs. two element stems 75–6
 - vs. long-distance reflexive 184
- appositive construction 3, 10, 44–5, 49–52, 54, 111
 - agreement marking 51
 - case-marking 51–2
 - intonation 45
 - use of articles 52
 - use of numerals 55–6
 - with kin terms and names 45–6
- Austronesian languages 104, 158
- Australian languages 7, 86, 103–4, 148, 191, 241, 247–8
- Bantu languages 159
- Binding Theory 18, 75
- case category 112–18
 - absolutive vs. ergative 5
 - accusative vs. ergative 106, 113, 115, 126–8, 133
 - fused affixes 114–15
 - nominative vs. accusative 5
 - pronoun vs. verb 16–17, 21–2
 - relation between speech roles 114
 - sagittal case-marking 113–15
- comitative 103, 105
- complementizer construction 187
- conceptual parts-of-speech 190
- conjunctive particle 165, 220, 224, 226, 250–5, 257–8, 261, 270–1
- continuum 12, 16–17
- coordinate construction 103–5
- definiteness
 - file-changing semantics 203
 - linguistic vs. extra-linguistic identity 11, 202
 - need for pragmatic definiteness 205
 - need for semantic definiteness 206, 214
 - noun-phrase vs. proform 202–3
 - nouns vs. demonstratives 47
 - persistence of semantic indefiniteness 207–8
 - pragmatic vs. semantic 203
 - problems with *such* 209–10
 - providing information 209
 - topical vs. non-topical antecedent 209
 - visibility condition 204
 - word-class distinctions 213–14

- diachronic
 - change 15
 - cline 17
 - direction 28
 - explanation 24, 129, 256, 262
 - source 26, 28, 61, 70, 86, 111, 196
 - stability 107
 - tendency 180
- deictic elements 25, 133, 156–7, 159, 168–9, 189, 190–1, 214, 221
- deictic roots 17
- deictic systems 178
 - person-oriented vs. distance-oriented 14, 140–1, 144–5
 - two-person vs. three-person languages 141–3
- deictic verbs 163
- deictic words 40, 77, 183
- demonstratives
 - absence of spatial distinction 179–80
 - diachronic tendencies 180–2
 - identifiers 193–4
 - other distinctions 180
 - person-oriented vs. distance-oriented 177
 - relation between forms 178–9
 - spatial distinctions 177
 - verbal 2, 9, 193, 213–14
- dependent vs. head marking 29
- direct vs. indirect speech 73
- disjunctive particle 165, 220, 224, 226, 228–9, 233, 250–5, 257, 261, 270, 276
- Dravidian languages 86, 148, 181–2, 242–3, 251–3, 261, 269
- East Asian languages 69, 111–12
- eliciting language
 - influence of 156, 196
- European languages 181, 262
- exclamatory sentences 4, 50
- free and bound pronoun languages 15–30
- focus marking
 - focus construction 243–4
 - focus particle 242–3
 - other focus positions 242
 - sentence-initial position 24–2
- freedom of speech 205
- function
 - anaphors and reflexives 81
 - articles 39
 - common nouns 39
 - modifiers 39
 - personal pronouns 6, 9, 15, 38–9
 - primary vs. secondary 33
 - proforms 6–7, 9
- function of proforms
 - combining the functions 175–6
 - demonstratives with interrogatives 187–8
 - demonstratives with relatives 189
 - interrogative-indefinites with relatives 189–90
 - overlap 196–8
 - single set 186–7
- general terms of proforms
 - adverbial distinctions 193
 - basic and non-basic concepts 170
 - cardinals vs. ordinals 193
 - human–non-human 192
 - manner and reason 170–1
 - number vs. amount 192
 - object and activity 170
 - purpose of their use 200–1
 - scale of categorial metaphors 170
 - single interrogative-indefinite pronoun 195
 - type of entity 192
 - use of case markers 195–6
 - vs. word-class distinctions 190–1
- gender distinction
 - agreement system 13–14
 - control of agreement 19
 - correlation with two-person languages 139–40
 - grammaticalization 111
 - personal pronouns 13
 - personal restriction 109–10
 - social distinctions 111
 - third person 14
- gradation 31, 113, 122, 130
 - free and pronoun languages 28
 - pronouns 12–13
- grammaticalization 111, 153, 164, 171, 173, 196, 238, 261
- indefinite pronouns
 - absence of specific–non-specific distinction 255–7
 - affirmative vs. negative 261
 - choice left to the addressee 260
 - conjunctive vs. disjunctive prefixes 250–1
 - explanation for derivation 252
 - free-choice 257–8
 - grammaticalization 261–2

- indefinite pronouns (*Contd.*)
 - knowledge of the speaker 260
 - specific vs. non-specific 252–5, 270
 - universal vs. non-universal 258–60
- indirect questions 227
 - absence of interrogativity 263–4
 - absence of question markers 264–5
 - use of indefinite pronouns 265–6
- individuating affixes 94
- Indo-European languages 104, 143, 147, 155, 159
- Indo-Aryan languages 18, 23, 86, 147, 181–2, 186, 191, 242, 253, 256, 262, 267–9
- interpretation of a phrase 216
- interrogative-indefinite
 - as an ‘ignorative’ pronoun 227–8
 - conjunctive and disjunctive particles 226–9
 - differentiating languages 182
 - distinct purposes 230–1
 - human–non-human distinction 183
 - monosemy vs. polisemy 227–9
 - non-differentiating languages 183
 - question operator + indefinite pronoun 228
 - Wh-movement 211
- interrogative pronouns
 - adnominal vs. pronominal 169–70
 - identity with indefinites 226
 - related to indefinites 226
 - sentence-initial occurrence 211
 - verbal 169–70, 193–5, 214
 - word-class distinctions 212–13
- interrogativity 231
 - affixing the pronoun 239–40
 - combining with focus 144–5
 - intonation 234–5, 245
 - mood 238
 - non-marking 246–8
 - particles 235–6, 245
- language acquisition 31
- language sample 29, 98, 108–9, 134–5, 139–41, 143, 145, 177–80, 177–80, 226–7, 278
- locative case 5
 - as a derivational suffix 173
 - for denoting possession 115–16
 - hierarchy of nominals 116
- logophoric pronouns
 - anti-logophoric 65–6
 - as denoting third person 60–1
 - diachotomous nature 70, 72
 - endophoric vs. exophoric 59
 - English translation 63
 - extended use 69–70
 - for third person 62–4
 - hierarchy of trigger predicates 72
 - included antecedent 71
 - obligatoriness 64–5, 72–3
 - person hierarchy 65
 - plural as conjunction 71–2
 - pronominal vs. verbal marker 62
 - source of information 70
 - speaker vs. addressee 59–60
 - verbal agreement 61, 70
- Mayan languages 104
- negative polarity contexts 221
- Niger-Congo languages 159
- non-configurational 17
- number category
 - animacy hierarchy 105–8
 - associative plural 106–8
 - conjoined forms 100–1
 - conjunction vs. plurality 95–9, 108
 - different affixes 98
 - distributive and collectives 107–8
 - heterogeneity 104–5
 - homogeneous vs. heterogeneous
 - plurals 93–4
 - inclusive dual for nouns 92
 - inclusive vs. exclusive 91–3, 95–101
 - minimal vs. augmented 101–3, 107
 - plurality vs. group 93
 - restricted to personal pronouns 96–7
 - suppletive vs. affixed form 97–8
- Oceanic languages 86, 143, 149, 182
- Papuan languages 143
- passive device 117
- personal pronouns
 - as Janus-like 128–30
 - as suffixes 39–40
 - definiteness 47–9
 - function 38–9
 - inclusion in the nominal category 52–6
 - non-use of modifiers 53
 - plural of approximation 59, 71, 93
 - prominence 121
 - referentiality 40
 - use of articles 48, 50
 - vs. proper names 40–1

- philosophers of language 208
- point of view 122
- politeness 30, 111–12, 194
- Polynesian languages 104, 143, 158
- polysynthetic 17–18, 24
- proforms
 - complex structures 163–73
 - demonstrative vs. interrogative-indefinite 171–2
 - dual structure 153–4
 - fusion of constituents 197–8
 - grammaticalization 171–2
 - linguistic vs. extralinguistic identity 202
 - occurring in paradigms 153–63
 - order of constituents 157–61
 - peripheral constituents 162–3, 172–3
 - scope 7
 - table of correlative pronouns 155
 - three types of functions 200
 - truncated systems 155–6
 - verbal 163, 169
- pronouns
 - absolute vs. qualificative 5
 - defining 1–4, 272–4
 - free vs. bound 16–17
 - interlocutory vs. substantive 5
 - logorific 43
 - reflexive 3
 - source 26
 - stability 25
 - unanalysability 25
- pronominal 5–6
- propositional attitude verbs 217
- Proto-Algonquian 117, 125
- Proto-Dravidian 100, 148, 181
- Proto-Indo-European 147, 182, 269
- Proto-Semitic 148
- Proto-Uralic 148, 181
- Proto-Uto-Aztecán 148
- Reciprocal meaning
 - absence of personal distinction 87
 - illusory coreference 85
 - reflexive vs. collective markers 86
 - use of indefinite expressions 86
 - vs. reflexive 85
- referent
 - dissociation 10, 38–40, 42
 - file cards 42–3
 - shift 10, 38
 - theme-rheme vs. given new 211–12
- referentiality
 - generic use 219–21
 - identifiable vs. non-identifiable 219
 - intention of the speaker 215–17
 - limited and unlimited extensivity 219, 221
 - noun phrase vs. proform 220
 - pragmatic vs. semantic 138, 215, 220–2
 - purpose 216
 - referential vs. non-referential 41, 217–18, 220–1
 - referring as a speech act 216–17
 - relevance of opaque context 217, 220
 - use of articles 218–19
- reflexive meaning
 - being clause-bound 83
 - extended uses 83
 - pronominal vs. verbal 79–81
 - restriction to third person 82–3
 - SELF-anaphors vs. SE-anaphors 81
 - vs. anaphoric 79–84
- relative pronoun
 - [+ INT] vs. [– INT] 268
 - functional similarity with anaphors 183
 - problem with traditional assumption 267–70
 - strategy 270–1
 - vs. correlative 186, 269–70
 - vs. demonstratives 267
 - vs. interrogative-indefinites 266–7
- scope
 - derivational distinctions 223
 - mathematical vs. natural language expressions 22–3
 - relative order of steps 222–3
 - sentences with quantifiers 223
 - use of distinct forms 224
- Semitic languages 13, 107
- shifters 3, 31, 39–40, 42, 50
 - personal pronouns vs. other expressions 40, 43
- source of information 43
- speaker's prominence
 - in pronominal forms 122
 - in verbal agreement 122–3
 - relative ordering of affixes 123
- speech act
 - two-way affair vs. group affair 134
- speech context
 - addressee in a group 45–6
 - exophoric vs. endophoric 58–9
 - oath taking 10–11, 44

- speech context (*Contd.*)
 - prototypical vs. non-prototypical 43
 - telephone conversation 43–4
 - written texts 44–5
- speech role 10
- switch-reference 84–5
 - continuity vs. discontinuity of events 84–5
 - same vs. different subject 84
- syncretism 23–4, 29

- theory of categorization
 - classical theory 1, 130
 - Prototype Theory 1–2, 13, 130
- theory of levels 205–6
- third person pronoun
 - distinct pronouns 133–4
 - function in language 149
 - identity with demonstratives 132–3
 - related to demonstratives 133, 135–8
 - two-person vs. three-person languages 14, 112, 134–49

- Tolkappiam 269
- tone 187, 265
 - high vs. falling 8
 - high vs. low 237, 264
- Tirukkural 269

- Uralic languages 181
- Uto-Aztec languages 148, 241–2

- verb-initial languages 157–61
- verbal form
 - synthetic vs. analytic 28
- vocative
 - intonation 46
 - as parenthetical 46

- West African languages 59, 64, 66
- word order 24, 158
- world-view 14–15