Typological Studies in Language 19:2

Approaches to Grammaticalization

VOLUME 2

edited by Elizabeth Closs Traugott and Bernd Heine

John Benjamins Publishing Company

APPROACHES TO GRAMMATICALIZATION II

TYPOLOGICAL STUDIES IN LANGUAGE (TSL)

A companion series to the journal "STUDIES IN LANGUAGE"

Honorary Editor: Joseph H. Greenberg
General Editor: T. Givón

Editorial Board:

Wallace Chafe (Santa Barbara)
Bernard Comrie (Los Angeles)
R.M.W. Dixon (Canberra)
John Haiman (St Paul)
Kenneth Hale (Cambridge, Mass.)
Bernd Heine (Köln)
Paul Hopper (Pittsburgh)
Margaret Langdon (San Diego)
Charles Li (Santa Barbara)
Johanna Nichols (Berkeley)
Andrew Pawley (Canberra)
Doris Payne (Oregon)
Frans Plank (Konstanz)
Jerrold Sadock (Chicago)
Gillian Sankoff (Philadelphia)
Dan Slobin (Berkeley)
Sandra Thompson (Santa Barbara)

Volumes in this series will be functionally and typologically oriented, covering specific topics in language by collecting together data from a wide variety of languages and language typologies. The orientation of the volumes will be substantive rather than formal, with the aim of investigating universals of human language via as broadly defined a data base as possible, leaning toward cross-linguistic, diachronic, developmental and live-discourse data. The series is, in spirit as well as in fact, a continuation of the tradition initiated by C. Li *(Word Order and Word Order Change, Subject and Topic, Mechanisms for Syntactic Change)* and continued by T. Givón *(Discourse and Syntax)* and P. Hopper *(Tense-Aspect: Between Semantics and Pragmatics)*.

Volume 19:2

Elizabeth Closs Traugott and Bernd Heine (eds)

APPROACHES TO GRAMMATICALIZATION

APPROACHES TO GRAMMATICALIZATION

VOLUME II
FOCUS ON TYPES OF GRAMMATICAL MARKERS

edited by

ELIZABETH CLOSS TRAUGOTT
and BERND HEINE

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

1991

Library of Congress Cataloging-in-Publication Data

Approaches to grammaticalization / edited by Elizabeth Closs Traugott and Bernd Heine.

p. cm. -- (Typological studies in language, ISSN 0167-7373; v. 19)
Contents: v. 1. Focus on theoretical and methodological issues -- v. 2. Focus on types of grammatical markers.
Includes bibliographical references and index.
1. Grammar, Comparative and general -- Grammaticalization. I. Traugott, Elizabeth Closs. II. Heine, Bernd, 1939- . III. Series.
P299.G73A6 1991
415--dc20 91-3650
ISBN 90 272 2901 5 (set)/90 272 2895 7 (Vol. I)/90 272 2899 X (Vol. II) (Eur.; hb; alk. paper)
ISBN 90 272 2902 3 (set)/90 272 2896 5 (Vol. I)/90 272 2900 7 (Vol. II) (Eur.; pb; alk. paper)
ISBN 1-55619-404-8 (set)/1-55619-400-5 (Vol. I)/1-55619-402-1 (Vol. II) (US; hb; alk. paper)
ISBN 1-55619-405-6 (set)/1-55619-401-3 (Vol. I)/1-55619-403-X (Vol. II) (US; pb; alk. paper)

© Copyright 1991 - John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

Table of contents

VOLUME II

Foreword xi

Introduction 1
Elizabeth C. Traugott and Bernd Heine

Section A: Verbal structure **15**

Back to the future 17
Joan L. Bybee, William Pagliuca, Revere D. Perkins

The emergence of perfective aspect in Indo-Aryan languages 59
Peter Edwin Hook

Section B: Argument structure **91**

Grammaticization of topic into subject 93
Masayoshi Shibatani

From V/2 to subject clitics: evidence from Northern Italian 135
John Haiman

The role of motivation in the emergence of grammatical categories: the grammaticization of subjects 159
Marianne Mithun

Verbs to promotional suffixes in Sahaptian and Klamath 185
Noel Rude

Grammaticalization of postpositions and word order in Senufo languages 201
Robert Carlson

Section C: Subordination **225**

From postposition to subordinator in Newari 227
Carol Genetti

The evolution of dependent clause morpho-syntax in Biblical Hebrew 257
T. Givón

Section D: Modality **311**

A quantitative perspective on the gramaticization of epistemic parentheticals in English 313
Sandra A. Thompson and Anthony Mulac

The grammaticization of the German modal particles 331
Werner Abraham

Section E: Spectra of grammaticalization patterns **381**

Areal and universal dimensions of grammatization in Lahu 383
James A. Matisoff

Ways to go in Rama: a case study in polygrammaticalization 455
Colette Craig

Grammaticalization and related changes in contemporary German 493
Christian Lehmann

VOLUME I

Foreword xi

Introduction 1
Elizabeth C. Traugott and Bernd Heine

Section A: General Method **15**

On some principles of grammaticization 17
Paul J. Hopper

On the gradualness of grammaticalization 37
Frantisek Lichtenberk

Serial verbs and the mental reality of 'event': Grammatical vs. cognitive packaging 81
T. Givón

Grammaticalization as retextualization 129
Johanna Nichols and Alan Timberlake

Section B: Directionality **147**

From cognition to grammar: Evidence from African languages 149
Bernd Heine, Ulrike Claudi and Friederike Hünnemeyer

The semantics-pragmatics of grammaticalization revisited 189
Elizabeth C. Traugott and Ekkehard König

The *de dicto* domain in language 219
Zygmunt Frajzyngier

The grammaticalization of rhetorical questions in Tamil 253
Susan C. Herring

Some grammaticalization changes in Estonian and their implications 285
Lyle Campbell

The last stages of grammatical elements; contractive and expansive desemanticization 301
Joseph H. Greenberg

Substrates, calquing and grammaticalization in Melanesian pidgin 315
Roger M. Keesing

ABBREVIATIONS

ABL	ablative
ABS	absolutive
ACC	accusative
ACT	active, actor
ADV	adverb
AF	actor focus
AGR	agreement
AOR	aorist
ART	article
ASSOC	associative
ASP	aspect
AUX	auxiliary
BEN	benefactive
CAUS	causative
COMP	complementizer
COMPL	complement
COMPAR	comparative
COND	conditional
CONJ	conjunction
COP	copula
DAT	dative
DEC	declarative
DEF	definite
DEM	demonstrative
DET	determinator
DIR	directional
DO	direct object
DS	directional subject
DU	dualis
DUB	dubitative
DUR	durative
EMPH	emphatic
ERG	ergative
EXCL	exclusive
EXCLAM	excalamative
FEM	feminine
FIN	finite
FOC	focus
FUT	future
GEN	genitive
GEN.OBL	general oblique preposition
GER	gerund
IMP	imperative
IMPERF	imperfect
IMPFV	imperfective
INCL	inclusive
INCONS	inconsequential
INDEF	indefinite
INF	infinitive
INFL	verbal inflection
INT	intensifier
INTR	intransitive
IO	indirect object
IRR	irrealis
LOC	locative
MASC	masculine
MED	medial
MID.VCE	middle voice

NEG	negative
NOM	nominative
NOMI	nominalizer
NONFUT	nonfuture
NONPAST	nonpast
NP	noun phrase
OBJ	object
OBL	oblique
PART	particle
PARTCP	participle
PARTV	partitive
PASS	passive
PAST	past
PERF	perfect
PFV	perfective
PL	plural
POS.PURP	positive purpose
POSS	possessive
PRE	verbal prefix
PREC	precaution/ negative purpose
PRED	predicate
PREP	preposition
PRES	present
PROG	progressive
PROH	prohibitive
PURP	purposive
Q	interrogative
QUOT	quotative
REAS	reason
RED	reduplicative
REFL	reflexive
REL	relative clause marker
REP	repetitive
SEQ	sequential
SG	singular
SS	same subject
STAT	stative
SUB	subordinate
SUBJ	subject
SUBJUNCT	subjunctive
TNS	tense
TOP	topic
TRANS	transitive
VB	verb
VOC	vocative
1SG	first person singular
1PL	first person plural
2SG	third person singular

Foreword

The papers in these volumes grow out of a symposium on grammaticalization organized by Talmy Givón, and held at The University of Oregon at Eugene in May 1988. The purpose of this symposium was to "present and discuss ... a broad range of empirical and theoretical issues concerning the genesis of morphosyntactic structure". We wish to acknowledge the inspiration provided by Talmy Givón's leadership of the symposium.

Many people helped bring these volumes into final shape. Especial thanks are due to John McWhorter and Gillian Ramchand for their assistance in commenting on and copy-editing papers, and to Ulrike Claudi and Friederike Hünnemeyer for preparation of the final materials.

Introduction

Elizabeth Closs Traugott and Bernd Heine
Stanford University and University of Cologne

1. INTRODUCTION

Although not everyone who uses the term would agree, "grammaticalization" as we define it refers to that part of the theory of language that focuses on the interdependence of langue and parole, of the categorial and less categorial, of the fixed and the less fixed in language. Grammaticalization is the linguistic process, both through time and synchronically, of organization of categories and of coding. The study of grammaticalization therefore highlights the tension between relatively unconstrained lexical expression and more constrained morphosyntactic coding, and points to relative indeterminacy in language and to the basic non-discreteness of categories.

These volumes present a number of diverse theoretical viewpoints on grammaticalization and give some detailed insights into specific grammaticalization phenomena in a number of languages. However, no attempt is made at completeness; for more comprehensive studies of the field of grammaticalization as a whole, see Lehmann (1985), Heine and Reh (1984), Hopper and Traugott (In progress), Heine et al. (1991). The diversity of perspectives is manifest in a number of ways. For example, disagreement about how to approach the subject starts with disagreements about what to call it. Some authors prefer "grammaticization" or even "grammatization" to "grammaticalization". There is more significant disagreement about whether grammaticalization is primarily a diachronic phenomenon to be studied from a "source and pathway" perspective, or primarily a syntactic, discourse-pragmatic phenomenon, to be studied from the point of view of fluid patterns of language use across time or at a synchronically segmented moment in time. In so far as there is any correlation between perspective and terminology, those who espouse the latter view of the subject tend to call it grammaticization. However, since the majority of participants used "grammaticalization", whatever

their perspective, we have chosen to use this term in the title and in this introduction. We have not imposed uniform terminology on the contributors, and both terms will be found in the volumes.

Together the papers suggest answers to a number of questions, some of which focus on testing the boundaries of grammaticalization, others of which test the assumptions of linguistic theory in general. We outline here some of the models of grammaticalization that the authors use, challenge, or develop, and some of the main issues that the papers highlight.

2. THE LEXICAL ITEM > MORPHEME MODEL

Early work on grammaticalization viewed it as a diachronic phenomenon. Although many examples were discussed by nineteenth century scholars (e.g. Bopp, 1816; Humboldt, 1825; Gabelentz, 1891), the term was apparently first used by Meillet (1948 [1912]). He defined it as the evolution of grammatical forms (function words, affixes, etc.) out of earlier lexical forms, and much subsequent work has been conducted within this framework. Later researchers have emphasized that morphemes may also arise out of other morphemes. For example, Lehmann, acknowledging Kuryłowicz (1965), says: "Under the diachronic aspect, grammaticalization is a process which turns lexemes into grammatical formatives and makes grammatical formatives still more grammatical..." (1985:303), and Heine and Reh define grammaticalization as: "...an evolution whereby linguistic units lose in semantic complexity, pragmatic significance, syntactic freedom, and phonetic substance..." (1984:15).

In the present volumes papers by Abraham, Bybee et al., Campbell, Carlson, Craig, Greenberg, Heine et al., Keesing, Lehmann, Matisoff, Rude, Traugott and König owe much to this perspective, although all have also been influenced to greater or lesser degree by the discourse > morphosyntax perspective outlined in 3.

3. THE DISCOURSE > MORPHOSYNTAX PERSPECTIVE

Alongside the lexical item > morpheme tradition which derives from Meillet, there has been a more recent tradition associated with Talmy Givón, Charles Li, Sandra Thompson and others that focuses on the "packaging" of discourse and evolution of syntactic and morphological structure through

fixing of discourse strategies. For example, Givón (1979:209) characterizes the process as one of cyclic waves involving

Discourse→Syntax→Morphology→Morphophonemics→Zero

cf. also Bybee (1985), Hopper (1988), Li and Thompson (1974), among others.

Papers in these volumes that develop this morphosyntactic perspective particularly clearly are those by Genetti, Givón on the development of dependent clauses in Hebrew, Haiman, Herring, Hook, Hopper, Lichtenberk, and Shibatani. It should be noted that in many cases the source item or items discussed by these authors is not lexical but morphological (a possibility clearly permitted by the definitions of grammaticalization cited in 2.).

4. THE GRAMMATICAL CODING PERSPECTIVE ON GRAMMATICALIZATION

So far we have reviewed issues of grammaticalization that are one way or another issues related to language change. Yet another perspective on grammaticalization asks not so much how grammatical coding comes into being, but what grammatical coding is possible typologically in a language or in languages, and how it is organized either at one period in time (Givón on serial verbs, Mithun, Thompson and Mulac) or panchronically (Frajzyngier, Nichols and Timberlake).

The answers to the question whether certain categories must be obligatorily grammaticalized challenge some of the fundamental assumptions some linguists bring to linguistic theory, for example, that there are universal "primitives" such as Subject, Object, Tense (cf. also DuBois, 1985).

5. THE RELATION OF DIACHRONIC GRAMMATICALIZATION TO LANGUAGE CHANGE

For those who see grammaticalization as a diachronic process, the question naturally arises: How can we distinguish grammaticalization from language change? The answer is that grammaticalization is a kind of language change, subject to certain general processes and mechanisms of change (see section 8.), and characterized by certain consequences such as changes in grammar. These are increased syntacticization in its early stages, and increased loss of morphosyntactic independence in its later stages, ultimately leading

to zero, i.e. increased morphologization, and phonologization. Like other changes, grammaticalization spreads gradually across linguistic contexts on the one hand (Lichtenberk, Nichols and Timberlake), and across social contexts on the other (cf. Labov, 1972, and much of the sociolinguistic literature on variation and language change in progress). But that does not mean that all change is grammaticalization. To take an obvious example, the semantic change from Middle English *bede* 'prayer' to its modern meaning of a small spherical object often used in necklaces ('bead'), is an example of change by metonymy, but not of grammaticalization. Papers in this volume that particularly address this issue are those by Hopper and Lehmann.

6. THE UNIDIRECTIONALITY OF GRAMMATICALIZATION

Firmly entrenched in diachronic perspectives on grammaticalization is the assumption of a cline of unidirectionality; or more specifically of a pathway that channels change through a limited number of structures that are minimally different from one another. This does not mean that all unidirectionality involves grammaticalization; for example, the well-known tendency for [sy] to become [š] but not vice versa is in itself not an instance of grammaticalization. However, if it results from morpheme boundary loss, then it is an instance of a stage in grammaticalization. The question is precisely what kinds of unidirectionality are necessary or at least typical of grammaticalization, and whether any kinds of reversals of grammaticalization are usefully considered as examples, albeit rare, of grammaticalization (see 7. below).

One kind of unidirectionality often associated with grammaticalization is increase in abstractness. This notion is appealed to in discussions of metaphorical change (Heine et al. on PERSON > OBJECT > SPACE > TIME > PROCESS > QUALITY, Traugott and König on temporal > causal > concessive). However intuitively attractive and widely agreed upon such a cline of abstractness may be, it is essentially a pre-theoretical notion. Certainly, increased abstractness does not in itself require grammaticalization. For example, the shift in meaning from *grasp* in the sense 'seize' to 'understand' is not a case of grammaticalization, although it is an example of what many linguists would consider increased abstractness in meaning.

Much of the literature on unidirectionality characterizes the development of grammatical from lexical meaning in terms of desemanticization, bleaching, and emptying or loss of semantic or pragmatic meaning (see the definitions

of grammaticalization in 2. above). Traugott and König challenge this concept and argue that in its early stages grammaticalization actually often involves an increase in pragmatic meaning (though semantic content, strictly speaking, may be reduced as pragmatic meaning increases). For example, a shift from temporal to concessive *while* does not involve primarily bleaching of semantic meaning, but rather increase of pragmatic significance pertaining to the speaker's attitude to what is being said. If this is correct, then we need to modify the standard view of semantic unidirectionality in the process of grammaticalization. It is not always a simple case of more > less semantic, but more often of an implicational hierarchy of the type semantic > pragmatic > less semantic-pragmatic. Note that although this is not a characterization of bleaching it is nevertheless unidirectional, since the claim is that one would not expect to find pragmatic > semantic > less semantic-pragmatic, e.g. concessive *while* > temporal. As Traugott and König indicate, the meaning changes outlined here for grammaticalization are characteristic of meaning change in general, and are therefore not diagnostic for grammaticalization (any more than bleaching is, for that matter).

Sometimes it has been proposed that the lexical item > morpheme and the discourse > morphosyntax perspectives come into conflict with respect to unidirectionality in the earliest stages of grammaticalization (Herring). However, there is actually no inconsistency when we consider that discourse presupposes lexicon; discourse uses lexical items in ways that endow them with pragmatic meaning, and if they have the properties salient to grammaticalization outlined above, and are used more frequently, they may well come to be syntacticized in the kinds of way Givón illustrates. The two formulae "lexical item > morpheme", "discourse > morphosyntax " can therefore be combined as: lexical item used in discourse > morphosyntax.

Another kind of unidirectionality often considered characteristic of grammaticalization is increase in bondedness. Typically, at the clause level independent clauses are combined, and a cline may develop from independent clauses through some kind of loose juxtaposition and coordination to subordination (cf. Wiegand, 1987). At the phrasal level, forms may become less free and more bound via grammaticalization; for example, postpositions become affixes (Craig, Greenberg). Although not confined to grammaticalization, one of its effects is that morphemes undergoing this process move away from cardinal categoriality, and in their later stages lose the ability to refer and to associate with the inflectional and derivational trappings of their morphosyntactic category (cf. Hopper and Thompson, 1984). This leads to the emergence of "linguistic hybrids" which show the properties of several morpheme classes.

Nouns which develop into adpositions, for example, lose nominal characteristics such as the ability to be marked for definiteness or number (Genetti), or verbs which are grammaticalized to case markers tend to lose the ability to be inflected for person, tense, aspect, and mood. Loss of categoriality may also lead to the kind of development of multiple grammaticalization chains in different functional domains that Craig discusses in connection with Rama.

Unidirectionality may suggest a single path of evolution. However, approaches referring to multiple functional domains (cf. Kemmer 1988) or to correlated processes seem closer to reality (cf. Heine and Reh, 1984; Lehmann, 1985; Croft, Forthcoming: Ch. 8). Heine and Reh suggest a tripartite classification of correlations of change: Semantic-pragmatic status, grammatical behavior, and phonological substance. Lehmann uses a bipartite classification according to paradigmatic and syntagmatic processes, i.e. according to the alternatives available on the one hand and the effect of linguistic context on the other. Lehmann's characterization is somewhat theory-bound and leads to the incorrect claim that grammaticalization involves reduction in scope (1985:309), despite the fact that the development of subordinators from pronouns or verbs of saying, for example, clearly involves expansion of the scope of these forms. However, its focus on correlations is of crucial importance. In these volumes, most of the papers alert us to correlations.

It is the phenomenon of correlatable parameters that allows us to project diachronic grammaticalization onto synchronic data. Consider for example, the grammaticalization of body part nouns. If arrayed simply as a case of OBJECT > SPACE > TIME > QUALITY the examples of *megbé* in Heine et al. would not qualify as a case of grammaticalization. They do so precisely because the semantic changes are correlated with a shift from nominal to prepositional structure, and with different syntactic privileges of occurrence.

7. COUNTEREXAMPLES TO UNIDIRECTIONALITY

The discussion so far has focused on unidirectionality, and what kinds of unidirectionality are characteristic of grammaticalization. Virtually nothing is exceptionless, and there are of course instances of change in languages that are counterexamples of tendencies that can be characterized as "less > more grammatical", "main clause > subordinate clause", etc. In these volumes the papers by Campbell and Greenberg explicitly raise counterexamples to unidirectionality (cf. also Matsumoto, 1988). The well-known phenomenon of

"lexicalization" is also a counterexample, for example when phonological changes result in morphological loss and the development of idiosyncratic lexical items, such as the English pairs *lie–lay*, *sit–set*, *stink–stench*, all of which have their origins in *i*-umlaut.

It is likely that all these examples are strictly speaking actually not cases of grammaticalization (although once they have occurred they may be subject to the generalization, reduction, loss, and other changes typical of grammaticalization). Rather, the examples Campbell and Greenberg cite can be regarded as instances of reanalysis (Greenberg actually refers to "reinterpretation which...goes beyond what is usually called reanalysis in discussions of grammaticalization"). Nevertheless, since they test the boundaries of what may be considered bona fide cases of grammaticalization, they provide an important perspective on the other papers in these volumes.

8. MECHANISMS OF LANGUAGE CHANGE RELEVANT TO GRAMMATICALIZATION

A number of mechanisms of language change have already been alluded to as being relevant to grammaticalization. This is hardly surprising if indeed grammaticalization is a subset of phenomena occurring in change. Of particular importance for these volumes are: Metaphorical transfer (Heine et al.), metonymic transfer (Traugott and König), reanalysis (Heine et al., Lehmann, cf. also Langacker, 1977), and analogy (Greenberg, Lehmann, Matisoff). Another mechanism is the incorporation of grammatical material from the substrate languages, as illustrated by evidence from Melanesian Pidgin (Keesing). All these are mechanisms that make change possible, but none are restricted to grammaticalization and *all* are independent of the unidirectionality associated with grammaticalization.

9. CONDITIONS LICENSING GRAMMATICALIZATION

A further set of questions for the diachronic perspective on grammaticalization pertains to what conditions have to apply for grammaticalization to take place, in other words, given that a form A exists, what is its potential for becoming grammaticalized, and how do we know when this is happening?

Experience with cross-linguistic studies of typology and of change have revealed extensive and robust evidence that only certain lexical classes are

likely to become grammaticalized. For example, it seems unlikely that *wall paper* would become grammaticalized. This is not to say that such a change is impossible, but rather that it is extremely unlikely. What we find in language after language is that for any given grammatical domain, there is only a restrictive set of lexical fields, and within them only a restricted set of lexical items, that are likely to be sources. For example, case markers, including prepositions and postpositions, typically derive from terms for body parts or verbs of motion; tense and aspect markers typically derive from specific spatial configurations; modals from terms for possession, or desire; middles from reflexives, etc. Furthermore, the paths of change are themselves highly restricted, and suggest step-by-step developments, not discrete leaps across semantic or pragmatic domains. Accounts of why these restrictions apply appeal to cognitive constraints. Whereas some have regarded these cognitive constraints as hard-wired (i.e. essentially arbitrary, cf. Bickerton, 1981), authors represented in these volumes regard them as functionally constrained by salience, communicative strategies, etc. For example, if one is attempting to express something abstract using the resources already available, or attempting to express communicative intent, also using resources already available, then the purposes to which one puts language in part determine the choices one makes. Therefore, one answer to the question concerning what potential a form A has for becoming grammaticalized is: This depends on whether its semantic content or the inferences one can draw from it serve the purposes of creating text, that is, of producing discourse. Ultimately, the question is, can the form be used to serve the metalinguistic purposes of text-construction in a natural and therefore easily understood way? It is hard to imagine a context in which *wall paper* could be used to serve such a purpose.

The fact that a lexical item A has potential for solving the problem of expression does not in itself predict that it is a candidate for grammaticalization. We need to look for additional factors. An example is *corner*, a term which is used metaphorically in temporal expressions (cf. *go round a corner in time*). However, *corner* does not appear to be grammaticalized into tense or aspect. Why might this be? It has been suggested that the operative constraint is perceptual salience: Topological spaces only are grammaticalized (Talmy, 1983; Sweetser, 1988). Although the expected end of a process of grammaticalization is usually zero, in fact there are "graveyards", such as *be*-clefts (Givón) and frozen idioms or lexicalizations (Hopper) that appear to block or at least delay the process of total loss (see especially Greenberg).

It would seem that the point at which grammaticalization stops may be determined by the typological characteristics of the language concerned. There

is, for example, a well established channel of grammaticalization leading from postpositions to nominal case inflections. To our knowledge, however, such a development has not been documented for languages of the analytic-isolating type, where grammaticalization is unlikely to lead to the development of inflectional morphology. Observations like these suggest that there may be typological language-internal constraints that block grammaticalization from proceeding beyond a certain point. Mithun suggests that some very specific factors such as the internal shape of the grammar of a particular language may make certain grammatical markers such as subject redundant and may therefore serve to block their development.

Given that a form A is a candidate for grammaticalization both because of its semantic context and its salience, a further condition has to apply for grammaticalization to take place: The form has to be used frequently. The more grammaticalized a form, the more frequent it is (Bybee et al., Givón on serial verbs, Hook). The seeds of grammaticalization are therefore in a correlated set of phenomena: Semantic suitability, salience, and frequency. Only the third actually leads to grammaticalization and hence to fixing, freezing, idiomatization, etc.

10. SOME QUESTIONS FOR THE FURTHER STUDY OF GRAMMATICALIZATION

Among a number of questions still to be answered is what the language-external motivation for grammaticalization may be. If semantic suitability, salience and frequency are among the prerequisites for grammaticalization to start, then the question still remains as to what motivates the beginning of the process. Is it discourse-pragmatic pressure, that is, the need to be informative and processable and expressive all at the same time (cf. Langacker, 1977), the phenomenon of gaps in grammatical paradigms or in the universe of abstract concepts, a "natural propensity" for signaling metalinguistic relations in non-lexical ways (cf. Bybee and Pagliuca, 1985), or some other factor or factors? These questions all point to linguistic problem-solving as a clue to motivations for grammaticalization. Any claim about problem-solving and functional purpose raises issues of teleology. Because grammaticalization does not necessarily happen in any given instance of potential grammaticalization, considerable caution needs to be taken in proposing teleological explanations. Clearly we do not know enough yet about communication to argue that "communicative necessity" motivates the development of grammatical cate-

gories. It is primarily through the study of what prevents grammaticalization, or simply fails to trigger it, that we can begin to hope to answer the question of motivation.

Also, what are the social and psychological dynamics that are present in the initial and final stages of grammaticalization? How far is it induced by individuals, and to what extent does it form an act of human creativity, rather than being the result of language-internal drifts or forces? Hardly any answers are available thus far to this type of question, despite extensive sociolinguistic research on the development of morphology. The question in most of this literature has been posed primarily in terms of variation and continua (e.g. Bickerton, 1975; Sankoff, 1980: Part II, Rickford, 1987), or of contact (e.g. Romaine, 1988), but for the most part not in terms of psychological or cognitive motivation (see, however, Slobin's (1985) suggestions regarding cross-linguistic operating principles in child language acquisition).

Yet another unsolved puzzle is what motivates the differential speed with which grammaticalization takes place in different functional domains. Observations on African languages suggest that some kinds of developments proceed faster than others. For example, new categories of tense and aspect have emerged within a relatively short period, and in some cases a new morphology evolving along the same grammaticalization pattern is already emerging, competing with the old one. The development of noun class systems or of verbal derivation, on the other hand, has been much more conservative: morphological paradigms found today can be reconstructed as having already existed in a similar form and function several millennia ago.

Another important question for any theory of language that grammaticalization poses is whether lexical morphemes are by nature different from grammatical ones. If it is true that diachronically grammatical forms ultimately arise out of lexical items (although their immediate sources are often other grammatical items), does this lead to the logical conclusion that they cannot be innovated, whereas lexical items can? Among the highly stable grammatical forms with no certain lexical origin are such forms as the Indo-European *to-* of the demonstrative pronoun (later further grammaticalized to a third person agreement marker). Are we justified in hypothesizing that this form was once derived from a lexical item, as the diachronic perspective on grammaticalization would suggest, or does it provide evidence that grammatical items can arise full-fledged, and if so, under what conditions? These and other questions remain for future empirical study.

11. ORGANIZATION OF THE VOLUMES

Arranging the many papers in a meaningful way turned out to be a difficult task, and the result is to some extent arbitrary. In view of the number and size of the contributions available we decided to divide the work up into two volumes. This division, too, is to some extent arbitrary.

Volume I focuses on theoretical and methodological issues, and is mainly concerned with developing the major themes for the study of grammaticalization outlined above.

Section A deals with some new perspectives on the development and structure of grammatical categories. Hopper discusses five principles which are relevant for describing grammaticalization. Lichtenberk draws attention to the role played by diachrony in explaining synchronic facts, and to gradualness in grammaticalization. By looking at the relationships between temporal packaging and information processing in serial verb constructions, Givón concludes that grammaticalization is not a gradual process, but rather an instantaneous one. Nichols and Timberlake develop a panchronic analysis of morphological realignments resulting from the tension between innovation and conventionalization.

In Section B, various perspectives on directionality are discussed. The first two argue for unidirectionality. The papers by Heine, Claudi and Hünnemeyer and by Traugott and König deal, among other things, with cognitive factors in directionality, especially the role of inference in metaphor and metonymy. Frajzyngier studies directionality of fit in grammar between the domain of the real world and that of speech. Unidirectional views on grammaticalization are challenged in a number of different ways. Herring argues for the independence of the development of grammatical elements out of lexical items and out of discourse-contexts, and the bi-directionality of some changes. Contributions by Greenberg and Campbell suggest that typically unidirectional processes may be reversed under certain conditions. And Keesing shows how areal influence may lead to the elaboration of grammatical structures through borrowing that is socially directional (from substrate to superstrate), but not linguistically so.

The papers in Volume II are for the most part detailed studies of distinct areas of language structure in particular languages or language families.

Section A focuses on the genesis of verbal categories such as tense and aspect. Bybee, Pagliuca, and Perkins summarize their cross-linguistic findings on future marking, and Hook discusses the notion of aspectogenesis in some Indo-Aryan languages.

Section B is concerned with various aspects of the development of argument structure. Shibatani resumes earlier discussion on the development of topic into subject, using evidence from Japanese and Philippine languages. And Haiman explores further the relationship between word order and the development of subject clitics. In a rather different vein, Mithun demonstrates why in some languages, specifically Cayuga (Iroquoian) and Selayarese (Austronesian), there was no need for the development of a subject category at all. In the final part of this section, Rude and Carlson show how non-subject marking arises out of verbal structures in Sahaptian and Klamath and in Senufo respectively.

Two different patterns in the development of subordinate structures are discussed in Section C. Genetti describes the transition from case postpositions to clausal subordinators in Newari (Tibeto-Burman). Givón reconstructs the transfer from relative clause structure to verbal complement subordination in Biblical Hebrew.

Section D is devoted to expressions of modality. Thompson and Mulac present the case of certain English "main clauses" (*I think (that)*) which are reinterpreted as epistemic phrases, and Abraham deals with the development of a set of modal particles in German.

Papers in Section E demonstrate a spectrum of cases of grammaticalization and a diverse array of factors at play. Matisoff investigates the development of a number of verbal and nominal particles, as well as of verbal concatenations in the Southeast Asian linguistic area. Craig shows how in Rama (Chibchan) one verb in different contexts develops multiple grammatical functions, one relating to argument structure, the other relating to tense, aspect and mood; she proposes the term "polygrammaticalization" for this kind of process. Lehmann discusses a number of emergent grammatical patterns in modern German which may or may not be instances of incipient grammaticalization, thereby pushing on the question of how one recognizes an instance of grammaticalization.

ACKNOWLEDGMENTS

This introduction is clearly inspired by all the participants in the symposium on grammaticalization. Particular thanks are due to Dan Slobin whose concluding remarks made the coordination of issues easier, and to Paul Hopper, Frank Lichtenberk, and Suzanne Romaine for comments on an earlier version.

REFERENCES

Bickerton, Derek. 1975. *Dynamics of a Creole System*. Cambridge: Cambridge University Press.

Bickerton, Derek. 1981. *Roots of Language*. Ann Arbor, MI: Karoma Publishers.

Bopp, Franz. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Frankfurt: Andreäische Buchhandlung.

Bybee, Joan. 1985. *Morphology. A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

Bybee, Joan and William Pagliuca. 1985. "Cross-linguistic comparison and the development of grammatical meaning." In *Historical Semantics and Historical Word Formation*, Jacek Fisiak (ed.), 59–83. Berlin: de Gruyter.

Croft, William. Forthcoming. *Typology and Universals*. Cambridge: Cambridge University Press.

DuBois, John W. 1985. "Competing motivations." In *Iconicity in Syntax*, John Haiman (ed.), 343–65. Amsterdam: John Benjamins.

Gabelentz, Georg von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel Nachfolger.

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.

Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.

Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Hopper, Paul J. 1988. "Emergent grammar and the *a priori* grammar postulate." In *Linguistics in Context: Connecting, Observations, and Understanding*, Deborah Tannen (ed.), 117–34. Norwood: Ablex.

Hopper, Paul J. and Sandra Thompson. 1984. "The discourse basis for lexical categories in universal grammar." *Language* 60:703–52.

Hopper, Paul J. and Elizabeth C. Traugott. In progress. *Grammaticalization*. Cambridge: Cambridge University Press.

Humboldt, Wilhelm von. 1825. "Über das Entstehen der grammatischen Formen und ihren Einfluss auf die Ideenentwicklung." *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin*:401–30.

Kemmer, Suzanne. 1988. The Middle Voice: A Typological and Diachronic Study. Unpublished Ph.D. dissertation, Stanford University.

Kuryłowicz, Jerzy. 1965. "The evolution of grammatical categories." *Esquisses Linguistiques* (2):38–54.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Langacker, Ronald W. 1977. "Syntactic reanalysis." In *Mechanisms of Syntactic Change*, Charles N. Li (ed.), 57–139. Austin: University of Texas Press.

Lehmann, Christian. 1985. "Grammaticalization: Synchronic variation and diachronic change." *Lingua e Stile* 20:303–18.

Li, Charles N. and Sandra A. Thompson. 1974. "Historical change and word order: A case study in Chinese and its implications." In *Historical Linguistics I: Syntax, Morphology, Internal and Comparative Reconstruction*, John M. Anderson and Charles Jones (eds), 199–218. Amsterdam: North-Holland.

Matsumoto, Yo. 1988. "From bound grammatical markers to free discourse markers: History of some Japanese connectives." In *Berkeley Linguistics Society, Proceedings of the Fourteenth Annual Meeting*, 340–51.

Meillet, Antoine. 1948 [1912]. "L'évolution des formes grammaticales." In *Linguistique générale et linguistique historique*, 130–48. Paris: Champion.

Rickford, John. 1987. *Dimensions of a Creole Continuum: History, Texts, and Linguistic Analysis of Guyanese Creole*. Stanford: Stanford University Press.

Romaine, Suzanne. 1988. "Pidgins, creoles, immigrant, and dying languages." In *Investigating Obsolescence. Studies in Language Contraction and Death*, Nancy Dorian (ed.), 369–83. Cambridge: Cambridge University Press.

Sankoff, Gillian. 1980. *The Social Life of Language*. Philadelphia: University of Pennsylvania Press.

Slobin, Dan I. 1985. "Cross-linguistic evidence for the language-making capacity." In *The Cross-Linguistic Study of Language Acquisition. Volume II. Theoretical Issues*, Dan I. Slobin (ed.), 1157–1256. Hillsdale, N.J.: Erlbaum Associates.

Sweetser, Eve. 1988. "Grammaticalization and semantic bleaching." In *Berkeley Linguistics Society, Proceedings of the Fourteenth Annual Meeting*, 389–405.

Talmy, Leonard. 1983. "How language structures space." In *Spatial Orientation: Theory, Research and Application*, Herbert Pick and Linda Acredolo (eds), 225–82. New York: Plenum Press.

Wiegand, Nancy. 1987. Causal Connectives in the Early History of English: A Study in Diachronic Syntax. Unpublished Ph.D. Dissertation, Stanford University.

SECTION A: VERBAL STRUCTURE

Back to the Future

Joan L. Bybee, William Pagliuca, and Revere D. Perkins
The University of New Mexico, The University of Wisconsin, and New York, New York

1. INTRODUCTION

In earlier works, the semantic development of future markers in certain well-known languages has been discussed — Fleischman (1982a) studies the development of futures in Romance languages, and in a much briefer study, Bybee and Pagliuca (1987) offer an analysis of the development of three futures in English.* Parallels in other unrelated languages are easy to find: Futures developing from the same lexical sources as in English and Romance languages are common (Ultan, 1978; Bybee and Pagliuca, 1987). Dahl's (1985) cross-linguistic study of grammatical meaning identified a universal prototype for future, claiming that the meaning signalled by future morphemes is similar cross-linguistically. In this paper, we go back to the future to test some hypotheses about grammaticization on a representative sample of the world's languages. In particular, we test the following hypotheses:

(1) That futures in all languages develop from a small set of lexical sources and that all future morphemes from a given source go through similar stages of development;

(2) That the semantic change in grammaticization is accompanied by formal reduction, whereby the morpheme loses its independence and may fuse to contiguous material.

The material upon which these hypotheses can be tested is a database constructed by the GRAMCATS project, which is an exhaustive survey of morphology associated with the verb in 75 languages. The languages were randomly selected from maximally unrelated genetic groups.[1] Our access to the languages is largely through reference grammars and partial descriptions.

We coded information about all of the closed class items that were fixed in position with respect to the verb.[2] The following types of information were included:

(1) *Meaning*: The separate senses conveyed by the gram are recorded as separate *uses*, and each use is characterized by one or more standardized meaning labels, called *meaning components*.

(2) *Form*: The formal properties of grams are recorded, as, for example, their phonetic shape, their position, whether they are bound, the existence of allomorphs, the type of conditioning for allomorphs, and so on.

(3) *Related strings*: Here we record whatever information is available about the historical relations of the gram, i.e. what other forms in the language it is homophonous with or similar to, whether they are grammatical or lexical, or any etymological information the author supplies.

Our general hypothesis is that there are only a small number of major grammatical categories or gram-types (such as 'past', 'future', or 'perfective') in the languages of the world, and that each of these develops historically via a small number of separate paths. Thus it has been found that futures — grams that have 'future' as one of their uses (hereafter *future grams*) — may develop from auxiliary constructions with the meanings of 'desire', 'obligation' or 'movement towards a goal' (Ultan, 1978; Bybee and Pagliuca, 1987). By postulating independent and unidirectional paths of development, we make predictions about the expression of grammatical meaning in languages. In particular, we predict that certain combinations of meanings will commonly be expressed by a single form, and that certain meanings are characteristic of early stages of development, while others are characteristic of later stages.

In the GRAMCATS sample to date, we have 129 grams which have 'future' as one of their meaning components, and these are distributed in 63 languages.[3] Analyzing the meanings and related strings of these grams, we come up with three major types of futures, each of which is represented by at least two separate paths of development, and a fourth, less common type:

(1) *Aspectual forms*: These are grams whose main function is to mark imperfective, and, in the rare case, perfective, but which also can be used for statements about future time.

(2) *Agent-oriented modalities*: Verbs or constructions with agent-oriented meaning, such as 'desire' or 'obligation' and, less commonly, 'ability', may develop into grams with future as one of their uses.

(3) *Movement towards*: Verbs or constructions that signal movement towards a goal may develop into future grams, with some differences discernible between those using 'come' and those using 'go'.

(4) *Temporal adverbs*: Adverbs indicating a time after the moment of speech or a reference time, e.g. *soon*, *then*.

We arrive at this classification into types by comparing the other meanings expressed by individual future grams, and by comparing their related strings. A detailed consideration of each of these types begins in section 3, but first we will make some general statements about the term 'future', what it stands for in actual language and how it is used in reference grammars.

2. WHAT IS FUTURE?

The future does not represent 'future time reference' in the same way that the past represents 'past time reference', either in the use of the term in reference grammars, or in the actual usage of future grams in languages. It appears that the prototypical use of future grams, as well as what authors of reference grammars mean when they use the term 'future', is to signal that *an assertion about future time* is being made, or in other words, to signal a *prediction* (Bybee and Pagliuca, 1987; cf. Coates, 1983; Dahl, 1985). Because a prediction is a type of assertion, future grams often do not occur in subordinate clauses, even where future time reference is clearly intended, as the English examples in (1) and (2) show:

(1) When I grow up, I want to be a pilot
*When I will grow up, I want to be a pilot
*When I'm gonna grow up, I want to be a pilot

(2) If he asks for it, I'll give it to him
If he asked for it, I would give it to him
*If he will ask, I'll give it to him

Note that the last sentence is acceptable only with the more agent-oriented reading of *will*, that of 'willingness'. Note also that an *if*-clause with a Past

Tense verb refers to a time that includes the future, albeit future hypothetical, but it still does not use the future gram.[4] This situation is typical of future grams (cf. Dahl, 1985): They only rarely occur in *if* and *when* clauses that refer to future time.[5] By contrast, a typical past tense does occur to signal past time in non-assertive clauses, such as *when*-clauses.

While prediction is the prototypical, and in our study, the defining, use for future grams, more than half of the future grams in our database have other uses as well. These other uses provide us with the clues to the semantic development of futures, telling us where the grams come from and how far they have advanced in their development. In the following, we examine the semantic development of future grams as indicated by these other uses, structuring our discussion in terms of the three most common types. Our goal will be to develop hypotheses about the semantic development of each of these types that can be tested on our cross-linguistic database. We begin with aspectual futures and follow with discussions of modality-derived and movement-derived futures.

3. ASPECTUAL GRAMS USED FOR FUTURE

The aspectual grams used for future in the database were identified by finding all the grams that had future as one of their meaning components, and in addition one of the following aspect or tense meaning components. The numbers indicate the number of co-occurrences with future: (See Appendix 1 for definitions of these meaning components.)

present	9	imperfective	7
continuous	9	perfective	3
habitual	7	past	3

We were looking here for cases of general perfective or imperfective grams that could be used as futures, as well as cases of presents with future uses, such as the English *'I go to Chicago tomorrow'*.

The prediction that this is a distinct (although not necessarily homogeneous) type from the others is borne out by the distribution of meaning components in these grams. These tense and aspect meanings tend to co-occur with one another in the same set of forms: The 44 aspect or tense meaning components co-occurring with future are clustered in only 24 forms. In fact, these meaning components tend to co-occur to the exclusion of other meaning components commonly found with futures which we will discuss

below, i.e. epistemic modalities, such as possibility and probability, and the agent-oriented modalities of desire, obligation and ability.

The aspectual grams that signal future naturally divide into two groups — those expressing imperfective meanings (such as present, continuous, habitual and imperfective) and those expressing perfective meanings (such as anterior and perfective). The former group comprises twenty grams in our database, while the latter contains only four.

For both groups, we suspect that the future reading is only obtained when the context warrants it; that is, the semantics of the gram does not explicitly contain the features 'prediction' or 'future time reference', but is much more general, and allows a future interpretation if it is supported by the context. For instance, among the perfective set, the Tahitian future anterior reading is only possible with a temporal clause that sets up the future time (e.g. *'When you return, I will already be dead'*).

Our evidence also provides support for the hypothesis that the original meaning of a gram determines the particular range of use it will have later in its history (Bybee and Pagliuca, 1987). The particular 'flavor' of future signalled by the imperfective grams in our corpus is different from that signalled by the perfectives. Except for the future anterior of Tahitian, all the perfectives (those of Abkhaz, Baining and Kanuri) signal immediate future only, suggesting that the non-past interpretation of a perfective, which cannot be present, is as close to the present as possible. On the other hand, none of the imperfectives has immediate future as a use. Further, all the 'expected futures' in the database (with the exception of one in Inuit) occur on this imperfective list. Expected or scheduled futures are the type we have in English sentences such as 'I go to Chicago tomorrow.'

The aspectual futures as a class differ from the other futures in their mode of expression: Three of them have zero expression, while none of the other 102 futures have zero as their only means of expression. Moreover, only 15% of the aspectual futures are polysyllabic, which, as we will argue later, suggests a generally high degree of grammaticization for grams in this class; in contrast, 40% of the non-aspectual futures are polysyllabic.[6] We predict a high degree of formal grammaticization for general perfectives and imperfectives because these meanings themselves evolve from more specific meanings through a long process (see Bybee and Dahl, 1989). We return in section 7 to a test of the hypothesis that aspectual futures have the expression properties of grams that have undergone a long evolution. For the moment, we simply display the distribution of expression types for aspectual vs. non-aspectual futures in Table 1, where the numbers indicate the percentage of

Table 1 *Modes of expression of aspectual and non-aspectual futures*

		Future type: aspectual	non-aspectual
Mode of expression:	zero	12	00
	monosyllabic	73	60
	polysyllabic	15	40

futures of each type that are expressed by zero, monosyllabic, and polysyllabic forms.

Table 1 shows the heavy concentration in monosyllabic expression for aspectual futures and their low incidence, relative to non-aspectual futures, of polysyllabic expression. Note also that the following implication holds between mode of expression and future type: If zero expression, then aspectual.

The aspectual path is different from the other paths, which we will discuss next, in that in the other cases, a morpheme or construction originally rich with lexical meaning is eroded or generalized into the grammatical meaning of future, whereas with the aspectual forms, the future meaning is just one interpretation of a much broader semantics. For those grams with zero expression, it is just one interpretation of a default meaning, which might result, for example, following the development of a past or perfective gram, the syntagmatic absence of which would then be interpretable as signalling present or imperfective.

4. MODALITY

It is often remarked that future grams tend to have modal uses (Chung and Timberlake, 1985). Our diachronic theory makes certain predictions about the nature of such modal uses, and how and when they appear upon the future paths. Before discussing the tests of these predictions on the cross-linguistic database, we first propose a division of the modality domain into three linguistically significant types.

If one works from a cross-linguistic database, one finds modalities of various sorts expressed grammatically in languages, but the classification of modalities usually put forth by philosophers (cf. Lyons, 1977), largely on the basis of English, is not adequate to describe their meanings, to classify them

into types or to understand how they change. Consequently, in Bybee and Pagliuca (1985) and Bybee (1985), we have proposed that the following three types of modality be distinguished:

i. *agent-oriented* (desire, obligation, ability, root possibility, permission)
ii. *epistemic* (possibility, probability)
iii. *speaker-oriented* (imperative, hortative, optative)

The usual distinction among modalities is between deontic and epistemic modalities, and it is often said that diachronically, deontic modalities develop into epistemic. We do not include deontic in our typology because it cuts across the modality domain in a way that is not cross-linguistically valid.

Deontic modality has to do with obligation and permission, and some authors associate deontic modality exclusively with directives, that is, utterances by which the speaker is doing something — issuing a command or granting permission — rather than with utterances in which the speaker is reporting on a situation (Palmer, 1979). While this supplies us with a classification of utterances, it does not supply us with a classification of the meanings of individual morphemes, since many of them, e.g. *must*, *have to*, *can*, *may*, are used both in directives and in reports with the same meaning. For instance, one can use *have to* to report a fact, as in (4), or to issue a command, as in (5).

(4) Todd has to go home at 5 on Thursdays.

(5) Todd, you have to go home now.

Such elements should be contrasted with grammatical elements which are used *only* in directives, such as imperatives, optatives and permissives. We might even want to say that the latter type, which we call *speaker-oriented* (because the *speaker* is doing something with the utterance) have little inherent semantic content, and that their principal function is to signal the type of speech act in which they occur.

Agent-oriented modality, in contrast, includes those elements that may be used to report on situations in which a modality is in effect. They predicate certain conditions on the agent with respect to the completion of the action or event described in the rest of the predicate, as in (4) and (6). Secondarily, as indirect speech acts, agent-oriented modalities may be used as directives for imposing obligations or granting permission, as in (5) and (7).

(6) I *can't* check books out. Only registered students *can* take out books

(7) You *can* come in now

Speaker-oriented modalities have the whole proposition in their scope and, rather than relating an agent to a predicate, relate the utterance to the speech situation by signalling what the speaker wants to accomplish with the utterance.

The linguistic reason for making this distinction is that agent-oriented modalities typically have a different expression type than speaker-oriented modalities. Specifically, as reported in Bybee (1985), agent-oriented modalities are not usually found as inflections on verbs, but speaker-oriented modalities commonly are. This is an indication that the agent-oriented modalities are less grammaticized than the speaker-oriented modalities.

A second problem with the term 'deontic' is that it includes only obligation and permission, that is, only the social modalities, those describing the social conditions on the agent. Many languages, however, also mark in similar ways other types of conditions on the agent, in particular the internal conditions of ability and desire. In addition, ability and desire undergo semantic change that is parallel to that which obligation undergoes.[7] In order to class all of these conditions together, we have proposed the broader term of agent-oriented modality.

We adopt the usual definition of epistemic modality as indicating 'the degree of commitment by the speaker to what he says' (Palmer, 1986:51).[8] Some examples from English include the use of *may* to indicate epistemic possibility in (8), and of *must* to indicate inferred probability in (9).

(8) I *may* have that book at home, but I'm not sure

(9) Jake *must* have missed the bus and that's why he's not here

Note that epistemic modality has in common with the speaker-oriented modalities propositional scope and the function of indicating what the speaker is doing with the utterance. Bybee (1985) finds epistemic modalities expressed inflectionally, often in a similar fashion to the speaker-oriented modalities, and proposes that these two should be called moods in contra-distinction to the agent-oriented modalities, which are not usually inflectional.

The prediction sense of future grams may also be considered an epistemic modality because it also has propositional scope and indicates 'the degree of commitment by the speaker to what he says'. That is, a prediction is an

assertion by the speaker that the state of affairs in the proposition will be true at some future time, or will be known to be true at some future time, as shown in (10) and (11).

(10) We shall no doubt live to see stranger things

(11) There's someone coming up the steps now with a blue coat — Oh, that'll be Martin

The third reason for choosing this classification is that from a diachronic standpoint it is more accurate to say that agent-oriented modalities develop into the other two than to say that deontic becomes epistemic.

There are a number of reasons for hypothesizing that agent-oriented modalities occur in the early stages of grammaticization: First, there are many documented cases where it is clear that agent-oriented modalities precede the other types. For instance, the earliest documented senses of the English modal auxiliaries are all agent-oriented: *Shall* meant 'to be obliged or destined'; *will* meant 'to want or desire'; *may* meant 'to be physically able'; and *can* meant 'to know (how to)'. Second, the semantics of agent-oriented modality is closer to lexical semantics, and, like lexical meaning, agent-oriented modality contributes to the propositional content of the utterance rather than operating on the whole proposition. Finally, agent-oriented modalities tend to have more co-occurrence restrictions than the other two types, since they require an animate agent and usually also an active verb.

Our data link three agent-oriented modalities to futures: Desire and obligation, which are richly attested, and ability, which is much less common. We are claiming that these modalities co-occur with future not because they develop later out of the future sense, but because they are present in the lexical antecedent of the future gram. Since we hypothesize that the only way for a future gram to have agent-oriented meaning is to have retained it from its original source, each future gram may have only *one* agent-oriented meaning. This prediction is borne out in our data — the future grams with agent-oriented meaning clearly fall into separate classes, that is, no one future gram has more than one agent-oriented sense.

On the basis of the classification of modality presented in this section and the documented diachronic development of future grams in languages such as English, we can postulate four stages in the semantic development of futures from modalities, which we will refer to as *future ages* (or FUTAGEs) 1, 2, 3 and 4. The semantic characteristics of each of these stages will be presented in the next section.

5. SEMANTIC AGE FOR FUTURES BASED ON MODALITY USES

In this section we will present and justify the four semantic ages for futures based on modality uses. All of the grams that are being classified for FUTAGE have 'future' as one their uses. That is, they are all used for prediction, or future time reference. Their classification into the four semantic ages is based on the uses they have in addition to their future use. In the following, we justify the criteria for the FUTAGEs using English examples; later on, we will test the hypothesized FUTAGEs on our cross-linguistic sample.

We have already mentioned that agent-oriented modalities are common sources for future grams. In Bybee and Pagliuca (1987) we show that developing grams often retain traces of their earlier lexical meaning, losing it only very gradually. The implication of this fact for futures from agent-oriented modalities is that if a future gram has a use expressing an agent-oriented modality, we may postulate that it is a young gram. Thus at FUTAGE 1 we place all the future grams that have agent-oriented uses. In this group are four grams expressing desire (found in Danish, Nimboran, Bongu and Tok Pisin), six grams expressing obligation (in Basque, Danish, Slave, Bari, Haka and Buriat), and two grams expressing ability (both in Cantonese). We consider these forms to be the youngest type of future. That is, we regard future grams with agent-oriented uses as being at the beginning of their development as futures.

The next stage also involves agent-oriented modality, but the particular modalities — intention, willingness and root possibility — have generalized from the more specific source meanings. In tracing the development of *shall* (from a verb meaning 'to owe' — an obligation source) and *will* (from a verb meaning 'to want' — a desire source) in English, Bybee and Pagliuca (1987) find that intermediate between the full agent-oriented senses of these auxiliary verbs and the prediction or future sense is a stage at which they are commonly used to express intention, especially of a first person agent. Such uses can still be found in the modern British English examples (12) thru (15) (from Coates, 1983).

(12) I'*ll* get you a map

(13) I'*ll* keep an eye open for it

(14) ...and I *shall* get to London as soon as I can

(15) She finally said 'don't argue with me, if you argue with me I *shall* put this phone down'

The intention sense may derive from either type of agent-oriented modality. This use develops because when a speaker states in first person what s/he *wants to* do or *has to* do, it can usually be inferred that the speaker *intends to* do just this. In fact, the use of modalities to state intentions is an important link in the chain of developments leading to future, because once the intention use becomes well-established, a further inference can be made: One can predict that the subject *will* do what s/he intends to do. Since the intention use appears after agent-oriented uses and before the prediction use, we classify grams with intention as a use as occurring at FUTAGE 2.

Root possibility refers to a type of agent-oriented possibility that commonly develops out of the ability sense, and is more general than simple ability. While ability describes enabling conditions internal to the agent, root possibility predicates more general enabling conditions which include those internal to the agent, but also conditions external to the agent, which may be physical or social. Modern English *can* is commonly used to express root possibility, as in (16) and (17).

(16) I know a place where I *can* get tickets, if you are willing to pay a little more

(17) *Can* you pick me up at three-thirty?

Since root possibility also encompasses social enabling conditions, it usually expresses permission also (see examples (6) and (7) above). In English, the permission and root possibility senses of *can* (and also *may*) developed out of the earlier ability sense. Thus we catergorize future grams with a root possibility sense as belonging to FUTAGE 2. Our database contains only two future grams with root possibility uses, in Tanga and Trukese, and in the latter case, the root possibility sense occurs only in the negative.

FUTAGE 3, the next stage in the semantic development, contains all those grams that have future as their only reported use. This stage represents the period during which the agent-oriented uses have been eroded, and no further uses have yet developed. Of course, at this point we can no longer tell if a particular gram has developed from a modality or from some other source, but at this stage, we claim, it does not matter. If the gram has evolved so as to signal prediction or future and nothing else, then semantically it belongs at this point.[9]

At the final stage postulated here, FUTAGE 4, are those grams with future as a use which also have epistemic uses (other than prediction) or

speaker-oriented modalities as uses. We have already mentioned that epistemic modalities develop out of the agent-oriented ones (Shepherd, 1982; Bybee and Pagliuca, 1985). Thus *may* used for epistemic possibility, as in example (8), develops out of earlier uses of *may* which were more like the uses of *can* in (16) and (17) (Shepherd, 1982; Bybee, 1988). Similarly, the use of *must* for epistemic probability is much later than its use for obligation.

We also propose that the use of future for imperative develops out of the prediction sense, and thus is a late use. One piece of evidence for this sequence of events is the fact that futures with imperatives as other uses in our database come from all the possible sources for futures — from desire, obligation, copula and movement verb futures. If imperative were an early use and developed only from a particular source (such as obligation), then we would expect it to be restricted to futures with that source. We propose that the imperative use of futures develops out of the reinterpretation of a secondary speech act: In a situation in which the speaker has authority over the addressee, a second person prediction, such as (18), is interpreted as a command:

(18) You will go to bed!

Thus futures with imperative uses are classified as belonging in FUTAGE 4.

Also in FUTAGE 4 are futures that are used in certain types of subordinate clauses, i.e. complement clauses and the protases of conditional sentences. Since the development of future sense from modalities involves the main clause functions of stating intentions and making predictions, future uses develop first in main clauses; only after a long period of development do they come to be used obligatorily in subordinate clauses. In fact, modalities that develop into futures often retain their older meanings in subordinate clauses, as, for example, *will* in Modern English protases, which conveys the sense of willingness rather than future:

(19) If you *will* help me, we can finish faster

Thus, in order for a future gram to come to be used obligatorily in such clauses without conveying any agent-oriented meaning, it must have already lost that meaning in other contexts, and must be quite old semantically.

To summarize, then, we postulate the paths of development shown in Figure 1, and classify the parts of the continuum according to four semantic ages.

Note that all of the grams we are classifying have future as a use, so that the ones with other uses as well actually occupy a stretch of the continuum rather than a single numbered position.

FUTAGE 1		*FUTAGE 2*		*FUTAGE 3*		*FUTAGE 4*
obligation desire	>	intention	>	future	>	probability imperative
ability	>	root possibility	>	future	>	possibility

Figure 1 *Stages of development and FUTAGE for modality senses*

6. FUTURE GRAMS FROM MOVEMENT VERBS

In this section we assign FUTAGE values to the future grams in our database that derive from movement verbs. We begin with some comments about the lexical semantics of the source constructions.

In the database are seventeen future grams whose lexical antecedents are identifiably verbs of movement; in eight cases, this verb is glossed as 'go', in nine as 'come'. The 'go'-derived futures occur in:

003 Margi (Afroasiatic, Chadic, East Chadic)
011 Cocama (Andean-Equatorial, Equatorial, Tupi)
018 Maung (Australian, Iwaidjan)
027 Atchin (Austronesian, Oceanic, N-Western New Hebrides)
043 Abipon (Ge-Pano-Carib, Macro-Panoan)
075 Mano (Niger-Kordofanian, Niger-Congo, Mande)
081 Zuni (Penutian, Penutian isolate)
088 Nung (Sino-Tibetan, Tibeto-Burman, Bodo-Naga-Kachin)

The 'come'-derived futures occur in:

003 Margi (Afroasiatic, Chadic, East Chadic)
014 Tucano (Andean-Equatorial, Tucanoan)
051 Danish (Indo-European, Germanic)
072 Mwera (Niger-Kordofanian, Niger-Congo, Benue-Congo)
073 Tem (Niger-Kordofanian, Niger-Congo, Gur)
074 Engenni (Niger-Kordofanian, Niger-Congo, Kwa)
075 Mano (Niger-Kordofanian, Niger-Congo, Mande)
080 Tojolabal (Penutian, Mayan)
089 Cantonese (Sino-Tibetan)

Taken together, then, in this sample, futures with sources in verbs of movement outnumber futures from desire, obligation or temporal adverb sources. Given the number of citations in the literature for *go-* and *come-*futures (Ultan, 1978; Marchese, 1986; Fleischman, 1982b; Heine and Reh, 1984; Bybee and Pagliuca, 1987), this is probably not remarkable. That a genetically balanced world-wide sample has resulted in an almost identical number of futures from each source, however, suggests that the sometimes alleged source asymmetry — that 'go' is more frequently a source of future than is 'come' (cf. Givón, 1973; Traugott, 1978; Fleischman, 1982b) — may be illusory. However, it should be observed that four of the languages with a 'come' source are Niger-Congo languages.

Since we find movement verbs as sources for future, pasts and progressives (see Bybee, Pagliuca and Perkins (In preparation)), we hypothesize that the semantics of movement is not sufficient in itself to give rise to the future sense. Rather, movement constructions that are sources for future grams actually signal that the subject is in the process of moving towards a goal. That is, along with movement as a component of meaning, the source of such futures includes an imperfective (or progressive) component and an allative component. Neither allativity nor imperfectivity need be overtly marked (as they are, for instance, in the English *be going to* construction), but we would not expect a motion verb with an ablative component or perfective marking to develop a future reading.[10]

Evaluating the hypothesis that the semantic components of allativity and imperfectivity are present in *come-* and *go*-derived futures is made difficult by the fact that relatively few of the lexical sources in our database are glossed with an explicit allative element such as 'towards', though *come*-derived futures seem more obliging than *go*-futures in this regard. Whether we can infer from this that the 'default' directional component for a generalized movement verb is allative, with ablative typically requiring an explicit marker, is not yet clear; we would prefer to evaluate such a possibility only after we know more about directional marking in general in the sample languages.[11] However, relatively clear cases include the Margi *come*-future auxiliary from a verb glossed as 'to approach/come near to', and Danish, with 'come' plus an allative preposition as parts of the *come*-future construction. Somewhat less clear is the case of the so-called Future Progressive in Tojolabal, the form of which may descend from a verb glossed as 'to enter'. The *come*-derived future in Mano, *nu pèá*, which is used as a future progressive and a venitive, consists of 'come' followed by the progressive marker *pèá*, which is 'perhaps identical' with the allative postposition *píè*, *pía* (Becker-Donner, 1965:37);

Mano's *go*-derived future, on the other hand, is simply identified with the verb 'go'. As a case in which evidence for an explicit allative component is lacking for a *come*-derived future but the expected aspectual restriction seems to be in force, we may cite Tem, whose future is built on *kɔni* 'to come', inflected for agreement and for non-past (Present or Remote Future) tense, which is expressed by *n*-, the full expression being:

[person/number + n + kɔni] [person/number + main verb].

Movement-derived futures do not provide the richness of semantic nuance that modality-derived futures do, and thus do not allow us the opportunity of making fine distinctions in semantic age. However, we can map movement-derived futures onto the FUTAGEs established in the last section by examining their other uses.[12]

It will be recalled that grams which express intention are classified as being at FUTAGE 2, and that intention uses can come from any of the agent-oriented modalities. Intention can arise in movement-derived futures as well, as can be seen in (20) and (21), examples of the *be going to* future from Coates (1983):

(20) We're not going to let you walk home on your own

(21) I'm going to give you something

Note that this use comes fairly directly from the literal meaning of 'the subject is on a path moving towards a goal'. The only necessary change is for the movement and path to be taken figuratively for the intention sense to arise.

Thus all futures used for intention are classified as FUTAGE 2, no matter what their lexical source.

Six out of the nine *come*-futures in our database have 'immediate future' rather than simple future as their use. We regard this as a more specific, semantically richer type of future than the simple future sense, and therefore classify it as being younger than the simple future. So all instances of immediate future (which are much more common for *come*-futures than for futures from any other source) are classified as FUTAGE 2, just one step younger than simple futures.

Other uses found with movement-derived futures, such as speaker-oriented, epistemic and subordinate uses, classify a gram as FUTAGE 4, just as they do for modality-derived futures. As we mentioned earlier, grams whose sources are unknown are classified according to this same scheme, given the meanings expressed in their other uses. Also classified in this way are the

Table 2 *Meaning components for each FUTAGE*

FUTAGE 1	*FUTAGE 2*	*FUTAGE 3*	*FUTAGE 4*
obligation	intention	future	probability
desire	root possibility		possibility
ability	immediate future		imperative
			use in complements
			use in protases

grams from the less common sources — the temporal adverbs, and a few grams derived from the meaning 'enter a state'. Table 2 displays the meaning components associated with each of the FUTAGEs.

To summarize, all of the non-aspectual future grams in our database can be classified as belonging in one of these four groups on the basis of the other meanings that they express, no matter whether we have information about their lexical sources or not. The aspectual futures are not classified in these four categories, because, on the basis of the other meanings they express, we hypothesize that the aspectual futures develop in an entirely different way: As we argued in Section 3, aspectual futures develop as imperfectives or perfectives, which happen to accomodate future readings, but they do not develop explicit future semantics. The grams that we have included in FUTAGEs 1 through 4 are grams that are undergoing similar developments, and include grams that are on the same paths. The aspectual futures are on an entirely different path. However, the range of meanings expressed by aspectual futures suggests to us that these grams are very advanced in their semantic development. We therefore predict that they are more grammaticized than explicit future grams, and thus assign them a semantic age as well; we will here call it FUTAGE A.

Having established a means of classifying individual grams according to how advanced their semantic development is, we turn now to the classification of the formal development of these grams in order to test our hypotheses concerning the universality of the paths of development for future and the parallel reduction of meaning and form.

7. FORM/MEANING COVARIATION

In Bybee (1985) it is argued that certain formal properties of affixes are related to the interaction of their semantic properties with those of the stem.

In particular, the degree of semantic relevance of the meaning of the affix to that of the stem correlates with the position of the affix and its degree of fusion with the stem. While in Bybee (1985) this principle is formulated in terms of the supercategories of aspect, tense, mood, person, number and so on, Bybee, 1986 shows that within a language affixes are not positioned according to membership in these supercategories, but rather correlations of meaning with form apply to individual affixes. That is, it is not true that languages tend to have all of their tense affixes in a single position, all of their aspect affixes in a single position, and so on. On the contrary, it is just as common to find, for example, a future affix in a different position from the past affix in the same language. The reason for this is that in most cases such affixes have developed independently of one another, and their positioning is governed by the source constructions from which the affixes arose diachronically, which in turn is governed by the general typological features of the language at that time.

Taking this historical perspective on the development of affixes (or grammatical morphemes in general) means that in addition to the formal properties predicted by semantic relevance, there may be formal properties predictable by the stage of development a gram has reached in any particular synchronic slice. Thus, intersecting the form/meaning correlation whose basis is the relevance principle will be a form/meaning correlation due to the degree of grammaticization of the form. We would then expect grams that are older — i.e. that have undergone more development — to be closer to the stem, more fused and shorter or more reduced in segmental material than younger grams of equal relevance.

Some findings of this nature are presented in Bybee (1985) and in Bybee and Pagliuca (1985), where it is reported that the agent-oriented modalities of obligation, permission, desire and ability, which are relatively close to lexical meanings, tend to occur in periphrastic expressions, while the more grammaticized epistemic and speaker-oriented moods can be either bound or free. Similarly, Dahl (1985) reports that in the tense/aspect domain the following correlations are found in his sample of 64 languages:

periphrastic	*bound*
progressive	imperfective
perfect	perfective
	past

Since there is evidence that progressives develop into imperfectives and perfects develop into perfectives or pasts, Bybee and Dahl (1989) argue that the process of grammaticization accounts for these correlations as well.[13]

In Dahl's (1985) data, the future grams are split between periphrastic and bound — 27 are periphrastic, 23 are bound. If form and meaning change in parallel in grammaticization, we would predict that the periphrastic ones are 'younger' and the bound ones 'older' semantically. In order to test this hypothesis, in the next section we establish three measures of 'degree of grammaticization of form' which we attempt to correlate with the semantic ages of futures established above.

8. DEGREE OF FORMAL GRAMMATICIZATION

The GRAMCATS database contains a number of variables that are designed to indicate varying degrees of grammaticization of form. For present purposes, we have chosen to divide these variables into three groups: Those that are indicative of the degree of fusion with the verb which the gram has undergone (FUSION), those that are indicative of the extent to which the gram has lost its autonomy and become dependent (DEPENDENCE), and those that indicate the extent to which the gram's phonological bulk has been reduced (SHORTNESS). In the next three subsections we explain how each of these measures is calculated.

8.1. Fusion with the verb

In Dahl's study, and to a large extent in Bybee (1985), whether a gram is considered an affix or not depends upon whether the orthographic conventions employed by the questionnaire respondent (in Dahl's study) or the author of the reference grammar (in Bybee's study) treat the gram as an affix (written together with the verb stem) or a free morpheme (written separately from the verb). Unfortunately, orthographic conventions only partially reflect the phenomenon of affixation, which is diachronically gradual and complex, involving several different factors. We have therefore attempted to construct a much finer measure of the degree of fusion of the gram to the stem by taking into account in our coding procedure for each form the following factors:

1. *Written bound*: Here we note the orthographic conventions of the author, by indicating Y for yes, N for no, H for hyphenated and S for sometimes written bound.

2. *Open class intervening*: Here we note whether (Y) or not (N) open class material can come between the gram and the main verb.

3. *Phonological process conditioned by stem*: Y is entered if the stem causes allomorphy of a phonological nature in the gram; N is entered if it does not.

4. *Lexical conditioning*: Y is entered if the lexical class of the verb (the conjugation class) conditions allomorphy in the gram; N is entered if it does not.

5. *Conditions stem change*: If the gram conditions a change in the verb stem, the following are entered: T for a change in tone, S for a stress change, V for a vowel change and C for a consonant change.

These five factors together allow us to calculate for each gram a figure that we call 'fusion with verb' by assigning numerical values to the variables entered for each field, as follows:

FUSION:

1.	Written bound:	N = 0; H or S = 1; Y = 2
2.	Open class intervening:	N = 1; Y = 0
3.	Phon. process cond'd by stem:	N = 0; Y = 1
4.	Lexical conditioning:	N = 0; Y = 2
5.	Conditions stem change:	N = 0; T or S = 1; V or C = 2

The numerical values are chosen so that a higher degree of grammaticization will be reflected in a higher score. Thus, for each variable, the higher value indicates greater grammaticization, with the following justification.

Assuming that grams develop from free lexical items, those that can be and are written bound are more grammaticized than those that are not, with hyphenated and variable representation assigned an intermediate value. 'Open class intervening' refers to the possibility of non-grammatical elements occurring between the gram and the verb stem. Such possibilities usually decrease as a gram fuses with the verb, although languages with incorporation are exceptions to this tendency. Phonological processes in the gram conditioned by the stem will occur only as the gram starts to become more fused with the stem. This would be an early indication of fusion. 'Lexical conditioning' refers to allomorphy conditioned by the lexical classification of the verb stem (i.e. conjugation class). Such allomorphy usually arises only after the affix has been attached to the stem long enough for phonological processes to lose their productivity and for reanalysis to occur. Under 'conditions stem change', tone or stress changes count less than consonant or vowel changes, since the

former may be predictable on a phonological basis, while we exclude that possibility for the latter. This scale ranges from 0 to 9, and, as we said above, a higher number indicates a greater degree of grammaticization.

To illustrate how this scale works, we apply it to two well-known English grams, which display typical differences in expression properties, the English Past Tense (*-ed*) and the English Future, *will*.

English Past Tense:

1. The Past Tense gram is written bound (score 2);
2. It allows no open class items to come between it and the verb stem (score 1);
3. It is affected by a phonological process conditioned by the stem, i.e. the devoicing found after voiceless consonants (*helped* [hɛlpt]) (score 1);
4. It has lexically conditioned allomorphy, since we would consider the stem changes of irregular verbs to be lexically conditioned (score 2); and
5. It conditions stem changes in some verbs, such as *kept* (score 2).

Total FUSION score: 8

Will:

1. *Will* is not written bound to the verb (score 0);
2. Open class items do intervene between *will* and the verb, in particular, manner adverbs (*I'll gladly help.*) (score 0);
3. There are no phonological processes conditioned by the verb stem (score 0);
4. There is no lexical conditioning — *will* does not vary according to which main verb it is used with (score 0);
5. *Will* does not condition any changes in the verb stem (score 0).

Total FUSION score: 0

The procedure produces the desired results, since *-ed* emerges with a much higher score than *will*, reflecting the fact that the former is a suffix and the latter is an auxiliary.

8.2. Dependence

The DEPENDENCE factors are kept separate from FUSION because there are cases in which a gram undergoes changes in form during grammatici-

zation which reduce its autonomy without resulting in fusion with the verb. We believe the lack of fusion is usually due to the position of the gram. This scale is called 'dependence' since it measures the extent to which the gram is losing its autonomy independently of whether or not it is fusing with the verb. For this measure, the following criteria are considered.

1. *The number of allomorphs which are not purely phonetically conditioned*: Entered as 0, 1, 2, 3 or >3.
2. *Phonological conditioning for allomorphs*: Possible values are Y and N.
3. *Morphological conditioning for allomorphs*: This refers to allomorphy conditioned by other grammatical morphemes, and the responses are Y and N.
4. *Suprasegmental reduction*: Here we note whether (Y) or not (N) the gram has stress or non-neutral tone; if stress alternates, A is recorded.

The numerical values are as follows:

DEPENDENCE:
1. Other allomorphs: 0=0; 1 or 2=2; 3=3; >3=4
2. Phonological conditioning: N=0; Y=1
3. Morphological conditioning: N=0; Y=2
4. Stress: N=2; A=1; Y=0

This scale ranges from 0 to 9, and again the higher number indicates a greater degree of grammaticization. The values assigned are justified as follows:

As a gram becomes more dependent upon surrounding elements, it develops allomorphs, especially due to phonological processes, some of which will be conditioned by other grammatical morphemes. Thus the number of allomorphs and the presence of phonological and morphological conditioning are indicative of greater dependence. Note that phonological conditioning could be controlled by the verb stem, indicating fusion with the verb, but since such conditioning is not necessarily verb-induced, we have included phonological conditioning under dependence. Since lexical items are usually stressed and grammatical morphemes usually not, we consider the lack of stress to indicate higher grammaticization. For tone languages we assume the comparable development would be the loss of high or non-neutral tone.

To illustrate this scale, we will again calculate the values for English *-ed* and *will*.

English Past Tense:

1. We count the Past Tense as having more than 3 allomorphs because it has [d] in addition to [ɨd], as well as several different vowel and consonant changes internal to the stem (score 4);
2. The [ɨd]/[d] alternation is phonologically conditioned (score 1);
3. There are no allomorphs conditioned by other grammatical morphemes (score 0);
4. The gram is always unstressed (score 2).

Total DEPENDENCE score: 7

Will:

1. If we count the reduced version of *will* [əl] as an allomorph, and also its form in contraction with *not*, then it has two allomorphs (score 2);
2. They are not phonologically conditioned, as far as we know (score 0)
3. The allomorphy in *won't* is morphologically conditioned (score 2);
4. *Will* is normally not stressed (score 2)

Total DEPENDENCE score: 6

On the DEPENDENCE scale, then, the Past Tense and *will* come out with similar scores, demonstrating that at least in some cases, FUSION and DEPENDENCE are separable measures.

8.3. Shortness

The third scale measures the extent of the segmental reduction of a gram, by starting with a value of 10 and subtracting points for each segment in the gram as follows:

SHORTNESS:
C = −1
V = −2
V: = −3

The resulting scale shows a higher number for a shorter form, and a lower number for a longer form. This non-iconic representation was decided upon so that all of the scales have values running in the same direction — a higher number for a higher degree of grammaticization.

The SHORTNESS scale includes another variable. Our coding procedure

requires us to record the longest allomorph as basic in most cases, but we also provided for the existence of a reduced allomorph, that is, one which appears to be the same as the basic allomorph, but with some segment(s) or feature(s) absent. Thus we add 2 points if a reduced allomorph exists. This scale ranges approximately from 0 to 10, although higher and lower values are possible.

Again, applying this procedure to the English Past Tense and to *will*, we obtain the following results:

English Past Tense:
The longest allomorph is [ɨd], which consists of one vowel (score −2) and one consonant (score −1); 3 from 10 leaves 7. This gram also has a reduced allomorph, [d], so we add two points, yielding a total SHORTNESS score of 9.

Will:
The longest allomorph is [wIl], which consists of a consonant (−1), a vowel (−2) and another consonant (−1); 4 from 10 leaves 6. Since there is also a reduced allomorph, we add 2, yielding a total SHORTNESS score of 8.

9. CORRELATIONS

Given these three measures of degree of grammaticization, we calculated values on each of these scales for each future gram in our database. We then attempted a correlation of these values with the division of future grams into the four semantic ages we presented in sections 4 and 5.

We predict that forms with higher FUTAGEs have significantly higher scores for the FUSION, DEPENDENCE, and SHORTNESS variables. The statistical test of this hypothesis may be seen in Tables 3 and 4. Table 3 shows the mean values for DEPENDENCE, FUSION and SHORTNESS for grams at each of the four FUTAGEs. Here we observe that the numerical measures of formal grammaticization rise as the FUTAGE rises. This is true for each FUTAGE in DEPENDENCE and FUSION, but for SHORTNESS, FUTAGE 2 shows a slightly lower score than FUTAGE 1, and FUTAGE 4 lower than FUTAGE 3. However, if FUTAGEs 1 and 2 are considered together and FUTAGEs 3 and 4 together, it is clear that 3 and 4 are more grammaticized than 1 and 2.

In Table 4 we show two ways to assess the significance of the correlations shown in Table 3, Kendall's Tau B and Gamma. The Tau-B statistic was used

Table 3 *Correlations of DEPENDENCE, FUSION and SHORTNESS by FUTAGE*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPENDENCE by levels of FUTAGE				
For Entire Population		3.1569	2.398	102
FUTAGE	1.00	2.0000	2.4863	12
FUTAGE	2.00	2.3750	1.9134	32
FUTAGE	3.00	3.7143	2.6521	35
FUTAGE	4.00	4.0000	2.7634	23
Summaries of FUSION by levels of FUTAGE				
For Entire Population		2.5098	1.7952	102
FUTAGE	1.00	1.5000	1.7838	12
FUTAGE	2.00	2.3750	1.5811	32
FUTAGE	3.00	2.5143	1.9000	35
FUTAGE	4.00	3.2174	1.7309	23
Summaries of SHORTNESS by levels of FUTAGE				
For Entire Population		6.3529	2.2719	102
FUTAGE	1.00	5.7500	1.8153	12
FUTAGE	2.00	5.5000	2.7001	32
FUTAGE	3.00	7.1429	2.0167	35
FUTAGE	4.00	6.6522	1.7738	23

Table 4 *Significance figures for DEPENDENCE, FUSION and SHORTNESS by FUTAGE*

Statistic	Value	Significance[14]
Statistics for DEPENDENCE by FUTAGE		
Kendall's Tau B	.25144	.0004
Gamma	.31420	
Statistics for FUSION by FUTAGE		
Kendall's Tau B	.22163	.0018
Gamma	.29530	
Statistics for SHORTNESS by FUTAGE		
Kendall's Tau B	.17975	.0062
Gamma	.22705	

because it has a statistical significance measure corresponding to it. The significance figure tells us what the probability is that the observed correlation occurred by chance. In the social sciences, any figure less than .05 (which means that the probability is 5 in one hundred) is considered statistically significant. By this measure the correlations of FUTAGE with each of the three types of formal grammaticization is highly significant.

Gamma was used because it has a proportional reduction in error interpretation; that is, the Gamma figure tells us how well we can predict the pairing of meaning and form if we know the FUTAGE and DEPENDENCE, FUSION and SHORTNESS scores for each gram. Whereas random pairings would produce approximately 50% error, knowing, for example, the FUTAGE and DEPENDENCE scores for a gram reduces that error by .31420, i.e. 31%.

Tables 3 and 4 show, then, that there is, as predicted, a highly significant correlation between the semantic classification we have imposed on the future grams and their formal properties. Moreover, these correlations are in the predicted direction — the semantic properties hypothesized to belong to older grams correlate with the formal properties that are accrued in the process of grammaticization. We conclude, then, that there is good reason to believe that future grams develop in very similar ways across unrelated languages, and that semantic and formal changes move in parallel in grammaticization.

10. THE ROLE OF LANGUAGE TYPOLOGY

It has long been observed that languages differ in their morphological types along just the formal parameters we have been discussing; in some languages grams are typically more fused, more dependent and shorter than in other languages. It is possible, then, that the form/meaning correlations we have presented in the preceding section may be affected by language type. There are two ways in which language type may interact with these correlations:

(1) Typology may interfere with the correlations. Thus, for example, if a language tends toward high fusion, then a semantically young gram may show higher fusion in that language than a semantically older gram in a less fusional language. In such a case, if we control for language type, then our form/meaning correlations will be stronger.

(2) Typology may contribute to the correlations. Thus, for example, a particular language may not only have more fused grams, but may also tend to have older grams. In this case, if we control for language type, then the correlations between form and meaning will not be as strong.

In order to control for language type, we standardize the FUSION, DEPENDENCE and SHORTNESS scores for each future gram. That is, we recalculate the scores so that e.g. a long gram in a language with mostly long grams will not be treated as long as a long gram in a language with mostly short grams. In order to calculate this standardized score, we derive the means for FUSION, DEPENDENCE and SHORTNESS for each language, giving us a quantification of the language type along our three parameters. Then a standardized score for each parameter for each future gram is derived by subtracting from the gram's raw score the mean score for the language and dividing by the standard deviation for the language. We can then apply the same methods used above for the raw scores to the standardized scores, as shown in Table 5.

The standardized scores look different from the raw scores because a gram that is less grammaticized than the mean for the language will have a negative value, and a gram that is more grammaticized than the mean a positive value. These values are often less than 1.0 because they are divided by the standard deviation. Thus they are actually measures of deviation from the mean for the language.

Note in Table 5 that, in general, the degree of formal grammaticization increases with increasing FUTAGE. There is a deviation from this trend only in FUTAGE 4 in DEPENDENCE and SHORTNESS.

Table 6 shows the tests of significance for these standardized scores. Note that the form/meaning correlation is still highly significant statistically, but that in comparison with Table 4, the degree of significance has gone down, indicating that language typology is contributing to the correlation. One way of putting it would be to say that some languages tend to have older grams — both semantically and formally — than other languages.

If it is the case that some languages tend to have grams that are more highly grammaticized than other languages, and if grammaticization involves both form and meaning, then it will also be the case that the traditional division of languages into isolating, agglutinative and fusional involves meaning as well as form. Of course, Sapir proposed that meaning was involved in morphological typology in 1921, but it has been impossible until now to test

Table 5 *Standardized DEP, FUSE and SHORT by FUTAGE*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPSTD by levels of FUTAGE				
For Entire Population		.0284	.9941	102
FUTAGE	1.00	−.2900	.9001	12
FUTAGE	2.00	−.2693	.7510	32
FUTAGE	3.00	.2720	1.0648	35
FUTAGE	4.00	.2382	1.1227	23
Summaries of FUSESTD by levels of FUTAGE				
For Entire Population		−.1104	.9390	102
FUTAGE	1.00	−.4488	.6517	12
FUTAGE	2.00	−.2171	.9318	32
FUTAGE	3.00	−.1573	.9328	35
FUTAGE	4.00	.2861	1.0104	23
Summaries of SHORTSTD by levels of FUTAGE				
For Entire Population		.0133	.8472	102
FUTAGE	1.00	−.2819	.6668	12
FUTAGE	2.00	−.2418	.8718	32
FUTAGE	3.00	.3363	.8824	35
FUTAGE	4.00	.0306	.7002	23

Table 6 *Significance figures for standardized scores*

Statistic	Value	Significance
Statistics for DEPSTD by FUTAGE		
Kendall's Tau B	.15384	.0104
Gamma	.18079	
Statistics for FUSESTD by FUTAGE		
Kendall's Tau B	.13985	.0161
Gamma	.16465	
Statistics for SHORTSTD by FUTAGE		
Kendall's Tau B	.15341	.0105
Gamma	.18026	

this hypothesis. However, since we have found a way of quantifying both meaning and form, we have been able to find support for this hypothesis. Another way of testing the hypothesis is to attempt to derive a typology of languages by correlating the formal parameters of DEPENDENCE, FUSION and SHORTNESS with the semantic age of the future grams of each language. Thus Table 7 shows the correlations between the semantic age of a gram and the means for each formal parameter for the languages in which the grams occur. Table 8 shows that we find statistically significant correlations such that semantically older future grams tend to occur in languages with more dependent, fused and short grams overall.

If what we find for futures is true of other gram-types as well, it would mean that morphological typology is not just a matter of form, but also of semantic content. That is, we might conclude that grams in isolating languages are *less* grammaticized — in both meaning and form — than grams in

Table 7 *Correlations of the means for DEPENDENCE, FUSION and SHORTNESS for each language with FUTAGE*

Variable	Value Lable	Mean	Std Dev	Cases
Summaries of DEPMEAN by FUTAGE				
For Entire Population		3.1435	1.2273	102
FUTAGE	1.00	2.7055	1.3699	12
FUTAGE	2.00	3.0128	1.1419	32
FUTAGE	3.00	3.1692	1.2065	35
FUTAGE	4.00	3.5146	1.2710	23
Summaries of FUSEMEAN by FUTAGE				
For Entire Population		2.6710	1.0673	102
FUTAGE	1.00	2.0475	1.2568	12
FUTAGE	2.00	2.5991	.8463	32
FUTAGE	3.00	2.7282	1.1422	35
FUTAGE	4.00	3.0093	1.0344	23
Summaries of SHRTMEAN by FUTAGE				
For Entire Population		6.3983	.9989	102
FUTAGE	1.00	6.3059	.8453	12
FUTAGE	2.00	6.1986	1.0267	32
FUTAGE	3.00	6.5143	1.0118	35
FUTAGE	4.00	6.5477	1.0243	23

Table 8 *Significance figures for the language means by FUTAGE*

Statistic	Value	Significance
Statistics for DEPMEAN by FUTAGE		
Kendall's Tau B	.12095	.0277
Gamma	.14270	
Statistics for FUSEMEAN by FUTAGE		
Kendall's Tau B	.14647	.0134
Gamma	.17290	
Statistics for SHRTMEAN by FUTAGE		
Kendall's Tau B	.11833	.0298
Gamma	.13983	

agglutinative languages, which are in turn less grammaticized than grams in fusional or inflectional languages.

The significant correlations between the means for DEPENDENCE, FUSION and SHORTNESS for each language and the semantic age of the future gram or grams in the language points to a dimension in grammaticization that needs much further investigation. However, it is important to remember that even when we control for language type the correlations between meaning and form predicted by our grammaticization hypotheses are still highly significant.

11. ASPECTUAL FUTURES

In section 3 we proposed that aspectual grams that have a future use are semantically very old, having gone through a long development before reaching the stage of signalling imperfective or present (Bybee and Dahl, 1989; Bybee, Pagliuca and Perkins, In preparation). We can test this hypothesis by comparing the measures of formal grammaticization for aspectual futures with those of the other futures. This comparison is in Table 9, where the aspectual futures are labelled as FUTAGE A. The significance figures are given in Table 10.

From Tables 9 and 10, it can be seen that aspectual futures do indeed display the properties of older grams, being more dependent, more fused and shorter than all the other FUTAGEs.

Table 9 *DEPENDENCE, FUSION and SHORTNESS by 5 FUTAGES*

Variable	Value Label	Mean	Std Dev	Cases
Summaries of DEPSTD by FUTAGE				
FUTAGE	1.00	−.2900	.9001	12
FUTAGE	2.00	−.2693	.7510	32
FUTAGE	3.00	.2720	1.0648	35
FUTAGE	4.00	.2382	1.1227	23
FUTAGE	A	.3379	1.0567	27
Summaries of FUSESTD by FUTAGE				
FUTAGE	1.00	−.4488	.6517	12
FUTAGE	2.00	−.2171	.9318	32
FUTAGE	3.00	−.1573	.9328	35
FUTAGE	4.00	.2861	1.0104	23
FUTAGE	A	.5590	1.0483	27
Summaries of SHORTSTD by FUTAGE				
FUTAGE	1.00	−.2819	.6668	12
FUTAGE	2.00	−.2418	.8718	32
FUTAGE	3.00	.3363	.8824	35
FUTAGE	4.00	.0306	.7002	23
FUTAGE	A	.8252	.9475	27

Table 10 *Significance figures for 5 FUTAGEs*

Statistic	Value	Significance
Statistics for DEPSTD by 5 FUTAGEs		
Kendall's Tau C	.19695	.0015
Gamma	.20241	
Statistics for FUSESTD by 5 FUTAGEs		
Kendall's Tau C	.23631	.0003
Gamma	.24384	
Statistics for SHORTSTD by 5 FUTAGEs		
Kendall's Tau C	.28829	.0000
Gamma	.29619	

12. CONCLUSION

We hope to have demonstrated how formal and semantic information on grammatical material in the languages of a large, carefully constructed sample of the world's languages can be profitably used in evaluating and testing hypotheses about grammaticization. In particular, analysis of the information in our database has allowed us to infer that the hypothesis of unidirectionality with respect to the evolution of future grams is, in large measure, confirmed. It has also allowed us to identify language typology as a significant factor in the grammaticization process. While the role of typology needs more investigation, even when the influence of typology is controlled for, the results provide strong corroboration of the hypothesis that there are certain universal pathways for the development of futures from different lexical sources. Some senses or uses, in particular, agent-oriented modality uses, occur early on in the evolution along a particular path, and are unique to that path. These uses, which are comparatively close to the lexical meaning of the source material, eventually erode and are gradually lost, giving way to uses which are semantically more general, and are the points of convergence for futures from distinct sources. Moreover, futures with only or predominantly 'late' uses are morphologically reduced relative to 'young' futures, suggesting a dynamic correlation of generalization of meaning with concomitant reduction in the formal expression of future grams. This correlation is shown to have strong statistical support, which suggests in turn that a rigorous evaluation of a general hypothesis, by which formal and semantic reduction proceed in parallel, is not only feasible but may in fact turn out positive.

NOTES

* We are very grateful to Daniel Devitt, Martin Haspelmath, Bernd Heine, Soteria Svorou, Sandra Thompson, Elizabeth Traugott, and Nancy Woodworth for detailed comments on this paper, and to the participants at the Eugene conference for helpful advice and information. The research on which this report is based was supported in part by NSF grant BNS 83–18262 and by a grant from the John Simon Guggenheim Foundation to the first-named author.

1. For a detailed description of the sampling algorithm, see Perkins (1984).

2. Two qualifications are necessary here. First, in languages in which much of the verbal morphology is fused into an auxiliary that is fixed in position with respect to the clause, i.e. in second position (e.g. Papago), we relaxed the rule that the grams had to be fixed in position with respect to the verb. Second, since there is a continuum between closed

and open class, we arbitrarily decided not to code any grams belonging to a class with more than twelve members.

3. A language entered our database with no future gram if it happened that the author of the grammar did not specifically mention 'future' in connection with any of the grams that met our coding criteria.

4. We follow the convention of capitalizing the first letter of the names of language-specific grams (which are proper names), while writing in lower case the 'meaning components' which are defined independently of any particular language. For definitions of meaning components referred to in this paper, see Appendix 1.

5. In our data, only one language has a future gram that is reported to occur in protases. However, since this is not something that would necessarily be mentioned in a reference grammar, it is good to have confirmation of this tendency from Dahl's questionnaire survey, where fewer than 20% of the futures in his 64 languages could occur in *when*, *if*, *even if* or *whatever* clauses.

6. The figure of 15% for polysyllabic aspectual futures includes the two aspectual futures of Alyawarra, a central Australian language, one in Baluchi, and an affirmative form in Udmurt (the Udmurt negative form, because it is built on the affirmative, was not counted as a separate form for Table 1). All four are polysyllabic, but the Alyawarra forms — the disyllabic suffixes *-ima* and *-iyla* — are perhaps only marginally so, since in this language, all words end in *a*, and this vowel is elided in connected speech (Yallop 1977:30).

7. Permission seems to be a special case, since in many known cases, it develops as just one of several readings of root possibility, usually by the generalization of a gram signalling ability (see Bybee and Pagliuca, 1985; and Bybee, 1988).

8. We are not, however, including here the closely-related domain of evidentials, which identify the information source the speaker relies on for his/her assertion. For a treatment of the grammaticization of evidentials cross-linguistically, see Willett (1988).

9. Positing a FUTAGE with 'future' as the only use does not imply that future grams are ever completely devoid of modal nuances. We would, however, suggest that the modal nuances are less salient at this stage. Furthermore, for some of our languages, a gram might be classified as having a future use but no other because the reference material did not attend to semantic detail. Such cases detract from the accuracy of our classification, and consequently from the correlations we seek to establish in the preceding sections. However, the fact that these correlations are still highly significant suggests that if the quality of data were higher, our results would be even better.

10. The case of Catalan is sometimes cited as a problem for a theory that makes universal predictions about the relation of grammatical to source meaning, since in Catalan a perfective past has developed from the verb 'to go' plus the infinitive. This periphrastic construction differs from the *go*-futures of Spanish and French, however, in lacking the allative preposition *a*. Furthermore, it does not use the regular present forms of the verb *anar*, but special forms, in all but 1st and 3rd Singular. The long versions of these special forms quite transparently consist of a stem *va* (the 3SG form) plus the inflectional perfective suffixes (Badía Margarit, 1951: 326ff.).

11. For information on directional marking in a 26-language subset of the GRAMCATS sample, see Svorou (1988).

12. We have one case each of a future gram also used for andative and venitive. An andative

conveys the sense that the subject is going towards some location away from the speaker in order to perform some activity, that is, it conveys the literal meaning of 'be going to'. A venitive conveys the meaning that a subject is coming to a location near the speaker in order to perform some activity, 'x is coming to' do something. However, most of the andatives and venitives in our database do not have future as a use. Bernd Heine (p.c.) reports that andative and venitive are usually derivational in African languages and occur in different types of constructions than movement-futures (see also Heine and Reh, 1984:142). Our data on andatives and venitives are thus in accord with this report. The situation parallels that of 'desire' components, which may become either futures or derivational desiderative markers.

13. For a finer-grained analysis of the correlation of meaning and form in the aspectual domain, see Bybee, Pagliuca and Perkins (In preparation).

14. Since our hypothesis predicts the direction of the association, the applicable statistical significance is a one-tailed significance figure and is so reported here and in the balance of the paper.

15. Three of the grams listed under FUTAGE 3 below are problematic. At the Eugene conference, Phil Young informed us that Alphonse (1956) is unreliable as a source for Guaymi, and graciously supplied us with more accurate source material. Although we recoded this language, the statistical correlations were, unfortunately, run on a version of the database that did not include the recoded future. Similarly, language 047 in the original sample and coding was Umbrian, which we later, for reasons of data quality, abandoned in favor of Latin, but, again, the statistics were run on the version of the database which contained the older forms. Finally, we have been unable to verify that the Kadugli form is in fact Kadugli rather than that of a closely-related dialect, and therefore cannot regard the form as representing an accurate coding. To our knowledge, there are no other errors of this type in the data presented here.

REFERENCES

Allen, Jerry and Janice Allen. 1965. *Halia Language Course*. Port Moresby, Papua New Guinea: Department of Information and Extension Services.

Alphonse, Ephraim S. 1956. *Guaymí Grammar and Dictionary with some Ethnological Notes*. Washington, D.C.: U.S. Government Printing Office [*Bureau of American Ethnology Bulletin* 162].

Anceaux, J.C. 1965. *The Nimboran Language: Phonology and Morphology*. 'S-Gravenhage: Martinus Nijhoff [*Verhandelingen van het Koninklijk Instituut voor Tall-, Land- en Volkenkunde* 44].

Badía Margarit, Antonio. 1951. *Gramática Histórica Catalana*. Barcelona: Noguer.

Barker, Muhammad Abd-al-Rahman and Aqil Khan Mengal. 1969. *A Course in Baluchi*, Vols. 1 and 2. Montreal: Institute of Islamic Studies, McGill University.

Barnard, J.T.O. 1934. *A Handbook of the Rawang Dialect of the Nung Language*. Rangoon: Government Printing.

Becker-Donner, Etta. 1965. *Die Sprache der Mano*. Wien: Hermann Böhlaus [*Österreichische Akademie der Wissenschaften, Philosophisch-Historisch Klasse*, Sitzungsberichte 245, Band 5].

Bell, F.L.S. 1977. *Tanga-English, English-Tanga Dictionary*. Sydney: University of Sydney [*Oceania Linguistic Monographs* No. 21].
Benton, Richard A. 1971. *Pangasinan Reference Grammar*. Honolulu: University of Hawaii Press.
Bright, William. 1957. *The Karok Language*. Berkeley: University of California Press [*UCPL* Vol. 13].
Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
Bybee, Joan L. 1986. "On the nature of grammatical categories." *ESCOL '85*:17–34.
Bybee, Joan L. 1988. "Semantic substance vs. contrast in the development of grammatical meaning." *Berkeley Linguistics Society* 14:247–64.
Bybee, Joan L. and Östen Dahl. 1989. "The creation of tense and aspect systems." *Studies in Language* 13:51–103.
Bybee, Joan L. and William Pagliuca. 1985. "Cross-linguistic comparison and the development of grammatical meaning." In *Historical Semantics and Historical Word Formation*, J. Fisiak (ed.), 59–83. The Hague: Mouton.
Bybee, Joan L. and William Pagliuca. 1987. "The evolution of future meaning." *Papers from the 7th International Conference on Historical Linguistics*, A.G. Ramat, O. Carruba, and G. Bernini (eds), 109–122. Amsterdam: John Benjamins.
Bybee, Joan L., William Pagliuca and Revere D. Perkins. In preparation. *The Origin and Development of Verbal Categories in the Languages of the World.*
Capell, A., and H.E. Hinch. 1970. *Maung Grammar: Texts and Vocabulary*. The Hague: Mouton.
Capell, A. and J. Layard. 1980. *Materials in Atchin, Malekula: Grammar, Vocabulary and Texts*. Canberra: The Australian National University [*Pacific Linguistics Series* D No. 20].
Caughley, Ross Charles. 1982. *The Syntax and Morphology of the Verb in Chepang*. Canberra: Australian National University [*Pacific Linguistics Series* B No. 84].
Chung, Sandra and Alan Timberlake. 1985. "Tense, aspect and mood." In *Language Typology and Syntactic Description*, Vol. III, T. Shopen (ed.), 202–58. Cambridge: University Press.
Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. London: Croom Helm.
Cortade, Jean-Marie. 1969. *Essai de grammaire Touareg (dialecte de l'Ahaggar)*. Alger: Institut de Recherches Sahariennes, Université d'Alger.
Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.
Der-Houssikian, Haig. 1980. *Tem Grammar Handbook*. Brattleboro, VT: For ACTION/Peace Corps.
Dyen, Isidore. 1965. *A Sketch of Trukese Grammar*. New Haven: American Oriental Society [*American Oriental Essay* No. 4].
Faust, Norma. 1971. "Cocama clause types." In *Tupi Studies* I, D. Bendor-Samuel (ed.), 73–105. Norman, OK: Summer Institute of Linguistics [SIL Publications in Linguistics and Related Fields No. 29].
Fleischman, Suzanne. 1982a. *The Future in Thought and Language*. Cambridge: University Press.
Fleischman, Suzanne. 1982b. "The Past and the Future: Are they *coming* or *going*?" *Berkeley Linguistics Society* 8:322–34.
Foreman, Velma M. 1974. *Grammar of Yessan-Mayo*. Santa Ana: Summer Institute of Linguistics [*Language Data Asian-Pacific Series* No. 4].
Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.

Furbee-Losee, Louanna. 1976. *The Correct Language: Tojolabal: A Grammar with Ethnographic Notes*. New York: Garland.
Givón, Talmy. 1973. "The time-axis phenomenon." *Language* 49:890–925.
Greenberg, Joseph H., Charles A. Ferguson and Edith Movavcsik (eds). 1978. *Universals of Human Language, Volume 3: Word Structure*. Stanford, CA: University Press.
Hall, Robert A. Jr. 1943. *Melanesian Pidgin English: Grammar, Texts, Vocabulary*. Baltimore: The Linguistic Society of America.
Hanke, A. 1909. *Grammatik und Vokabularium der Bongu-Sprache (Astrolabei, Kaiser-Wilhelmsland)*. Berlin: Georg Reimer [*Archiv für das Studium deutscher Kolonialsprachen*, Band VIII].
Harries, Lyndon. 1950. *A Grammar of Mwera*. Johannesburg: Witwatersrand University.
Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.
Hewitt, B.G. 1979. *Abkhaz*. Amsterdam: North-Holland [*Lingua Descriptive Studies* Vol. 2].
Hoffmann, Carl. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.
Householder, Fred W., Kostas Kazasis and Andreas Koutsoudas. 1964. *Reference Grammar of Literary Dhimotiki*. Bloomington: Indiana University [*IJAL* 30.2 Part II].
Iorss, Martial T. 1963. *Grammaire Tahitienne*. Papeete, Tahiti: Imprimerie Officielle.
Koefoed, H.A. 1958. *Danish*. London: The English Universities Press [*Teach Yourself Books*].
Kwok, Helen. 1971. *A Linguistic Study of the Cantonese Verb*. Hong Kong: Center of Asian Studies, University of Hong Kong [*Center of Asian Studies Occasional Papers and Monographs*, No. 3].
Lane, George M. 1898. *A Latin Grammar for Schools and Colleges*. Reprinted, New York: Greenwood Press.
Leslau, Wolf. 1945. *Short grammar of Tigre*. New Haven: American Oriental Society [*Publication of the American Oriental Society, Offprint Series* No. 18].
Leman, Wayne. 1980. *A Reference Grammar of the Cheyenne Language*. Greeley: University of Northern Colorado Museum of Anthropology [*Occasional Publications in Anthropology, Linguistic Series* No. 5].
Li, Paul Jen-Kuei. 1973. *Rukai Structure*. Ann Arbor: University Microfilms.
Lister-Turner, Rev. R. and Rev. J.B. Clark. n.d. *A Grammar of the Motu Language of Papua*. 2nd edn, Percy Chatterton (ed.). Sydney: Government Printer.
Love, J.R.B. 1932. "An introduction to the Worrora language, Part II." *Journal of the Royal Society of Western Australia* 18:13–22.
Lukas, Johannes. 1937. *A Study of the Kanuri Language: Grammar and Vocabulary*. London: Oxford University. (Reprinted 1967 by Dawsons of Pall Mall, London, for the International African Institute.)
Lyons, John. 1977. *Semantics*. 2 vols. Cambridge: University Press.
Marchese, Lynell. 1986. *Tense/aspect and the Development of Auxiliaries in Kru Dialects*. Dallas: Summer Institute of Linguistics.
Milne, Mrs. Leslie. 1921. *An Elementary Palaung Grammar*. Oxford: Clarendon Press.
Nadzhip, E.N. 1971. *Modern Uigur*. Translated by D.M. Segal. Moscow: Nauka.
Najlis, Elena Lidia. 1966. *Lengua Abipona*, Tomo I. Buenos Aires: Centro de Estudios

Lingüísticos, Universidad de Buenos Aires [*Archivo de Lenguas Precolombinas* Vol. 1].
N'Diaye, Genevieve. 1970. *Structure du Dialecte Basque de Maya*. The Hague: Mouton.
Newland, A.G.E. 1897. *A Practical Hand-book of the Language of the Lais as Spoken by the Hakas and other Allied Tribes of the Chin Hills (commonly the Baungshè dialect)*. Rangoon: Government Printing.
Newman, Stanley. 1965. *Zuni Grammar*. Albuquerque: University of New Mexico Press [*University of New Mexico Publication in Anthropology* No. 14].
Palmer, F.R. 1979. *Modality and the English Modals*. London: Longman.
Palmer, F.R. 1986. *Mood and Modality*. Cambridge: University Press.
Perevoshchikov, P.N., et al. 1962. *Grammatika sovremennogo udmurtskogo jazyka: fonetika i morfologija*. Izhevsk: Udmurtskoe Knizhnoe Izdatelstvo.
Perkins, Revere D. 1984. "GRAMCATS sampling procedure." ms, SUNY at Buffalo.
Platt, J.T. 1972. *An Outline Grammar of the Gugada Dialect, South Australia*. Canberra: Australian Institute of Aboriginal Studies.
Poppe, Nicholas N. 1960. *Buriat Grammar*. Bloomington and The Hague: Indiana University and Mouton [*Indiana University Publications Uralic and Altaic Series* Vol. 2].
Prost, Gibert R. 1962. "Signalling of transitive and intransitive in Chacobo (Pano)." *IJAL* 28:108–18.
Rascher, Matthäus. 1904. "Grundregeln der Bainingsprache." *Mitteilungen des Seminars für Orientalische Sprachen* 7:31–85.
Renck, G.L. 1975. *A Grammar of Yagaria*. Canberra: Department of Linguistics, Australian National University [*Pacific Linguistics Series B* No. 40].
Rice, Keren. 1989. *A Grammar of Slave*. The Hague: Mouton.
Sapir, Edward. 1921. *Language*. New York: Harcourt Brace and World.
Sharpe, Margaret C. 1972. *Alawa phonology and grammar*. Canberra: Australian Institute of Aboriginal Studies [*Australian Aboriginal Studies* No. 37].
Shepherd, Susan C. 1982. "From deontic to epistemic: an analysis of modals in the history of English, creoles, and language acquisition." In *Papers from the 5th International Conference on Historical Linguistics*, Anders Ahlqvist (ed.), 316–23. Amsterdam: John Benjamins.
Shipley, William F. 1964. *Maidu Grammar*. Berkeley: University of California Press [*UCPL* Vol. 41].
Sorensen, Arthur P. Jr. 1969. The Morphology of Tucano. Columbia University dissertation.
Spagnolo, L.M. 1933. *Bari Grammar*. Verona: Missioni Africane.
Stevenson, Roland C. 1956–7. "A survey of the phonetics and grammatical structure of the Nuba mountain languages, with particular reference to Otoro, Katcha and Nyimang." *Afrika und Übersee* 40:73–84, 93–115; 41:27–65, 117–52, 171–96.
Subrahmanyam, P.S. 1971. *Dravidian Verb Morphology (a Comparative Study)*. Annamalainagar, Tamilnadu: Annamalai University [Department of Linguistics Publication No. 24].
Svorou, Soteria. 1988. The Experiential Basis of the Grammar of Space: Evidence from the Languages of the World. SUNY at Buffalo dissertation.
Taylor, Douglas. 1956. "Island Carib II: Word-classes, affixes, nouns and verbs." *International Journal of American Linguistics* 22:1–44.

Thomas, Elaine. 1978. *A Grammatical Description of the Engenni Language*. Dallas: Summer Institute of Linguistics [*SIL Publication in Linguistics* No. 60].
Traugott, Elizabeth C. 1978. "On the expression of spatio-temporal relations in language." Greenberg et al. (eds), 1978:369–400.
Turner, Glen D. 1958. Jivaro: Phonology and Morphology. Indiana University dissertation.
Ultan, Russell. 1978. "The nature of future tenses." Greenberg et al. (eds), 1978:83–123.
Vandame, R.P. Charles. 1963. *Le Ngambay-Moundou: phonologie, grammaire et textes*. Dakar: Ifan [*Mémoires de l'Institut Français d'Afrique Noire* No. 69].
Wacke, K. 1930. "Formenlehre der Ono-sprache (Neuguinea)." *Zeitschrift für Eingeborenen-Sprachen* 21:161–208.
West, Birdie. 1980. *Gramática Popular del Tucano*. Bogotá: Ministerio de Gobierno, Instituto Lingüístico de Verano.
Willett, Thomas. 1988. "A cross-linguistic survey of the grammaticization of evidentiality." *Studies in Language* 12:51–97.
Wilson, W.A.A. 1961. *An outline of the Temne Language*. London: School of Oriental and African Studies, University of London.
Winfield, W.W. 1928. *A Grammar of the Kui Language*. Calcutta: The Asiatic Society of Bengal [*Biblioteca Indica* Work No. 245].
Yallop, Colin. 1977. *Alyawarra: An Aboriginal Language of Central Australia*. Canberra: Australian Institute of Aboriginal Studies.
Yates, Warren G., and Souksomboun Sayasithsena. 1970. *Lao Basic Course* I. Washington, D.C.: U.S. Foreign Service Institute.

APPENDIX 1: GRAMCATS Meaning Components

ability:	The agent of the verb has the mental or physical ability to complete the action of the main verb.
admonitive:	The command constitutes a warning; "you had better not..."
almost:	Nearly; only a little less than.
andative:	The agent moves away from the deictic center in order to do something, literally, "to be going to do something".
anterior:	The situation occurs prior to reference time, and is relevant to the situation at reference time.
certainty:	The speaker is emphasizing that the proposition is true.
continuous:	A single situation is viewed as in progress, as maintained over a period of time.
counterfactual:	The proposition describes an unreal or imagined situation that could have been true but was not.
desire:	The agent of the verb desires or wants to complete the action of the main verb.

expected: The situation is to occur in the near future; what is scheduled to happen; qualifies *future*.

future: The situation takes place after the moment of speech; the speaker predicts that the situation in the proposition will hold.

gnomic: The situation described in the proposition is generic; the predicate has held, holds and will hold for the class of entities named by the subject.

habitual: The sitution is customary or usual, repeated on different occasions over a period of time.

hypothetical: The situation is unreal or imagined, but one that could be true.

immediate: A meaning component that may be combined with other tense components to indicate a situation not simultaneous with the moment of speech, but very close to it:
immediate future: What is about to occur.
immediate past: Occurring immediately before the moment of speech.

imperative: The utterance is a direct command.

inceptive: The action or event begins.

indefinite: The situation took place or will take place at some non-specific time or place.

inferred certainty: The speaker infers from evidence that the proposition is true.

intention: The agent intends to carry out the action of the main verb.

iterative: The action is repeated on one occasion.

narrative time: The use of a form with no other *past* uses for reporting a past narrative.

past: The situation occurred before the moment of speech.

perfective: The situation is viewed as bounded temporally. It cannot be simultaneous with the moment of speech; in the non-past it is sometimes interpreted as future.

permission: The agent is allowed to complete the action of the main verb.

possibility: The speaker is indicating that the situation described in the proposition is possibly true.

present: The situation occurs simultaneously with the moment of speech.

probability: The speaker is indicating that the situation described in the proposition is probably true.

progressive: The action takes place simultaneouly with the moment of reference; 'to be in the process of...'.

purpose: The clause states the purpose for which the main clause action is taking place.

root possibility: It is possible for the agent to carry out the action of the main verb; i.e. s/he is able *and* external conditions allow it.

sequence-first: A marker occurring on the first verb of a two-verb sequence indicating that the situation described by the first verb is prior to that described by the second.

venitive: The agent moves towards the deictic center in order to do something, literally, "be coming to do something".

APPENDIX 2: Future grams in the GRAMCATS database grouped by assigned semantic age (FUTAGE).[15]

FUTAGE 1

lng	lng – name	form	shape	source
002	Basque	19	-tu bear	N'Diaye 1970:211f.
051	Danish	4	skulle	Koefoed 1958:192
051	Danish	5	ville	Koefoed 1958:192
058	Nimboran	6	d	Anceaux 1965:58f.,71,79,112
059	Bongu	9	aʁ	Hanke 1909:52f.;66;87
059	Bongu	10	aʁ...m	Hanke 1909:53;87
068	Slave	50	gó?o	Rice 1989:415f
076	Bari	4	kɔ	Spagnolo 1933:105
089	Cantonese	3	hó nàng	Kwok 1971:77f.
089	Cantonese	4	wui	Kwok 1971:77
093	Buriat	21	xa	Poppe 1960:61f.
094	Tok Pisin	7	lajk	Hall 1943:32f.

FUTAGE 2

lng	lng – name	form	shape	source
001	Inuit	17	niar	Fortescue 1984:274f.,325
001	Inuit	44	qqajar	Fortescue 1984:285
001	Inuit	65	sussaa	Fortescue 1984:292
001	Inuit	89	llarumaar	Fortescue 1984:296
003	Margi	42	ləgəri	Hoffman 1963:222
008	Tigre	22	STCH	Leslau 1945:7f.
011	Cocama	6	utsu	Faust 1971:55
014	Tucano	28	gɨ + ti + gɨ + wé	West 1980:74f.
014	Tucano	32	gɨ + ti	West 1980:36
014	Tucano	35	gɨ + s a	Sorensen 1969:270
026	Motu	8	ai	Lister-Turner and Clark n.d.:12,16f.
026	Motu	36	b	Lister-Turner and Clark n.d.:17
028	Halia	17	moto	Allen and Allen 1965:31

FUTAGE 2 (*continued*)

lng	lng – name	form	shape	source
028	Halia	19	ow	Allen and Allen 1965:31
034	Pangasinan	10	na	Benton 1971:134
052	Yagaria	25	o	Renck 1975:97f.
060	Yessan-Mayo	32	nin	Foreman 1974:144f.
068	Slave	11	de + y + 0	Rice 1989:588
068	Slave	51	gha	Rice 1989:417
072	Mwera	34	ci	Harries 1950:98
072	Mwera	35	cika	Harries 1950:98
073	Tem	14	ngɔni	Der-Houssikian 1980:64f.
075	Mano	3	nu pèá	Becker-Donner 1965:37f.
076	Bari	5	tu	Spagnolo 1933:105
076	Bari	23	dc'ya	Spagnolo 1933:186
080	Tojolabal	4	ʔoh	Furbee-Losee 1976:129,135
085	Chepang	38	dhaŋ	Caughley 1982:108,101,124
085	Chepang	41	kheʔ	Caughley 1982:199
086	Haka	10	lai	Newland 1897:34f.,40,468
086	Haka	16	de mung	Newland 1897:36,378
089	Cantonese	20	jau	Kwok 1971:114
091	Udmurt	4	u-...-y	Perevoshchikov et al. 1962:200-202

FUTAGE 3

lng	lng – name	form	shape	source
001	Inuit	18	jumaar	Fortescue 1984:275,325
003	Margi	37	ra	Hoffman 1963:212–217
006	Touareg	15	ed	Cortade 1969:38ff.
009	Cheyenne	63	hta	Leman 1980:191
009	Cheyenne	66	stse	Leman 1980:191
011	Cocama	2	á	Faust 1971:42
013	Jivaro	29	tat	Turner 1958:64,71
019	Worora	14	STCH	Love 1932:14
030	Tanga	5	rokok	Bell 1977:xvi–xxvii
030	Tanga	12	ku	Bell 1977:xvi–xxi
033	Trukese	4	saq	Dyen 1965:24
035	Rukai	6	ay	Li 1973:156f.
039	Abkhaz	48	zaa	Hewitt 1979:176
040	Guaymi	8	ai	Alphonse 1956:27–41
043	Abipon	7	am	Najlis 1966:37
046	Karok	15	aviš	Bright 1957:124
047	Latin	9	bi	Lane 1898:149,273ff.
051	Danish	6	få	Koefoed 1958:192
052	Yagaria	23	s	Renck 1975:94

FUTAGE 3 (*continued*)

lng	lng – name	form	shape	source
052	Yagaria	24	g	Renck 1975:95f.
056	Ono	4	ke	Wacke 1930:165
063	Baining	3	i-	Rascher 1904:62f.
069	Kadugli	19	ara	Stevenson 1956–7:60
072	Mwera	36	jiya	Harries 1950:99
074	Engenni	13	TONE	Thomas 1978:70
074	Engenni	25	si	Thomas 1978:68
075	Mano	5	ḷo	Becker-Donner 1965:40f.
076	Bari	3	dé	Spagnolo 1933:105
084	Lao	1	si	Yates and Sayasisthsena 1970:19f.,161
088	Nung	8	di	Barnard 1934:19,21,28f.
088	Nung	12	hka	Barnard 1934:21,29,33
091	Udmurt	3	-o-	Perevoshchikov et al. 1962:200ff.
092	Uigur	27	i	Nadzhip 1971:123
093	Buriat	3	uuža	Poppe 1960:56ff.
094	Tok Pisin	1	bajmbaj	Hall 1943:90,27

FUTAGE 4

lng	lng – name	form	shape	source
001	Inuit	16	ssa	Fortescue 1984:274f.,64–7
002	Basque	5	-ko da	N'Diaye 1970:203
012	Chácobo	23	tiari	Prost 1962:116
014	Tucano	13	some	Sorensen 1969:155,184
014	Tucano	34	gi + sa	West 1980:34
017	Alyawarra	9	iya	Yallop 1977:54f.
018	Maung	3	wana	Capell and Hinch 1970:74;67ff.
018	Maung	7	ji	Capell and Hinch 1970: 67ff.,78ff.
020	Alawa	30	yi	Sharpe 1972:87ff.
026	Motu	7	bain	Lister-Turner and Clark n.d.: 11,18
027	Atchin	3	p-	Capell and Layard 1980:75ff. et passim
027	Atchin	10	ro	Capell and Layard 1980:75ff.
033	Trukese	3	qe	Dyen 1965:24
039	Abkhaz	49	p'	Hewitt 1979:176f.
050	Modern Greek	17	θa	Householder et al. 1964:105f.
051	Danish	18	komme til at	Koefoed 1958:18;183;185
060	Yessan-Mayo	14	iti	Foreman 1974:40f.
063	Baining	4	di-	Rascher 1904:62f.
077	Ngambay	6	a k-	Vandame 1963:97–9
081	Zuni	3	ʔanna	Newman 1965:37

FUTAGE 4 (*continued*)

lng	lng – name	form	shape	source
081	Zuni	5	šuk$^{\omega}$a	Newman 1965:38
082	Maidu	42	mak	Shipley 1964:47
085	Chepang	12	ca$^{?}$	Caughley 1982:49,104–111, 92f.

FUTAGE A

lng	lng – name	form	shape	source
003	Margi	32	a	Hoffman 1963:190–6
006	Touareg	2	STCH	Cortade 1969:29f., 189f.,90
008	Tigre	21	STCH	Leslau 1945:6f.
010	Island Carib	16	-ba	Taylor 1956:6,23f.,20, 42f.,xxiv
015	Gugada	3	ninj	Platt 1972:29
017	Alyawarra	4	ima	Yallop 1977:49,52f.
017	Alyawarra	5	iyla	Yallop 1977:51f.
023	Palaung	7	di	Milne 1921:35,47,68
023	Palaung	8	chang	Milne 1921:68f.
025	Tahitian	3	ua	Iorss 1963:62f.
025	Tahitian	4	e	Iorss 1963:63
034	Pangasinan	5	on	Benton 1971:133ff.,196f.
035	Rukai	4	REDUP	Li 1973:156ff.,267f.,177,208
039	Abkhaz	44	ZERO	Hewitt 1979:173
040	Guaymi	11	tau	Alphonse 1956:36f.
042	Kui	2	in	Winfield 1928:60–85; Subrahmanyam 1971:278f.
049	Baluchi	1	ZERO	Barker and Mengal 1969:129f.
049	Baluchi	9	əga bu	Barker and Mengal 1969:233f.
050	Modern Greek	7	ZERO	Householder et al. 1964:115ff.
051	Danish	1	-r	Koefoed 1958:180f.,185
063	Baining	5	sa	Rascher 1904:62ff.
071	Temne	10	tə	Wilson 1961:25
078	Kanuri	1	in	Lukas 1967:35f.,48,50f.
078	Kanuri	7	i	Lukas 1967:45ff.
091	Udmurt	1	-içk-o-	Perevoshchikov et al. 1962: 197–200
091	Udmurt	2	u-...-içk-y	Perevoshchikov et al. 1962: 197–200
091	Udmurt	5	-i-	Perevoshchikov et al. 1962:202ff.

The Emergence of Perfective Aspect in Indo-Aryan Languages

Peter Edwin Hook
University of Michigan

1. INTRODUCTION

Many studies of grammaticization are concerned with the more dramatic instances of semantic change in which what is at an early point in time an independent lexical item later appears in the guise of a function word or bound morpheme, filling some general, often quite abstract role in the grammatical system of its language. Frequently such a metamorphosis involves what is commonly referred to as "semantic bleaching": The sense of an item may become more general by losing features of meaning (Sweetser, 1988:390). However, the lexical items which are commonly found to participate in grammaticization tend to have quite general senses to begin with. They have few features to lose. The bleaching of such items should be reflected in sudden changes in their grammatical roles. How are we to square this essentially saccadal view of semantic change with the graduality of grammaticization reported by Lord (1973), Givón (1975) and others? In the present study I show that gradual change in the optionality of the use of a grammaticized element is at least as important as change in the meaning of the element itself.

2. GENERAL PROPERTIES OF THE INDO-ARYAN COMPOUND VERB

With the exception of some dialects of Shina every contemporary Indo-Aryan language or dialect has a set of auxiliary (or "vector") verbs homophonous with members of its inventory of basic lexical verbs. As full lexical verbs these express a change in location or posture, or an action that entails such

a change: GO, GIVE, TAKE, THROW, LET GO, GET UP, COME, STRIKE, SIT, FALL, etc. A *compound verb* (CV) comprises the finite form of one of these following a non-finite or stem form of a main or primary verb. In (1h) and (2h) *kar* and *ho* are the non-finite forms of primary verbs, while *diyaa hotaa* and *jaatii* are the finite forms of the vectors *de* GIVE and *jaa* GO (examples marked *h* are from Hindi-Urdu):

(1h) *agar mAI ne darvaazaa band kar diyaa ho-taa...*[1].
if I ERG door shut make GAVE be-CTF...
'If I had closed the door...' (Vaid, 1970:25)

(2h) *baat vahII xatam ho jaa-tii*
thing there over become GO-CTF
'The matter would have ended right there.' (Vaid, 1970:25)

For every CV there is almost always a corresponding non-compound or simple verb (SV) in which the auxiliary is absent and the primary verb bears the desinential affixes:

(1h′) *agar mAI ne darvaazaa band kiyaa ho-taa...*
if I ERG door shut made be-CTF...
'If I had closed the door...'

(2h′) *baat vahII xatam ho-tii*
thing there over become-CTF
'The matter would have ended right there.'

The compound verb is an innovation in Indo-Aryan. It is first attested in Buddhist Pali writings dating from the early centuries of the Common Era and composed in Sri Lanka: In the Jātaka Tales (Hendriksen, 1944:136):

(3p) *mayā (varo) gahitakam ka-tvā thāpito*
by-me boon in-reserve make-CP PUT
'I have kept (the wish you granted me) in reserve.'

and the philosophical works of Buddhaghosa (example (4p) is from *sumangalavilāsinī*:180[2]):

(4p) *...yam yam virujjhati tam tam ujukam ka-tvā denti*
...what what conflicts that that straight make-CP GIVE
'(True scholars) correct whatever is contradictory.'

It is almost completely absent in Sanskrit, the language that is taken to be the common ancestor of Hindi-Urdu and Marathi, even from Sanskrit

texts composed in modern times, where we might expect it to occur due to interference from the mother tongues of the authors. There are one or two ambiguous instances from late texts or undated ones like the *śuktasaptati*:

(5s) *tato makṣikā+ uḍḍi-ya gatā*
then fly fly-CP WENT
'Then the fly flew away.' (cited by Tikkanen, 1987:176)

It may have come into Indo-Aryan as a calque on parallel (and more ancient) structures in Dravidian or in Altaic (compare (6t) with (4p) and (7o) with (5s)):

(6t) *araiyolai cey-tu koTu-ttu viTu-taka...*
proclamation make-CP GIVE-CP leave-IMP
'...prepare the proclamation and issue (it).'
(Tamil inscription, circa 550 AD, in Zvelebil, 1964)

(7o) *(kül tigin), uč-a bard-iGiz*
Kul Prince fly-CP WENT-2PL
'(Prince Kul, you) died.'
(Orkhon inscription, circa 650 AD, in Tekin, 1968[3])

In nearly all the languages that have the compound verb, it turns out to be structurally homologous with serial verb constructions expressing tightly sequenced actions:

(8h) *darvaazaa band kar jaanaa* (cf. *kar* in (1h))
door shut do go-IMP
'(Be sure to) close the door on your way out.'

(9p) *dvāram vari-tvā pavisati* (cf. *-tvā* in (3p) and (4p))
door open-CP enters
'He opens the door and enters.'

It is tempting to propose that compound verbs developed historically out of serial verb constructions. (Notice, for example, that (5s) and (7o) are equally interpretable as having serial or compound verbs.) However, the origin of the compound verb is not an issue that I pursue here.

In Indo-Aryan, the CV auxiliaries or vectors exhibit a greater or lesser degree of semantic bleaching. The difference in meaning which stems from their presence as opposed to their absence also varies from one language to another. For some languages it can be shown that the CV/SV opposition has

become part of a system of regular semantic contrasts, a part of the verb paradigm, functioning to express perfective (versus aspectually unspecified) conceptualization of actions and events. For others it can be shown that this has not (yet) happened.

Assuming that the direction of change is the same in all of Indo-Aryan, we can treat the characteristics of CV systems in various languages as representing successive phases along a developmental continuum[4]. In the next section I describe the functions of Hindi-Urdu's highly evolved CV system, which is one of those in which the CV/SV opposition has emerged as a marker of perfective aspect. Since the perfective value of the CV in Hindi-Urdu is something which I have treated elsewhere in substantial detail (Hook, 1978; 1974), I give only a brief demonstration of it in the next section.

3. THE ARYAVARTAN COMPOUND VERB AND PERFECTIVE ASPECT

"Aryavarta" is the traditional name for the heartland of the Aryan occupation of the South Asian subcontinent. It corresponds roughly to the non-mountainous portions of the modern Indian state of Uttar Pradesh and it happens to be the location of languages in which the CV/SV opposition has achieved its most complete integration into the regular, paradigmatically organized system of semantic contrasts that govern the morphology of the verb (although I hesitate to describe that oppositon as inflectional). Although many NIA languages other than Hindi-Urdu are spoken in Aryavarta (viz. Kauravi, Braj, Kannauji, Awadhi, Bhojpuri), the CV/SV systems in all of them exhibit nearly identical structural and functional properties. (This is in marked contrast to some CV systems in NIA languages outside the area.)

Applying diagnostic tests to the Hindi-Urdu CV, we can demonstrate that (among other functions) it stands in opposition to the corresponding simple verb (SV) in a system of contrasts that include: 1. relative tense; and 2. perfective aspect.

1. In a context that explicitly contrasts the time of occurrence of a pair of actions, the CV must be used to express the prior action; and the SV, the later one:

(10h) *jab tak aap yahAA aae us ne mujhe ciTThii de*
when by you here came he ERG me letter give
dii thii
GIVEN had
'By the time you came here he had given me the letter.'

(11h) *jab tak aap ne mujhe ciTThii dii vo yahAA aa*
when by you ERG me letter gave he here come
gayaa thaa
WENT had
'By the time you gave me the letter he had come here.'

The distribution of compound and simple forms in exx. (10h) and (11h) is controlled by the relative location in time of the two actions expressed: In them the CV vs. SV opposition functions to co-express relative tense[5].

2. Independent of (and at times in conflict with) this, the CV/SV opposition also functions to express perfectivity in the sense which that term has in Slavistics: Viewing an action or process as (summed into) a whole vis-à-vis some specific moment or interval of time. This rather cumbersome definition[6] has the merit of accommodating two divergent meanings of the perfective found in languages where it is a paradigmatic category: Completion of action and inchoation of state or activity. However, discussions of the semantics of perfectivity in languages that do not treat it as a regular category of the verb (for example, English) often lead to unsatisfyingly discursive explanations with recourse to visualizations and impressionistic analogies when such explanations fail[7]. How can we assure accurate identification of perfectivity in a language like Hindi-Urdu which is only distantly related to Slavic and in which the morphological means (prefixation and suppletion) used in Slavic for its expression are completely absent?

Russian or Czech-speaking Indologists commonly identify Hindi-Urdu CV's as being perfectives[8]. To buttress their intuitive judgements with explicit procedures I have developed independent diagnostic tests. For instance, in aspectological studies of Russian it has been observed (Forsyth, 1970:258–61 and 297) that clauses which are dependent on expressions of fear and anxiety show a preference for perfective forms of the verb which borders on being categorical:

(12r) *mat' bojalas' kak-by eë syn ne zabolel*
mother feared lest her son NEG took-sick-PFV
'The mother was afraid that her son might get sick.'

There is a similar (although not quite so univocal) preference for perfective forms in clauses dependent on expressions that mean 'until' ((13r) from Forsyth, 1970:133):

(13r) *ona budet ubajukivat' rebënka poka on ne zasnët*
she will rock child while he NEG sleeps-PFV
'She will rock the baby till it goes to sleep.'

These asymmetries in the distribution of aspectual forms in Russian are closely paralleled in Hindi-Urdu[9]:

(14h) *mujhe Dar thaa ki kahII tum use ciTThii na de do*
to-me fear was that lest you him letter NEG give GIVE
'I was afraid that you might give him the letter.'

(15h) *tum yahAA Thaharo jab-tak vo tumhE ciTThii na de de*
you here stay until he you letter NEG give GIVE
'Wait here until he gives you the letter.'

Applying these same tests to Marathi we find that its CV/SV opposition does not regularly express either relative tense or perfectivity[10]:

(10m) *tu ye-Nyaa-caa agodar tyaa-na ma-laa patr dila hota*
you come-INF-GEN before he-ERG me-to letter gave was
'By the time you got here he had given me the letter.'

(11m) *tu patr de-Nyaa-caa agodar haa itha aalaa hotaa*
you letter give-INF-GEN before he here came was
'By the time you gave me the letter he had come here.'

(14m) *ma-laa kaaLji hoti ki tu tyaa-laa patr de-Sil*
me-to anxiety was that you him-to letter give-FUT
'I was afraid that you might give him the letter.'

(15m) *to tu-laa patr de-i paryant itha thaamb*
he you-to letter give-INF until here wait
'Wait here until he gives you the letter.'

These data show that there is a statable difference in the functions of the CV/SV opposition for at least one pair of Indo-Aryan languages. At this point, the following objection could be raised: Granting that the CV/SV opposition in Marathi does not express a perfective/non-perfective aspectual contrast does not mean that CV/SV does not have some other general paradigmatic function in Marathi. It may be that the CV systems in Hindi-Urdu and Marathi are merely different and not representatives of successive stages on a path of grammaticization. In the next section I adduce evidence to show that the particular functional differences I have demonstrated in this section correlate with a number of other differences between Hindi-Urdu and Marathi which not only exclude the "merely-different" hypothesis but in fact

provide the basis for an explanation in terms of grammaticization for the systematic differences that we have just seen.

4. GRAMMATICIZATION AND ASPECTOGENESIS

4.1. Text frequency

Semantic parameters that commonly gain entry into a system of paradigmatic contrasts associated with some major word class (e.g. noun or verb) are those which are encountered frequently. Even though it is not required that such meanings be of the same degree of importance or salience (or even decidability) for every token or instance of the word class with which they are associated, it is nonetheless true that they tend to be meanings that frequently recur. Meanings that are relevant to only some restricted subset of a word class (say, color with respect to nouns or use/non-use of tools with respect to verbs) usually find expression as independent lexical items. Conversely, lexical items that occur only rarely in texts or which are severely restricted with respect to what other lexical items they may occur with are less likely to take on paradigmatic semantic or grammatical values. In short, an item's text frequency is an important correlate of its potential for grammaticization.

Surveying the hundreds of languages and dialects which belong to the Indo-Aryan family, one encounters wide variation in the text frequency of the compound verb (see Table 1). While there is some variability that depends on speaker (or writer), speech situation (or genre), social class, etc., the most important factor by far is geographical. Controlling for speech situation, subject matter and educational level, one still finds sharp differences in rates

Table 1 *Relative text-frequency of compound verbs in NIA*[11]

Shina (Gilgit)	0	
Kashmiri	1	--
Marathi	3	------
Gujarati	6	----------
Bengali	7	----------
Marwari	8	----------
Hindi-Urdu	9	----------

of CV occurrence in texts. Hindi-Urdu seems to have about as frequent a level of CV occurrence as one can find in Indo-Aryan. In Hindi-Urdu we find roughly one CV in every eight clauses. Since the Hindi-Urdu CV is only half as common in non-finite forms, limiting the count to clauses with finite forms in them brings the ratio up to one in seven. For Marathi the corresponding figures are one in thirty-two and one in twenty-four[12].

4.2. Internal privileges of co-occurrence

In a language like Marathi in which the overall text frequency of compound verbs is low, some vectors seem to retain a greater fraction of their original lexical meaning. For instance, the use of the auxiliary GIVE is limited to either situations in which outward movement is imparted to the patient of an action[13]:

(16m) *tyaana katsraa Taak-un dilaa*
he-ERG trash throw-CP GAVE
'He threw out the trash.'

or in which the result of the action is outwardly vectored, that is, in which it makes sense to think of the action as being done for another's (as opposed to being done for the doer's) benefit:

(17m) *maajhaa saaThi he kaam kar-un de-Sil kaa*
my-OBL sake this job do-CP GIVE-FUT Q
'Will you do this job for me?'

In a language like Hindi-Urdu in which the compound verb is of much more frequent occurrence, the connection of the auxiliary to its primary is harder to explain or predict in terms of the senses of its homphonous counterpart among the main verbs. Of course, combinations like those in (16h) and (17h) are not excluded[14]:

(16h) *us ne kacraa phEk diyaa*
he ERG trash throw GAVE
'He threw out the trash.'

(17h) *kyaa mere liye ye kaam kar d-oge*
Q my sake this job do GIVE-FUT
'Will you do this job for me?'

However, we cannot so easily give an explanation of GIVE's occurrence in

(18h) where the lexical meaning of GIVE is orthogonal to the meaning of the verb phrase as a whole:

(18h) *mAI ne uske hoThO ko...TaTol-naa...Suruu kar diyaa*
I ERG her lips DAT feel-INF start do GAVE
'I began to feel her lips.' (Vaid, 1970:13)

nor in (19h) where the semantic element of outward orientation is inherent in the meaning of the primary verb and the vector appears to be redundant:

(19h) *aaxirkaar us ne mujh ko paise de diye*
finally he ERG me to money give GAVE
'At last he gave me the money'.

Even more problematic is any explanation in lexical semantic terms of compounds like those in (20h) and (21h) in which the literal meaning of the auxiliary verb seems to cancel or contradict the meaning of the primary:

(20h) *idhar aa jaao*
here come GO
'Come here!'

(21h) *is mE us ne teraa naam le diyaa*
this in he ERG your name take GAVE
'He implicated you in this.'

In Marathi, the corresponding combinations are either not acceptable or (in the case of (20m)) can be interpreted only as expressing two separate actions[15]:

(18m) *mi ti-tse oTh tsaatsap-na suru kela* (**kar-un dila*)
I-ERG her lips touch-INF start did do-CP GAVE
'I began to feel her lips.'

(19m) *SevaTi tyaa na ma-laa paise dile* (**de-un dile*)
finally he ERG me-to money gave give-CP GAVE
'Finally he gave me the money.'

(20m) *ikaDa ye* (**ye-un dzaa*)
here come come-CP GO
'Come here!'
(*ye-un dzaa* can only mean 'Come here and then go!')

(21m) *hyaacaat tyaa-na tudzha naaw ghetla* (**ghe-un dila*)
this-LOC he-ERG your name took took-CP GAVE
'He implicated you in this.'

If we compare entire texts in Hindi-Urdu with similar ones in Marathi, we find striking differences not only in the frequency of the CV but in the meanings of the primary verbs that appear in CV form. Sorting them into classes we observe that a smaller range of primaries show up as CVs in Marathi and that these are semantically more concrete: As a result of the action expressed in CV form some tangible object moves from one place to another (or is disposed of altogether) or a sentient being experiences a change of inner state. Primaries in Hindi-Urdu CV's belong to a wider range of semantic categories. In particular, verbs of communication (in which no tangible object moves) and verbs expressing a change in state which is not psychological are quite commonly found in CV form (see Table 2).

Table 2 *Semantic classes of primaries occurring in CV's.*[16]

Displacement	42%	---------- (Mar)
or disposal:	27%	---------- (H-U)
Change of	17%	--------- (Mar)
psychic state:	9%	----- (H-U)
Sensation or	8%	---- (Mar)
perception:	4%	-- (H-U)
Mental action:	7%	---- (Mar)
	11%	------ (H-U)
Communication:	3%	-- (Mar)
	17%	--------- (H-U)
Creation/change	13%	------- (Mar)
of state:	24%	---------- (H-U)
Other:	10%	----- (Mar)
	8%	---- (H-U)

Since these figures give the ratio of individual categories of CV's out of a total pool of one hundred CV's in both Marathi and Hindi-Urdu, they skew the absolute difference between the two languages in the probability of CV expression by a factor of at least three or four hundred percent. By comparing the ratio of CV to total verb forms for each class of primaries we can better see the actual differences between the two languages (see Table 3).

These figures raise many interesting questions about the process of aspectogenesis, more than I can pursue here. Let us examine the most striking disparity: In Hindi-Urdu verbs of communication quite commonly occur in CV form; in Marathi, hardly ever. Now, in Hindi, the CV has perfective aspectual value and alternates with the SV in a systematic way. In Marathi,

Table 3 *Ratio of CV to total by semantic class of primary.*[18]

Displacement	10%	----- (Mar)
or disposal:	44%	---------- (H-U)
Change of	8%	---- (Mar)
psychic state:	8%	---- (H-U)
Sensation or	4%	-- (Mar)
perception:	8%	---- (H-U)
Mental action:	4%	-- (Mar)
	10%	----- (H-U)
Communication:	2%	- (Mar)
	20%	---------- (H-U)
Creation/change	8%	---- (Mar)
of state:	30%	---------- (H-U)

it does not. What bearing does this have on the difference between the two in the likelihood of a CV manifestation for their verbs of communication?

The difference can be understood if we recognize that perfective aspect has a more redundant semantic value for verbs of communication than it has in other categories. For instance, if we apply the Aristotelian taxonomy of inherent aspectual properties (as extended and refined by Z. Vendler and David Dowty) to broad classes of verbs:

(22)

	−culmination:	+culmination:
−activity:	State	Achievement
+activity:	Activity	Accomplishment

We observe that most verbs of communication *in context* are successful Accomplishments (with extremely short periods of activity leading up to their culminations or completions):

(23m) *"saaDe caar" ti na sAAgitla*
half four she ERG said
"'4:30," she said.' (Chirmule, 65)

(24h) *mAI xaamoS rah saktii thii lekin mAI ne kah*
I silent stay be-able was but I ERG say
diyaa "hAA"
GAVE yes
'I could have kept quiet but I said, "Yes."' (Vaid, 29)

If we assume that innovations enter linguistic systems in order to resolve communicative problems (Traugott, 1988:414), it follows that they would take root first in domains where they are maximally informative. A look at the adverbs and the direct object NP's which occur in construction with Marathi CV's shows that CV's in Marathi often function to express the fullness or completeness of actions denoted by verbs inherently unspecified with regard to completeness:

(25m) *to purtaats gOndhaLun gelaa*
he completely confuse WENT
'He became completely confused.' (Chirmule, 60)

(26m) *sagLe kaagad mi phaaRun Taakle*
all papers I tear THREW
'I tore up all the papers.' (Chirmule, 69)

(27m) *ek dzaaNiv sagLyaa haal madhe sarsarun geli*
an awareness entire hall in ripple WENT
'A feeling (of anxiousness) spread through the entire (examination) hall.' (Chirmule, 24)

Over time as the frequency of explicit marking of completeness grows, predicates expressing actions and events which are complete by their very nature are drawn into the domain of the compound verb. We see this in the CV manifestation in Hindi-Urdu of prototypical Achievements such as 'find', 'catch', 'arrive', and 'begin' which in Marathi are always SV (compare also (18h) and (18m)):

(28h) *ek aadmii ne apnii patnii ko ain mauqe par pakaR liyaa*
one man ERG his wife DAT right moment at catch TOOK
'A man caught his wife red-handed...' (Vaid, 16)

(28m) **ekaa maaNsaa na baayko laa ain veLe-s pakaDun ghetla*
one man ERG wife DAT right time-at catch TOOK
'A man caught his wife red-handed.'

(29m) *ekaa maaNsaa na baayko laa ain veLe-s pakaDla*
one man ERG wife DAT right time-at caught
'A man caught his wife red-handed.'

Verbs of articulate communication, especially those that introduce direct quotation, are close to being prototypical Achievements and hence are rarely found in Marathi in CV form. In fact, if we classify the CV's in the Appendix according to the inherent aspectual value of the proposition they are in (subtracting the effect of their vectors), we find nearly 70% of the Hindi-Urdu sample have inferrable culminations while only a little over 50% of the Marathi sample do.

This difference becomes even more striking if we set aside a second type of CV which is much more common in Hindi-Urdu than in Marathi. These are CV's made from Activities or States which rarely lead to culminations:

(30h) *to mAI Saayad khisiyaa-kar hAs detaa*
then I maybe embarrassed-CP laugh GAVE
'Then maybe I would have laughed in frustration.' (Vaid, 12)

(31h) *jin se baad-mE vo kAAP uThaa kartii thii*
which from later she tremble RISE doing was
'...which later on she would begin to tremble from.' (Vaid, 20)

(32h) *is kalpanaa ke Dar mE Duubaa mAI so gayaa*
this idea GEN fear in drowned I sleep WENT
'Sinking in fear at this thought I fell asleep.' (Vaid, 23)

Here the presence of a vector delimits the *inception* of the Activity or State denoted by the primary, not its culmination. Comparison may be made with the category of "procedurals" (*sposob dejstvija*) as elaborated by Slavic aspectologists. (30h–32h) do not have CV counterparts in Marathi:

(30m) *mag mi kadaacit laadz-un haslo asto*
then I perhaps embarrassed-CP laughed had
'Then maybe I would have laughed in frustration.'

(31m) *jaancaa muL-e nantar ti kAApat ase*
whose root-LOC later she tremble would
'...which later on she would begin to tremble from.'

(32m) *hyaa kalpane caa bhiti-t buD-un mi dzhoplo*
this thought GEN fear-in drown-CP I slept
'Sinking in fear at this thought I fell asleep.'

Thus, as CV frequency in Indo-Aryan increases overall, it shows proportionately greater increases in semantic domains where the processes leading

to culmination of action or event are very short or non-existent ones and in others where there is no culmination at all. It is the emergence of these secondary classes of CV that universalizes the CV-construction to the extent necessary for it to participate in a regular paradigmatic contrast:

(33)	Pre-perfective phase	Perfective phase:
		Achievement CV's
		Near-achievement CV's
	Accomplishment CV's	Accomplishment CV's
		Activity CV's (Procedurals)

4.3. External privileges of co-occurrence

Paradoxically, in CV-rich languages like Hindi-Urdu there are many more syntactic environments that exclude the compound verb altogether than there are in CV-poor languages. Thus, Marathi with its less frequent and less grammaticized CV auxiliaries permits them with negatives and semi-negatives (viz. 'only'):

(34m) *tyaa na katsraa Taak-un dilaa naahi*
he ERG trash throw-CP GAVE not
'He did not throw out the trash.'

(35m) *haa katsraa jaan na ts Taak-un dilaa*
this trash John ERG only throw-CP GAVE
'John was the only one who threw out this trash.'

in construction with modal and phasal verbs:

(36m) *aadz haa katsraa Taak-un de-u Sak-Sil kaa* (modal)
today this trash throw-CP GIVE can-FUT Q
'Can (you) throw out this trash today?'

(37m) *ti ann vaaDh-un ghe-u laagli* (phasal)
she food dish-out-CP TAKE-INF began
'She began to dish out the food.' (Chirmule, 60)[18]

and in the forms of past and conjunctive participles:

(38m) *vaaDh-un ghet-lela ann ... thaNDgaar dzhaala hota*
serve-CP TAKEN-PART food cold become was
'The food dished out had become cold.' (Chirmule, 62)

(39m) *bhaambaa-un dzaa-un to tsaatsar-at mhaNaalaa...*
fall-apart-CP GO-CP he stammer-ing said
'Losing his composure he stammered out...' (Chirmule, 55)

while Hindi-Urdu, where the compound verb is from three to four times more frequent than it is in Marathi, does not:

(34h) *us ne kacraa nahII phEkaa* (**nahII phEk diyaa*)
he ERG trash not threw not throw GAVE
'He didn't throw out the trash.'

(35h) *ye kacraa jaan ne hii phEkaa*
this trash John ERG only threw
'John was the only one who threw out this trash.'

(**ye kacraa jaan ne hii phEk diyaa*
this trash John ERG only throw GAVE)

(36h) *kyaa aaj ye kacraa phEk sak-oge*
Q today this trash throw can-FUT
'Can (you) throw out this trash today?'

(**kyaa aaj ye kacraa phEk de sak-oge*
Q today this trash throw GIVE can-FUT)

(37h) *vo khaanaa paros-ne lagii* (**paros de-ne lagii*)
she food serve-INF began serve GIVE-INF began
'She began to dish out the food.'

(38h) *parosaa huaa khaanaa ThaNDaa ho gayaa thaa*
served PART food cold become GONE was
'The dished out food had become cold.'

(**paros diyaa/liyaa huaa khaanaa...*)[19]
serve GIVEN/TAKEN PART food

(39h) *hairaan ho-kar us ne haklaa-te-hue kahaa*
astonished become-CP he ERG stammer-PRES-PART said
'Losing his composure he stammered out...'

(**hairaan ho jaa-kar us ne haklaate hue kahaa...*)
astonished become GO-CP

These reduced privileges of external co-occurrence are not in contradiction with, but are rather a consequence of, the greater freedom in internal

co-occurrence relations shown by the Hindi-Urdu compound verb as compared to Marathi's. They follow from the presence of general across-the-board paradigmatic values which the CV/SV opposition has acquired in CV-rich languages like Hindi-Urdu. For example, the inability of CV's in Hindi-Urdu to appear as dependents of phasal expressions (cf. 37m and 37h) is a direct consequence of its perfective aspectual value. Other languages which have explicit means for indicating the perfective also do not allow it to occur as a complement of a phasal:

(40r) *on načinaet čitat' predloženie (*načinaet pročitat')*
he begins read-IMPFV sentence begins read-PFV
'He begins to read the sentence.'

(41g) *arxisa na matheno elinika* (Modern Greek)
began-1SG that learn-1SG-IMPFV greek
'I began to learn Greek.'

(**arxisa na matho elinika*
began-1SG that learn-1SG-PFV greek)

In a CV-poor language like Marathi, where each vector preserves more of its lexical semantics, it plays no role in any system of general semantic oppositions that determines its presence or absence. That is, it has a tendency to form a tighter bond with its primary and to co-occur with it in all environments open to the primary itself. For the CV-construction to participate in a system of contrasts which sweeps nearly every lexical verb into its domain, there must be a possible CV form for nearly every lexical verb. This is not the case in Marathi, especially for intransitive primaries. Since the compound forms that do occur in Marathi do not have paradigmatic values, it is not surprising that they are less sensitive to their grammatical environments and may occur in a variety of contexts (e.g. (34m) through (39m)) which are closed to their structural counterparts in Hindi-Urdu.

5. OTHER POSSIBLE CONCOMITANTS OF GRAMMATICIZATION

5.1. Specialization

In his paper for this volume Paul Hopper proposes some other candidates for phenomena that indicate grammaticization. One of these is *specialization*:

As a set of items grammaticizes, some subset of them are distinguished as the least marked, most prototypical exemplars of the acquired grammatical function. In the present instance, we should expect either a smaller number of vectors in Hindi-Urdu or some subset of them to have a higher ratio of occurrence than their counterparts in Marathi. Indeed, GO, the vector that occurs the most frequently in both languages, accounts for 44% of the Hindi-Urdu sample but only 32% of the Marathi sample. The five most important vectors in the Hindi-Urdu sample account for 92% of the total; in Marathi, 82%. The Marathi sample has CV's formed with 14 different vectors; the Hindi-Urdu sample, only ten.[20]

5.2. Morphological processes

In languages in which the compound verb is both rarer and semantically more transparent we should see more morphological processes with respect to which it is internal. For example, in Shina-Gurez and in Kashmiri, where the compound verb is seldom encountered in running texts and where compounds like those in (18h) and (19h) are not found, the causativizing morpheme is added to the compound verb as a whole[21]:

(42g) *se-s y-ow SiSaa phuT-@@*[22]
he-ERG THROW-PAST glass break-CP
'He broke the glass.'

(43g) *se-s y-ay-ow myU dastU SiSaa phuT-@@*
he-ERG THROW-CAUS-PAST my hand glass break-CP
'He had me break the glass.'

(44k) *yi chu su laDki yus darivaazi khuul-yith dy-aav-yi*
this is that boy who door open-CP GIVE-CAUS-FUT
'This is the boy who will have the door opened (for us).'

whereas in Hindi-Urdu the vector verb is always external to the morpheme of causation:

(42h) *us ne gilaas toR diyaa*
he ERG glass break GAVE
'He broke the glass.'

(43h) *us ne mujh se gilaas tuR-vaa diyaa*
he ERG me from glass break-CAUS GAVE
'He had me break the glass.'

(44h) *ye vo laRkaa hai jo darvaazaa khul-vaa de-gaa*
this that boy is who door open-CAUS GIVE-FUT
'This is the boy who will have the door opened (for us).'

The causative in Indo-Aryan is a lexically conditioned morphological process whose input is an intransitive or a transitive stem and whose output is a stem which is higher in transitivity by one degree. Not every verb undergoes it. In particular, modal auxiliaries and certain primaries with quasi-modal properties ('come', 'go', 'be', 'become') cannot be made causative (except by suppletion) in any modern Indo-Aryan language[23]. While vectors THROW and GIVE have been grammaticized in Shina-Gurez and Kashmiri as well as in Hindi-Urdu, they have not progressed so far in the former toward becoming purely auxiliary as they have in the latter. Hence they can still remain within the scope of the causative in those two languages[24]. Although this phenomenon does not happen to distinguish Hindi-Urdu from Marathi, it may be regarded as parallel to the greater flexibility in external privileges of co-occurrence which we found in the Marathi CV in comparison with the Hindi-Urdu CV.

6. SUMMARY OF FINDINGS

In this paper we have seen that the CV construction in a CV-rich language like Hindi-Urdu has poorer privileges of occurrence than it does in a CV-poor language like Marathi. This difference is understandable if we think of the less grammaticalized CV auxiliaries in Marathi as being themselves more like lexical verbs while the more grammaticized ones in Hindi-Urdu are less like their homophonous primaries, more like modals. As such they do not occur in places where other modals (e.g. *sak* 'can', *caahiye* 'should' and *paR* 'have to') do not occur while the CV auxiliaries of Marathi (in general) may occur wherever a homophonous counterpart among primary lexical verbs occurs.

On the basis of the characteristics of the CV in Indo-Aryan, I propose that there is a correlation among at least the following three concomitants of grammaticization: (1) An increase in textual frequency, (2) a universalization of privileges of occurrence with respect to a major word class and, (3) a reduction in privileges of occurrence vis-à-vis some inventory of constructions.

7. IMPLICATIONS FOR THE GENERAL STUDY OF GRAMMATICIZATION

In a contrastive study of the CV systems of Hindi-Urdu and Marathi, one is dealing with the grammaticization of a set of auxiliary verbs that is only in part correlatable to a process of semantic bleaching or loss of original lexical sense. While it is true that the Hindi-Urdu vectors *de* GIVE and *jaa* GO have traveled somewhat further along the path of lexical emptying than their Marathi counterparts, that difference in itself is not sufficient to explain the three to four hundred percent difference in textual frequency or the assumption of paradigmatic aspectual value by the CV/SV opposition in Hindi-Urdu.

Just as important as change in the meanings of vectors are changes in the optionality of their use to express the conceptual completeness of an action or event. In Marathi they are not used with primary verbs whose meanings in and of themselves indicate conceptual completeness, while in Hindi-Urdu they appear especially frequently with just such primaries. In Marathi they are not used in clausal contexts that guarantee a completive conceptualization of an action or event. In Hindi-Urdu they must be used in such contexts.

There are two general principles of communication which are invoked in explanations of grammaticization. One, "say no more than you must" (Traugott, 1989:50) is in direct conflict with the other: "Every speaker wants to give the fullest possible expression to what he means" (Lehmann, 1985:315). The present instance, which exemplifies use of explicit means to express meanings that are usually recoverable from context or which are inherent in co-occurring lexical items, seems clearly to support the second principle. But that does not mean that every event of grammaticization can be explained as an instance of Lehmann's aphorism. Speakers may increase the expressivity of what they say by moving in the opposite direction, by inviting the hearer to infer from silence what the speaker might have said overtly[25].

Most discussions of grammaticization (Traugott, 1988 and Sweetser, 1988 are good examples) focus on the grammaticized element as the locus of the semantic or functional change. That approach may be less productive in cases like that of the spread of the CV construction through Indo-Aryan languages. There it is not just the meaning of the grammaticized element itself that has changed. The meaning of the *absence* of the grammaticized element has also changed. In CV-poor Marathi that absence has no conventional interpretation that a hearer can rely on: He or she must infer from context

or knowledge of the world in what way the speaker conceives of an action or event. In Hindi-Urdu the absence of a vector has assumed significance: To the degree that its presence *even where it is redundant* has become an obligatory marker of perfectivity, its absence has come correspondingly closer to having a conventional interpretation of imperfectivity. Thus, we are dealing not just with the grammaticization of a set of overt linguistic items but with the complementary grammaticization of zero.

ABBREVIATIONS

CP	conjunctive participate
NIA	New Indo-Aryan
CTF	counter-to-fact

ACKNOWLEDGEMENTS

* I gratefully acknowledge the organizational and the financial support of the American Institute of Indian Studies, the U.S. Office of Education (Fulbright-Hays) and the Horace H. Rackham School of Graduate Studies (University of Michigan) for studies in the field which led to the findings presented here. I am also indebted to the insightful comments and suggestions of friends and colleagues, in particular, to those of John Myhill and Madhav Deshpande. A portion of the data used in this paper was presented at the Fourteenth Annual Meeting of the Berkeley Linguistics Society in February of 1988 and published in its *Proceedings*.

NOTES

1. The transcription system used for Marathi, Hindi-Urdu, Kashmiri, Shina-Gurez, and other NIA languages is the one generally found in the Indo-Aryan linguistics literature with the following exceptions: Contrastive length in vowels is indicated by doubling the symbol (not by a macron or colon); nasality in vowels, by capitalization (not by tilde or following capital *N*); retroflexion of apical stops and flaps, by upper case letters (not by sublinear dots); and palatal fricatives, by capitalization of the symbol for the corresponding alveolar fricatives. Data from Russian, Pali, Sanskrit, Tamil, and Orkhon Turkic are in traditional transcriptions.
2. Madhav Deshpande provided this example.

3. These inscriptions came to light some 250 years ago in what is now the central part of Mongolia.

4. While many more specifically diachronic studies need to be undertaken, research on older forms of Bengali and Hindi indicates that the construction is becoming more common with time (Zbavitel, 1970, Kapp, 1972).

5. I have used the term *co-express* because the past perfect also expresses anteriority of action here. Other examples that pattern like (10h) and (11h) exist in which perfect tenses cannot be used (example (Ah) is from Nespital, 1981:311):

   (Ah) *is-se pahle ki vo tumhE piiTe biicbacaav kar dEge*
   this before that he you beat rescue do GIVE
   'Before he can hit you they'll come to your rescue.'

6. It derives from Comrie (1976:16) and Timberlake (1985).

7. See, for instance, the "point" versus "blob" discussion in Comrie (1976:17–18).

8. Notably, Pořízka (1967–69) and Barannikov (1956).

9. Parallels exist in secondary aspectual functions, too. In both Russian and Hindi-Urdu, non-perfective forms in the imperative are felt to be more polite:

   (A) teacher: *vojdite v klass* *klaas mE aa jaa-iye*
   enter-PFV in class class in come GO-IMP
   'Come into the classroom!'
   (B) host: *vxodite* *andar aa-iye*
   enter-IMPFV inside come-IMP
   '(Please) come in!'

   Russian data from Isačenko (1954) as cited in Babby (1964:10).

10. This is not to say that the use of the compound verb in Marathi has nothing to do with these semantic categories. On the contrary, the expression of anteriority, for example, is facilitated in (Am) by the use of a CV:

    (Am) *mi aadhi haa ker kaaDh-un gheu. mag tu vaats.*
    I first this dust remove-CP TAKE then you read
    'First let me sweep. Then (sit down to) read.'

    But a CV is not required in constructions like (Am). In fact, for many primaries (especially intransitives) there are no compound forms available even when the environment is such as to favor the use of one:

    (Bm) *aapaN aadhi sTeSanaa-laa pohotsu. mag fon-karu.*
    we first station-DAT arrive then telephone
    'Let's first get to the station and then phone.'

11. The table I present here shows relative degrees of CV-text frequency on the basis of several hundred translations of the first paragraph of the Parable of the Prodigal Son published in Grierson's *Linguistic Survey of India.* A map of higher resolution incorporating a grid of 250 kilometers by 250 kilometers is in Hook, 1977. I suspect that the picture may have changed substantially since Grierson's time.

12. These figures are derived from counts made of CV use in modern short stories written in the 1970's by two highly educated writers: K.B. Vaid (Hindi-Urdu) and S.V. Chirmule

(Marathi). The stories are depictions of contemporary urban educated life written in a realistic style. Percentages are the ratios of CVs to total finite or non-finite verb forms:

(A)	Finite CV/Total:	Non-finite CV/Total:	Total CV/Total:
Hindi-Urdu:	64/463 14%	10/138 7%	74/601 12%
Marathi:	19/479 4%	4/282 1.4%	23/761 3%

13. Marathi data confirmed by Madhav Deshpande.

14. Hindi-Urdu data confirmed by Afroz Taj.

15. Is it possible to explain the different constraints on the occurrence of vectors GO and GIVE in the two languages as a difference in metaphorical mappings from lexical to grammatical meaning such that the Marathi mappings preserve more details of lexical meaning, and the Hindi-Urdu ones, fewer? For instance, how many of the Marathi CV's that have GO as vector (see Appendix) involve a disappearance, departure, diminution, or movement away of their subjects? While it is not always easy to decide in every case, approximately 60 percent of Marathi CV's made with GO may be interpreted as one or another kind of going away. For Hindi-Urdu CV's the corresponding figure is only 40 percent. The other 60 percent involve approach (as in 20h), coming to mind (*yaad aa jaa* 'remember'), coming to exist (*garbh ho jaa* 'conceive a child'), intensification (*baRh jaa* 'increase'), or entry into a new state (*xaraab ho jaa* 'turn bad') or activity (*so jaa* 'fall asleep'). While the physical dimensions of going are not invoked in such occurrences, some "topological" (cf. Sweetser, 1988; Talmy, 1988; Brugman and Lakoff, 1988) elements of the lexical meaning of GO are preserved (i.e., in going to completeness, or going from one state to another, or going further along a scale). Although they are proportionately less frequent, similarly topological uses are made of *dzaa*, the corresponding Marathi vector (e.g. *bharun dzaa* 'fill up', *harkhun dzaa* 'become happy', *avaghaDun dzaa* 'become inconvenienced'). It is conceivable that a sufficiently detailed taxonomy of the topological senses of GO and GIVE might allow one to order or partially order them along a scale or in some implicational hierarchy that predicts their paths of desemanticization and metaphorical extension. That is a much more ambitious project than what is possible to undertake here.

16. These figures were obtained by collecting one hundred CV's in the order in which they occurred from short stories written by two contemporary authors. See Appendix for details.

17. These figures were obtained by counting both SV's and CV's in a given semantic category for each language until a total of at least fifty was reached.

18. According to at least one native speaker (Madhav Deshpande), the ability to occur as an infinitive dependent on a phasal does not include CV's made with vector *Taak* THROW. This may mean that this vector has progressed further than others to becoming a marker of grammatical perfectivity.

19. Past participial CV's are found in Hindi-Urdu poetry:

(A) *vahII kahII mAI bhii... bhaag cale logO mE bhaagtaa*
there somewhere I too flee MOVE-PART people in fleeing
'There somewhere (am) I, too, fleeing among the fleeing...'
From part 3, "cambal kii ghaaTii," Muktibodh, 1964:246.

20. Of course, this correlation cannot be extrapolated back in time indefinitely. At some point

near the very beginning of the grammaticization process, the number of vector verbs must have been less than at later times.

21. This phenomenon was first noted by Vijay K. Kaul in his 1985 dissertation on the compound verb in Kashmiri. Shina-Gurez data from Abdul Ghaffar of Taubat and Muzaffarabad.

22. @ represents a long mid to high central vowel.

23. Kalasha may constitute an exception. See Bashir (1988:183) for an example (34-a) of the causative form of *hi-* 'become' as part of a complex predicate *wal hi-aw-* 'have (someone) guard'. Another example has the derived transitive form of *i-* 'come':

(Ak) *a gamb'uri gand'uryak i-'e-m-dai*
I-NOM flower fragrance come-CAUS-1S-PRES
'I am smelling the flower.' (Analysis has been normalized.) (Bashir, 1988:202)

The verb *par-'ik* 'to go' also has a derived transitive form *par-'ek* 'to make overflow' (E. Bashir, personal communication).

24. What look like counterexamples are found in Marwari, a relatively CV-rich language spoken in western Rajasthan:

(A) *baapjii, AAkhiyau mIIc li-raav-au*
father eyes close TAKE-CAUS-IMP
'Father, please close (your) eyes.' (Bahl, 1972:50)

But here the "causative" morpheme *-raav* is functioning only as a mark of honorification in a partially grammaticalized system of social deixis. See Magier, 1982 for discussion.

25. For instance, the use in "hip" English of simple verbs in place of phrasals like *blow up* or *tick off*:
(Ae) *We were on our way to the mess hall when the ammo dump blew.*
(Be) *He gave her a B minus. Was she ever ticked!*

APPENDIX: CV DATA

This appendix consists of two sets of one hundred CV's from Marathi and Hindi-Urdu collected seriatim from short stories by S.V. Chirmule and K.B. Vaid and grouped here according to the semantic class of their primaries (see Table 2 above).

I. Data from Marathi. (*d* = *dhaage* 'Threads'; *e* = *ekaa sTeSan maasTaraaci goST* 'A Stationmaster's Story'; *p* = *paaik* 'Pike'; *t* = *Traepidzh* 'Trapeze'; *a* = *aniket* 'Homeless'; or approximately 48 pages of text.)

Vectors:	*dzaa* GO 33	*ghe* TAKE 15	*ye* COME 13
	Taak THROW 13	*de* GIVE 8	*soD* LEAVE 4
	Thev PUT 4	*bas* SIT 2	*ghaal* PUT IN 2
	raah STAY 2	*kaaDh* TAKE OUT 1	*nigh* GO OUT 1
	laav PUT ON 1	*thabak* STOP 1	

Primaries (intransitive 53, transitive 47):

1.m. Displacement, destruction, or disposal:

d.1 *hiskaavun ghetli* 'snatched away'
d.1 *Thevun dili* 'put'
d.2 *vaaDhun gheu (laagli)* '(began) to dish out'
d.3 *chaaTun Taaklaay* 'plucked out; shed'
d.4 *vaaDhun ghetlela...* 'dished out'
d.5 *oDhun ghetlaa* 'pulled; drew toward'
d.6 *misaLun gelelyaa...* 'intertwined'
d.7 *(khiDkyaa) band karun ghe* 'close (the windows)'
d.9 *(pharaaTe) maarun Thevle* 'flick on (splotches)'
d.10 *deun Taak* 'give'
d.10 *(ker) kaaDhun ghete* 'remove, sweep up (the dust)'
d.11 *Taakun dili* 'threw out'
d.11 *phaaDun Taakle* 'ripped up'
d.11 *dzhaaDun Taaklaa* 'flung off'
e.1 *kaDelaa sarakvun deto* 'pushes (them) to one side'
e.2 *neun ghaalaa* 'make (it) crash'
e.2 *caDhun yeNaaryaa...* 'climbing'
e.3 *vyaapun raahilyaa (saarkha)* '(as if it) had pervaded'
e.3 *baandhun ghaatlyaa (saarkhe)* '(as if) wrapping up...'
p.4 *(gaaDi) ulthun deNyaatsaa* '..to blow up (a train)'
p.5 *tsaDhun aale* 'climbed up'
p.6 *nighun geli* 'left (the scene)'
p.7 *paLun gelaa* 'ran away'
t.4 *phaaDun Taakle hote* 'had ripped up'
t.7 *jhirpun gelele...* 'soaked up, absorbed into'
t.8 *(svatahaalaa) gurphaTun ghyaava* 'to snuggle up'
t.9 *tsaalun yaavaa* 'should come over'
t.10 *phaaDun Taakaavet* 'should tear up'
t.10 *bhirkaTun dyaavet* 'flick away'
t.12 *laavun ghetle* 'closed'
a.1 *(paaTi) Tekvun Thevli* 'leaned (his slate against...)'
a.3 *nighun gelaa* 'left; went out'
a.3 *haaklun laavit hotaa* 'was driving off'
a.5 *Thevun ghe* 'keep'
a.7 *utarun gelaa* 'got off'
a.7 *kaaDhun dila* 'took out'
a.7 *nighun geli* 'left; exited'
a.7 *bharun aalele...* 'overcast'
a.7 *aDakvun dila* 'hung up'
a.8 *bharun yet...* 'filling up'
a.11 *bharun yet...* 'filling up'
a.11 *(paaNi) piun Taakla* 'drank up (the water)'

2.0.m. Change of psychic state (emotion):
- d.2 *gOdhaLun gelaa* 'became confused'
- e.7 *(tyaalaa) aswasth karun soDto* 'makes (him) uneasy'
- e.8 *aswasth karun soDNaari...* 'disquieting'
- p.6 *bhaambaavun dzaaun...* 'having lost his composure'
- t.3 *(kaalvaakaalav) houn geli* 'he felt a pang of fear'
- t.3 *sukhaavun dzaate* 'becomes pleased'
- t.8 *vyaakLun gele* 'became agitated'
- t.8 *utsambaLun aali* 'overflowed (with emotion)'
- a.4 *bhaDbhaDun yeta* 'become overwhelmed'
- a.8 *harkhun gelaa* 'became happy'

2.1.m. Change of psychic state (consciousness):
- d.9 *gaDhun gelaa* 'became absorbed'
- d.11 *gaLun gelyaa (saarkha)* '(as if) fainting'
- e.7 *(tyaatsa man) vyaapun raahte* 'invades (his mind)'
- t.12 *ramun geli* 'became engrossed'
- a.2 *(dzhop) uDun geli* '(sleep) disappeared'
- a.7 *rangun gelaa* 'became absorbed'
- a.11 *ghaabrun gelaa* 'got scared'

3.0.m. Sensation or perception (inner source):
- a.9 *gudmarun Taakit...* 'tickling'
- a.9 *veTh vaLun aalyaa (saarkha)* '(as if) she felt a cramp'
- a.11 *ur bharun aalaa* 'felt suffocated'

3.1.m. Sensation or perception (outer source):
- e.1 *utsambaLun yeNaare* 'upward rushing (sounds)'
- e.3 *paDun thabakto* 'falls (within earshot)'
- t.1 *(DoLe) dipun dzaatil* 'dazzling'
- t.6 *sarsarun geli* 'went rippling'
- a.3 *disun aala* 'was visible'

4.0.m. Mental action (involuntary):
- p.8 *tarLun geli* 'flashed (through his mind)'
- t.5 *tarLun gelaa hotaa* 'had flashed (through his mind)'
- a.5 *kaLun aali* 'understood'
- a.5 *visrun gelaa* 'forgot'

4.1.m. Mental action (voluntary):
- t.3 *aakhun ghetlaa hotaa* 'had planned out (her life)'
- t.4 *(samzut) karun ghetli* 'took (it into his head)'
- t.7 *visrun dzaaylaa (havaa)* '(necessary) to forget'

5.0.m. Communication (inarticulate):
- p.1 *(daat) aavLun ghet...* 'gnashing (his teeth)...'
- t.5 *(mansokt) raDun ghetla* 'cried (his fill)'

5.1.m. Communication (articulate):
a.3 *sAAgun Taakaava* 'should tell'

6.0.m. Creation or change of physical state:
d.8 *pura karun Taakto* '(I) will complete'
d.9 *puraa karun Taakto* '(I) will finish'
d.9 *badalun Taaku yaa* 'let's change (the arrangement)'
e.1 *bharun gelela (aakaaS)* 'an overcast (sky)'
e.4 *avagaDun gelyaati* 'gotten big'
e.4 *dhaaraN karun baslelyaa...* 'assuming (a severe look)'
e.4 *aa vaasun baslela...* 'gaping'
e.6 *(ghaamaana) DabDabun gelelaa...* 'dripping (with sweat)'
t.6 *sazvun dyaaylaa* 'to decorate'
t.6 *aaTun dzaaNaara...* 'shrinking'
a.7 *avaghaDun geliy* 'has gotten uncomfortable'
a.8 *avghaDun gelelaa...* '(pregnant and) uncomfortable'

6.1.m. Creation or change of non-physical state:
e.7 *nhaaun nighte* 'becomes suffused (with light)'
p.6 *kabarstaan karun soDlay* 'have made (it a graveyard)'

7.0.m. Other (temporal displacement):
t.9 *(tiin) vaadzun gele* 'it's past (three)'

7.1.m. Other (social):
d.3 *(bhAADaN) ukrun kaaDhtes* '(you) pick (a fight)'
t.4 *(don caar haat) karun Taakaavet* 'should marry'

7.2.m. Other (writing):
e.1 *Tipun gheto* 'notes, writes down'
t.2 *lihun ghei* 'used to write down'

7.9.m. Other (unclassified):
p.7 *(tahSil) dumdumun soDlaa* 'made (the district) echo'
t.8 *asaa koN laagun gelaay* 'Who does he think he is?'
a.3 *(jikiri) karun Thevalyaat* 'created (problems)'
a.8 *(ghoTaaLe) karun Thevalet* 'made (a complete mess of)'

II. Hindi-Urdu data from *t* = *trikoN* 'Triangle'; *b* = *bhuut* 'Ghosts'; *a* = *avasar* 'Opportunity'; or approximately 15 pages of text.)

Vectors: *jaa* GO 44 *de* GIVE 17 *le* TAKE 14
uTh RISE 10 *Daal* THROW 7 *aa* COME 4
cal MOVE 1 *rakh* PUT 1 *baiTh* SIT 1
dhamak DROP 1

Primaries (intransitive 63, transitive 37):

1.0.h. Displacement, destruction, or disposal:
t.1 *dabaa diyaa ho* 'had pressed'
t.1 *Daal lii ho* 'had stuck (something) in'

t.1 *toR Daalaa ho* 'had broken'
t.2 *aa dhamaktii* 'had showed up'
t.3 *haT jaane kaa...* '..to move away'
t.3 *kas liyaa* 'pressed, hugged'
t.3 *(cuur-cuur) kar dUU* 'break (into smithereens)'
t.4 *(use) kas liyaa thaa* 'had hugged (her)'
t.5 *(mAI) kas lii gaii thii* '(I) had been embraced'
t.5 *(pati) aa jaae* '(husband) may show up'
t.6 *(guldastaa...) de maaraa thaa* 'had dashed (a bouquet)'
b.4 *gum ho gaye hAI* 'have disappeared'
b.4 *nikal gayaa* 'left; went out'
b.4 *cale jaaEge...* 'will go away'
b.5 *maar Daalne (kii baat...)* '(the idea of) killing'
b.6 *Takraa gaII* 'collided'
a.1 *(darvaazaa) band kar diyaa hotaa* 'had shut (the door)'
a.1 *lauT jaataa* 'would have gone back'
a.1 *gum ho jaatii* 'would have disappeared'
a.2 *(koii aur) aa jaae to* '(if someone else) came (then)'
a.2 *sarak gaii ho* 'had moved away'
a.2 *muR gayaa* 'turned away'
a.2 *lauT aayaa* 'came back'
a.2 *band kar diyaa* 'closed'
a.2 *pii liyaa* 'drank up; swallowed'

1.1.h. Change of posture:
a.1 *baiTh jaaiye* 'have a seat'
a.1 *baiTh gayaa* 'sat down'
a.1 *baiTh jaaiye* 'have a seat'

2.0.h. Change of psychic state (emotion):
t.4 *bhaRak uThtii* 'would have gotten angry'
b.2 *kAAp uThaa kartii thii* 'used to tremble (with fear)'
b.3 *(mujhe) akhar uThaa thaa* 'had irritated (me)'
b.5 *taRap uThte hOge* 'must be tormented with frustration'
a.1 *cauk uThii* 'started (in fright)'
t.5 *(mujhe) garmaa diyaa hogaa* 'must have excited (me)'
t.5 *(uttejit) ho uThaa thaa* 'had gotten (excited)'

2.1.h. Change of psychic state (consciousness):
b.5 *so jaane se pahle* 'before falling asleep'
b.5 *so gayaa* 'went to sleep'

3.0.h. Sensation or perception (inner source):
t.1 *jhanjhanaa uThaa thaa* 'had tingled'
t.2 *(jisam behaal) ho hii jaataa hai* '(body) gets (horny)'

3.1.h. Sensation or perception (outer source):
b.1 *mahak uThtaa thaa* 'used to become redolent (with)'

a.2 *(fon) baj uThaa ho* '(phone) had suddenly rung'

4.0.h. Mental action (involuntary):
t.6 *(ghaTnaa) yaad aa gaii thii* '(event) had come to mind'
a.1 *yaad aa jaane se...* 'from remembering...'

4.1.h. Mental action (voluntary):
t.2 *(hisaab) kar rakhaa ho* 'had kept (an account)'
t.3 *andaazaa lagaa lenaa* 'to assume'
t.4 *viSvaas kar lene kaa (abhinay)* '(pretend) to believe'
t.4 *(dayaa) kar dii ho* 'had taken (pity)'
t.7 *sviikaar kar liyaa* 'accepted'
b.1 *faislaa kar liyaa thaa* 'had decided'
b.1 *(Sak) maar Daalne ke lie* 'to kill (my doubt)'
b.2 *bardaaSt kar letaa thaa* 'used to tolerate'
b.4 *vo ubhar aaii* 'she recovered (her wits)'

5.0.h. Communication (inarticulate):
t.2 *hAs dii hotii* 'had laughed'
t.2 *hAs detaa* 'had laughed'
t.3 *cillaa uThne kaa...* 'to scream'
t.7 *(muskaraahaT) khil aatii hai* '(a smile) breaks out'
a.1 *muskraa detii* 'would smile'

5.1.h. Communication (articulate):
t.1 *(mere mUh se) nikal gayaa* '(I) blurted out'
t.2 *(jhuuTh) bol diyaa ho* 'had told (a lie)'
t.2 *(javaab) de diyaa hotaa* 'had given (an answer)'
t.2 *kah diyaa hotaa* 'had said'
t.3 *(kaa) dam bhar lene ke saath* 'gasping out (that)'
t.3 *bak Daalne kii...* 'to babble nonsense'
t.7 *bataa Daalne kaa...* 'to tell'
b.2 *(Gazal) gaa Daalne ke...* 'to sing (a ghazal)'
b.4 *bataa diyaa* 'told'
b.5 *sunaa Daaltii ho* 'may narrate'
a.2 *kah diyaa* 'said'
a.3 *(sab) bataa diyaa* 'told (everything)'

6.0.h. Creation or change of physical state:
t.3 *(AAkhE) band ho gaII* '(eyes) closed'
t.3 *(AAkhE) na khul jaatII* '(eyes) had not opened'
t.6 *(garbh) ho gayaa* 'conceived'
t.6 *sAbhal jaanaa* 'to get dressed'
b.3 *(muskraahaT) simaT aaii* '(smile) shrank'
a.1 *(muskraahaT) gir jaatii* '(smile) would have faded'
a.2 *(AAkhE) mUUd lII* 'closed (eyes)'

6.1.h. Creation or change of non-physical state:
t.2 *khatm ho jaatii* 'would have ended'

t.2 *badal jaataa* 'would have changed'
t.3 *(tanaav) baRh gayaa* '(tension) increased'
t.3 *(raakh) ho gayaa hogaa* 'must have turned (to ash)'
t.4 *(aadat) ban gaii hai* 'has become (a habit)'
t.5 *ekras ho hii jaate hAI* 'to become monotonous'
t.5 *pathraa jaaegaa* 'will freeze, turn to stone'
t.7 *badal jaaegaa* 'will change'
t.7 *aam ho jaaegaa* 'will become routine'
t.8 *(bedaaG) bac gayaa hUU* 'I came out (unscathed)'
b.4 *(hairaanii) baRh gayii thii* '(astonishment) had grown'
b.5 *aadii ho jaate hAI* 'become accustomed'
a.1 *(baat) khatam ho jaatii* '(matter) would have ended'
a.1 *(...rux) badal jaataa* 'taken a different (turn)'
a.1 *(xaamoSii...) ho jaatii hai* '(silence...) falls'
a.2 *(diivaar saa) ho gayaa thaa* 'had become (like a wall)'
a.2 *xaraab ho gayaa* 'turned bad'

7.0.h. Other (temporal displacements):
t.5 *(samay) biit jaataa hai* '(time) goes by'
b.3 *(ek muddat) biit calii* '(a while) passed'

7.3.h. Other (phasals):
t.3 *(Suruu) kar diyaa* 'began (to...)'

7.4.h. Other (instantaneous emergence):
t.5 *(hamE) pakaR le* 'should catch (us)'
t.6 *(un ko) pakaR liyaa thaa* 'had discovered (them)'

7.9.h. Other (unclassified):
t.2 *Tal jaataa* 'would have avoided'
a.2 *(na jaane kyaa) kar baiThtii* '(God knows what) I'd have done.'

REFERENCES

Axmaker, S., A. Jaisser and H. Singmaster (eds). 1988. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Babby, Leonard H. 1964. Aspect Selection in the Negated Imperative Compared to the Affirmative in Modern Russian. MA thesis, Harvard University.

Bahl, K. C. 1972. *On the Present State of Modern Rajasthani Grammar*. Jodhpur: Rajasthani Shodh Sansthan.

Barannikov, A. P. and P. A. Barannikov. 1956. *Khindustani: Khindi i Urdu*. Moscow: Foreign Literature Pub. House.

Bashir, E.L. 1988. Topics in Kalasha Syntax: An Areal and Typological Perspective. PhD dissertation, University of Michigan.

Brecht, Richard. 1985. "The form and function of aspect in Russian." In Flier and Brecht, 1985:9–34.
Brugman, Claudia and George Lakoff. 1988. "Cognitive typology and lexical networks." In *Lexical Ambiguity Resolution*, G. Cottrell and S. Small (eds). Palo Alto: Morgan Kaufman.
Buddhaghosa's *sumaṅgalavilāsinī*. 1974. M. Tiwary (ed.). Nalanda: Nava Nalanda Mahavihara.
Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
Chirmule, S.V. 1967. *SriSillak*. Bombay: Mauz Prakshan Grih.
Comrie, B. 1976. *Aspect*. Cambridge: Cambridge University Press.
Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel.
Flier, M. S. and R. D. Brecht (eds). 1985. *Issues in Russian Morphosyntax*. Columbus, OH: Slavica.
Forsyth, J. 1970. *A Grammar of Aspect: Usage and Meaning in the Russian Verb*. Cambridge: Cambridge University Press.
Givón, Talmy. 1975. "Serial verbs and syntactic change: Niger-Congo." In *Word Order and Word Order Change*, C. Li (ed.). Austin: University of Texas Press.
Givón, Talmy. 1984. *Syntax: A Functional-Typological Introduction*. Amsterdam: John Benjamins.
Grierson, G. A. 1919. *Linguistic Survey of India*. Calcutta: Govt. of India (1968 reprint by Motilal Banarsidass).
Hendriksen, Hans. 1944. *Syntax of the Infinite Verb Forms of Pali*. Copenhagen: Einar Munksgaard.
Hook, Peter E. 1974. *The Compound Verb in Hindi*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
Hook, Peter E. 1977. "The distribution of the compound verb in the languages of North India and the question of its origin." *International Journal of Dravidian Linguistics* 6:336–51.
Hook, Peter E. 1978. "Perfecting a test for the perfective: Aspectual parallels in Russian, Lithuanian, Modern Greek, Hindi and Pashto." *University of Michigan Papers in Linguistics* 2:89–104.
Hook, Peter E. 1987. "Linguistic areas: Getting at the grain of history." In *Festschrift for Henry Hoenigswald*, George Cardona and Norman Zide (eds), 155–168. Tübingen: Gunter Narr Verlag.
Hook, Peter E. 1988. "Paradigmaticization: A Case Study from South Asia." In Axmaker et al. 1988:293–303.
Hopper, Paul. Volume I. "On some principles of grammaticization."
Isačenko, A.V. 1954. "O imperative v ruštine a slovenštine — Využitie vidovych odtienkov v imperative." *Sovetska jazykoveda* 4.
Kapp, Dieter B. 1972. *Das Verbum paraba in seiner Funktion als Simplex und Explikativum in Jayasis Padumavati*. Wiesbaden: Verlag Otto Harrassowitz.
Kaul, Vijay K. 1985. The Compound Verb in Kashmiri. Unpublished PhD dissertation, Kurukshetra University.
Lehmann, C. 1985. "Grammaticalization: Synchronic variation and diachronic change." *Lingua e Stile* 20:303–318.
Lord, C. 1973. "Serial verbs in transition." *Studies in African Linguistics* 4(3):269–96.
Magier, David. 1982. "Marwari honorifics." *South Asian Review* 6:160-173.

Masica, Colin P. 1976. *Defining a Linguistic Area: South Asia.* Chicago: University of Chicago Press.
Maslov, Iu. S. 1984. *Očerki po Aspektologii.* Leningrad: Izdatel'stvo Leningradskogo Universiteta.
Merrill, Peter. 1985. "Universal quantification and aspect in Russian." In Flier and Brecht, 1985:58–72.
Muktibodh, Gajanan Madhav. 1964. *cAAd kaa mUh teRhaa hai.* Varanasi: Bharatiya Jnanpith Prakashan.
Nespital, Helmut. 1981. *Das Futursystem im Hindi und Urdu.* Wiesbaden: Franz Steiner Verlag.
Pořízka, Vincenc. 1967–69. "On the perfective verbal aspect in Hindi (Some features of parallelism between New Indo-Aryan and Slavonic languages)." *Archív Orientální* 35:64–88, 208–231; 36:233–251; 37:19–47, 345–364.
Sweetser, E. 1988. "Grammaticalization and semantic bleaching." In Axmaker et al. 1988:389–405.
Talmy, Leonard. 1988. "Force dynamics in language and cognition." *Cognitive Science* 12(1).
Tekin, Talât. 1968. *A Grammar of Orkhon Turkic.* Bloomington: Indiana University [*Uralic and Altaic Series*, 69].
Tikkanen, Bertil. 1987. *The Sanskrit Gerund: A Synchronic, Diachronic and Typological Analysis.* Helsinki: Finnish Oriental Society [*Studia Orientalia*, 62].
Timberlake, Alan. 1985. "The temporal schemata of Russian predicates." In Flier and Brecht, 1985:35–57.
Traugott, E.C. 1988. "Pragmatic strengthening and grammaticalization." In Axmaker et al. 1988:406–416.
Traugott, E.C. 1989. "On the rise of epistemic meanings in English: An example of subjectification in semantic change." *Language* 65:31–55.
Vaid, Krishna Baldev. 1970. *duusre kinaare se...* (*From the other shore...*). Delhi: Radhakrishna Prakashan.
Vendler, Zeno. 1967. "Verbs and times." In *Linguistics in Philosophy*, 97–121. Ithaca: Cornell University Press.
Zbavitel, Dušan. 1970. *Non-finite Verbal Forms in Bengali.* Prague: Czechoslovak Academy of Sciences.
Zvelebil, Kamil. 1964. *Tamil in 550 AD: An Interpretation of Early Inscriptional Tamil.* Prague: Czechoslovak Academy of Sciences.

SECTION B: ARGUMENT STRUCTURE

Grammaticization of Topic into Subject

Masayoshi Shibatani
Kobe University

1. INTRODUCTION

While the notion of the grammaticization of topics into subjects has been widely alluded to in the literature, detailed examination of the process in progress has been lacking (Li and Thompson, 1976; Givón, 1979; Mallinson and Blake, 1981). One notable exception is Givón (1976), who offers an explicit account of how topic is incorporated into a clausal structure and becomes subject. Givón's oft-quoted formulation is as follows:

(1)

TOPIC-SHIFT STRUCTURE			NEUTRAL (RE-ANALYZED)	
The man, he came		→	*The man he-came*	
TOP	PRO		SUBJ	AGR

Givón's study, which deals with the grammaticization patterns of agreement morphemes among Bantu languages, however, was primarily concerned with the rise of agreement morphology from topic-anaphoric pronouns rather than the rise of a grammatical subject per se.[1] Furthermore, there are those languages that appear to have grammaticized topics into subjects without ever involving concomitant reanalysis of topic-anaphoric pronouns as agreement markers.

This paper examines much more general grammaticization processes whereby topics are turned into subjects. The scope of our investigation covers several typological groups of languages. Because of the need to have a clear understanding of a grammatical topic, we first look at Japanese, which makes clear distinctions between topic sentences and topicless sentences. English represents a language in which the topic and the subject functions have converged on an agentive nominal (of a transitive clause), whereas the deep ergative language Dyirbal exemplifies the languages in which the two functions have converged on a patient nominal. In other words, English is an instance

of those languages in which "the nominative marks a grammaticalized topic", and Dyirbal those in which "the absolutive represents a grammaticalized topic" (Mallinson and Blake, 1981:109).

Philippine languages, on the one hand, show a transition from those languages in which a grammatical topic is not an obligatory clausal constitutent (e.g. Japanese) to those in which a topic has been grammaticized to the extent that it has become an obligatory constituent of major clause types. On the other hand, Philippine languages bridge a gap between accusative languages and ergative languages. Philippine languages and ergative languages together form a group of languages that pose a challenge for an account on the rise of split syntax. In the past, a great deal of attention has been focused on the understanding and explanation of split-case marking (Silverstein, 1976; Wierzbicka, 1981), but little has been said on the rise and nature of split syntax. Below, we offer a hypothesis concerning the rise of split syntax in terms of the grammaticization patterns of topics into subjects.

While the traditional definition of grammaticization has been largely confined to the changes from lexical to grammatical items (see the introduction to the present volume by Traugott & Heine), we feel that syntactic phenomena are equally amenable to the concept of grammaticization. Indeed, the formation of grammatical categories such as topic and subject appears to involve processes similar to the grammaticization processes leading to the establishment of grammatical categories such as tense and aspect, though specific lexical morphemes may not be involved. In other words, the concept of grammaticization becomes even more useful by not confining itself to origins in the lexical domain. To show that this is the case is one of the goals of this paper.

2. SUBJECT AND TOPIC

In the Western grammatical tradition, there have been two distinct notions that recur in characterizing the grammatical subject. One is a notional rendering of the subject defined in the tradition of Western philosophy, namely "the subject is what you talk about" (Jespersen, 1924:146) or "the subject is what we are talking about" (Chafe, 1976:43). This is in actual fact closer to a notional definition of the topic of a sentence, but the philosophers opted for the term subject, although, as we see below, they were in fact dealing with what is represented by the topic category in a language like Japanese. Grammarians could have done better rather than blindly adopt the philosophers' term, but they were handicapped in that Western languages did

not present a clear case calling for two distinct categories of topic and subject.

The other notion central in the characterization of the subject is that of actor or agent. In traditional grammar, the notion of subject as an agent figures prominently only when an agentive nominal is not chosen as a surface grammatical subject as in the passive construction, where a distinction is often drawn between the grammatical subject, referring to the surface subject, and the logical subject referring to the agentive nominal—see Jespersen (1924). However, Bloomfield (1933) prefers the term "actor-action" in characterizing a typical English sentence construction over the term "subject-predicate," which he finds "superfluous" (173) in the case of English. And more recently Dixon (1979) proposes a universal definition of subject as {A, S}: namely, a union of the agentive nominal of a transitive clause {A} and the obligatory nominal of an intransitive clause {S}.

The above two distinct notions—subject as a topic and subject as an actor/agent—used in the characterization of the category subject continue to figure prominently in the definition of subject in contemporary studies. Thus, Keenan's (1976) universal subject property list includes statements such as: "b-subjects [basic subjects] are normally the topic of the b-sentence [basic sentence]" (318) and "b-subjects normally express the agent of the action, if any" (321). Likewise, for Comrie (1981) the prototypical subject is the intersection of agent and topic.

Contrary to these views, our view on the matter is that the concepts of the grammatical topic and that of the grammatical subject need to be clearly separated at the theoretical level, for languages such as Japanese and Korean make a clear distinction between them affording separate grammatical constructions involving them. Whether a language, like English, merges them into a single category or not is but one grammaticization pattern whereby the topic converges on the agentive nominal. While this is a frequently observed pattern of grammaticization of topics into subjects, there are those languages that show a different grammaticization pattern with the result of a syntactic organization profoundly different from languages such as Japanese and English. In order to see these different realizations of the topic and the subject categories, it is essential that we have a better understanding of the natures of the topic and the subject by examining them in those languages that have the clearest manifestations of these categories.

2.1. On the nature of a grammatical topic

The topic construction in Japanese has been a target of intense study both within Japan and abroad. While the correlation between the *wa*-marked

Japanese topic and old information, as pointed out by Chafe (1970) and elaborated on by Kuno (1973), has had the effect of shifting emphasis of research from more grammatical to more pragmatic or discourse oriented aspects of the topic construction, we find that the fundamental role of the particle *wa* and the nature of the entity marked by this particle are best understood in terms of abstract cognitive styles of apprehension of a state of affairs and in terms of their grammatical expressions. We believe that an understanding of these matters is prerequisite to a better understanding of the pragmatics of the topic construction. Our discussion of the Japanese topic below draws heavily on the works carried out within the tradition of Japanese grammatical studies.[2]

Within the Japanese grammatical tradition, the earliest remarks on the topic particle *wa* appear in the 18th century in the works of such grammarians as Toganoi Michitoshi (*Teniha Abikizuna* 1770) and Fujitani Nariakira (*Ayuhishô* 1778). The key terms employed in these studies are "emphasis" and "separation". While a number of interpretations of the early grammarians' characterizations of the function of *wa* are possible, and indeed further interpretations are necessary in order to appreciate their implications, the essential ideas are clear enough. Namely, *wa* separates an entity and has the effect of making judgment emphatic.

The notion of "emphatic judgment" must be interpreted carefully, since the presence of emphasis is not always clear in every topic sentence. Indeed the notion of emphasis can be clearly appreciated only when a topic sentence is embedded in the context of contrast. When no contrasting proposition exists or is implied, there is little emphasis one can detect in a topic sentence. In this case, the notion of emphasis needs to be interpreted much more abstractly. Namely, when we make a categorial judgment such as "The sun rises in the east" and "Man is mortal," we are isolating or singling out these propositions from other possible ones, and thus the contrast is inherent in this kind of judgment, but it becomes apparent only when a parallel or contrasting proposition exists. In the words of Sansom (1928:258), "It [*wa*] marks, however, not an emphasis modifying a propostion, but an emphasis inherent in every proposition." Thus, the so-called contrastive *wa* and thematic *wa* (Kuno, 1973) are basically one and the same thing; the topic as well as the entire topic containing proposition are inherently contrastive, but the contrast becomes pronounced only when there is a contrasting factor in the context.

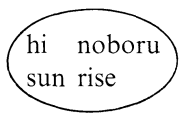
The notion of separation is multiply ambiguous. In the broadest sense, it echoes the above discussion on the emphasis of judgment; namely what is

expressed by a topic sentence as a whole is isolated from other possible propositions. In one of the narrow senses, the notion of separation applies to an item marked by the particle *wa*. In this sense, the particle separates or singles out the reference of the preceding noun phrases from other entities. It is in this sense that W. G. Aston remarked in his *A Grammar of the Japanese Written Language* (1872:51) that: "*Wa* is used not only to isolate or single out one or more objects from a number, but to contrast one object with another."

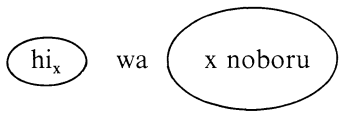
Whereas Aston's sense of separation emphasizes the separation of an entity from other general objects, there is another sense of separation that can be ascertained in the early grammarians' writings. That is, *wa* separates an entity from what constitutes the entire proposition. What is separated is a unit of existence corresponding on the one hand to the subject in metaphysics, a central entity of what is to be known, and on the other to the subject of judgment in traditional logic.

To continue in philosophical terms, what is to be known or judged is a state of affairs (*Sachverhalt*). *Wa*, as a separating particle in the last sense, isolates an entity from this state of affairs to be known; that is, it has the function of setting up two terms—one a central entity of what is to be known, and the other a predication describing the state of affairs to be known. (2) is a schematic representation of the process of isolation effected by *wa*.

(2) a. State of affairs to be known



b. Separation of a topic



c. Resulting topic sentence
Hi wa noboru
sun TOP rise
'The sun rises.'

An important point to be recognized here is that the particle *wa* functions as both separating and connecting particle. The mode of judgment underlying the topic expression is that of distinguishing two terms of a state of affairs to be known and that of affirming (or negating) the connection of the two

terms. Again, in the words of Sansom (1928:258), "the mental process by which any logical proposition is formed consists of two stages, first an analysis and then synthesis." This "connecting" aspect of *wa* is likened to the logical and grammatical copula by a number of grammarians, e.g. Yamada (1908) and Sansom (op. cit.). Sansom (op. cit. 258) notes that: "There does not seem to be any fundamental difference between the function of *wa* in *hi wa atsushi* ['fire is hot'] and the function of 'is' in 'fire is hot'." That is, *wa* in the form of the proposition {A *wa* B} has as much of the connecting function as the copula in {A *is* B} as it has the function of separating A and B. Thus, the topic construction in Japanese, in fact, closely corresponds to the subject-predicate paradigm in the traditions of Western philosophy and grammar. And this is precisely how the Western grammarians obtained the characterization of the subject as "what you talk about", which also fits as a notional characterization of the topic.

So far so good, since there is a clear correspondence between the Japanese topic and the subject in the Western grammatical tradition; it is merely a matter of terminology. However, this parallelism breaks down immediately when we come to examine another way of apprehending a state of affairs. Instead of being analyzed in the manner underlying a topic sentence, a state of affairs can be apprehended wholly as an unanalyzed unit or in terms of a one-term representation. In Japanese, the one-term expression of (2a) is (3), where the particle *ga*, an earlier genitive particle, marks the subject in Modern Japanese.

(3) *Hi ga noboru.*
sun NOM rise
'The sun rises.'

In Japanese the distinction between a topic sentence and a topicless sentence is reflected not only in terms of particles, but also structurally. That is, the notion of separation discussed above entails a structural expression, and this aspect of the idea of separation also has been noted by a number of grammarians. For example, S. R. Brown (1863:xxxiii) noted that: "*wa* ... separates the words before it from those following it..." Though the flow of articulation needn't be actually cut off after *wa*, Brown's observation points out an important structural property of a topic sentence.

The structural effect of the separating function of *wa* has been discussed in terms of how immediately the topic and the subject nominals are related to the following verbal material. In a topicless sentence, a clause is more tightly organized (reflecting the unanalyzed material of judgment) and the

subject shows a tighter bondage with the predicate, disallowing the insertion of certain types of epenthetic clauses. This is not the case when *wa* occurs. The topic is often separated from its lexical predicate by a fair number of intervening elements. The favorite examples among Japanese grammarians illustrating this point, beginning with Yamada (1908), involve the contrasting forms in (4).

(4) a. [*tori ga tobu toki*] *naku*
bird NOM fly time cry
'X cries when the bird flies.'

b. *tori wa* [*tobu toki*] *naku*
bird TOP fly time cry
'The bird, when (it) flies, cries.'

As indicated by the bracketing, the topic in (4b) forms a discontinuous constituency with the final predicate, whereas the subject in (4a) does not allow such an interpretation. Again, we observe that the subject-predicate bipartite constituency posited in the Western philosophical and grammatical traditions is reflected accurately in the structural property of the Japanese topic sentence. What this means is that the subject-predicate analysis of the Western tradition has been concerned with the properties of a topic sentence rather than a topicless sentence. But, unfortunately, Western languages do not readily show a formal distinction between a topic sentence and a topicless sentence; as a result an important distinction has not been made.[3] The term subject and the subject-predicate analysis have been applied indiscriminately even to a topicless sentence by those dealing with languages in which a topic and a subject have merged structurally, or, in our terms, languages in which a topic has been grammaticized into a subject.

We shall now explicate the difference between a topic sentence and a topicless sentence in terms of their use and in terms of the types of judgment underlying them. The topic sentence *hi wa noboru* 'the sun rises' and the topicless counterpart *hi ga noboru* 'the sun rises' differ substantially with regard to the context in which each is used appropriately. The topic sentence is a generic statement regarding a general property of the sun. Specifically, this kind of sentence cannot describe an ongoing event just witnessed or a state just discovered.[4] It is the topicless sentences that describe those events and states that have been just witnessed. However, in such a context, topicless sentences are normally uttered with some kind of exclamation, just like the English forms such as *There rises the sun!* or *Look, the sun rises!*

We can now talk about the difference between the topic sentence and the topicless sentence in terms of the difference in the type of judgment involved. A generic statement of the kind that the sentence *hi wa noboru* 'the sun rises' represents involves what can be called "experiential judgment", whereas a topicless sentence depicting a witnessed event or state involves "perceptual judgment". Thus, the topic sentence *hi wa noboru* can be uttered even when the sky is cloudy; the statement is backed up by the speaker's past experience with the sun. On the other hand, the topicless utterance *hi ga noboru* is not possible when the sun is not in fact rising at the moment of the utterance.

Just as a state is expressed by the use of the copula in English and many other languages, there is a strong affinity between a stative expression and the topic construction in Japanese such that states are normally expressed in topic sentences. This seems to be due to the difference between states and activities. While activities are categorical, states are relative, and whereas only a perceptual judgment suffices in the apprehension of an activity, an experiential judgment is called for in the case of a state. Here again, we see some parallelism between the topic particle *wa* and the copula in English and other languages. And one might also realize that Western philosophers normally discuss a stative expression when they talk about the subject-predicate paradigm.

Finally, let us briefly touch upon the discourse or textual roles that topic and topicless sentences play. Contrary to the highly plausible view that the grammatical topic (e.g. the *wa*-marked topic nominal) indicates a discourse topic, no direct correlation obtains between the two. While a discourse topic can be referred to by the *wa*-marked nominal in a Japanese narrative, and it is often so marked, *wa*-marking is determined by other factors than the thematicity of a discourse topic. Indeed, the recent study by Clancy and Downing (1987:24) draws the conclusion that: "*wa*-marking is not necessary to establish thematic status, nor does *wa*-marking, when it appears, necessarily indicate that the participant in question is thematic..."

Clancy and Downing's conclusion also indicates that the grammatical topic and old/given information is not correlated in a one-to-one fashion. Whereas the topic is typically old/given (in the sense of Chafe, 1976), a nominal with the old/given status is not always realized as a grammatical topic. The correlation of the grammatical topic and old/given information stems from the requirement underlying the experiential judgment; namely, what is isolated as a topic is a unit of existence, and as such the existence of the referent of a topic must necessarily be presupposed.

In order to ascertain the textual roles of the topic and topicless sentences, again the fundamental differences between the two types of sentences must be recognized. Topicless sentences present witnessed events and states as straightforwardly as they are perceived without analyzing their parts and without reflecting the speaker's experiential judgment regarding the relationship between the analyzed parts. Topic sentences, on the other hand, separate out topics, which are experientially judged in terms of their relationship to the rest of the sentence. Thus, whether a given nominal is a discourse topic or not and whether it is old/given information or not, it will not be made a grammatical topic when an event involving its referent is to be presented in the manner associated with a topicless sentence. For example, when an event is presented in a subordinate clause as subsidiary, background information, no grammatical topic is established. Likewise, when an event is introduced as something that is objectively witnessed, no topic sentence obtains.

In the text, the grammatical topic functions as a cohesive device. A long stretch of topicless sentences presents a series of events as if each is witnessed afresh. Such a sequence of topicless sentences is disfavored from the textual consideration of cohesion. The grammatical topic functions as a powerful cohesive device that relates an event to the preceding event in such a way that the new event is presented as a further development of the preceding event by way of sharing the topic with it. Thus, the actual use of the topic sentence is pragmatic in nature rather than grammatical. However, it is only by understanding the grammatical aspect of the construction that its pragmatic function can be assessed in a proper perspective.

2.2. On the nature of a grammatical subject

We shall now briefly discuss the nature of a grammatical subject so that it can be contrasted with the grammatical topic. A grammatical subject is a syntactic category resulting from the generalization of an agentive participant. That is, an agentive participant {A} of a transitive clause is the archetypical subject, from which one of the traditional definitions of subject as an agent/actor obtains.

Languages differ in the extent that this generalization has been achieved. That is, languages show varying degrees of aligning non-agentive participants according to the syntactic patterns of the agentive nominal. English is subject-prominent in that it has pushed this generalization to the most extreme limit.[5] Thus, in English not only an experiencer and a natural force, but also a possessor, an instrumental, and even a location, a time and means are treated

like an agentive nominal, while many languages have not extended the agent patterns to some of these semantic roles. Indeed many languages of the world do not treat the following subjects in the same way as English which aligns them with the agentive subject:

(5) a. David killed Goliath
b. John saw Bill
c. I like beer
d. The recent earthquake killed 150 persons
e. John has three children
f. This key won't open the door
g. This tent sleeps five
h. The 1930's saw the world at the brink of disaster
i. £100,000 won't buy a decent flat in London anymore

Another manifestation of subject prominence in English is the use of expletive elements, *it* and *there*, in ambient expressions, existential expressions, and passives, where many other languages do without having them in subject position.

(6) a. It rains
b. There is a boy in the garden
c. It is believed that John is honest

Languages are thus different with respect to the status of the subject category. This is easy to see if one compares English and other languages. For example, in Spanish and German many of the subject nominals in (5) are not treated like an agentive subject. The Spanish counterpart of (5c) has the experiencer in the dative: *Me gusta la cerveza*, and the direct translation of (5g) in German results in an ungrammatical sentence: **Dieses Zelt schläft fünf*. Also German does not require an expletive subject in the context where the V2 requirement is satisfied: e.g. *Gestern wurde getanzt* 'There was dancing yesterday'. These facts show that the category of subject in English is much more firmly established than the German counterpart. That is, English has grammaticized the subject category to the full extent, such that, on the one hand, many different semantic roles can be chosen as subjects, and, on the other, every (tensed) clause requires a subject.

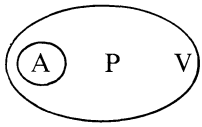
The subjects of intransitive clauses {S}, which include both agentive and patient participants, are also treated like an agentive subject of a transitive clause in English and many other languages. But again, certain languages do not generalize here, making a distinction between agent nominals and patient

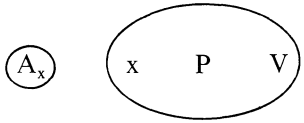
nominals of intransitive clauses. The so-called active-type languages are a case in point, but again even a familiar language like German treats an experiencer of certain verbs differently from an agentive subject; e.g. *Mich hungert* 'I am hungry,' *Mich friert* 'I am freezing'.

Thus, our view on the subject is that: (a) it is a syntactic category resulting from the generalization of an agent over other semantic roles, (b) languages vary as to how far this generalization has taken place; i.e. the grammatical status of subject differs from one language to another, and therefore, (c) the subject is not necessarily a universal category.[6]

For those languages that grammaticize the subject category, syntactic asymmetry is observed between subject and other non-subject nominals such that the members of the agent-based subject category {S, A} have syntactic primacy over other nominal categories.

Whether it is A or S, a subject is an argument of a lexical predicate, and its presence and semantic properties are determined by the predicate. However, this is not the case with the topic, which is a category established by additional cognitive processes as described in the preceding section.[7] Thus, while a subject is an obligatory sentence element (for most predicates), a topic is not. As the schematic representation below makes clear, whereas a subject is related to a lexical predicate as its argument, a topic is only indirectly related to a lexical predicate; indeed, for a topic the real predicate is the "remaining clause," from which the notion of topic being what is being talked about obtains. It is this indirect connection between a topic and a lexical predicate that permits certain unique expressions seen in topic-prominent languages. Thus, the double-"subject" constructions exemplified in (7c) and "illogical" equation sentences of the (7d)-type exploit the basic properties of the topic construction. The semantico-syntactically defined subject is uniquely determined by the lexical predicate, and as such only one such subject is permitted per clause, and a semantic connection between subject and predicate is lexically determined, as opposed to the looser topic-comment connection that can be pragmatically determined, as in (7d). The diagram in (7) schematically summarizes the ways in which the categories of subject and topic are defined:

(7) a. Subject 

b. Topic 

c. Double "subject" constructions
Japanese
Zoo wa hana ga nagai.
elephant TOP nose NOM long
'An elephant has a long trunk.'

Chinese
Dā xiàng bízi cháng.
big elephant nose long
'An elephant has a long trunk.'

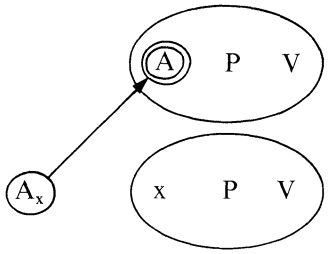
d. "Illogical" equation sentences
Japanese
Boku wa unagi da.
I TOP eel COP
'(Lit.) I am an eel.'
'(In a restaurant) I will order an eel (dish).'

Chinese
Tā shì ge Rìběn nǚrén.
he be CL Japanese woman
'(Lit.) He is a Japanese woman.'
'(In an appropriate context) His servant is a Japanese woman.'
(Chao, 1968:71)

In languages like English, in which a subject and a topic are not formally distinguished, the notional definitions of subject as an actor and of topic as "what is being talked about" have been appropriated to the "subject". This confounding situation is avoidable in principle in languages like Japanese, in which the two categories are distinguished. But here too, a typical situation of A's being chosen as a topic, as in (7b), have confused many researchers. At any rate, it is now clear that the definition of a prototypical subject as the intersection of agent and topic (Keenan, 1976; Comrie, 1981) represents a typical situation (in accusative languages) in which an agent is chosen as a topic, as in Japanese, or in which the categories of agentive subject and topic have merged, as in English.

In the subsequent discussion, we will be viewing the situation in English as an instance of the grammaticization of a topic into a subject, which can be schematically represented as in (8), where the arrow indicates the process of integration of a topic as a clausal argument overlapping the semantico-syntactically defined subject, namely the A argument.

(8)



3. FROM TOPIC TO SUBJECT

In order to ascertain the diachronic possibility of the change of a topic into a subject as envisioned above, we need to find a language in which the change in question is in progress. Our contention is that Philippine languages provide a transitional stage where a grammatical topic is being grammaticized (or reanalyzed) as a subject. The object of our inquiry is the status of the so-called topic constituent in Philippine languages that are typically marked by the prepositional marker *ang* (for common nouns) and *si* (for personal names) as in Tagalog and Cebuano. For our purpose, we must show, on the one hand, that the topic constituent in these languages retains sufficient characteristics of a grammatical topic; e.g. that it shares a number of important characteristics with the Japanese grammatical topic, and, on the other, that the Philippine topic has acquired a number of characteristic subject properties exhibited by the semantico-syntactically defined subjects {A, S}. Our data primarily come from Cebuano, which boasts the largest number of native speakers among Philippine languages, while Tagalog and other Philippine (-type) languages are occasionally referred to.

3.1. Philippine topics as grammatical topics

A first indication that the so-called topic constituent in Philippine languages is, or has origins in, a grammatical topic is its wide possibility of selection. Whereas the subjectivization via passive derivation severely limits the scope of its application to a direct object and in some languages to a direct object and an indirect object, the Philippine topic can be any nominal constituent within the governing domain of a predicate. Thus, not only goals (objects) and recipients, but also locations, directions, as well as instrumentals can be realized as a topic. The following is by now a familiar paradigm, where the four classes of verbal marking are known as: Actor focus (AF) in

(9a), Goal focus (GF) in (9b), Directional focus (DF) in (9c-g), and Instrumental focus (IF) in (9h). (Unless otherwise indicated, all examples are from the dialect of Cebuano spoken in Negros Oriental. The word in the translation corresponding to the topic form is indicated by the boldface.)

(9) a. *Ni-hatag si Juan ug libro sa bata.*
AF-give TOP GOAL book RECIP child
'**Juan** gave a book to the child.'

b. *Gi-hatag ni Juan ang libro sa bata.*
GF-give ACT TOP book RECIP child
'Juan gave **the book** to the child.'

c. *Gi-hatag-an ni Juan ug libro ang bata.*
D-give-F ACT GOAL book TOP child
'Juan gave the book to **the child**.'

d. *Gi-palit-an ni Juan si Maria ug libro.*
D-buy-F ACT TOP GOAL book
'Juan bought a book for **Maria**.'

e. *Gi-adtu ni Juan ang estasionan.*
DF-go ACT TOP station
'Juan went to **the station**.'

f. *Gi-pritu-han ni Juan ug isda ang kaha.*
D-fry-F ACT GOAL fish TOP pan
'Juan fried a fish in **the pan**.'

g. *Gi-duan ni Juan ang bata.*
DF-play ACT TOP child
'Juan played with **the child**.'

h. *I-hiwa ang kostilyo ug mangga ni Juan.*
IF-cut TOP knife GOAL mango ACT
'Juan will cut a mango with **the knife**.'

This wide variety of the argument type with which the Philippine topic can be construed is a major characteristic that is shared by the topic in Japanese and, as noted above, it contrasts sharply with the subjectivization accompanying the passive conversion. Topicalization differs from passivization in another respect. Whereas passivization is an agent defocusing mechanism (see Shibatani, 1985), topicalization does not defocus agent — an agent

remains an integral constituent in the patient (Goal) and other oblique topic constructions. Again, Philippine topicalization and Japanese topicalization show parallelism with regard to this characteristic, though Philippine languages, unlike Japanese, do not have an independent passivization process.

Secondly, Philippine topics are selected from among referential nominals; those indefinite, unmentioned nominals will not be chosen as topics (but see below). Thus, (9a) and (9b) above contrast with regard to the referential status of the object; *ang libro* in (9b) must be translated as "the book', whereas the nontopic *ug libro* in (9a) has the indefinite goal marker *ug*. The referentiality is an important property of the topic in Japanese as well; the topic has a specific reference that has been established linguistically or extralinguistically. This property contrasts with the subject, which can be indefinite as in Japanese and English.

A third consideration that leads us to believe that Philippine topics originate from grammatical topics has to do with the distribution of the topic sentences. As detailed in section 2, the basic function of a topic construction is to make an experiential judgment about a particular entity; as such it occurs only as a main clause. The Japanese topic construction thus does not occur in the nominalized clause, nor does it occur in subordinate clauses that introduce background information. For example, the forms in (10) with a topic are all ill-formed:

(10) a. [*Taroo* **wa*/*ga* *kita*] *koto*
TOP/NOM came that
'that [Taro came]'

b. [*Taroo* **wa*/*ga* *kita*] *toki*
TOP/NOM came time
'when [Taro came]'

c. [*Taro* **wa*/*ga* *kuru*] *nara*
TOP/NOM come if
'if [Taro comes]'

As also pointed out in section 2, reporting of a newly discovered event or state and simple recounting of events are done in topicless sentences.

Philippine topic sentences follow these distributional restrictions to a certain extent. First, while topic sentences occur as subordinate clauses of many types, clausal nominalization excludes a topic constituent, as shown in (11).

(11) a. *pag-dagan *si/ni Juan*
NOMI-run TOP/GEN
'running by Juan'

b. *pag-hatag *si/ni Juan ug libro sa bata* (cf. 9a)
NOMI-give TOP/GEN GOAL book DIR child
'giving of a book to the child by Juan'

The prefix *pag-* also introduces a subordinate clause with the meaning of 'when...' and in this use it also excludes a topic as in (12).

(12) *pag-abut *aku/naku sa balay, ...*
NOMI-arrive I(TOP)/I(GEN) at home
'when I arrived home, ...'

The nominalized clauses formed by the prefix *pag(ka)-* or *ka-* are used in two other expressions; exclamatory expressions and the recently completed aspect—the latter in Cebuano takes the form of *mau pa(y)* followed by a nominalized clause.

(13) a. *Ka-gwapu *ang/sa imung balay!*
NOMI-beautiful TOP/GEN your house
'How beautiful your house is!'

b. *Ka-taas *si/ni Juan!*
NOMI-tall TOP/GEN
'How tall Juan is!'

c. *Ka-bait *si/ni Maria!* (Tagalog)
NOMI-kind TOP/GEN
'How kind Maria is!'

(14) a. *Ka-liligo lang *siya/niya.* (Tagalog)
NOMI-bath just he(TOP)/he(GEN)
'He has just bathed.'

b. *Mau pay pag-abut *ang/sa manga bisita.*
be still NOMI-arrive TOP/GEN PL visitor
'The visitors have just arrived.'

Both these expressions are of the type that presents newly discovered events or states, for which the topic constructions are not used in Japanese either. The restrictions on the occurrence of topic constructions in other

Philippine(-type) languages also show a certain degree of parallelism to those imposed on the Japanese topic construction. Lee (1964) reports that, in Maguindanao, a topicless form marked by the prefix *ka-* is restricted to subordinate clauses. Whereas topic sentences occur as both main and subordinate clauses in general, they too appear to be restricted in distribution in certain types of subordinate clauses. Thus, a clause introduced by *endai* 'as soon as,' allows only topicless forms.

In Melanau, a Philippine-type language spoken along the coastal belt of Sarawak in Borneo, what Clayre (1973) calls the Verb Focus form contains no topic, whereas regular clauses have a topic, which is identified in the verbal marking in a manner analogous to Philippine languages. According to Clayre, in the Verb Focus sentence the action is what is being talked about. In narrative such a topicless sentence appears typically in a series of rapid actions. Again, this is reminiscent of the use of topicless sentences in Japanese in recounting a series of discrete events (see section 2). Thus, Philippine (-type) topics are avoided in those contexts in which the Japanese topic too is generally avoided.

We have now seen that Philippine topics and the Japanese topic show similarities in three areas: 1. wide possibility of topic selection, 2. the referential status of the topic nominal, and 3. the distributional restrictions on topic sentences. These similarities can be explained naturally under the assumption that Philippine topics were once genuine grammatical topics with the functions similar to those of the Japanese topic. We shall now examine those properties of Philippine topics that lead us to believe that they have been grammaticized into subjects to a considerable extent.

3.2. Philippine topics as subjects

One of the clear indications for the conclusion that Philippine topics have been far more grammaticized than the Japanese topic is that they have become an obligatory constituent of the major sentence types. As pointed out above, there are clausal types in which grammatical topics do not appear among Philippine languages (see (11)–(14)), and their distribution reveals their original function, which is still maintained by the Japanese topic. However, the distribution of Philippine topics contrasts markedly with that of the Japanese topic, which is optional in the sense that there are a large number of main clauses that do not have topic constituents.

In Philippine languages, in addition to the nominalized clauses, which are used for exclamatory expressions and in the recently completed aspect,

two additional types of clauses do not have a topic constituent. One is meteorological expressions and the other is the existential construction, shown in (15):

(15) a. *Nag-ulan na.*
AF-rain now
'It's raining now.'

b. *Naay maayong libro sa tindahan.*
exist good book at store
'There is a good book at the store.'

Other than these types of clauses, all Philippine clauses must have a topic constituent, though both topic and nontopic constituents under appropriate discourse contexts may not actually surface. This obligatory status of Philippine topics is a first indication that topic is becoming grammaticized as subject — a nominal constituent that is obligatory in a subject-prominent language like English.

In some Philippine languages, grammaticization of topic as subject seems to cause pressure to have a topic even where a topic is not normally expected. In the beginning of a folktale, for example, a non-referential actor is first introduced by the existential clause, which, as noted above, does not include a topic constituent. Thus:

(16) *Diha kadtoy usa ka bata nga ginganlag Juan Pusog.*
exist once one child LN named
'Once there was a child named Juan Pusog.'
(LN = linker for a subordinate clause)

Here, the non-referential *usa ka bata* 'one child' is not marked by *ang*, and the clause contains no topic. However, often enough a non-referential actor is topicalized directly. Thus, another folktale begins with the form in (17), which contains the topic form of *usa ka magtiyaon* 'one couple'.

(17) *Nanganak kadto ang usa ka magitayon ug bata nga walay*
give birth once TOP one couple GOAL child LN without
kamot ug tiil.
hand and leg
'Once **a couple** gave birth to a child without limbs.'

Among the eight tales in *Cebuano Folktales 2* (University of San Carlos, Cebu City, 1977), four began with nontopic existential sentences, three with

sentences involving non-referential topics, and one with a clause containing a proper name as the topic. Notice that the occurrence of non-referential topics is not confined to the beginning of a folktale. Quantified non-referential nominals functioning as topic seem to be fairly common in Cebuano, e.g. (18a), though some non-quantified, non-referential topics are also observed as in (18).

(18) a. *Nalabyan niya ang usa ka mananangot.*
pass by he(ACT) TOP one tuba gatherer
'He passed by **a tuba gatherer.**'

b. ... *iyang na-kita ang usa ka sinaw nga awa ...*
he GF-see TOP one shiny LN fish
'He saw **a shiny fish.**'

c. *Mipadayon pagligidligid si Kapuroy ug nasugatan niya*
continue rolling TOP and meet he(ACT)
ang manunukdok ug gapas
TOP worker GOAL cotton
ug ang mamumugan ug mais.
and TOP planter GOAL corn
'Kaproy continued rolling and happened to meet **a cotton worker** and **a corn planter**.'

Thus, the grammaticization of topic as an obligatory clause element — subject — has had the effect of lifting the original requirement that the topic be referential. This semantic weakening or dissociation from a pragmatic constraint and the acquisition of a more solid grammatical status (in the sense that the topic constituent has become an obligatory clausal element for the major clause types) show that Philippine topics have been undergoing a typical grammaticization process.

The obligatory nature of Philippine topics is not the only evidence that indicates the change of their grammatical status from topic to subject. A genuine grammatical topic like the Japanese topic does not participate in syntactic processes as a grammatical subject does. In the case of Japanese, for example, a topic does not control such phenomena as reflexivization, subject honorification, quantifier float, binding of a gap in a coordinated clause, and binding of a gap in a subordinate clause that are all controlled by a subject. Only those topics that bind a subject gap (or a variable in subject position) control these phenomena in disguise. Philippine topics show a

marked contrast in this respect in that they actively participate in syntactic phenomena. Indeed, Philippine topics control some of those phenomena that are typically controlled by a subject in other languages, indicating that they behave like a subject even syntactically.

Schachter (1976; 1977) has pointed out that those phenomena that are controlled by subjects in many other languages are divided into two categories in such a way that one group is controlled by an Actor nominal and the other by a topic nominal. A closer examination reveals that in fact three groups of phenomena obtain: a. those that are controlled by an Actor, b. those controlled by both Actor and topic, and c. those that are controlled by a topic. They are summarized in (19) (see Schachter, 1976; 1977; and Shibatani, 1988 for details):

(19) a. *Actor-controlled phenomena*
Reflexive binding
Controlling of an Actor gap in the second coordinated clause
Deletion in imperatives

b. *Phenomena controlled by both Actor and topic*
Gap in the second coordinated clause
Controlling of a gap in a subordinate clause
Gap in a subordinate clause
Topicalization into sentence initial position

c. *Topic-controlled phenomena*
Relativization
Focus construction
(Focussed) Wh-questions (see below)
Quantifier float
Controlling of a gap in the *samtang* 'while'-clause
Gap in the *samtang*-clause
Raising out of the *nga* 'that' subordinate clause

Notice that since an Actor-topic is the intersection of Actor and topic, it controls all these phenomena. Those that are categorized under the Actor-controlled phenomena refer to those that are controlled by a non-topic Actor, and those under the topic-controlled phenomena are the ones controlled by a topic irrespective of its semantic role.

Although the phenomena controlled by the three groups of nominals are detailed in Shibatani (1988), some illustrations here might be helpful for those unfamiliar with the Philippine situation. The examples in (20) and (21)

illustrate situations (19a) and (19b), respectively. For situation (19c), see (30)–(35).

(20) Actor-controlled phenomena
'Reflexive binding'
a. *Ni-ingnan si Juan ni Pedro tongod sa iyang kaugalingon.*
AF-talk TOP/ACT GOAL about OBL his self
'**Juan**$_i$ talked to Pedro about himself$_i$.

b. *Gi-ingnan ni Juan si Pedro tongod sa iyang kaugalingon.*
GF-talk ACT TOP/GOAL about OBL his self
'Juan$_i$ talked to **Pedro** about himself$_i$.'
(Non-Actor topic does not bind the reflexive form.)

'Controlling of an Actor gap'
a. *Ni-bunal si Juan ni Pedro ug ni-lakaw Ø.*
AF-hit TOP/ACT DIR and AF-left
'**Juan** hit Pedro and left.' (Ø = Juan)

b. *Gi-bunal-an ni Juan si Pedro ug ni-lakaw Ø.*
D-hit-F ACT TOP/DIR and AF-leave
'Juan hit **Pedro** and left.' (Ø = Juan)
(Non-Actor topic does not control the Actor gap in the second coordinated clause, whereas non-topic Actor controls the gap.)

'Imperative deletion'
a. *Hatag ug libro ni Juan.*
give GOAL book DIR
'Give a book to Juan.' (Deletion of Actor-topic)

b. *I-hatag ang libro ni Juan.*
IF-give TOP book DIR
'Give **the book** to Juan.' (Deletion of non-topic Actor)

c. *Hatag-an ug libro si Juan.*
give-D GOAL book TOP/DIR
'Give a book to **Juan**.' (Deletion of non-topic Actor)

(21) Phenomena controlled by both Actor and topic
'Gap in a subordinate clause'
a. *Gusto ni Juan nga mu-tutdlo Ø Maria.*
want ACT LN AF-teach
Lit. 'Juan desires that Ø teach Maria.'
(Actor-topic gap; Ø = Juan)

b. *Gusto ni Juan nga tutdlo-an Ø si Maria.*
want ACT LN teach-GF TOP
Lit. 'Juan desires that Ø teach **Maria**.'
(Non-topic Actor gap; Ø = Juan)

c. *Gusto ni Juan nga mu-tutdlo Ø si Maestro.*
want ACT LN AF-teach TOP teacher
Lit. 'Juan desires that **the teacher** teach Ø.'
(Non-topic, non-Actor gap; not controllable, Ø = someone)

d. *Gusto ni Juan nga tutdlo-an Ø sa Maestro.*
want ACT LN teach-GF ACT teacher
Lit. 'Juan desires that the teacher teach Ø.'
(Non-Actor topic gap; Ø = Juan)
'Topicalization into sentence initial position' (See (22)–(23))

Now, assuming that Schachter (1977) is correct in separating subject properties into those that are reference-related and those that are role-related, the above situation in Cebuano can be understood as follows. As a topic is integrated into a clause and begins to acquire the subject status, it begins to assert its subject status by first involving itself with phenomena in which the referential status of a particular nominal is in question. This is only natural since a topic itself is a referential nominal. Relativization, focus construction, and focused Wh-questions are thus first phenomena in which a topic qua grammaticized subject gets itself involved (see section 7 below for a more detailed discussion of this point). As a topic continues to assert itself as a grammatical subject, it encroaches on those phenomena that are controlled by the erstwhile subject — Actor (agent). When the process is in progress we expect to find an area in which the original Actor subject retains its superiority over other nominals. The phenomena in (19b) constitute an area of clash, and (19a) a domain where the Actor is still in control.

The above scenario describes the Philippine situation as reaching an advanced stage of the grammaticization process of turning a topic into a subject. The plausibility of the scenario is shown by the fact that Cebuano and other Philippine languages have developed new topic constructions related but distinct from the ones that we have been discussing. Thus, Philippine languages exhibit the full cycle of the change; the change of a topic into a subject and the rise of a new grammatical topic again.

The new topic construction places a topic in sentence initial position. Such a topic is formally identical with the sentence internal topic in many

languages, but in Tagalog it is followed by the particle *ay*. The first example in (22) is from *Cebuano Folktales 2*, and the others from linguistic papers.

(22) a. Cebuano
Si Maria mi-duol sa babaye ug mi-tabang pagdala
TOP AF-approach DIR woman and AF-help carrying
sa bantal.
GOAL bundle
'**Maria** approached the woman and helped carrying the bundle.'

b. Tagalog
Ang sulat ay tinanggap ko kahapon.
TOP letter arrive I yesterday
'**The letter** arrived yesterday.'

c. Kapampangan (Mirikitani, 1972)
Ing lalaki, seli ne(na+ya) ing mangga king tindahan.
TOP boy buy he+it TOP mango at store
'The boy, he bought **the mango** at the store.'

The initial *ang*/*si*-marked topic in Cebuano and Tagalog are obtained by shifting the clause internal topic into initial position. However, a non-topic actor may be made an initial topic, as in (22c). The Kapampangan form contains two topics — the initial *ing lalaki* 'the boy' and the internal *ing mangga* 'the mango'. A similar pattern is observed in Cebuano as well, leading to the conclusion that the initial topic is accessible to both internal topic slot and non-topic actor slot. Furthermore, the initial topic construction allows topic sentences unique to topic-prominent languages such as Japanese, Korean, and Chinese (see section 2.2). Namely, the initial topic need not be an argument of the lexical predicate. Thus, the examples in (23), which involve extra topic constituents, are also possible.

(23) a. Cebuano
Si Juan dako ang iyang ulu.
TOP big TOP his head
'Juan, **his head** is big.'

b. Tagalog
Si Pedro ay malaki ang bahay.
TOP big TOP house
'Pedro, **his house** is big.'

Though the forms like (23) are among the favorite topic constructions in Japanese and other topic-prominent languages, the status of the initial topic construction in Philippine languages does not seem to be well established yet, and its grammatical role is not clear at the moment (but see Shibatani, 1988). That it is a new and marked construction is indicated by the Schachter-Otanes characterization that it is found in formal style and is "more common in writing, lectures, sermons, etc. than ... in ordinary conversation." (1972:485) My survey of Cebuano data indicates that the Schachter-Otanes characterization seems to be generally true. While there is a large number of sentence initial topic forms in *Cebuano Folktales 2*, conversational data contain a far smaller amount.

The fledgling status and the uncertainty of the role of the initial topic construction notwithstanding, it is clear that Philippine languages have developed a genuine topic construction anew. Our contention is that it was motivated by the grammaticization of the old topic form. We shall now turn to the question of topic selection pattern. We consider difference in the pattern of topic selection to be a crucial factor that determines whether a given language shows a consistent accusative syntax like English and Japanese or a split syntax like Philippine languages and some ergative languages.

4. TOPIC SELECTION PATTERNS

Philippine topics differ from the Japanese topic in yet another respect. In Japanese, an agent is a favorite choice for a topic, and choosing a patient topic requires a far more elaborate context. This is not so among Philippine languages; a patient (or Goal) is preferably chosen as a topic. Thus, the preferred pattern of topic selection is accusative, {S, A}, in Japanese, Korean, and Chinese, while it is ergative, {S, P}, in Philippine languages. The contrast can be seen in the translations of the made-up passage in (24):

(24) a. Jiji and Annie were good friends
b. Annie's birthday was approaching
c. Annie always liked to read
d. She was very happy whenever she got books as presents
e. There was a book that Annie liked at the store
f. So, Jiji gave the book to Annie

The most natural translations of this passage will be (25a–f) in Japanese and (26a–f) in Cebuano:

(25) **Japanese**

a. *Jiji to Annie wa nakayosi no tomodati desita*
b. *Annie no tanzyoobi ga tikazuite imasita*
c. *Annie wa itumo hon o yomu no ga suki desita*
d. *Purezento ni hon o morauto kanozyo wa itumo yorokobimasita*
e. *Annie no suki na hon ga mise ni arimasita*
f. *Sokode, Jiji wa Annie ni sono hon o agemasita*
'So, **Jiji** gave the book to Annie.'
f′. ??*Sokode, sono hon wa Jiji ga Annie ni agemasita*
'So, the book, Jiji gave to Annie.'

(26) **Cebuano**

a. *Si Jiji ug si Annie maayong mga higala*
b. *Ang adlaw nga tanaohan ni Annie muabot*
c. *Si Annie gustong magsigi basa ug libro*
d. *Malipayon siya kun marigalohan ug libro*
e. *Dinhay libro sa tindahan nga nagustohan ni Annie*
f. *Busa, gihatag ni Jiji ang libro sa Annie*
'So, Jiji gave **the book** to Annie.'
f′. *Busa, nihatag si Jiji ni Annie sa libro*
'So, **Jiji** gave the book to Annie.'

In the Japanese translation, (25f), with the agent topic, forms quite a natural sequence, despite the fact that both book and Jiji have referentially prominent status and even though "the book" is made a local discourse topic by the preceding sentence. On the other hand, (25f′) is quite infelicitous, just as its English translation (or the passive sentence *So, the book was given to Annie by Jiji*) would be if it replaced (24f).

The Philippine situation differs markedly from the situations in Japanese (and English) in that it is most likely that Goal-topic form (26f) will be chosen over Actor-topic form (26f′). There is some variation across different languages. In Cebuano, either (26f) or (26f′) will be equally possible, though the former is more favored — it is the version that appears spontaneously in translation. In Tagalog, however, only (26f) is permissible. That is, across Philippine languages a Goal-topic is favored when both Actor and Goal are referential, but in some languages, e.g. Tagalog, the choice of a Goal-topic is made obligatory under the same circumstance. Rules of thumb for topic selection among Philippine languages that capture this situation have been

summarized as in (27) in Shibatani (1988) (cf. Naylor, 1975; McFarland, 1978):

(27) *Rules of thumb for topic selection in Philippine languages*
(i) if there is only one referential nominal, then make it a topic,
(ii) if both Actor (agent) and Goal (patient) are referential, opt for the Goal-topic, and
(iii)if the choice is between referential Actor and referential Directional, then give preference to the Actor-topic.

The variation between Cebuano and Tagalog pertain to Rule (ii); in Cebuano the rule is optional, whereas it is obligatory in Tagalog. As noted above, there is some irregularity such that even a non-referential Goal is sometimes chosen as a topic. But it is this rule that is responsible for the difference in the pattern of topic selection between Japanese (Korean, and Chinese) and Philippine languages. The preference of an agent(Actor)-topic in the former group of languages and the preference of a Goal(patient)-topic in the latter are statistically reflected in the occurrence of the respective type of topic.

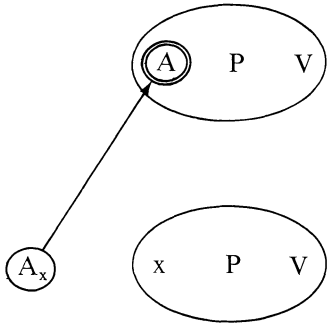
The survey (1964) by the National Language Institute in Tokyo on modern written Japanese indicates that, overall, almost 56 percent of subjects (Actor) are topicalized (i.e. marked by the topic marker *wa*), while only 10 percent of the objects (Goal) are topicalized. According to my small count of 71 topic-containing transitive clauses in written material, 64 (90%) were agent(Actor)-topic, and 7 (10%) were patient(Goal)-topic. Philippine languages show stronger overall inclination toward the Goal-topic. Cooreman, Fox and Givón's (1984) count indicate that 76% of the transitive clauses in their Tagalog corpus were Goal-topic sentences, while my own count shows that 46% of the transitive clauses of the Cebuano corpus were Goal-topic sentences. The difference here seems to be due to the fact that, as noted above, Tagalog applies the second rule of topic selection more stringently than Cebuano.

5. GRAMMATICIZATION PATTERNS AND THE RISE OF SPLIT SYNTAX

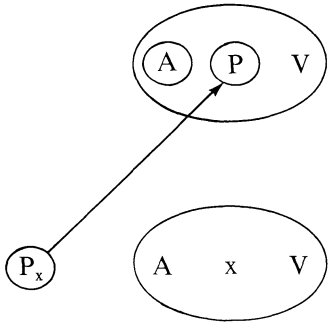
We have observed above that topic selection patterns differ between Japanese (Korean, and Chinese) and Philippine languages. Specifically, while the former group of languages show the accusative pattern, {S, A}, the latter

show the ergative pattern {S, P}. We shall now examine the consequence of this difference in the process of grammaticization of topics into subjects. When a language prefers the accusative pattern of topic selection, as in Japanese, the topic and the (semantico-syntactically defined) subject converge on the same nominals, namely {S, A}. However, when a language prefers the ergative pattern of topic selection like Philippine languages, the topic and the subject are realized in different arguments of a transitive clause, the subject being {S, A} and the topic selection displaying the ergative pattern of {S, P}. When these two different types of language grammaticize topics into subjects, there arises a profound difference in the way syntax is organized. This is caused by the fact that in the consistently accusative language, there will be only one subject, whereas in the language that subjectivizes a patient topic, there will be two competing subjects, as indicated by the schematic representation of the two grammaticization patterns in (28):

(28) a. Accusative grammaticization pattern



b. Ergative grammaticization pattern



The syntactic consequence of the difference in the above grammaticization patterns is that whereas the agent-topic language will maintain a consistent syntax centering around A, the patient-topic language may give rise to split syntax in such a way that while certain syntactic phenomena remain controlled by A, some others yield to the newly grammaticized patient topic. This is exactly what has happened in the Philippine languages. When an Actor (agent) nominal is chosen as a topic, the syntax is consistently controlled by the Actor-topic, for this is a situation in which the semantico-syntactically defined subject and the grammaticized topic qua subject converge on an agentive nominal. However, when the Goal-topic occurs, the syntax is split along the line of the division of phenomena, as summarized in (19), and results in a situation where an issue of identifying the subject arises (Schachter, 1976). Our contention is that split syntax in ergative languages arises in the same manner.

6. SPLIT SYNTAX IN ERGATIVE LANGUAGES

While the idea that the absolutive nominal in an ergative language represents a grammaticized topic is fairly widely entertained (cf. Wierzbicka, 1981; Mallinson and Blake, 1981), evidence for it is far more indirect than the case of Philippine languages, where the so-called topic nominal is more firmly correlated with the referential status of the nominal in question, and where the distribution of the topic sentence shows a similar pattern that is displayed by the prototypical grammatical topic seen in Japanese (recall the earlier discussion on this). Although the topic role of the absolutive nominal is often discussed in terms of the general notion of topicality and its correlation with the split case-marking (Wierzbicka, 1981), the notion of topicality defined along the lines of Givón's hierarchy or Silverstein's seems to be too broad a concept for our purpose. While the realization of a grammatical topic is broadly correlated with these hierarchies, e.g. 1st person agent is the favorite topic choice in Japanese, evidence from Japanese indicates that a grammatical topic is correlated with a much narrower, locally determined concept; namely that of the referentiality of a nominal constituent. Thus, as shown by Clancy and Downing (1987), the topic phrase in Japanese does not correlate with either the nominals with high topicality or a discourse topic in a direct manner. In dealing with ergative languages, therefore, we must be careful; rather than correlating the absolutive with the general notion of topicality, for which all languages appear to align an agentive nominal, we must seek a

narrower pragmatic correlation of the absolutive, if we are to establish a connection between the absolutive nominal and a grammatical topic.

In most ergative languages, a topic is far more grammaticized in the sense that even an indefinite, non-referential nominal is put in the absolute case, just as the English subject is hospitable to a non-referential nominal as well. However, certain languages show a far stronger correlation between the absolutive nominal and referentiality. A case in point is Eskimo, in which the ergative-antipassive alternation appears to be strongly correlated with the referentiality of the patient nominal in question. Kalmar (1979:118) shows the contrast in (29), where the definite patient is realized in the absolutive and the indefinite patient in the oblique *-mik* form:

(29) a. Ergative clause
inu-up *qimmiq-Ø* *taku-v-a-a*
person-ERG dog-ABS see-IND-POLYP-3/3
'A/the person saw **the dog**.'

b. Antipassive clause
inuk-Ø *qimmiq-mik* *taku-v-uq-Ø*
person-ABS dog-OBL see-IND-MONOP-3
'A/the person saw **a dog**.'

Though the correlation between the absolutive case and the referentiality of the nominal is not absolute (since a non-referential agent can occur in either the ergative or the absolutive case), such a correlation seems to hold generally with regard to a patient nominal. This is reminiscent of the pattern of the patient topic in Philippine languages — when a patient is referential, it is preferably chosen as a topic.

Even in those languages in which non-referential patient nominals freely occur in the absolutive, some vestige of the correlation between the absolutive and referentiality is seen. Thus, Dixon (1980:447) notes that in Dyirbal, a language which has subjectivized the absolutive patient to a high degree, the highly referentially prominent demonstrative determiner *giy-ginya* 'this' occurs only in absolutive case form. In order for this form to qualify an agentive nominal, which is normally realized in the ergative, antipassive conversion must apply so that the agentive nominal will be put in the absolutive case.

While these correlations support the view that the absolutive in ergative languages has (had) associated with it the same kind of pragmatic conditions that is associated with a grammatical topic, perhaps the strongest evidence

for the claim that in ergative languages a grammatical topic converges on the patient nominal in the manner of grammaticization depicted in (28b) above derives from the pattern of split syntax observed among ergative languages.

Contrary to the claim that ergativity is "a comparatively trivial fact about morphology" (Anderson, 1976:17), a number of languages have been shown to organize syntax partially according to the ergative parameter, whereby S and P are treated alike to the exclusion of A. The least equivocal are two language groups: the Australian group, for which Dyirbal (Dixon, 1972) and Yidiny (Dixon, 1977) are most thoroughly investigated, and the Mayan group, which includes Jacaltec (Craig, 1977), Quiche (Larsen and Norman, 1979), and Mam (England, 1983). All these languages exhibit split syntax, for in all of them there are certain syntactic phenomena, e.g. imperatives, reflexives, that operate accusatively, i.e. A rather than P is involved. Indeed, ergative syntax obtains to a lesser extent in Mayan languages and perhaps also in Yidiny. What interests us most are those phenomena that operate ergatively. Cross-linguistic comparison reveals that certain phenomena are more likely to organize ergatively than others. Mayan languages are characterized by organizing the three phenomena of relativization, Wh-question formation, and clefting according to the ergative parameter. The favorite ergative phenomena in Dyirbal and Yidiny include relativization and some other subordinate clauses, which involve deletion of a coreferential nominal of the subordinate clause.

Now, it so happens that relativization, Wh-question formation, and clefting are among those phenomena that crucially refer to the topic nominal in Philippine languages. We have mentioned earlier that the fact that the topic is involved in these phenomena has to do with the referential status of the topic nominal. Indeed, a few linguists (Kuno, 1976; Schachter, 1976; 1977; and Bresnan and Mchombo, 1987) have argued that topics underlie relative pronouns or relativized nominals. Bresnan and Mchombo also show that clefting also involves topics. And then, certain languages involve a cleft construction for Wh-question formation, leading to a situation where Wh-question formation also refers to a topic. Let us first examine how the topic — at least the Philippine-style topic — is involved in these phenomena in Cebuano.

That only a topic can be relativized upon is shown by the ungrammaticality of (30c) and (31b):

(30) a. *Ni-basa ang bata ug libro.*
AF-read TOP child GOAL book
'**The child** read the book.'

HEAD TOPIC GAP
b. ***ang bata*** *nga* [*ni-basa* *Ø ug* *libro*]
child LN AF-read GOAL book
'the child who read a book'

HEAD NONTOPIC GAP
c. ****ang libro*** *nga* [*ni-basa ang bata* *Ø*]
book LN AF-read TOP child
(intended as) 'the book the child read'

(31) a. *Gi-basa sa bata ang libro.*
GF-read ACT child TOP book
'The child read **the book**.'

HEAD NONTOPIC GAP
b. ****ang bata*** *nga* [*gi-basa Ø ang libro*]
child LN GF-read TOP book
(intended as) 'the child who read the book'

HEAD TOPIC GAP
c. ***ang libro*** *nga* [*gi-basa sa bata* *Ø*]
book LN GF-read ACT child
'the book the child read'

In (30c) the Goal nominal *ug libro* has been relativized without having become a topic. The grammatical counterpart is (31c), which has been derived from the relative clause in which the Goal nominal has been made a topic, as indicated by the GF prefix in the verb and by the absence of any other topic nominal in the clause. The relationship between grammatical (30b) and ungrammatical (31b) is the same.

Now consider the following examples of the cleft (or focus) construction:

FOCUS TOPIC GAP
(32) a. ***Ang bata*** *ang* [*ni-basa* *Ø ug* *libro*]
child AF-read GOAL book
'It is the child that read a book.
The one that read a book is the child.'

FOCUS NONTOPIC GAP
b. ****Ang libro*** *ang* [*ni-basa ang bata Ø*]
book AF-read TOP child
(intended as) 'It is the book that the child read.'

(33) a. **Ang bata ang* [gi-basa ang libro]
child GF-read TOP book
(intended as) 'It is the child that read the book.'

b. *Ang libro ang* [*gi-basa sa bata*]
book GF-read ACT child
'It is the book that the child read.'

Again, the nominal coreferential with the focus — the clause initial *ang* nominal — is deleted in the following (headless relative) clause, but it must be the topic. (32b) and (33a) are ungrammatical because the coreferential nominals deleted are not topics.

Wh-questions, at least those questioning the arguments, are built upon the cleft construction such that the questioned zero nominal within the headless relative clause that is bound by an interrogative pronoun must be the topic. Thus, Wh-question formation exhibits the same grammaticality pattern as clefting[8]:

FOCUS TOPIC GAP
(34) a. ***Kinsa*** *ba ang* [*ni-basa* Ø *ug libro*]?
who Q AF-read GOAL book
'Who read a book?/Who is it that read a book?/The one who read the book is who?'

FOCUS NONTOPIC GAP
b. ****Unsa*** *ba ang* [*ni-basa ang bata* Ø]?
what Q AF-read TOP child
(intended as) 'What did the child read?'

(35) a. **Kinsa ba ang*[*gi-basa ang libro*]?
who Q GF-read TOP book
(intended as) 'Who read the book?'

b. *Unsa ba ang* [*gi-basa sa bata*]?
what Q GF-read ACT child
'What did the child read?'

Exactly the same patterns obtain in Mayan languages with regard to the same phenomena. Only S and P, which trigger the absolutive verbal affixes, can be relativized upon straightforwardly; if A is to be relativized upon, it must be turned to a derived S via antipassivization. Thus in Quiche forms (36b) and (37b), involving absolutive S and P, are satisfactory, but (37c)

involving an ergative A is ill-formed; instead (37d) with a derived S is called for.

(36) a. *x-Ø-tzaaq-ik* *lee achih*
COMPL-3SABS-fall-PF the man
'The man fell.'

HEAD ABS GAP
b. *x-Ø-inw-il* ***lee achih*** [*x-Øtzaaq-ik Ø*]
COMPL-3SABS-1SERG-see the man COMPL-3SABS-fall-PF
'I saw the man who fell.'

(37) a. *x-Ø-u-ch'ay* *lee ixoq* *lee achih*
COMPL-3SABS-3SERG-hit the woman the man
'The man hit the woman.'

HEAD
b. *x-Ø-inw-il* ***lee ixoq***
COMPL-3SABS-1SERG-see the woman
ABS GAP
[*x-Ø-u-ch'ay Ø* *lee achih*]
COMPL-3SABS-3SERG-hit the man
'I saw the woman who the man hit.'

HEAD
c. **x-Ø-inw-il* ***lee achih***
COMPL-3SABS-1SERG-see the man
ERG GAP
[*x-Ø-u-ch'ay* *lee ixoq Ø*]
COMPL-3SABS-3SERG-hit the woman
'I saw the man who hit the woman.'
(O.K. with the meaning of 'I saw the man who the woman hit.')

HEAD
d. *x-Ø-inw-il* ***lee achih***
COMPL-3SABS-1SERG-see the man
ABS GAP
[*x-Ø-ch'ay-ow* *lee ixoq Ø*]
COMPL-3SABS-hit-ANTI the woman
'I saw the man who hit the woman.'

The cleft construction in Quiche parallels that of Philippine languages in that the clause following the focused nominal must contain a gap bound by the focused nominal that triggers the absolutive verbal affix. No ergative triggering gap can occur, as evidenced by the ill-formedness of (38c) below, which can be rendered grammatical by converting the gap into the absolutive triggering form via antipassivization:

FOCUS ABS GAP

(38) a. *aree* ***lee achih*** [*x-Ø-tzaaq-ik* *Ø*]
FOC the man COMPL-3SABS-fall-PF
'It was the man who fell.'

b. *aree lee ixoq* [*x-Ø-u-ch'ay* *lee achih*]
FOC the woman COMPL-3SABS-3SERG-hit the man
'It was the woman who the man hit.'

FOCUS ERG GAP

c. **aree* ***lee achih*** [*x-Ø-u-ch'ay* *lee ixoq Ø*]
FOC the man COMPL-3SABS-3SERG-hit the woman
'It was the man who hit the woman.'
(O.K. with the meaning of 'It was the man who the woman hit.')

d. *aree lee achih* [*x-Ø-ch'ay-ow* *lee ixoq*]
FOC the man COMPL-3SABS-hit-ANTI the woman
'It was the man who hit the woman.'

Wh-questions in Quiche are also built upon the cleft construction. They are derivable by replacing the focused nominals by interrogative pronouns; as such, they also require the absolutive gap in the clause following an interrogative pronoun. Thus, by replacing the focused pronouns in (38) by the interrogative pronoun *jachin* 'who', we obtain the paradigm in (39) that parallels the relativization and the clefting patterns:

FOCUS ABS GAP

(39) a. ***jachin*** [*x-Ø-tzaaq-ik* *Ø*]
who COMPL-3SABS-fall-PF
'Who fell?'

b. ***jachin*** [*x-Ø-u-ch'ay* *lee achih*]
who COMPL-3SABS-3SERG-hit the man
'Who did the man hit?'

FOCUS ERG GAP

c. ***jachin*** [*x-Ø-u-ch'ay* *lee ixoq Ø*]
who COMPL-3SABS-3SERG-hit the woman
'Who hit the woman?'
(O.K. with the meaning of 'Who did the woman hit?'

d. *jachin* [*x-Ø-ch'ay-ow* *lee ixoq*]
who COMPL-3SABS-hit-ANTI the woman
'Who hit the woman?'

This striking parallelism between the Philippine relativization, clefting, and Wh-questions and the Quiche counterparts strongly indicates that the Philippine topic and the Quiche absolutive are functionally equivalent, which in turn argues for the claim that the Quiche absolutive is a grammaticized topic.[9]

Among the Australian languages (e.g. Dyirbal, Yidiny, Kalkatungu) that show split syntax, relativization and analogous subordinate constructions make crucial reference to the absolutive gap in the subordinate clause (see Dixon, 1980). Just for the sake of completeness, let us examine the relativization pattern in Dyirbal. As the following data reveals, the ergative A gap of a relative clause cannot be properly bound by the relative head (40b); but once it is turned into the absolutive form via antipassivization (40c), relativization upon it will succeed (40d).

(40) a. Ergative construction
bayi *yurri* *banggul* *yarranggu* *bagan*
DET.ABS kangaroo.ABS DET.ERG man.ERG spear
'The man speared a kangaroo.'

HEAD ERG GAP

b. ****bayi*** ***yarra*** [*bagan-ngu bayi* *yurri* *Ø*]
DET.ABS man.ABS spear-REL DET.ABS kangaroo.ABS
banaganyu (reconstructed)
return
'The man who speared a kangaroo is returning.'

c. Antipassive construction
bayi *yarra* *bagalnganyu* *bagul* *yurrigu*
DET.ABS man.ABS spear.ANTI DET.DAT kangaroo-DAT
'The man speared a kangaroo.'

HEAD ABS GAP

d. *bayi yarra* [*bagal-nga-ngu bagul yurrigu Ø*]
DET.ABS man.ABS spear-ANTI-REL DET-DAT kangaroo.DAT
banaganyu
return
'The man who speared the kangaroo is returning.' (Dixon, 1972)

Among Australian languages Dyirbal has grammaticized the absolutive as a subject to an extremely high degree such that even the coordinate or paratactic subject deletion applies to the (absolutive) patient nominals. While this is a favorite example for showing ergative syntax, the phenomenon of cross-sentential coreference deletion seems to involve A in most languages following the pattern of topic continuity (DuBois, 1987). Indeed, among Australian languages that show partial ergative syntax, coordinate subject deletion generally operates on the accusative pattern, and even Dyirbal contains a fair amount of cross-clausal coreference deletion involving A (Tsunoda, 1985; Cooreman, Forthcoming). A comparative study above (cf. Van Valin, 1981) indicates that ergative syntax starts with (or clings to, depending on whether ergative syntax is expanding or receding) the constructions such as relativization and clefting that involve a referentially prominent nominal. Miyaoka (1986) confirms this, as in Yukip Eskimo only relativization shows any trace of ergative syntactic patterns in the otherwise accusative syntactic organization of the language. That the absolutive plays such a role of referentially prominent nominal in ergative languages with split syntax strongly points to the plausibility that a grammatical topic has converged on the patient nominal in ergative languages, just as the favorite pattern of convergence is between the topic and the Goal (patient) nominal in Philippine languages (see diagram (28b)).

7. SUMMARY

This paper started off with a discussion of the grammatical topic in Japanese, which shows the most stable topic construction, the early use of which is seen already in the writings of the early 8th century. There is sufficient similarity between the Japanese topic and Philippine topics to warrant the assumption that they are essentially of the same nature. However, the fact that Philippine topics show a number of characteristic subject properties

indicates that they have been grammaticized into subjects to a large extent. The syntactic phenomena controlled by the grammaticized topics include those that are controlled by the absolutive patient nominal (of a transitive clause) in those ergative languages that show syntactic split, indicating that Philippine topics and the absolutive in ergative languages show a functional similarity. The chain of reasoning leads us to conclude that the absolutive in at least some ergative languages is a grammaticized topic.

Whether a language shows a split syntax or not depends on the topic selection pattern and the subsequent grammaticization pattern of a topic. The grammaticization of a topic in those languages that favor agentive topic selection results in the convergence of a topic and a semantico-syntactically defined subject on the A nominal, and as a result no syntactic split arises. However, when those languages that favor patient topic selection grammaticize the topic, the A subject and the P topic compete for grammatical primacy. While in many languages the A subject is strong enough to retain its superiority over the P topic such that no syntactic split arises, in certain languages the grammaticized P topic manages to assert itself in the syntactic domain. When this latter situation arises, syntactic split will result.

Even those languages like English where the grammaticization of topics into subjects results in the convergence of these categories on an agentive nominal, there is a certain syntactic consequence. In this case, the topic effect of the subject may exert influence toward the avoidance of an indefinite subject, unfit for the topic. Thus, English tends to avoid an indefinite subject construction, e.g. *A boy is reading a book in the garden*, in favor of an existential expression, e.g. *There is a boy reading a book in the garden*. Chinese, which merges a topic and a subject when a definite nominal is involved, also shows this tendency, while in languages like Japanese and Korean, where the topic and the subject are clearly kept apart, there is no pressure for avoiding an indefinite subject. Thus, the grammaticization of topics into subjects has syntactic repercussions of one kind or another in all types of language.

ABBREVIATIONS

PRO	pronoun
CL	classifier
DF	directional focus
RECIP	recipient

IF	instrumental focus
NOMI	nominaliser
GF	goal focus
IND	indicative
POLYP	polypersonal (affix)
MONOP	monopersonal (affix)
3sABS	3rd person singular absolutive
ANTI	antipassive

ACKNOWLEDGEMENTS

I am grateful to Bernd Heine, Gilbert Lazard, and Tasaku Tsunoda for their comments on an earlier version of this paper. Special thanks are due to Gillian Ramchand and Elizabeth Traugott, who offered useful comments in improving the original manuscript. I have also profited from discussions with Takanori Hirano and Curtis McFarland on Philippine languages. This work was supported in part by a grant from the Japanese Ministry of Education and Culture given to the specially promoted project "Theoretical and empirical studies of properties of Japanese in terms of linguistic universals" (No. 60060001) headed by Professor Kazuko Inoue.

NOTES

1. See the recent paper by Bresnan and Mchombo (1987) for further support.
2. For a fuller discussion on the treatments of *wa* within this tradition, see Shibatani (1990).
3. But see Kuroda (1972; 1976) on the discussion of the Brentao-Marty theory of judgment, which appears to be an attempt to capture a kind of distinction being discussed here.
4. Notice that the same holds for the English equivalent in that the normally uttered sentence *The sun rises* cannot be used in describing the witnessed event of the sun's rising.
5. Even Bloomfield (1933), who uses the term "actor-action" instead of "subject-predicate", generalizes the term "actor" to include a non-agentive subject like *John* in *John fell.*
6. See Kibrik (1979) and Mithun (Volume II) for languages that appear not to have grammaticized the subject category.
7. See Li and Thompson (1976) for a further discussion on the difference between subjects and topics.
8. (34a) and (35a) are reconstructed by the author while all other Quiche forms come from Larsen and Norman (1979).

9. For the correlation of the topic and the relativization and clefting in Chichewa, see Bresnan and Mchombo (1987).

REFERENCES

Anderson, J. 1976. "On the notion of subject in ergative languages." In Li (ed.) 1976:1–24.

Aston, W. 1872. *A Grammar of the Japanese Written Language*. Nagasaki.

Bloomfield, L. 1933. *Language*. New York: Holt.

Bresnan, J. and S.A. Mchombo. 1987. "Topic, pronoun, and agreement in Chichewa." *Language* 63:741–782.

Brown, S.R. 1863. *Colloquial Japanese*. Shanghai: Presbyterian Mission Press.

Chao, Y.R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chafe, W.L. 1970. *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

Chafe, W.L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view." In Li (ed.) 1976:25–56.

Clancy, P. and P. Downing. 1987. "The use of *wa* as a cohesive marker in Japanese oral narratives." In *Perspectives on Topicalization: The Case of Japanese 'wa'*, J. Hinds et al. (eds), 3–56. Amsterdam: John Benjamins.

Clayre, I. 1973. "A preliminary note on focus and emphasis in Melanau —a language of coastal Sarawak." *Lingua* 32:239–253.

Comrie, B. 1981. *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

Cooreman, A. Forthcoming. "Ergativity in Dyirbal discourse."

Cooreman, A., B. Fox and T. Givón. 1984. "The discourse definition of ergativity." *Studies in Language* 8:1–34.

Craig, C. 1977. *The Structure of Jacaltec*. Austin: University of Texas Press.

Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1977. *A Grammar of Yidiny*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1979. "Ergativity." *Language* 55:59–138.

Dixon, R.M.W. 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

Du Bois, J.W. 1987. "The discourse basis of ergativity." *Language* 63:805–855.

England, N. 1983. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

Fujitani, N. 1778. *Ayuhishô*. (Repr. in Fukui 1938.)

Fukui, K. (ed.). 1938. *Kokugogaku taikei* 15. Tokyo: Kôseikaku.

Givón, T. 1976. "Topic, pronoun, and grammatical agreement." In Li (ed.) 1976:149–188.

Givón, T. 1979. *On Understanding Grammar*. New York: Academic Press.

Jespersen, O. 1924. *The Philosophy of Language*. London: George Allen and Unwin.

Kalmar, I. 1979. "The antipassive and grammatical relations in Eskimo." In Plank (ed.) 1979:117–144.

Keenan, E.L. 1976. "Towards a universal definition of 'subject'." In Li (ed.) 1976:303–334.
Kibrik, A.E. 1979. "Canonical ergativity in Daghestan languages." In Plank (ed.) 1976:61–78.
Kuno, S. 1973. *The Structure of the Japanese Language*. Cambridge, Mass: MIT Press.
Kuno, S. 1976. "Subject, theme, and the speaker's empathy — A reexamination of relativization phenomena." In Li (ed.) 1976:417–444.
Kuroda, S.-Y. 1972. "The categorial and thetic judgment: Evidence from Japanese." *Foundations of Language* 9:153–185.
Kuroda, S.-Y. 1976. "Subject." In *Japanese Generative Grammar*, M. Shibatani (ed.), 1–16. New York: Academic Press [*Syntax and Semantics* 5].
Larsen, T.W. and M. Norman. 1979. "Correlates of ergativity in Mayan grammar." In Plank (ed.) 1979:347–370.
Lee, E.W. 1964. "Non-focus verbs in Maguindanao." *Oceanic Linguistics* 3(1):49–58.
Lehmann, W.P. 1976. "From topic to subject in Indo-European." In Li (ed.) 1976:445–457.
Li, C. 1976 (ed.). *Subject and Topic*. New York: Academic Press.
Li, C. and S. Thompson. 1976. "Subject and topic: A new typology of language." In Li (ed.) 1976:457–490.
Mallinson, G. and B.J. Blake. 1981. *Language Typology*. Amsterdam: North-Holland.
McFarland, C. 1978. "Definite object and subject selection in Philippine languages." *Studies in Philippine Linguistics* 2(1).
Mirikitani, L. 1972. *Kapampangan Syntax*. Honolulu: University of Hawaii Press.
Mithun, M. Volume II. "The role of motivation in the emergence of grammatical categories: the grammaticization of subjects."
Miyaoka, O. 1986. "Ergativity in Eskimo." *Gengo Kenkyu* 90: 97–118.
National Language Institute. 1964. Report 25. *Gendai-zasshi 90-syu no yoji*. Tokyo.
Naylor, P.B. 1975. "Topic, focus, and emphasis in the Tagalog clause." *Oceanic Linguistics* 14(1):12–79.
Plank, F. 1979. "Ergativity, syntactic typology and universal grammar." In Plank (ed.) 1979:3–39.
Plank, F. (ed.). 1979. *Ergativity: Towards a Theory of Grammatical Relations*. London: Academic Press.
Sansom, G. 1928. *An Historical Grammar of Japanese*. London: Oxford University Press.
Schachter, P. 1976. "The subject in Philippine languages: Topic, actor, actor-topic, or none of the above." In Li (ed.) 1976:491–518.
Schachter, P. 1977. "Reference-related and role-related properties of subject." In *Grammatical Relations*, P. Cole and J. Sadock (eds), 279–306. New York: Academic Press [Syntax and Semantics 8].
Schachter, P. and F. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.
Shibatani, M. 1985. "Passives and related constructions: A prototype analysis." *Language* 61:821–848.
Shibatani, M. 1988. "Voice in Philippine languages." In *Passives and Voice*, M. Shibatani (ed.), 85–142. Amsterdam: John Benjamins.
Shibatani, M. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.
Silverstein, M. 1976. "Hierarchy of features and ergativity." In *Grammatical*

Categories in Australian Languages, R.M.W. Dixon (ed.), 112–171. Canberra: Australian Institute of Aboriginal Studies.
Toganoi, M. 1770. *Teniha Abukizuna*. (Repr. in Fukui 1938.)
Tsunoda, T. 1985. "Ergativity and coreference in Warrungu discourse." *Nagoya Working Papers in Linguistics* 1:143–165. University of Nagoya.
Van Valin, R.D. 1981. "Grammatical relations in ergative languages." *Studies in Language* 5(3):361–394.
Wierzbicka, A. 1981. "Case marking and human nature." *Australian Journal of Linguistics* 1:43–80.
Yamada, Y. 1908. *Nippon Bunpôron*. Tokyo: Hôbunkan.

From V/2 to Subject Clitics: Evidence from Northern Italian

John Haiman
Macalester College

INTRODUCTION

In this essay, I want to discuss a specific example of a kind of grammaticalization which has achieved the status of a virtual cliché: The degeneration of personal pronouns into agreement marking affixes on finite verbs. By looking carefully at the mechanics of this change in a small number of well-documented and closely related languages, we can gain some insight into the mechanics of this common change, and perhaps account for the fact that some languages are more likely to undergo it than are others. The languages in question are the Northern Italian or Gallo-Italian dialects to the North of the La Spezia-Rimini isogloss, all of which diverge from both standard Italian and standard French in consistent ways.

It is now widely recognized that the familiar distinction between Perlmutter's "type A" languages like standard French, and "type B" languages like standard Italian (Perlmutter, 1971: chapter 5) is not exhaustive. That is, in addition to type A languages, which require overt personal pronoun subjects even for subjectless verbs (cf. ***il*** *pleut "it* is raining"), and type B languages, which never do (cf. *Ø piove* "it is raining"), there are "intermediate languages" of two different sorts.

Haiman (1974) dealt with languages like Icelandic, medieval French and German, and even modern German, arguing that in these languages, or in particular constructions within these languages, the superficial subject pronouns characteristic of Type A languages appeared only when they were required to keep the finite verb in principal clauses in second position. The pattern can be illustrated with a vestigial contrast in modern English:

(1) a. *There* are three books on the table.
b. On the table (there) are three books.

A strong claim about type A languages is that only those languages which have or had the V/2 constraint ever become "type A". The present-day independence of word order and the appearance of superficial subject pronouns (most pronounced in French and English) is viewed as the outcome of a number of historical changes.

That superficial pronoun subjects are motivated by word order requirements is an old idea: Haiman (1974) cites Beneš (1962), who explicitly correlates the appearance of "dummy *es*" with the requirements of verb-second order in modern German. Kuen (1957:302) notes that in medieval French and German manuscripts, subject pronouns are attested prior to the attested loss of personal subject-marking desinences on the verb, and refers to Thurneysen (1892) for the idea that V/2 required the presence of a subject pronoun "when no other sentential constituent stands before the verb" (ibid.304). In his classic survey of personal pronoun subjects in the Lombard dialects, Spiess (1956:5fn) cites Darmesteter (1897: sect. 390), who notes that in medieval French, the personal pronoun subject was spelled out "only when the verb or object pronoun would otherwise stand at the head of the clause". For more recent works which cite an extensive prior literature on this correlation, cf. Harris (1978) for French, and Breivik (1983) for Germanic languages.

Recent generative studies of Northern Italian dialects, notably Brandi and Cordin (1981), have drawn attention to another class of "intermediate languages". Brandi and Cordin showed that in the dialects of Florence and Trento, there is subject pronoun doubling. The same phenomenon can be observed in the dialects of Padua (Benincà, 1983), Venice (Benincà and Vanelli, 1982), Genoa (Browne and Vattuone, 1975), Fassa (Elwert, 1943), Badiot Ladin (Pizzinini and Plangg, 1966), Ampezzo (Appollonio, 1930), Bologna (Kuen, 1957), Ravenna (Gregor, 1972), and the Friuli (Marchetti, 1952:143 et passim). For general surveys, cf. Rohlfs (1949/1972: 169–79), Spiess (1956), Kuen (1957), Renzi and Vanelli (1983), Vanelli (1984a,b), and Benincà (1986).

Subject pronoun doubling means that all verbs (including impersonal, semantically subjectless meteorological verbs) occur with obligatory subject pronouns. The distribution of these subject pronouns corresponds roughly to the distribution of subject pronouns in type A languages.

(2) a. *(Te) tu parli* (Fiorentino: Brandi and Cordin, 1981)
you you speak

b. **(Te) Ø parli*
you speak

(3) a. *(Ti) te vien* (Padovano: Benincà, 1983)
you you come
b **(Ti) Ø vien*
you come

(4) a. *(Jo) o feveli* (Furlan: Marchetti, 1952)
I I say
b. **(Jo) Ø feveli*
I say

(5) a. *(El) no l se fida* (Fassa: Elwert, 1943)
he not he self trusts
'He does not dare.'
b. **(El) no Ø se fida*

(6) *(Me) a voj ben* (Romagnol: Gregor, 1972)
I I want well
'I am fond (of ...)'

From the description given by Browne and Vattuone (1975), it seems that Zeneyze is intermediate in both senses: Superficial subjects are obligatory only in certain constructions, but in those constructions, they resemble the superficial pronouns of Type A languages. Thus, Zeneyze personal pronouns, like standard Italian personal pronouns, are likely to occur in subject position only under contrastive stress. As far as these pronouns are concerned, Zeneyze is a Type B language. However, there exists a most uncharacteristic (for a Type B language) dummy pronoun *u*, roughly corresponding to English "there", German *es*, and the like, which gives Zeneyze a Type A look:

(7) *U vene u Zorzu*
he comes he George
'George is coming.' (thetic)

A similar construction is of course attested in all the languages exemplified in (2) to (6), along with other indices of their apparent Type A status.

What Browne and Vattuone call "U-insertion" occurs only in thetic or topicless sentences, *where the entire sentence conveys new information* (ibid. 138). Such sentences are characterized by VX order. (For what seems to be entirely similar distribution of the dummy pronoun *a* in Padovano, cf. Benincà (1983:34): "*a* would seem to mark syntactically the so-called 'entirely new phrase' which indicates that there is nothing in topic position, that is, that

there is no pragmatic *topic*." In Padovano also, such sentences have VX order.)

As the sentence fragments above very clearly demonstrate, the verb may in addition occur with optional (disjunctive?) pronominal subjects. The latter, like subject pronouns in Type B languages, seem to occur only when they are focussed.

So, are these Northern Italian dialects Type A or Type B languages? They seem to be both, and the answer we choose depends on which pronouns we are looking at.

Cordin and Brandi show that in Fiorentino and Trentino, in a number of ways, the obligatory subject pronouns are syntactically part of the INFL or AGR node, rather than true subject markers. Other arguments can be added to theirs for the other Northern Italian dialects: The synchronic status of obligatory subject markers in all the languages under discussion here is that of bound clitics rather than nominal arguments. Typologically, then, the Northern Italian dialects are impeccably and consistently Type B languages, just as is the standard language. There is no need to complicate the Type A/Type B distinction or the pro-drop parameter. With this conclusion, Brandi and Cordin are content to consider the matter closed.

However, from a diachronic perspective, these dialects seem to be counterexamples to the strong claim that only Verb-second languages become type A languages. Bound clitics are not — at least not diachronically — generated *ex nihilo*. At some previous stage in the languages, it is almost certain that they were "full" pronouns, with argument status (cf. Meillet, 1912; Meinhof, 1936; Bally, 1942; and Givón, 1976, 1979). If their linear distribution at that time resembled their distribution now, then at some earlier time the Northern Italian dialects *were* type A languages. And presently, at least, these dialects exhibit standard Italian word order: Crucially, there is no V/2 constraint. So, the Northern Italian dialects are type A languages without the V/2 constraint.

But this was not always the case in Northern Italian, as it was not always the case in medieval French. In the thirteenth century, word order in attested Northern Italian dialects was similar to that of medieval French, and so was the appearance of superficial pronoun subjects: Word order was V/2, and personal pronoun subjects appeared only when their absence would have led to V-initial word order (Spiess, 1956:17). The correlation between word order and the appearance of the subject pronoun is still transparent in Zeneyze and Padovano, where *u* or *a* occurs typically at the beginning of sentences with VX order. (It is somewhat less transparently recoverable in languages like non-standard French, Furlan and Romagnol, which exhibit a "one preverbal

clitic constraint". I will return to this below.) In some way, subject pronoun doubling was the outcome of V/2 ordering, just as the Type A language phenomenon was the outcome of this ordering. What is not yet clear is how this change occurred, and what factors favoured it in Northern Italy.

After reviewing the current syntactic status of the subject clitic pronouns in Section 1, I will try to show why and how their reduction took place (and why this reduction qualifies as grammaticalization) in Section 2.

1. THE CLITIC STATUS OF OBLIGATORY SUBJECT PRONOUNS

Clitics exhibit the properties of both bound affixes and independent words. The only property of independent words which obligatory subject pronouns still have seems to be that they undergo subject-verb inversion in questions (cf. Brandi and Cordin, who explicitly invoke this criterion of wordhood.) Thus, for example, in Furlan (Marchetti, 1952:143), we observe a paradigm: subject clitic-verb order as in (8) for assertive sentences, and verb-subject clitic order as in (9) for interrogative sentences:

(8) a. *Jo o feveli*
b. *O feveli jo*
'I speak' (assertive)

(9) a. *Jo feveli — o?*
b. *Feveli — o jo?*
'Do I speak?' (interrogative)

The full pronoun *jo* occurs on either the left or the right margin of the verbal complex in both statements and questions. In its freedom to appear either sentence-initially or sentence-finally, irrespective of sentential mood, it still reflects its origins as a dislocated constituent. The obligatory subject marker *o*, on the other hand, is sensitive to whether the sentence is an assertion or a question. This kind of sensitivity to syntactic information is characteristic of clitics in Romance, and distinguishes them from bound affixes, whose status as prefixes or suffixes is not subject to perturbation. (Compare Vanelli, 1984a:283 fn., Linder, 1987:94–5 for the Northern Italian distribution of this feature.) Crucially, in non-standard French (Lambrecht, 1981) and in some varieties of Furlan, this last vestige of word status is missing. Subject pronouns are always prefixes, irrespective of sentential mood:

Ou tu vas ? (NSF: Lambrecht, 1981)
'Where are you going?'

(11) *I pensi (tu) k a vol vinji: ?* (Furlan: Passariano)
you think you that she wants to-come
'Do you think she wants to come?'

In all other respects, subject markers pattern like bound agreement marking morphemes, in all the Northern Italian dialects.

Most obviously, they co-occur not only with pronominal, but also with lexical subject nominal expressions:

(12) a. *in om al veve doi fis* (Furlan)
a man he had two sons
'A man had two sons.'

b. *la strade e va ju a plomp*
the road she goes down steeply
'The road goes down steeply.'

c. *une vore pocis e an timp di ciala:*
a real few they have time to look
'Only a few have time to look.'

(13) a. *Da doman e les stades oshorades les vatges*
of morning are they become fed the cows
'In the morning, the cows are fed.'

b. *tagn d'agn a la pa osta fomena?*
how-many years has she then your wife
'How old is your wife, then?'
(Badiot Ladin: Pizzinini and Plangg, 1966)

(14) a. *U Zorzu u vene* (Zeneyze: Browne and Vattuone, 1975)
he George he comes

b. *A Katajning a vene*
she Catherine she comes

c. *I vezhing i venju*
they neighbours they come

(15) *dov e l al professaur* (Bolognese: Kuen, 1957)
where is he the professor
'Where is the professor?'

(16) a. *duta ra me biancheria r e fata de bona tera*
all the my linen she is made of fine cloth
'All my linen is made of fine cloth.'

b. *chel contadin l a bona tera inz i so ciampe*
that farmer he has good earth in the his fields
'That farmer has good soil in his fields.'
(Ampezzano: Appollonio, 1930)

(17) a. *la colpa l e vostra* (Romagnol: Gregor, 1972)
the fault she is yours
'It is your fault.'

b. *cosa dira l la mama?*
what will-say she the mama
'What will mama say?'

In the same way that obligatory subject pronouns cooccur with lexical NP, they also cooccur with the relative pronoun {che}, as in the Furlan and Fassa phrases of (18):

(18) a. *dut ce che al e gno* (Furlan)
all that which he is mine
'all that which is mine'

b. *doj peres pitshuj ke i me prea l arberk* (Fassa)
two poor little-boys who they me ask the shelter
'two poor little boys who ask me for shelter'

This too is general throughout Northern Italian, cf. Rohlfs (1949/72:176). It is unlikely that {che} in these examples can be analyzed as an extra-clausal complementizer, since (a) it typically occurs in different nominative and accusative forms (cf. Rohlfs ibid. 233), and (b), the accusative form *does* replace *object* pronouns. (The data in fact are consistent with the hypothesis that "subject {che}" is a relative pronoun, while "object {che}" is a homophonous complementizer, but a detailed analysis belongs elsewhere.) For the complementary distribution of "object {che}" and object pronouns, witness the Furlan:

(19) a. *la curtisse che al veve piardut*
the knife that he had lost

b. *il sace che i doi vecios puartavin su lis spalis*
the bag that the two old-men carried on the shoulders
'the bag that the two old men carried on their shoulders'

(The examples of (19) are from Gregor, 1982.)

A natural temptation is to think of these *NP+pronoun* collocations as examples of left dislocation, as in "the one-l lama, he's a priest". (As noted, the word order of the optional full pronouns in languages like Furlan still reflects this status. So too does their morphology. Vanelli (1984:285) points out that the free and optional pronouns tend to be similar to the *oblique* or disjunctive forms, and probably therefore derive from them historically. See, for example, the Romagnol sentence (6).)

If this interpretation were generally true (as it seems to be for non-standard French, cf. Lambrecht, 1981), then there would be no impediment to analyzing the underlined pronominal morphemes as referential nominal arguments. However, there are at least two arguments against this interpretation in the Northern Italian dialects.

The first, due to Brandi and Cordin, is a co-occurrence restriction: Indefinite NPs (including negatives like "no one") cannot be topics, or co-occur as left-dislocated elements with resumptive topic. However, indefinite NP and negatives co-occur freely with the clitic pronoun subjects, as the following sentences from Trentino and Fiorentino show:

(20) a. *Nisun l ha dit niente.* (Trentino)
no one he has said nothing
'No one said anything.'

b. *Nessun gl ha detto nulla* (Fiorentino)
no one he has said nothing
'No one said anything.'

This structure is impossible for both standard Italian and all varieties of current French. In all other respects to be enumerated, however, non-standard French marches with the non-standard Italian dialects.

Another argument against the analysis of clitics as resumptive pronouns is that — like agreement markers in general, and under the same conditions — they frequently *fail to mark agreement with the NP which they supposedly*

resume. By the same token, unlike lexical pronouns, they fail to mark the gender of the NP for which they stand.

Roughly speaking, agreement between subject and predicate, or between between object and predicate, operates obligatorily from left to right but only optionally in the opposite direction. Again, we can model the contrast in English:

(21) a. *Two mice are (*is) hiding in your drawer*
(subject precedes agreement-marking verb)

b. *There's (are) two mice hiding in your drawer*
(subject follows agreement-marking verb)

"Grammatical" English, of course, demands agreement with even a postposed subject in sentences like (21b). But, as even a purist will admit, (21b) is possible with unmarked singular agreement, and (21a) is not. (For a discussion of the phenomenon in Swiss Romantsch, cf. Linder, 1982; for Italian, cf. Rohlfs, 1949/1972:448.)

Some examples from Northern Italian:

(22) a. *l e venju la vivano* (Fassa: Elwert, 1943)
he is come MASC.SG. the witch FEM.SG.
"There came the witch."

b. *chi e lo po i ozitegn* (Moena: Plangg, 1983)
who is he then the Occitans MASC.PL.
'Who are the Occitans?'

c. *da doman vegn l oshore les vatges*
of morning becomes he fed MASC.SG. the cows FEM.PL.
'The cows are fed in the morning.'
(Badiot: Pizzinini and Plangg, 1966. cf. (10a), with agreement)

d. *gl e venuto delle ragazze* (Fiorentino: Brandi
he is come MASC.SG. some girls FEM.PL. and Cordin, 1981)
'There came some girls.'

e. *u vene a Katajning* (Zeneyze: Browne and
he comes she Catherine Vattuone, 1975)
'Catherine is coming.'

f. *agnere l e sta ra sagra inz el nosc paes.*
yesterday he is been the feast FEM.SG. in the our county
'Yesterday we had our county feast.'
(Ampezzano: Appollonio, 1930)

g. *al era una volta una fameja* (Furlan: Iliescu, 1972)
he was one time a family FEM.SG.
'There was once a family.'

h. *e chenta una turtureina* (Romagnol: Gregor, 1972)
he sings a turtle dove FEM.SG.
'A turtle dove is singing.'

Paradoxically, non-agreement in the examples of (22) can be motivated if the morphemes in question are analyzed as grammaticalized agreement markers. It is anomalous if they are analyzed as referential (anaphoric or cataphoric) pronouns. (I will return to my reasons for making this claim in the concluding section of the essay.) The third way in which obligatory subject pronouns differ from "true" or referential subject NPs with argument status is in their position relative to the negative morpheme {no}. Subject NPs invariably precede this morpheme: Obligatory subject pronouns in at least some dialects (among them Fassa, Padovano, Trentino, Furlan) either obligatorily or conditionally follow it. Thus, Fassa

(23) *El no l se fida*
'He doesn't dare.'

motivates the structure proposed by Brandi and Cordin, where the obligatory subject marker *l* is part of the verbal inflection. (In Fiorentino, the negative may precede 2sg., but must follow all third person obligatory pronoun subjects. The potential for permutation, Brandi and Cordin argue, following Rizzi (1986), provides further evidence for the constituency of subject pronouns within the verbal complex, but this argument strikes me as counterintuitive: We could accept it only if the order of morphemes within a word were subject to rearrangement, while the order of words within a phrase were fixed. If anything, the opposite seems to be the case, cf. Perlmutter, 1971:100.)

Fourth, obligatory subject pronouns seem to resist coordination reduction. The prevailing pattern is illustrated by the following examples from Furlan:

(24) a. *al vent dut e al va lontan*
he sells everything and he goes far-away

b. *quant ch al vigni a se e al cognosse l predi*
when he came to himself and he recognized the priest

By way of contrast, English, although a Type A language, exhibits coordination reduction in sentences like the normal translations of (24):

(25) a. *he sells everything and goes far away*

b. *when he came to himself and recognized the priest*

Compare the entirely similar behavior of Surselvan Romantsch, another Type A language very closely related to the Northern Italian dialects:

(26) a. *Ins selegra ed ei satisfatgs ...*
PRO rejoices and is satisfied
'We rejoice and are satisfied...'

b. *auters savessen e duessen imitar quella initiativa*
others should-know and should imitate that initiative
'Others should be able to, and indeed should in fact imitate this initiative.'

Curiously, Lambrecht (1981) notes that while standard French is like English and Surselvan, non-standard French agrees with Northern Italian:

(27) a. *il mange et boit comme un cochon* (standard)

b. *i mange et i boit comme un cochon* (non-standard)

Sublexical material *may* be deleted under identity in coordinate constructions (cf. Kiparsky, 1968, Haiman, 1983), but this is much more constrained than deletion of independent words. The failure of CR in (24), as in (27b) provides further evidence in support of the sublexical status of obligatory subject pronouns. (For the same argument in Trentino, cf. Brandi and Cordin (1981), duly rehashed in Bouchard, 1982:407; Rizzi, 1986:402.)

Finally, in spite of what has been said so far, and is also the prevailing opinion (cf. Rohlfs, 1949/72:169, 174), "obligatory" subject pronouns are not *entirely* obligatory, but may be omitted in the presence of certain object or negative clitics. The rules vary greatly from dialect to dialect.

In some dialects of Furlan, obligatory subject markers are not used in persons other than 2sg. when either object pronouns or the negative particle

no precede the verb (Marchetti, 1952: 145). In others, all subject pronouns except 2sg. may be omitted when the verb is preceded by any object or reflexive clitic (Gregor, 1975:107). The textual data from Romagnol provided in Gregor (1972) seem to support a similar constraint in that dialect, and Lambrecht (1981) notes that non-standard French also allows subject clitic pronouns to be absent where an object pronoun also precedes the verb. Illustrative of this mutual exclusion are Furlan examples such as those of (28):

(28) a. *denant del ciar si e viodut un sflandor*
in-front of-the cart self is seen a flash
'A flash of light appeared before the cart.'

b. *so pari lu viot.*
his father him saw
'His father saw him.'

It is this mutual exclusion, rather than coordination reduction, which accounts for the absence of obligatory subject marking pronouns in examples such as those of (29):

(29) a. *e veve squasi 35 agn e di zovine si jere fate ciacara*
she had about 35 years and of young self was made to-talk
'She was about 35, and had gotten herself talked about in her youth.'

b. *al clame un servidor e j domande ce ch al ere*
he calls a servant and him asks what that it was
'He calls a servant and asks him what was the matter.'

This sensitivity to the presence of prefixed verbal clitics suggests that obligatory subject markers belong in the verbal clitic complex. To analyze them in this way makes it possible to state a relatively simple "clitic constraint" at least for Furlan, and possibly Romagnol and non-standard French as well:

(30) All finite verbs must occur with at least one prefixed clitic in statements.

Note that (30), like constraints on the distribution of Zeneyze *u* and Padovano *a*, may still reflect, in a rather obvious way, a correlation between word order and the V/2 constraint. In an earlier stage of the language, where all the present clitics were free-standing words, the presence of any one of them would keep the verb from sentence-initial position.

(On the other hand, it must be admitted that this argument is not unambiguous. For example, it seems that in Venetian and Padovano, 2sg. is always obligatory, but third person forms may be omitted in the presence of a full NP (which may be a disjunctive personal pronoun, an interrogative pronoun, a relative pronoun, or a lexical NP), cf. Benincà, 1983; Benincà and Vanelli, 1984:167. Here, the sensitivity of the obligatory subject morpheme to the presence of full subject NP would seem to argue for its status as an independent NP. We will return to these messy data, and the arguments they provide for the synchronically transitional status of the clitic subject pronouns, in the second section of this essay.)

There are, then, numerous reasons for believing that subject marking "pronouns" are really inflectional or agreement markers in the Northern Italian dialects today, and well on the way to achieving the same status in non-standard French.

One last observation: While there is dialect variation in the degree of subject clisis which the dialects have undergone, all the dialects seem to agree in some special treatment of the 2sg. This is the subject pronoun which has advanced the furthest in reduction to clitic status. In Furlan, Venetian, and Padovano, it alone can never be omitted; in Fiorentino (which permits the absence of 1sg.) and Trentino (which permits the absence of first person, and 2pl.), it is one of the many which can never be omitted; in Fiorentino, it is one of the many which follows the negative prefix.

Granting the plausible syntactic change

$$\text{pronoun } \# \text{ V} > \text{clitic} + \text{V}$$

which has been defended and illustrated in many languages, there are various indications that 2sg. pronouns led the way.

2. STAGES ON THE PATH OF GRAMMATICALIZATION

That agreement affixes frequently originate as personal pronouns has been a truism at least since Meillet (1912). Nevertheless, the current situation in the Northern Italian dialects is presently unique among the Romance languages (although there is evidence that subject pronouns in French are beginning to move along the same path now, cf. Kayne, 1975; Lambrecht, 1981, 1988). What particular conditions distinguish the Northern Italian dialects from standard Italian, and from the closely related Swiss dialects of Romantsch?

First, there is agreement that Northern Italian, unlike central and southern dialects, but very much like French and Romantsch, came under considerable Germanic influence. Helty (1975) speaks of a "Westfrankish superstrate influence" on the Gallo-Italian dialects (which would date to the eighth century conquests of Charlemagne), but Germanic occupation of Northern Italy dates back even further, to the invasions of the Bavarians and the Lombards in the sixth century at least.

Among the attested syntactic results of this contact is the prevalence of V/2 order in the only Northern Italian dialects with a continuous written tradition going back to the fourteenth century, those of Lombardy (Spiess, 1956).

It is therefore tempting to see a correlation between subject doubling and V/2 order. The written records agree, however, that no such doubling was attested in thirteenth or fourteenth century Lombard, or even fifteenth century Furlan. In fact, in written Furlan, obligatory *tonic* subject pronouns are solidly attested no earlier than the seventeenth century (Iliescu, 1972:149). And, of course, nothing similar exists in modern French or Swiss Romantsch, which also have a history of being V/2 languages under Germanic influence, and are now typical type A languages.

So what marked Northern Italian off from the closely related and typologically similar Romance dialects to the North? In particular, what separated Ladin and Furlan from Surselvan, when they are all allegedly dialects of the same Rhaeto-Romance language? And what is special about the 2sg. form? In this section, I want to demonstrate the plausibility of the hypothesis that reduction of the pronoun to a bound affix originated *in inverted word order*. That is, unlike Givón (1979), I think that the degeneration of pronoun subjects to clitics (and ultimately, to completely bound verbal agreement markers), originates not in left-dislocated structures like

My old man, he rides with the Angels.

but rather in interrogative and other VS structures like

Does he your old man ride with the Angels?

The bulk of this section will be devoted to supporting my claim.

2.1. The status of the 2sg.

The pioneering status of the 2sg. pronoun on the steep descent from pronoun to agreement marker is explained by the hypothesis that reduction

occurred in inverted word order: This is the order in which 2sg. pronouns usually occur, for the simple pragmatic reason that (when dealing with matters of fact rather than value judgments) one asks questions rather than makes statements concerning one's interlocutor.

As inverted word order is characteristic of questions, in all the dialects investigated here, the most frequent or typical subject of a sentence with inverted word order is one with a 2sg. subject. On the hypothesis that reduction occurred first in inverted word order, this has predictably led to egregious reduction of the inverted 2sg. pronoun, leading to its syntactic reinterpretation or deletion in many cases.

In the Gorizian dialect of Furlan, the 2sg. and 2pl. atonic subject pronoun affixes may not only co-occur with the preverbal subject pronouns in the interrogative or inverted word order, but also in the direct or assertive word order as well: That is, structures like

(31) *pjardis-tu*
lose 2SG. 2SG.
'you lose'

occur irrespective of word order: Clear evidence that the one-time pronoun *tu* has completely lost its one last trace of syntactic identity as a noun phrase, its sensitivity to the rule of subject-verb inversion (Frau, 1984:113). It has been reinterpreted as part of the extended agreement marker of the 2sg. This, of course, is exactly analogous to what has happened in other languages like German and even Old English, where the 2sg. *-t* enlargement could have no other source.

The reduced morpheme *t(u)* still maintains a vestigial presence as a pronoun. This ghostly survival accounts for "2 deletion" (Haiman, 1971) in Germanic inversions like

(32) *Hast killed the Jabberwock?*

which are totally productive in impeccably Type A languages like German, and in the Rhaeto-Romance dialects of Switzerland (cf. Table 1, below). But mainly, the reduced morpheme is now simply part of the agreement marker, and, as such, cooccurs with full pronominal 2sg. subjects. My thesis is that Old English 2sg. *-(es)t*, still vestigially attested in archaic forms like

(33) a. *Thou watches-t*

b. *Thou shal-t*

and the like, exemplify the exact same subject doubling as the Northern Italian dialects exhibit in the sentences of (9), and that here, too, the 2sg. has led the way (and no other pronoun as yet has followed).

Ettmayer (1903:50fn) provides evidence for another reinterpretation which is no less indicative of the reduction of the 2sg. pronoun. The interrogative or inverted paradigm in many Northern Italian dialects was restructured on the basis of the 2sg. so that *-t* became a marker of the interrogative mood. In Verona, Vicenza, and Trento, 2sg. interrogative *fe-tu* 'are you' was reinterpreted as

(34) *fe — Ø— t(e)*
be 2sg. interrogative
'Are you?'

That is, the same structure, with the same "external syntax", was presumably reanalyzed by speakers with the erstwhile 2sg. desinence *-t(e)* reinterpreted as a non-person-specific marker of the interrogative mood.

Evidence for this reinterpretation is provided by the analogical extension of the *-t(e)* suffix to other persons and numbers as a sign of the interrogative:

(35) 1sg *fon-te* 'am I'
1pl *fen-te* 'are we'

This restructuring demonstrates that in inverted word order, the 2sg. pronoun is reduced to the point that it is not perceived as a pronoun at all.

It is tempting to speculate that the French liaison /t/ of inverted word order, now limited to the 3sg., is not an etymologically motivated restoration of Latin 3sg. *-t*, but a relic of the same reinterpretation process. Note in particular the non-standard example from Vendryes, 1914, cited by Lambrecht, 1981:

(36) *i l a — ti pu attraper le gendarme le voleur?*
he him has ? been-able to-catch the cop the thief

The verbal suffix *-ti* is now a general marker of the interrogative in Canadian French as well. Whatever the original source of *-ti*, example (36) shows an extension of it. A 2sg. source is at least as plausible as a hypercorrect "restoration" of a final consonant which had been absent from the spoken language for several hundred years by the time that it first appears in written thirteenth century French.

2.2 Some dialects have doubling only in inverted word order.

Our best evidence for the hypothesis that reduction originated in inverted order comes from a comparison of the various Gallo-Italian dialects which are known in the literature as Rhaeto-Romance. The major *apparent* split is between the Swiss RR dialects (known collectively as Romantsch) and the Italian RR dialects (which include Fassa, Badiot, Ampezzano, and Furlan, from among the examples we have already discussed).

The Italian RR dialects, as we have seen, are similar to other Northern Italian dialects like Romagnol, Zeneyze, Padovano, and Trentino, in allowing clitic doubling. The Swiss dialects, on the other hand, are typically described as standard Type A languages like French: One subject pronoun and only one is necessary for every finite verb. Like medieval French, and like the Ladin dialects of Badia and Gardena in the Dolomitic Alps, the Romantsch dialects are still subject to the borrowed V/2 constraint (cf. Nay (1965) for Surselvan, Thoeni (1969) for Surmeiran, Scheitlin (1962) for Puter, and Arquint (1964) for Vallader).

Closer examination of the Swiss dialects reveals that all but Surselvan have two sets of subject pronouns: The first set, which Surselvan also possesses, is of *tonic* pronouns, clearly having lexical status. The second, attested in Surmeiran, Puter, and Vallader, is of bound *atonic* pronouns which are affixes on the verb. This second set occurs only in inverted word order, and is obligatory there. An examination of the morphology of these atonic pronouns is instructive in showing, first, the close kinship between most Swiss and Northern Italian dialects, and, second, the familiar reduction of the 2sg. (and here, also 2pl.) inverted pronominal forms to zero:
The standard grammars from which these forms are drawn state that atonic subject pronouns in RR languages exist in complementary distribution with the tonic forms as mentioned:

$$\textit{tonic pronoun} \# V \sim V + \textit{atonic pronoun}$$

However, Linder (1987:4–12) shows that the relationship between the two is much more interesting. Atonic subject pronouns do occur preverbally, and are here still in complementary distribution with the tonic forms and other lexical NP like the relative pronoun. Postverbally, however, atonic pronouns may cooccur with full subject pronouns, as in

(37) a. *chantains — a (nus)*
sing-1pl. we we (Puter)

Table 1 *Postverbal Atonic Subject Pronouns*

	Surmeiran	*Puter*	*Vallader*	*Fassa*	*Padovano*	*Furlan*
1.	a	i	a	e i	i	V
sg.2.	—	—	—	te ~ tu	to	tu ~ ti
3m.	l	e	l	el ~ al	el	(V)l
f.	la	e	la	ala	la	V ~ l
n.	i	e	a			
1.	a	a	a	(i)	i	V
pl.2.	—	—	—	(i)	u	V
3m.	ilj	e	a	i ~ aj	li	V ~ aj
f.				les ~ ales	le	V ~ as

(Atonic pronouns are defined as those which are incapable of carrying contrastive stress.)

b. *pigliain — a (no)*
take-1pl. we we (Vallader)

Here, reduction of form has plainly led to reduction of syntactic function: The erstwhile pronoun has lost its status, and functions only as a verbal ending. It is important to notice that this status reduction has occurred — at this point — only in inverted word order. Preverbally, the fact that atonic and tonic pronouns are mutually exclusive suggests that both are viewed as members of the same syntactic category.

Finally, it is notable that Surselvan, which does not have atonic pronouns in inverted word order (although it exhibits "2 deletion" there), has not developed or retained such atonic pronouns elsewhere. (Given the presence of atonic subject pronouns in seventeenth century Surselvan, it is generally agreed that these forms were lost as a result of German influence.)

Reduction originates in the inverted word order, and we do not know why: Nevertheless, the data allow us to state an implicational relationship like

(38) If a language has atonic subject pronouns in direct order, it also has them in inverted word order.

We may now predict that clitic doubling takes place only in those languages and dialects which (a) have inverted word order, primarily in questions, and which (b) exhibit reduction in this inverted word order. Reduction of subject pronouns in inverted word order is then an indispensable prerequisite for the reinterpretation described herein.

We provide an outline of the conjectured changes involved in this development, with the languages in which each stage is attested given alongside. If our changes were chronologically ordered, the future history of any language could be read by looking down:

Table 2 *Stages from Type A to pronoun absorption*

Stage 1:	standard type A lg.	(standard French, English)
Change 1:	V#2sg. > V + 2sg.	(Old Eng., German, Surselvan, Gorizian, Trentino, Vicenzan, Venetian)
Change 2:	V#pronoun > V + pronoun	(Surmeiran, Puter, Vallader)
Change 3:	V#pronoun > V + pronoun # (NP)	(Puter, Vallader)
Change 4:	pronoun#V > pronoun + V	(beginning in Surmeiran, Puter, Vallader)
Change 5:	pronoun + V > (NP) # pronoun + V	(Northern Italian)
Change 6:	pronoun + V > agreement marking prefix + V	(NSF, ?Furlan)

It is important to observe that incorporated pronouns, by our reckoning, are no longer sentential arguments: Therefore all stages after stage 1 represent Type B languages of one kind or another. (If the philological hypothesis about the genesis of Type A languages is correct, then the history of Northern Italian, deriving from medieval Lombard and, ultimately, from Latin, represents a kind of cycle: Type B (Latin) yields to Type A (medieval Lombard), which finally develops into another Type B (Padovano, Furlan, etc.).

Clearly, the changes are not strictly chronological: For example, nonstandard French, having failed to undergo changes (3) or (5), has nonetheless undergone change (6). On the other hand, Old English underwent the first change, which modern English has reversed, while seventeenth century Surselvan passed through change (2), which modern Surselvan has also reversed.

3. CONCLUSIONS

Why call any of this grammaticalization? There are two kinds of change which are always associated with grammaticalization: Phonetic reduction

and semantic bleaching or generalization. Phonetic reduction takes place as a lexeme is incorporated into a more complex construction, losing its integrity as a word. Semantic reduction, or bleaching, occurs as a morpheme loses its intention: From describing a narrow set of ideas, it comes to describe an ever broader range of them, and eventually may lose its meaning altogether.

The two frequently but not invariably go together. For example, the 2pl. pronoun V has undergone some grammaticalization in the European languages in that it now denotes politeness as well as plurality. In a number of Romance languages, the plurality is emphasised by the addition of some word like **alteri* (French *autres*, Italian *altri*): As long as the plurality is expressed by a phonologically autonomous word, semantic bleaching is inhibited. *Voialtri* does not yet, in any Italian dialect, denote politeness. Conversely, our first clue that grammaticalization of this morpheme has occurred will be when some dialect weighs in with *voialtri* as the polite V form.

The motivating force for both reduction and bleaching, I believe, is frequency of occurrence, which may in some cases cause the one kind of change without the other. A familiar example of semantic bleaching without phonetic reduction is the development of the interrogative into an indirect imperative. Like most indirect speech acts, "whimperatives" are motivated by the demands of politeness. But if a question can function as a polite imperative, why are the following three sentences not equally acceptable in this role?

(39) a. Can you tell me the time?
b. Is it the case that you can tell me the time?
c. Are you able to tell me the time?

Searle (1985:49) points out that differences like these are a major problem for a general theory of indirect speech acts. But it seems perfectly clear that the most frequent locution, that of a., is the one which has a broader range of semantic functions. If grammaticalization can indeed be identified by the association of a single form with a greater range of possible meanings, then whimperatives are examples of partial grammaticalization. And frequency of use favours structures like (39)a. for semantic bleaching, but has not resulted in any obvious phonetic reduction.

The phonetic reduction of the subject pronoun clitics in Northern Italian is illustrated by a comparison of their form with the full pronouns. But we have not yet emphasized how this phonetic reduction goes hand in hand with bleaching.

Recall, however, that these obligatory subject pronouns frequently fail

to mean what they should, and appear, as in the examples of (18), in the *unmarked* form. They do so in exactly those contexts where agreement markers fail to mark agreement.

Let's define as grammatical a morpheme whose distribution is determined not (entirely) by its correspondence with ideas or reality, but rather by purely formal conventions. (Compare grammatical gender with natural gender, tense markers with adverbs of time, and polite greetings or whimperatives with true information-seeking questions). The cultural analogy is with a ritual, where frequent repetition also leads to automatic, "insincere" performance, in which utterance meaning diverges from literal sentence meaning. (The distinction between utterance meaning and sentence meaning is made as such by Searle (ibid.) and illustrated by him primarily with reference to the phenomenon of metaphorical speech. Among the kinds of speaking which he has failed to enumerate where sentence and utterance meaning diverge is ritual speech and clichés. It may be, however, that these are profitably dealt with under the heading of metaphor. In particular, any grammaticalized morpheme is conceptually parallel to a dead metaphor, a kind of cliché whose original force has been lost through repetition.)

Then obligatory pronouns prove their grammaticalized status as clichés in failing to do the one thing they are supposed to do, which is to mark agreement. Like nearly meaningless conventions everywhere, their presence is dictated not by the speaker's meaning or by considerations of personal style, but by the etiquette of grammar.

They are well on their way to grammatical cipherhood, although the next stage on their semantic evolution is obscure. In the words of the Soviet joke, however, we may say that like this year's average harvest (worse than last year's, and better than next year's), there is nowhere for these morphemes to go but down.

REFERENCES

Appollonio, Bruno. 1930. *Grammatica del dialetto Ampezzano*. Trento: Arti Grafiche Tridentum.

Arquint, Jachen. 1964. *Vierv ladin*. Cuoira: Lia Rumantscha.

Bally, Charles. 1942. *Linguistique générale et linguistique française*. Bern: Francke.

Beneš, Eduard. 1962. "Die deutsche Wortfolge von der Mitteilungsperspektive her betrachtet." *Philologica Pragensia* 2:6–19.

Benincà, Paola. 1983. "Il clitico *a* nel dialetto Padovano." In Pellegrini 1983:25–35.
Benincà, Paola. 1986. "Punti di sintassi comparata dei dialetti italiani settentrionali." In *Raetia antiqua et moderna*, G. Holtus and K. Ringger (eds), 457–78. Tübingen: Max Niemeyer.
Benincà, Paola and Laura Vanelli. 1982. "Appunti di sintassi veneta." In *Guida ai dialetti veneti*, M. Cortelazzo (ed.), 4:7–38. Padova: Istituto di Glottologia e fonetica dell'Università di Padova.
Benincà, Paola and Laura Vanelli, 1984. "Italiano, Veneto, Friulano: Fenomeni sintattici a confronto." *Rivista Italiana di Dialettologia* 8:165–93.
Bouchard, Marc. 1982. On the Content of Empty Categories. Ph.D. thesis, Massachusetts Institute of Technology.
Brandi, Patricia and Lucia Cordin. 1981. "Dialetti e italiano: Un confronto sul parametro del soggetto nullo." *Rivista di Grammatica Generativa* 6:33–87.
Breivik, Leif. 1983. *Existential "There"*. Bergen: Department of English, University of Bergen.
Browne, Wayles and Bartolo Vattuone. 1975. "Theme-rheme structure and Zeneyze clitics." *Linguistic Inquiry* 6:136–40.
Darmesteter, Arsène. 1897. *Cours de grammaire historique de la langue française:* 4me partie, syntaxe. Paris: Payot.
Elwert, W. Theodor. 1943. *Die Mundart des Fassa-Tals*. Heidelberg: Carl Winter. [*Wörter und Sachen,* n.f. Beiheft 2].
Ettmayer, Karl von. 1903. *Bergamaskische Alpenmundarten*. Leipzig: Reisland.
Frau, Giovanni, 1984. *I dialetti del Friuli*. Udine: Società Filologica Friulana.
Givón, Talmy. 1976. "Topic, pronoun, and grammatical agreement." In *Subject and topic*, Charles Li (ed.). New York: Academic Press.
Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
Gregor, Douglas Barton. 1972. *Romagnol: Language and Literature*. New York: Oleander Press.
Gregor, Douglas Barton. 1975. *Friulan: Language and Literature*. New York and Cambridge: Oleander Press.
Haiman, John. 1971. "Targets and paradigmatic borrowing in Romantsch." *Language* 47:797–808.
Haiman, John. 1974. *Targets and Syntactic Change*. The Hague: Mouton.
Haiman, John. 1983. "On some origins of switch-reference marking." In *Switch-Reference and Universal Grammar*, John Haiman and Pamela Munro (eds), 105–28. Amsterdam: John Benjamins.
Harris, Martin. 1978. *The Evolution of French Syntax: A Comparative Approach*. London: Longmans.
Helty, G. 1975. "Westfrankische Superstrateinflüsse auf die Galloromanische Syntax." *Romanische Forschungen* 87: 413–27.
Iliescu, Maria. 1972. *Le frioulan*. The Hague: Mouton.
Kayne, Richie. 1975. *French Syntax*. Cambridge: MIT Press.
Kiparsky, Paul. 1968. "Tense and mood in Indo-European syntax." *Foundations of Language* 4:30–56.
Kuen, Heinz. 1957. "Die Gewohnheit der mehrfachen Bezeichnung des Subjekts in der Romania und die Gründe ihres Aufkommens." In Reichenkron et al. (eds) 1957:293–326.
Lambrecht, Knud. 1981. *Topic, Antitopic, and Verbal Agreement in Non-standard French*. Amsterdam: John Benjamins.

Lambrecht, Knud. 1988. "Presentational cleft constructions in spoken French." In *Clause Combining in Discourse and Grammar*, John Haiman and Sandra A. Thompson (eds). Amsterdam: John Benjamins.
Linder, Karl. 1982. "Die Nichtübereinstimmung von finitem Verb und nachgestelltem Subjekt bei (Genus und) Numerus im Raetoromanischen Graubündens." In *Fakten und Theorien: Festschrift für Helmut Stimm*, S. Heinz and Ulrich Wandruszka (eds), 147–62. Tübingen: Gunter Narr.
Linder, Karl. 1987. *Grammatische Untersuchungen zur Charakteristik des Rätoromanischen in Graubünden*. Tübingen: Narr.
Marchetti, G. 1952. *Lineamenti di grammatica friulana*. Udine: Società filologica friulana.
Meillet, Antoine. 1912. "L'évolution des formes grammaticales." In *Linguistique historique et linguistique générale*, 130–49. Paris: Champion.
Meinhof, Carl. 1936. *Die Entstehung flektierender Sprachen*. Berlin: Reimer.
Nay, Sep. 1965. *Bien di, bien onn*. Disentis: Lia Rumantscha.
Pellegrini, 1983. *Scritti in onore di Giovan Battista Pellegrini*. Pisa: Pacini.
Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt.
Pizzinini, A. and Guntram Plangg. 1966. *Parores ladines*. Innsbruck: Leopold-Franzens Universität. [*Romanica Aenipontana*, 3].
Plangg, Guntram. 1983. "Emphatische Konstruktionen im Zentralladinischen." In Pellegrini, 1983:305–17.
Reichenkron, G. et al. (eds.) 1957. *Syntactica und Stylistica: Festschrift für Emil Gamillscheg zum 70. Geburtstag*. Tübingen: Niemeyer.
Renzi, Lorenzo and Laura Vanelli, 1983. "Il pronome soggetto in alcune varietà romanze." In Pellegrini 1983:121–45.
Rizzi, Luigi. 1986. "On the status of clitic subjects in Romance." In *Studies in Romance Linguistics*, Osvaldo Jaeggli and Carmen Silva-Corvalan (eds), 391–419. Dordrecht: Foris.
Rohlfs, Gerhard. 1949/1972. *Historische Grammatik der Italienischen Sprache. Band II*. Bern: Francke.
Scheitlin, Walter. 1962. *Il pled Puter*. Samedan: Uniun dals grischs.
Searle, John. 1985. *Expression and Meaning*. Cambridge: Cambridge University Press.
Spiess, Federico. 1956. *Die Verwendung des Subjekt–Personalpronomens in den Lombardischen Mundarten. Bern: Francke. [Romanica Helvetica 59]*.
Thoeni, Gion. 1969. *Rumantsch-Surmeir*. Chur: Lia Rumantscha.
Thurneysen, Rudolf. 1892. "Zur Stellung des Verbums im Altfranzösischen." *Zeitschrift für romanische Philologie* 16.
Vanelli, Laura. 1984a. "Pronomi e fenomeni di prostesi vocalica nei dialetti italiani settentrionali." *Revue de Linguistique Romane* 48:281–95.
Vanelli, Laura. 1984b. "Il sistema dei pronomi soggetto nelle parlate Ladine." In *Das Romanische in den Ostalpen*, D. Messner (ed.), 147–60. Vienna.

The Role of Motivation in the Emergence of Grammatical Categories: the Grammaticization of Subjects

Marianne Mithun
University of California, Santa Barbara

INTRODUCTION

Our search for forces motivating the grammaticization of particular categories has uncovered two main kinds of factors: Cognitive and communicative. Speakers tend to develop linguistic structures that mirror their cognitive structures in order to have efficient tools for the expression of thought. The ongoing activity of communication hones these structures: Patterns that are used frequently become routinized and automatic. If all human beings have similar cognitive structures and communicative needs, we might expect the kinds of categories grammaticized in different languages to be the same. Certain grammatical categories do appear in language after language, yet one of the most intriguing aspects of language is the fact that grammatical systems are far from universal.

Understanding the processes of grammaticization involves not only discovering which categories tend to be grammaticized in languages, but also why these are not grammaticized in every language. Cognitive structures may set up predispositions for the grammaticization of certain categories rather than others, but they are not mandates. Whether a preferred category is grammaticized or not can be affected by several factors. First, it may depend simply on chance. Second, it may depend on cultural preoccupations. For Karok speakers, for example, the basis of spatial orientation was apparently the Klamath River. Reference to the river must have appeared sufficiently often in everyday speech to become routinized, ultimately yielding a rich system of locative suffixes with such meanings as 'hence upriverward', 'hither from upriver', 'hence across a body of water', 'hither from across a body of

water', 'hence downriverward', 'hither from downriver', etc. (Bright, 1957: 95). In many other cultures, orientation in terms of a river is so rare that it will probably always remain expressible only lexically. Third, whether or not a category is grammaticized at a particular moment must depend upon the shape of the current grammar. The formation of new grammatical categories is motivated or hindered by the contours of the existing grammatical system.

As is well known, it was once proposed that all languages are "deeply accusative": Whatever the organization of their morphological case marking, their syntax is always accusatively based (Anderson, 1976). Some grammatical theories are still based on this notion. Of course the basic grammatical category in accusative systems is that of subject. In accusative systems, if any case is formally unmarked, it is normally the subject. If any case is required in all clauses, it is the subject. The citation form of nouns is usually the subject form. The function of subjects is clear: They are essentially grammaticized clause topics. Accusative case organization is accordingly prevalent in constructions involving clauses and combinations of clauses: Syntactic structures. In fact, many languages with alternative case categories in their morphologies do exhibit accusatively organized syntax. The formal grammaticization of a subject category is not universal among languages, however. Some languages exhibit no evidence of an accusative base, even in their syntactic patterns, as has been demonstrated by Li and Thompson (1976), Dixon (1979), and Bossong (1984) among others.

The grammaticization of a subject category is not always a matter of chance. Whether or not a particular language develops this category can be affected by structural features elsewhere in the grammar. One of these is the presence of a full set of obligatory, morphologically bound pronouns, in at least three persons and two cases, based on case distinctions other than nominative versus accusative. In what follows, the case systems of two different types of non-accusative languages will be examined: One agent/patient and one ergative/absolutive. Neither has formally grammaticized a subject category within either the morphology or syntax. It will be shown that this lack of grammaticization is related to common characteristics of their grammars that diminish the motivation for such a development.

1. CODING

The grammaticization of a case system in a language is most obvious in the coding, i.e. formal marking of case categories. The case suffixes on nouns

in Latin and Russian, for example, provide clear evidence of a subject category in those languages. The forms of their pronouns provide similar evidence: Pronouns referring to subjects have different forms from those referring to objects, possessors, etc.

In Cayuga, an Iroquoian language, formal case marking is concentrated in the pronominal system. All verbs contain pronominal prefixes referring to their core or primary arguments. The forms of the prefixes reflect an agent/patient or active case system. Intransitive verbs contain a single pronoun whose form reflects the semantic role of the single core participant. When an event involves only an agent, the pronominal prefixes have the forms in (1). The data cited here come from the speech of Reginald Henry of Six Nations, Ontario.

(1) Cayuga intransitives: Agents

kǫhté·kye?s '***I*** am leaving'	*katkę́hęhs* '***I*** get up'
sǫhté·kye?s '***you*** are leaving'	*satkę́hęhs* '***you*** get up'
hahté·kye?s '***he*** is leaving'	*hatkę́hęhs* '***he*** gets up'

When the event involves only a patient, a different set of pronominal prefixes is used. Compare the pronouns in (2).

(2) Cayuga intransitives: Patients

akáht?aǫh '***I*** have gotten full'	*akétshahni?k* '***I*** am afraid'
sǫ́htá?ǫh '***you*** have gotten full'	*satsháhni?k* '***you*** are afraid'
hohtá?ǫh '***he*** has gotten full'	*hotsháhni?k* '***he*** is afraid'

Transitive verbs in Cayuga contain complex pronominal prefixes referring to both agent and patient. (Agent and patient components of most of the transitive prefixes have become fused morphologically.)

(3) Cayuga transitives: Agents and patients

aské·kę?	'***you*** saw ***me***'	*aha·ké·kę?*	'***he*** saw ***me***'
ahehsé·kę?	'***you*** saw ***him***'	*ahé·kę?*	'***I*** saw ***him***'
ashé·kę?	'***you*** saw ***her***'	*asha·kó·kę?*	'***he*** saw ***her***'

(Cayuga contains many more prefixes than those cited here, around sixty in all. Pronouns distinguish singular, dual, and plural number, first, second, and third person, exclusive and inclusive in first person, and masculine, feminine, and neuter gender in third person, in addition to agent and patient case.) The classification of primary participants into agents and patients is semantically based, but it is fully grammaticized. Some agents may seem semantically more

active than others: A person who cuts a rope seems more agentive than one who sees a rope, but they are categorized equivalently as agents by the grammar. A rope that is cut seems more affected than one that is only seen, but both are classified as patients grammatically.

These prefixes are not simple agreement markers: They are full referential pronouns, representing the core arguments themselves. Verbs containing them constitute grammatically complete clauses on their own. When separate noun phrases are also present in a clause, the pronominal prefixes remain in place. The noun phrases function essentially as appositives to the pronouns and bear no case marking (Jelinek, 1984; Mithun, 1986a; 1986b).

(4) Cayuga clauses with appositive noun phrases

Ha-hní	***neʔ so·wa·s***
MASC.AGT-bark.HABITUAL	***the dog***
he-bark-s	***the dog***
'The dog is barking.'	

Hy-ʔaní	a-***hé·-kę-ʔ***
3/2-father	AOR-**1/MASC**-see-PUNCTUAL
your father I saw ***him***	
'I saw your father.'	

Selayarese, an Austronesian language of Indonesia, also shows no subject category in its formal marking. Like Cayuga, it contains obligatory bound pronouns, but these operate on an ergative/absolutive basis. The Selayarese data cited here come from the speech of Hasan Basri of South Sulawesi.

The single argument of an intransitive verb is represented by an absolutive pronominal enclitic. Note that the forms of the absolutive enclitics are the same whether they refer to agents or patients.

(5) Selayarese intransitives: Absolutive enclitics

*aʔlúmpaʔ**a***	'***I*** jump'
*aʔlúmpak**ko***	'***you*** jump'
*aʔlúmpak**ki***	'***we (incl)*** jump'
*aʔlúmpak**kaŋ***	'***we (excl)*** jump'
*aʔlúmpaʔ**i***	'***he/she/they*** jump'
*máŋŋaŋ**a***	'***I*** am tired'
*máŋŋaŋ**ko***	'***you*** are tired'
*máŋŋaŋ**ki***	'***we (incl)*** are tired'
*máŋŋaŋ**kaŋ***	'***we (excl)*** are tired'
*máŋŋaŋ**i***	'***he/she/they*** are tired'

In transitive verbs, the absolutive enclitic refers to the semantic patient. The semantic agent is indicated by an ergative prefix.

(6) Selayarese ergative prefixes and absolutive enclitics

kuísse?i	'***I*** know him'	*laísse?**a***	'he knows ***me***'
muísse?i	'***you*** know him'	*laíssek**ko***	'he knows ***you***'
ri?ísse?i	'***we (incl)*** know him'	*laíssek**ki***	'he knows ***us (incl)***'
toísse?i	'***we (excl)*** know him'	*laísse**kaŋ***	'he knows ***us (excl)***'
laísse?i	'***he/she/they*** know(s)	***him/her/it/them***'	

(Neither number nor gender is distinguished in Selayarese pronouns. The inclusive form also functions as a formal second person pronoun. Verbs unmarked for tense are used to describe both present and past events.)

In Selayarese, as in Cayuga, verbs with bound pronouns constitute grammatically complete clauses in themselves. When separate noun phrases are also present in the clause, the bound pronouns remain. The noun phrases are unmarked for case.

(7) Selayarese clauses with appositive noun phrases

*Máŋŋaŋ-**i** **hásaŋ** aló-nni*
tired-**3.ABS** **Hasan** day-this
he is tired **Hasan** today
'Hasan is tired today.'

La-*pállu-**i*** ***berasá-ñjo*** ***i-hásaŋ***
3.ERG-cook-**3.ABS** **rice-the** **Mr-Hasan**
he cook **it** **the rice** **Mr. Hasan**
'Hasan cooked the rice.'

Neither Cayuga nor Selayarese thus exhibits any evidence of a subject category in its formal case marking. Morphological case marking is agent/patient based in Cayuga and ergative/absolutive based in Selayarese.

2. BEHAVIOR

Evidence of the grammaticization of a category need not be limited to formal markers. If certain kinds of constituents all behave grammatically in a particular way or if they trigger certain grammatical patterns, and no others do, this group is often said to constitute a grammatical category. The hypothesis that all languages are "deeply" or syntactically nominative/accusative was proposed on this basis.

The most common syntactic areas in which evidence of the grammaticization of subject has been sought are in imperatives, coordination, and subordination.

2.1. Commands

Imperative constructions have been taken as evidence of the grammaticization of a subject category in many languages. It has been suggested that in English, for example, subjects pattern as a group in that they are systematically deleted from imperatives. Second person subject pronouns are omitted from all commands, whether intransitive (___ eat!) or transitive (___ eat your beans!).

Cayuga offers no equivalent evidence of a grammaticized subject category, because pronominal prefixes are present in every verb, including imperatives. Note the second person pronouns in the commands in (8).

(8) Cayuga imperatives with bound pronouns

***S**-an-ahá-owe·k,*	*s-ak-ya?t-awí-?t*	*hni?*
2.AGT-SEMI.REFL-crown-cover	**2.AGT**-body-encircled-CAUS	too
you put on your hat	***you*** put on your coat	too

'Put on your hat and coat.'

Selayarese formal imperatives are structurally equivalent to those in Cayuga. In formal commands, the bound pronouns remain, whether the commands are transitive or intransitive.

(9) Selayarese formal imperatives with bound pronouns

*Máŋe-**ko***
go-**2.ABS**
you-go
'Please go.'

*Bérasa **mu**-pállu-**i***
rice **2.ERG**-(TRANS).cook-**3.ABS**
rice ***you*** cook ***it***
'Please cook the rice.'

Informal commands, however, may appear without pronouns. Absolutive pronouns may be omitted from intransitive verbs, and ergative pronouns from transitive verbs.

(10) Selayarese informal commands without pronouns
Máŋe(-**Ø**)
(**Ø**)go
'Go!'

(**Ø**-)Pállu-i
(**Ø**-)(TRANS).cook-3.ABS
'Cook it!'

These informal commands might at first be taken as behavioral evidence of a subject category in Selayarese. The category represented here is not subjecthood, however, but agency or control, as pointed out by Dixon (1979), Anderson (1979), and others. Just as in English, commands can only be felicitously addressed to someone in a position to carry them out. When the second person is not in control, Selayarese verbs with second person pronouns constitute statements, not formal commands. If the pronouns are omitted, the result is not an ungrammatical informal command, but a simple verb stem, pragmatic nonsense.

(11) Selayarese pragmatic pattern: Control
máŋŋaŋko 'you are tired' *máŋŋaŋ* 'to be tired'
bássoroko 'you are full' *bássoro* 'to be full'

Selayarese commands provide no more evidence of a grammaticized subject category than their English counterparts. In both languages, second persons in control, not subjects, are omitted from informal commands. This represents a pragmatic pattern rather than a syntactic restriction.

2.2. Coordination

A second area in which behavioral evidence of a particular case organization is sometimes sought is clause coordination. Dixon (1972) proposed that in Dyirbal, for example, clauses sharing a coreferent nominal can be conjoined only if that nominal appears in absolutive case in both clauses. This restriction was considered evidence of a grammaticized absolutive category.

The omission of coreferent nominals from coordinate constructions has also been cited as evidence of certain case categories in various languages. In English conjoined sentences, for example, the subject of the second clause may be omitted if it is coreferent with the subject of the first. Speakers can thus say either of the two sentences below.

(12) a. *He picked up his suitcase and he left*
b. *He picked up his suitcase and left*

This pattern is restricted to subjects. English speakers do not omit the second of two direct objects.

(13) **He picked up his suitcase and carried out*

The noun phrase omitted from the second clause must be coreferent with the *subject* of the first clause, not some other constituent. There is no doubt from the sentence below about who left.

(14) *He kissed her and left.*

Cayuga exhibits no syntactic restrictions on the argument structure of coordinate clauses. Any pragmatically relevant clauses may be conjoined. Furthermore, since all Cayuga verbs contain pronominal prefixes, there is no deletion of coreferent noun phrases under coordination. The pronominal prefixes in the verbs in (15) refer to the same second person singular agent, but neither can be omitted.

(15) *Cayuga conjunction with coreferent agents*

*ę-t-**hs**-é-ʔ*	*kiʔ*	*kyę·ʔ*	*waʔtsí*
FUT-CISLOC-**2.SG.AGT**-GO-PUNCTUAL	just	then	later
you will come	just	then	later

*ę-t-**hs**-háwi̱-ht-ahk.*
FUT-CISLOC-**2.SG.AGT**-carry-CAUS-INST-PUNCTUAL
you will carry with you
'Why don't you just come on over later, then, and bring it along.'

Selayarese, like Cayuga, exhibits no syntactic restrictions on the argument structure of coordinate clauses. Bound pronouns are also not omitted under conjunction. In (16) for example, pronouns remain with each clause, even though they are coreferent.

(16) Selayarese conjunction

Maŋe-i	**ku**-*alle*	*anák-ku,*	*nampaʔ-**a***	*mulíam*	*maŋe*
go-3.ABS	**1.ERG**-take	child-1.POSS	then-**1.ABS**	go	toward
go her	***I*** take	my daughter	then ***I***	go	toward

ri sápo nampaʔ-a tínro.
to house then-**1.ABS** sleep
to house then ***I*** sleep
'I picked up my daughter, went home, and went to sleep.'

All would be subjects in English, as can be seen from the translation.

Since neither Cayuga nor Selayarese contains syntactic restrictions on the argument structure of conjoined clauses, and no pronouns are omitted from conjoined clauses, conjunction provides no evidence of the grammaticization of a subject category in either language.

2.3. Nominalized clauses

In many languages, clauses are subordinated by nominalization. This is often accomplished with possessive forms. In English nominalized clauses, the subject appears as a possessor whether it functions semantically as an agent or patient, and whether the clause is intransitive or transitive.

(17) *I left* (Agent of intransitive)
[My leaving] disturbed him

(18) *I was sick* (Patient of intransitive)
[My being sick] slowed things down

(19) *I threw out the comics* (Agent of transitive)
[My throwing out the comics] was a mistake

Cayuga exhibits no comparable morphological nominalization of clauses. In Selayarese, however, nominalization with possessives is the most common subordinating device.

Basic possession in Selayarese is shown by pronominal suffixes on nouns, as in (20).

(20) Selayarese possessive pronominal suffixes

*asúŋ**ku***	'**my** dog'
*asúm**mu***	'**your (familiar)** dog'
*asún**ta***	'**our (inclusive)** dog'
*asúm**ba***'	'**our (exlusive)** dog'
*asún**na***	'**his/her/their** dog'

The same set of possessive suffixes is used in nominalized clauses. When the nominalized clause is intransitive, the possessor is of course the single participant, the absolutive. (The absolutive enclitic appears on the first element of the clause.)

(21) Selayarese intransitive clause nominalization
*Tájaŋ-***a** *rínni [saŋgéŋ-ku mulíaŋ].*
await-**1.ABS** at-this [until-**1.POSS** return]
await **me** here [until **my** return]
'Wait here until I return (until my return).'

When the clause is transitive, the possessor is the semantic patient, again the absolutive.

(22) Selayarese transitive clause nominalization
*Tájaŋ-**a** rínni [saŋgén-**na** **ku-**ʔúppa ammám-mu].*
await-**1.ABS** at-this [until-3.POSS 1.ERG-find father-**2.POSS**]
await **me** here [until **his** **I** find your father]
'Wait here until I find your father (until his finding).'

Subordination by nominalization thus provides no evidence of a grammaticized subject category in either Cayuga or Selayarese. In Cayuga, it yields no evidence of any category at all, and in Selayarese, it is consistently absolutive based.

2.4. Complex constructions

Evidence of a subject category has been sought in some languages in patterns sometimes referred to by "equivalent noun phrase deletion" and "raising". It has been proposed that in English for example, the subject of a complement clause is omitted if it is coreferent with the subject of its intransitive matrix clause.

(23) *Sally hopes [Hilda will win]*
(24) *Sally hopes [to win]*

Subjects of complements are also assumed to be omitted if they are coreferent with the object of a transitive matrix clause. In the sentence below, the performer must be Hilda.

(25) *Sally invited Hilda [to perform]*

Cayuga exhibits no coreferent deletion at all, since all verbs contain obligatory pronominal prefixes referring to their agents and/or patients. All of the first person agent pronouns must remain with their verbs in the sentence in (26), even though they would be coreferent subjects in English.

(26) Cayuga complement of intransitive
A·yę́·ʔ k-ą́hska·né·-s o-thowʔe-ké
INFER **1.AGT**-want-HAB N-cold-LOC
seems **I** wish at it's cold

h-á·-k-e·-ʔ,
TRANSLOC-OPT-**1.AGT**-GO-PUNC
I would go there

*a·**-k**-ę-nat-áę-hna-ʔ*
OPT-**1.AGT**-SEMI.REFL-camp-lay-PURP-PUNC
'I want to go up north **I** would go camp to camp.'

All pronominal prefixes also remain in place when the matrix clause is transitive.

(27) Cayuga complement of transitive
*ę-**kashe**-hǫká-ǫ-ʔ hniʔ*
FUT-**2/3.PL**-invitation-give-PUNC too
you will invite **them** too

*ę-t-**kaǫ**-te-kh-ǫny-áhne-ʔ hwis-hǫ·ʔǫh.*
FUT-CISLOC-**3.PL**-REFL-meal-make-PURP-PUNC five-DISTR
they will come eat five-ish
'Invite them to come eat around five o'clock, too.'

Selayarese differs in an interesting way from Cayuga. When both the matrix and the complement clause are intransitive, the complement verb may contain no clitic, as in (28).

(28) Selayarese intransitive complement
*Aʔjáñji-**a** la-mináhaŋ.*
promise-**1.ABS** FUT-(INTR).follow
I promise will follow
'I promise to follow.'

This might at first appear to exemplify coreferent subject deletion. When the coreferent noun phrase is the agent of a transitive complement clause (ergative), however, no omission is possible.

(29) Selayarese transitive complement
*A?jáñji-**a*** *la-**ku**-pináhaŋ-**ko***
promise-**1.ABS** FUT-**1.ERG**-follow-**2.ABS**
I promise **I** will follow **you**
'I promise to follow you.'

Only coreferent absolutives can be omitted.

A second kind of subordinate construction, sometimes called "raising", has also been cited as evidence of a subject category in some languages. According to raising analyses, underlying subjects of complement clauses appear as derived subjects or objects of matrix clauses.

(30) *It [John is happy] seems → John seems happy*

(31) *I saw it [that she left] → I saw her leave*

(Underlying subjects of subject complements appear as derived matrix subjects. Underlying subjects of object complements appear as derived matrix objects.)

As usual, Cayuga agent and patient pronouns remain with their verbs, regardless of syntactic context.

(32) Cayuga complement of intransitive
*T-**ka**·k-ǫ́·t* *kye·?ǫ́* *hne·?* *o·nę́* *ę-hs-ní·nǫ-?*
CISLOC-N-necessary INFER CONTR NOW FUT-**2.AGT**-buy-PUNC
it is necessary guess but now **you** will buy
'But you'll have to buy that (the lard).'

(33) Cayuga complement of transitive
*A-**kǫ**·-ké-?* *sǫ·té?* ***s**-yę́thw-áhsǫ-h*
AOR-**1/2**-see-PUNC last night **2.AGT**-plant-DISTR-STATIVE
I saw **you** last night **you** were planting things
'I saw you planting last night.'

Selayarese presents a slightly different picture. Constructions like the negative sentence below have been analyzed as the result of raising in some related languages. It has been posited that the pronoun 'I' originates with the second verb 'tired', then is raised to the first, 'not'.

(34) Selayarese complex predicate
*Géle-**a*** *máŋŋaŋ*
not-**1.ABS** tired
I not tired
'I am not tired.'

Similar constructions can be found with other intransitive verbs in Selayarese.

(35) Selayarese complex predicates
A?ra?-**a** *ŋ-ánre*
want-**1.ABS** INTR-eat
I want eat
'I want to eat.'

*Máŋe-**i*** *n-río.*
go-**3.ABS** INTR-bathe
he go bathe
'He went to take a bath.'

The first verb may be transitive.

(36) Selayarese transitive matrix
*La-jáñjaŋ-**ki*** *?-bisára*
3.ERG-see-**1.INCL.ABS** INTR-talk
he saw **us** talk
'He saw us talk.'

When the second verb is transitive, however, the "raised" pronoun refers to the semantic patient. The construction involves absolutives, not subjects.

(37) Selayarese transitive complements
*Géle-**a*** *la-jáñjaŋ.*
not-**1.ABS** 3.ERG-see
not-**me** he see
'He didn't see me.'

*A?rak-**ko*** *ku-jáñjaŋ.*
want-**2.ABS** 1.ERG-see
want-**you** I see
'I want to see you.'

*Máim-mo-**ko*** *ku-saré-aŋ.*
finish-PERF-**2.ABS** 1.ERG-give-DAT
have-**you** I give it
'I have already given it to you.'

The two patterns described here could be interpreted as absolutive-based syntactic processes, whereby coreferent absolutive arguments of complements are deleted and raised respectively. Absolutive pronouns in Selayarese are

actually enclitics rather than suffixes. They are suffixed to the first element of clauses, not simply to verbs. The first word is often a verb, but not always; note the absolutive enclitic on the conjunction *nampa* 'and then' in (16) and on *saŋgen* 'until' in (21). It is not surprising that they should have scope over the entire complex clause.

2.5. Relativization

The formation of relative clauses in many languages is sensitive to case, often to the category of subject. In some languages, for example, heads of relative clauses must be their subjects. Neither Cayuga nor Selayarese exhibits any such restrictions.

In Cayuga, strict formal identity between the head and constituents of the relative clause is not necessary. In (38), 'someone' would normally be considered third person feminine-indefinite singular, but the associated verb 'marry' has a first person plural pronominal prefix.

(38) Cayuga relativization without formal identity

Thę?	*ni?* *t'-áǫ*	*a·-**ke**-tshę́i-?*	***sǫ**ká·?a*
not	*just* NEG-OPT	OPT-**1.AGT**-find-PUNC	**someone**
not	just impose	**I** would find	**someone**

*a·-**yǫ**khní-nya·k.*
OPT-**1.DU.PAT**-marry
we two would marry
'I can't find anyone to marry me.'

Even when they are equivalent, all pronouns remain with their verbs in both the matrix and relative clauses. (The feminine-indefinite singular is used for a people.)

(39) Cayuga relativization with formal coreference

Tho	*ti?*	*ni·-yo-ht*	*nę·kyę*
there	just	so-it-is	this

'And that is the way

*a-k-**yako**-ihw-áęt-a?-s.*
AOR-DU-**FEM.PAT**-matter-lay-INCH-DAT
one decided
they decided,

nę·kyę ne ǫ·kwe ***e-ná·kre?***
this the person **FEM.SG.AGT**-dwell.STAT
this the person **one** lives
the people who were living there.'

In Selayarese, an optional prefix *to-* 'who' on the verb of the relative clause indicates that the head is human, *nu-* that it is non-human. There are no case restrictions on the shared constituent in either the matrix or relative clause. The bag in the sentence below functions as an absolutive in the matrix clause and as a locative in the relative clause.

(40) Selayarese relativization with locative head
Ku-*?úppa-mo-**i*** *tási (nu)-**mu**-panaro-íñjo dóe?*
1.ERG-find-PERF-**3.ABS** bag REL-**2.ERG**-put-the money
I found **it** bag the **which** you put money
'I found the bag in which you put money.'

The same bound pronouns appear in relative clauses as in independent clauses, just as in Cayuga.

(41) Selayarese bound pronouns under relativization
Ku-*?úppa-mo-**i*** *ana?-ána?*
1.ERG-find-PERF-**3.ABS** child-child
I have found **him** child

*(to)-**la**-lukka?-íñjo-**i*** *asúŋ-ku*
that **3.ERG**-steal-the-**3.ABS** dog-1.POSS
the who *he*-stole it my dog
'I have found the child that stole my dog.'

Relativization thus provides no evidence of a grammaticized subject category in either Cayuga or in Selayarese.

3. WHY CAYUGA AND SELAYARESE NEED NOT GRAMMATICIZE A SUBJECT CATEGORY

It is clear why so many languages have grammaticized the notion of clause topic into a category of subject. Topic-comment structures are among the most frequently occuring in natural discourse in any language. Identification of the clause topic is crucial to understanding the point of most speech. The codification of this discourse category into a grammatical category is thus well motivated. Yet neither Cayuga nor Selayarese exhibits evidence of

the grammaticization of a subject category at all, even in syntactic constructions. This is no accident. Certain structures already present in their grammatical systems reduce the motivation for such grammaticization.

As is well known, speakers do not choose clause topics at random. They select one participant as a starting point, then say something about it (Chafe, 1976). This starting point is normally that participant with which the speaker chooses to identify. Speakers generally prefer first person topics to second, and second to third. They prefer humans to nonhumans, and agents to nonagents. Accordingly, the choice of subject, the grammaticized clause topic, usually depends upon person, humanness, and agency.

Most of these distinctions are already obligatorily specified in every clause in Cayuga and Selayarese. Both languages distinguish first, second, and third person within the bound affixes. Cayuga overtly distinguishes not only agents and patients, but humans and nonhumans within pronominal forms. Selayarese distinguishes agents and patients when not otherwise obvious: In transitive clauses. The obligatoriness of just these pronouns insures that the clause topic, the participant that would usually be the subject in an accusative language, is overtly represented in every Cayuga and Selayarese clause.

A subset of core participants in each language is unrepresented by the bound pronouns. Each of these subsets consists of participants that do not normally function as clause topics.

In Cayuga, all agents are represented by pronouns. All human patients are also represented. The transitive pronominal prefix *k-he-* 'I/her' for example, refers to a first person agent (*k-*) and a feminine patient (*-he-*).

(42) Cayuga transitive pronoun with human patient
E-ksaʷ-ʔá *a-**khé**·-kę-ʔ.*
FEM.child-DIM AOR-**1.AGT/FEM.PAT**-see-PUNC
little girl **I** saw *her*
'I saw a/the girl.'

Nonhuman patients are represented except in one context: When an agent is also present. The pronominal prefix *k-* 'I(/it)' contains no overt mention of the inanimate patient. (The following *e* is epenthetic.)

(43) Cayuga verb with human agent and nonhuman patient
Ka-nǫhs-ó·t *a-**k**-é·kę-ʔ.*
N-house-stand.STAT AOR-**1.AGT**-see-PUNC
house **I** saw
'I saw a/the house.'

Note that the same prefix is used in intransitive verbs.

(44) Cayuga verb with human agent only
*A-**k**-at-awę́-hna-ʔ.*
AOR-**1.AGT**-SEMI.REFL-bathe-PURP-PUNC
'I went swimming.'

Nonhuman patients are not normally selected as clause topics when agents are also involved, exactly the situation presented by transitive verbs.

In Selayarese transitive clauses, agents are represented by ergative pronouns and patients by absolutive pronouns, whether they are human or not.

(46) Selayarese agent and patient of transitive
Ku**-hálliʔ-**i *sapó-ñjo.*
1.ERG-buy-**3.ABS** house-the
I buy **it** the house
'I bought the house.'

If the patient of a transitive is indefinite, however, it is not represented by a pronoun. The verb is then formally intransitive, and the agent is absolutive.

(47) Selayarese semantic transitive with indefinite patient
*M-máli-**a*** *sápo.*
INTR-buy-**1.ABS** house
I buy house
'I bought a house.'

Indefinite patients of semantic transitives are not normally topical. Topics, the starting points of predications, usually represent established information, a point of departure for other, new information. Indefinites are by definition not established. Furthermore, transitive verbs have agents, and agents are usually preferred over patients as topics.

Thus in both Cayuga and Selayarese, all arguments that commonly function as topics are obligatorily represented by morphologically bound pronouns in every clause.

3.2. Discourse-level topicality: Point of view

Although subject selection in accusative languages is normally correlated with such features as person and agency, the ultimate choice is not simply the result of a mechanical calculation of these features in each clause. On occasion, patients of transitive events can function as subjects in languages

like English. The deviation from usual subject choice is motivated by the structure of larger segments of discourse.

Subjects establish an orientation. Speakers do not ordinarily shift this orientation with each new clause; they tend to establish a topic or point of view and retain it over a certain stretch of discourse. (See among others Givón, 1983.) This topic may not function as an agent in every single clause of the general discussion. It may even function as the patient of some event involving another agent. At this point the speaker must decide whether to shift the general orientation of the discourse to the vantage point of the other agent, or to continue from the vantage point of the original topic. Constructions like English passives permit speakers to maintain topic continuity by categorizing patients as subjects.

Because speakers normally retain the same topic over a series of clauses, hearers expect topic continuity until there is a reason for a reorientation. When there is a shift in vantage point or scene, this is usually indicated not only by a shift in subject, but by other devices as well that overtly introduce new topics and alert hearers to the changes.

Cayuga and Selayarese both contain ample stylistic devices for alerting hearers to discontinuity and for foregrounding and backgrounding primary participants. They contain special presentative constructions to introduce new participants. They contain devices such as shifting constituent order, demonstratives, and derivational morphology to indicate the discourse roles of partipants. In both languages, full noun phrases are used to alert hearers to topic shifts.

Consider for example the Cayuga passage in (47) from the cosmology legend. A woman has fallen through a hole in the sky, and as she falls, numerous thoughts go through her mind. Suddenly she sees some birds flying around. The discourse topic then shifts from the woman to the birds. The shift is indicated by several devices. The initial particles in the first line signal a reframing. The birds, who had just been introduced ('and then she noticed some birds') reappear at the beginning of the second line. The ordering of this noun for 'birds' early in the clause, before the verb, serves to foreground it, indicating that it represents significant new information, in this case, a shift in topic. The proximate demonstrative *nę·kyę* 'this' indicates that the birds occupy center stage.

(47) Cayuga
Ne·ʔne? o·nę nę·kyęh
it·is the now this

t?itę?shǫ́·?ǫ teyotiya?towéhtǫh atkatiya?to·wéht
birds they are thinking they thought

ne? nǫ ne? nę·kyę a·kenat?enyę·tę́?
the now the this they should try

aǫsa·kotiy?ata·kó?
they would help her

'Now then at this time these birds were thinking that they should try to help her.'

On some occasions, events may imply the involvement of agents that are peripheral to the discourse as a whole. Cayuga contains devices for conveying such events by means of verbs whose only argument is the patient. The passage cited above continues:

*ahkwíkwa? nę·kyę a·-**yakó**-nǫhnya?-k*
that·not this OPT-**FEM.PAT**-hurt-PUNC
that not this **she** would get hurt

ne·? k?ishę́ nę·kyę ne? a·-yakó·-nyo·-?.
it·is maybe this that OPT-**FEM.PAT**-kill-PUNC
or else this that she would get killed
'so that she would not get hurt or get killed.'

The verbs 'hurt' and 'kill' normally involve agents, but in this context their identity was unknown and not worthy of mention.

Cayuga thus contains not only obligatory mention of all clause topics in every verb, it also contains elaborate devices for foregrounding and backgrounding participants according to their roles within the discourse. Sets of particles identify major discourse breaks, shifts in scene. Full noun phrases and word order identify important new participants. Demonstratives point out the locus of center stage. The lexicon and verbal morphology permit the elimination of peripheral agents from consideration as topics.

Selayarese contains similar devices for tracing discourse topics. Significant new participants are commonly introduced with a presentative construction based on *rie?* 'there is/was'. A story about a policeman opened as follows.

(48) Selayarese presentatives
*Ri táuŋ 1979, **ríe?** pulísi ap-pa-katiánaŋ*
in year 1979 ***there*·is** policeman INTR-CAUS-pregnant

ri Silájara
in Selayar

'In the year 1979, there was a policeman who made somebody pregnant in Selayar.'

The story continued as the policeman promised to marry the pregnant woman, then went home, ostensibly to get money for the wedding. Once home, however, he married someone else. The pregnant woman's relatives were outraged. When three of them came forward to seek revenge, they were introduced as follows.

(49) Selayarese presentatives
Saŋgén-na ***ríe***ʔ*-mo* *tállu,*
finally-3.POSS **there·is**-PERF three

totarúna táŋgoŋ *la-*ʔ*-bálasa.*
teenager guarantee FUT-INTR-avenge

'At last there were three of them, teenagers who guaranteed they would take revenge.'

The teenagers became the new topic of the narrative. They awaited the policeman outside of a movie theater and attacked him.

Shifts in scene and topic are often signaled in Selayarese, as in other languages, by the use of full noun phrases. Once established, discourse topics are referred to by pronouns only, until there is a shift to another topic, at which point a full noun phrase is again used. Note the appearance of the noun phrases underlined in the passage below. The speaker had just said that after about a month at home the policeman returned to Selayar.

(50) Selayarese full noun phrases for topic shift
*La-laŋŋeré*ʔ*-na* ***to-toá-na*** ***to-bahiné-ñjo,***
3.ERG-hear-3.POSS **who-old-3.POSS** **who-woman-the**

*lako ríe*ʔ*-mo,* *muliam pulisí-ñjo,*
that there·is-PERF return policeman-the

máŋe-mo-i *póle* *la-kutáʔnaŋ,*
go-PERF-3.ABS again 3.ERG-ask

*sikuraía na-*ʔ*á*ʔ*ra*ʔ*-i* ʔ*-bóntiŋ.*
when when-want-3.ABS INTR-marry

Míŋka ***pulisíñjo*** *to-ríe*ʔ*-mo-i* *bahinén-na*
but **policeman-the** who-there·is-PERF-3.ABS woman-3.POSS

ka ri maŋé-na ri kampón-na,
because at go-3.POSS to village-3.POSS

géle-i maŋe ŋéra doe? pa?bóntiŋ
not-3.ABS go ask·for money wedding

míŋka máŋe-i ?-bóntiŋ.
but go-3.ABSINTR-marry

'When the **woman's parents** heard that the policeman had returned, they again went to ask him when he wanted to get married. But *the policeman* already had a wife, because when he went to his village, he didn't go to ask for wedding money, but he went to get married.'

In English, patients of transitives are often encoded as subjects when the agent is indefinite or nonhuman. English speakers thus rarely say *A car hit me* or *A dog bit me*. Speakers are more likely to assume their own point of view as a point of departure: *I was hit by a car* or *I got bitten by a dog*. Selayarese has no equivalent construction. In a normal context, ordinary transitive verbs would be used.

(51) Selayarese transitives with nonhuman agents
La-*lúpuru-****a*** *óto.*
3.ERG-hit-**1.ABS** car
it hit **me** car
'I was hit by a car' = 'A car hit me'

La-*kókko?-****a*** *?ásu.*
3.ERG-bite-**1.ABS** dog
it bit **me** dog
'I was bitten by a dog' = 'A dog bit me'

Since person is clear from the pronouns, no information is lost.

Noun phrases can be foregrounded if they are of unusual importance. The sentences below with initial noun phrases would be used contrastively, as in 'It was a *car* that hit me, not a truck' or 'It was a *dog* that bit me, not a bear', or as the answers to questions 'What hit you?' or 'What bit you?'.

(52) Selayarese fronting for contrast
óto la-lúpuru-a
car **3.ERG**-hit-**1.ABS**
car **it** hit **me**
'*A car* hit me.' (not some other vehicle)

ásu ***la-****kókko?-****a***
dog **3.ERG**-bite-**1.ABS**
dog **it** bit **me**
'**A dog** bit me.'

Like Cayuga, Selayarese contains a device for eliminating unimportant agents from the set of primary participants. A derivational prefix *ta-* forms verbs indicating lack of control. Some non-control verbs have control counterparts.

(53) Selayarese non-control verbs with control counterparts

péla?	'throw (something) away'
tappéla?	'get lost'
pabáŋka	'surprise (someone)'
ta?báŋka	'be surprised'
sámbaŋ	'catch (someone or something)' (as string catching foot)
tassámbaŋ	'trip' (intransitive)

The prefix is derivational. It forms new lexical items. Many verbs containing the prefix have no counterparts without it.

(54) Selayarese non-control verbs without control counterparts

ta?béssolo	'slip'	(no **béssolo*)
ta?dó?do?	'be sleepy'	(no **dó?do?*)
tattó?ro	'stumble'	(no **tó?ro*)

The non-control verbs above are intransitive. (The glottal stop, which assimilates to following voiceless consonants, is an intransitive marker.) The absolutive pronouns that appear with these derived intransitives represent semantic patients.

(55) Selayarese absolutive patients of non-control verbs

*ta-p-péla-****a***	*ta-?-béssolo-****s***
NON.CONTROL-INTR-lose-**1.ABS**	NON.CONTROL-INTR-slip-**1.ABS**
'**I** got lost'	'**I** slipped'

*ta-ʔ-báŋkaʔ-**a***
NON.CONTROL-INTR-surprise-**1.ABS**
'**I** am surprised'

*ta-ʔ-dóʔdoʔ-**a***
NON.CONTROL-INTR-sleepy-**1.ABS**
'**I** am sleepy'

*ta-s-sámbaŋ-**a***
NON.CONTROL-INTR-trip-**1.ABS**
'**I** tripped'

*ta-t-tóʔro-**a***
NON-CONTROL-INTR-stumble-**1.ABS**
'**I** stumbled'

The prefix *ta-* can function as a discourse device for eliminating unimportant agents by creating lexical items that do not include these agents among their core arguments. The effect can be similar to that of passivization in English.

(56) Selayarese counterparts of English passive

ku-pállu-i	*ta-p-pállu-**i***
1.ERG-(TR)cook-**3.ABS**	NON.CONTROL-INTR-cook-**3.ABS**
'**I** cooked it'	'**It** is cooked'

ku-sássa-mo-i
1.ERG-(TR)wash-PERF-**3.ABS**
'**I** have washed it'

*ta-s-sássa-mo-**i***
NON.CONTROL-INTR-wash-PERF-**3.ABS**
'**It** has been washed'

The effect of this device can be seen by comparing the two sentences below containing relative clauses. In the first, the agent is overt. In the second, the agent is bypassed by the use of the derived verb 'get stolen'.

(57) Selayarese use of non-control for backgrounding agents

ku-ʔúppa-mo-i	*ásu*	*(nu)-**la**-lúkkaʔ-íñjo*	*i-Báso.*
1.ERG-find-PERF-**3.ABS**	dog	that-**3.ERG**-steal-the	Mr.-Baso
I have found **it**	dog	that **he** stole the	Mr. Baso

'I have found the dog that Baso stole.'

ku-ʔ*úppa-mo-****i*** *ásu (nu)-ta-ʔ-lukka*ʔ*-íñjo.*
1.ERG-find-PERF-**3.ABS** dog that-**PAT-INTR**-steal-the
I have found **it** dog the that was stolen
'I have found the dog that was stolen.'

4. CONCLUSION

The motivation behind the codification in so many languages of the discourse role of clause topic into a grammatical category of subject is clear. Topic-comment structures are among the most frequent discourse patterns in any language, and identification of the topic, the starting point of a predication, is crucial to its understanding.

Yet neither Cayuga nor Selayarese has grammaticized a subject category in any area of the language. Their formal case marking systems reflect the grammaticization of subtly different distinctions from the clause topicality grammaticized in accusative systems, and their syntactic patterns are sensitive to the distinctions of these systems or none at all. Cayuga grammatical constructions are based either on agent and patient categories or exhibit no case restrictions whatsoever. Selayarese constructions are based on ergative and absolutive categories, control, or no case category at all. The lack of a grammaticized subject category in these languages is not an accident. All participants that function as clause topics are already obligatory represented in every clause by overt morphologically bound pronouns. The features that usually enter into the selection of subjects in accusative languages, namely person and agentivity, are systematically distinguished on the pronouns.

In situations where the clause topic does not correlate with usual person, agentivity, and animacy preferences, other factors guarantee that topicality remains clear. Morphological devices are used to eliminate nontopical agents from the set of core arguments of a predicate. Cues to discourse-level topicality abound in these languages in the systematic use of demonstratives and word order, as well as in the distribution of full noun phrases. As in all languages, hearers generally assume that the topic will remain constant until it is overtly shifted. Although topics themselves are not expressed by a single grammatical category in these languages, shifts in topic, situations worthy of special attention, are signalled by a rich set of devices.

Obviously topicality does not play a lesser discourse role in Cayuga and Selayarese than in languages like English. Cayuga and Selayarese speakers

normally speak as coherently as speakers of other languages, exhibiting the same kinds of topic continuity. The patterns have simply not been formally grammaticized to the same extent. Full grammaticization of a special subject category would be superfluous, given the shapes of the existing grammars. There is thus considerably less motivation for its development. A full understanding of the processes of grammaticization must involve not only a knowledge of the possible sources of grammatical markers and categories, but also of the factors that motivate their development, especially their potential interaction with the grammatical systems in which they emerge. Such interaction is bound to be complex, involving multiple factors and varying degrees of motivation or hindrance. Yet as we discover more about these interactions, we have much to gain in our overall understanding of the forces that mold grammatical systems, of why languages are as they are.

ABBREVIATIONS

AGT	agent
CISLOC	cislocative
DIM	diminutive
INCH	inchoative
INTER	inferential
N	neuter
OPT	optative
PAT	patient
SEMI-REFL	semi-reflexive

REFERENCES

Anderson, John. 1979. "On being without a subject." Bloomington, Indiana: Indiana University Linguistics Club.

Anderson, Stephen. 1976. "On the notion of subject in ergative languages." In *Subject and Topic*, Charles N. Li (ed.), 1–23. New York: Academic Press.

Bossong, Georg. 1984. "Ergativity in Basque." *Linguistics* 22:341–392.

Bright, William. 1957. *The Karok Language*. Berkeley: University of California Press.

Chafe, Wallace L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view." In *Subject and Topic* Charles N. Li (ed.), 25–56. New York: Academic Press.

Cooreman, A., Fox, B. and Givón, T. 1984. "The discourse definition of ergativity." *Studies in Language* 8:1–34.

Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland.* Cambridge: Cambridge University Press.

Givón, Talmy (ed.). 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study.* Amsterdam: John Benjamins.

Jelinek, Eloise. 1984. "Empty categories, case, and configurationality." *Natural Language and Linguistic Theory* 2:39–76.

Mithun, Marianne. 1986a. "Disagreement: The case of pronominal affixes and nouns." In *Proceedings of the Georgetown University Round Table Conference on Languages and Linguistics*, James Alatis and Deborah Tannen (eds), 50–66. Washington D.C.: Georgetown University Press.

Mithun, Marianne. 1986b. "When zero isn't there." *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, K. Nikiforidou et al. (eds), 195–211. Berkeley: Berkeley Linguistics Society.

Verbs to Promotional Suffixes in Sahaptian and Klamath

Noel Rude
University of Oregon

INTRODUCTION

Nez Perce, Sahaptin, and Klamath are American Indian languages of the southern plateau region of the northwestern United States. Together Nez Perce and Sahaptin comprise the Sahaptian language family. Although Sahaptian and Klamath are related genetically, this paper presumes no particular subdivision within the greater Penutian phylum. In the past these languages have interested linguists mostly for phonological reasons. Nez Perce vowel harmony, beginning with Aoki (1966), has attracted considerable attention. And Bright (1982) cites thirteen references for Klamath, all of which deal with phonology. This paper, however, is concerned with the historical development of one peculiarity in the syntax of these languages. These languages have special voice constructions for the promotion of nonpatient case roles to object. The Nez Perce constructions have been described in Rude (1986a; 1986b; 1988a), and Rude (1988b) offers insight into the Klamath system. In all the languages these grammatical constructions are morphologically based and not dependent on word order. Objects are marked by NP suffixes and verbal agreement, and verbal suffixes code the case role of nonpatient objects. The purpose of this paper is to comment on the evolution of the system, and to show a verbal origin for the verbal suffixes. The process discussed — the derivation of verbal affixes from verbs — entails the structural merger of two clauses into one (i.e. two verbs become one) with an associated change in transitivity of the resultant single clause. It also includes the functional change from an earlier coding of a simultaneity of events to a later marking of objects of discourse/pragmatic importance.[1]

1. BENEFACTIVE/GENITIVE OBJECTS

In the Sahaptian languages a benefactive may be coded by an oblique NP case marker. In Nez Perce, for example, this is the suffix *-'ayn*, as illustrated in (1). The NP object, *walásna* 'knife', is case marked with the object suffix *-na*.

(1) *walás-na páa-ni-ya 'áayato-**'ayn***
knife-OBJ 3SUBJ.3OBJ-make-PAST woman-BEN
'He made a knife *for* the woman.'

A benefactive argument, however, is usually "shifted" to become the object.[2] The benefactive NP is then case marked as an object and the patient NP is unmarked. In each of the languages the benefactive case role is then coded in the verb, in the Nez Perce example (2) by the suffix ***-a'n***.

(2) *wálc páa-ny-**a'n**-ya 'áayato-na*
knife 3SUBJ.3OBJ-make-BEN-PAST woman-OBJ
'He made the woman a knife.'

The same verbal morphology which would otherwise code the patient in a transitive construction is now sensitive to the shifted benefactive. For example, when an object is 1st or 2nd person, the verb prefixes the same subject agreement markers as are otherwise found on intransitive verbs. In (2) above the verbal prefix is the transitive *páa-* which codes agreement with 3rd person subjects and 3rd person objects (in this case the 3rd person objectified benefactive). But in (3) *páa-* has been discarded in favor of *hi-*(here *ha-* by assimilation) because of the 1st person objectified benefactive. The prefix *hi-* otherwise codes the 3rd person subjects of intransitive verbs.

(3) *wálc **ha**-any-áa'n-ya 'íine*
knife 3NOM-make-BEN-PAST me
'He made me a knife.'

The same construction which codes benefactive objects also marks genitive objects. For example, (3) above could also mean, 'He made my knife'. Context or pragmatics disambiguates. For details, see Rude (1986b).

An analysis of the various allomorphs of the Sahaptian benefactive/genitive verbal suffixes reveals a complex consisting of the past marker *-e/-a* plus either of two morphemes which, as we shall see below, are of verbal origin. The two suffixes are phonemically *-'eni* and *-'ey* in Nez Perce and *-ni*

and *-y* in Sahaptin. Their occurrence is determined phonologically or dialectally.[3]

1.1. The Sahaptian benefactive *-'eni*

Nez Perce *-'eni* and Sahaptin *-ni* are identical in form to Nez Perce *'eni* 'give' and Sahaptin *ni* 'give'. The coding of datives and benefactives by the verb *give* is quite common in the world's languages.[4]

1.2. The Sahaptian/Klamath benefactive *-*'ey/-*oy.*

Though the Sahaptian verbal suffix *-*'eni* is identical to the Sahaptian verb 'give', the form *-*'ey* is not readily identifiable in Sahaptian. But in Klamath, a genetically related language of Oregon, dative/benefactive objects are marked by the verbal suffix *-ii*, a likely cognate of Sahaptian *-*'ey*.[5] And the Klamath dative/benefactive suffix *-ii* evidently derives from the Klamath verb *oy* 'give'. In Klamath the sequence *oy* regularly becomes *ii* when not in initial syllable. For example, *-oyGi* 'up' with the prefix *g-* 'go volitionally' is *goyGi* 'go up', but with the prefix *ngat'* 'jump' we have *ngat'iiGi* 'jump up'.

Thus, though Sahaptian *-*'eni* clearly derives from the verb for 'give' within Sahaptian, *-*'ey* would appear to be cognate with the verb for 'give' which survives in Klamath.[6]

2. LOCATIVE GOALS

In Sahaptian and Klamath several directionals or locatives may be coded alternatively as oblique NPs or object NPs and, like benefactive objects, code the semantic role by verbal suffix. The subsections below treat five such verbal suffixes.

2.1. The Sahaptian directive *-*éwe*

The two sentences (4) and (5) illustrate the Nez Perce ability to code the same intransitive proposition either by structurally intransitive or transitive clauses. In (4) the locative goal of the intransitive verb *páayn* 'arrive' is coded as an oblique NP marked by the locative suffix *-pe*, and the verb has prefixed the intransitive 3rd person subject agreement marker *hi-*. The subject NP *'áayat* 'woman' is not case marked (as is expected in intransitive clauses). In

(5) the same intransitive proposition is coded by a clause with transitive morphology; the locative goal has suffixed the object marker *-ne*, the subject the ergative marker *-m*, and the verb has prefixed the transitive 3rd person subject plus 3rd person object marker *pà-*. The semantic role of the object is coded in the verb by the suffix *-óo*.

(4) *'iníit-**pe*** *hi-páayn-a* *'áayat*
lodge-LOC 3NOM-arrive-PAST woman
'The woman arrived at a/her lodge.'

(5) *'iníi-**ne*** ***pà**-payn-**óo**-ya* *'áayato-**m***
lodge-OBJ 3SUBJ.3OBJ-arrive-LOC-PAST woman-ERG
'The woman arrived at a lodge.'

The same verbal suffix also marks shifted goals in transitive clauses, as illustrated in example (7).

(6) *'áayato-na* *páa-'naxpayk-a* *'iníit-**pe***
woman-OBJ 3SUBJ.3OBJ-bring-PAST lodge-LOC
'He brought the woman to [his] lodge.'

(7) *'áayat* *pà-'naxpayk-**óo**-ya* *miyóoxato-**na***
woman 3SUBJ.3OBJ-bring-LOC-PAST chief-OBJ
'He brought the woman to the chief.'

The Nez Perce verbal suffix *-úu* (*-óo* by vowel harmony) decomposes into the frozen past marker *-e* and the copula *wée*. A regular phonological rule converts the sequence *ewe* into *uu* and *awa* into *oo* in Nez Perce. Proof that *-úu* reconstructs as *-*éwe* is supplied by those Sahaptin dialects which do not share this phonological rule. In Umatilla Sahaptin the directive suffix is *-áwa(n)*:

(8) *awkú* *wílaps-in* *pà-winan-á**wa**n-a*
then sturgeon-ERG 3SUBJ.3OBJ-go-LOC-PAST
'Then the sturgeon went to him.'

2.2. The Klamath directive *-otan*

Directive goals in Klamath are marked by the verbal suffix *-ot(a)n* and its allomorphs, e.g. *-tan* in (9) which is from Stern (1951).

(9) *čoy hok w'ak'a hoot-**tan**-ank lilhanks-as*
and the little coyote run/jump-LOC-PERF deer-OBJ
'And the little coyote having jumped onto the deer...'

The Sahaptian verbal suffix -**ten* meaning 'go (there) in order to ...' is a likely cognate. Also possibly related are the Sahaptin future *-ta*, and the Nez Perce verb *tin'-* '(plural) walk' and the Klamath verb *din-* '(plural) run'. Table 1 schematizes the varied grammaticalizations (as verbal suffixes) of 'go' and 'be' in Sahaptian-Klamath.

Table 1 *Grammaticalization of 'go' and 'be'*

	**ten* 'go'	**we* 'be'
Nez Perce	*-ten* 'go in order to ...'	*-úu* DIR*
Sahaptin a.	*-tan* 'go in order to ...'	*-úu/áwa* DIR
b.	*-ta* FUT	
Klamath	*-otan* DIR	*-wapk* FUT

*DIR = Directive

2.3. The Nez Perce locative *-c'aa*

In Nez Perce locatives with the sense 'on' or 'over' can be coded by the verbal suffix *-c'aa*, as in (11). In (10) the same proposition is coded as an intransitive clause; the intransitive 3rd person *hi-* is prefixed to the verb and the NP is case marked as an oblique locative.

(10) *hi-watiki-s-a miya'ác-**pa***
3NOM-step-PROG-SG.NOM child-LOC
'He is stepping *on/over* the child.'

(11) *pó-otix-**c'a**-s-a miya'ás-**na***
3SUBJ.3OBJ-step-LOC-PROG-SG.NOM child-OBJ
'He is stepping *over* the child.'

It is highly likely that the suffix *-c'aa* 'on or over the object' is either derived from or shares a common morpheme with the verb *c'áa'k* 'hang'.

2.4. The Sahaptian locative -**éče*[7]

In Nez Perce the verbal suffix *-éece* marks objects for the locative sense of 'upon' or 'onto', e.g.

(12) *kaa yoq'op-**né*** *'e-'yeqik-**éece**-m* *kúus-ki*
and that very-OBJ 1SUBJ.3OBJ-pour-LOC-DIR water-INSTR
'And I have poured water *upon* that very one' (Phinney, 1934:2.5)

The cognate suffix in Sahaptin is *-ásha*, e.g.

(13) *pà-xa-watik-**ášа**-m*
2SUBJ.1OBJ-unwittingly-step-LOC-DIR
'You stepped on me' (Jacobs, 1931:200)

Nez Perce *-éece* and Sahaptin *-ásha* suggest a Proto-Sahaptian *-*éče* (see Aoki, 1963 for sound correspondences), which minus the frozen past morpheme yields **če*. For a possible verbal source, cf. the Klamath *c-* '(singular) sit' and *cii* 'stay, live, dwell' (see Barker, 1963).

2.5. The Nez Perce locative *-'aatk*

Although the verbal suffix *-'aatk* does not itself make transitive an intransitive verb, it does provide semantic modification of the object in an already transitive clause, i.e. 'as the object passes the subject'. In (14) the allomorph *-atk* is suffixed to the transitive stem *n* 'say/tell'.

(14) *koná náaqs-nim qíiwn-im* *hi-náas-n-**atk**-a*
there one-ERG old man-ERG 3NOM-PL.OBJ-say-pass-PAST
'Then an old man said to them *as they passed him*, "..."
(Aoki, 1979:19:25)

When *-'aatk* modifies a directive goal, *-úu* is also suffixed:

(15) *koná xáxaas-nim pú-uyexy-**uu-'atk**-a*
there grizzly-ERG 3SUBJ.3OBJ-rush out-LOC-pass-PAST
'Then grizzly rushed out *at him as he passed by*.'
(Phinney, 1934:1:11–12)

The Nez Perce verbal suffix *-'aatk* 'as the object passes by' derives from the verb *'áat* 'go out'. The *k* element is the suffix described in Aoki (1970:99).

2.6. The Nez Perce locative *-úukini*

Like *-'aatk*, *-úukini* also seems only to modify an existing object. It imparts the meaning 'as the object approaches', as in (16). (In Nez Perce any

verb which could imply a dative goal, as *wehín* 'bark', is automatically transitive and the dative is the object.)

(16) *kaa ciq'áamqal-m wáaqo' pè-wehn-**úukiny**-e*
and dog-ERG now 3SUBJ.3OBJ-bark-LOC-PAST
'And the dog now barked at him *as he approached.*'
(Phinney, 1934:480:3)

The suffix *-úukini* decomposes into the frozen past morpheme *-e* and the verb *wéwkuni* 'meet'.

3. ABLATIVE OBJECTS

In Nez Perce ablative NPs are marked by the suffix *-ki'nix*, as in (17).

(17) *hi-pa-ws-páyx-toq-a meqséem-**ki'nix***
3NOM-PL.NOM-journey-arrive-back-PAST mountain-ABL
'They journeyed back *from* the mountains.' (Phinney, 1934:41:6)

Although ablatives are most often coded as oblique NPs in Nez Perce, they can occur as objects, as in (18) which is taken from Aoki's *Nez Perce Grammar*. Although in this example the ablative object does not surface as an NP, its object status is proved by the verbal prefix *nas-* which marks plural objects. The ablative semantic role is coded by the verbal suffix *-áapiik*.

(18) *kaa hi-**nas**-waka'yk-**áapiik**-s-a*
and 3NOM-PL.OBJ-fly-ABL-PROG-SG.NOM
'And she is flying away *from* us.' (Aoki, 1970:97)

If the suffix *-áapiik* also derives from a verbal source, the best candidate is *sáapiik* 'sharpen', which probably had an earlier meaning something like 'remove'. Morrison (manuscript) lists a Warm Springs verb *shápik-* 'put down'.

4. ASSOCIATIVE OBJECTS

In Nez Perce the associative case role is most often coded by the NP suffix *-iin* (and its allomorphs) which also marks duals. Associatives have a higher degree of participation in a state or event than do other semantic

obliques, and in Nez Perce this is realized by the fact that human NPs marked by *-iin* also require plural subject-verb agreement, as in (19).

(19) *láwtiwaa-**yiin** hi-túuqi-s-ix*
friend-ASSOC 3NOM-smoke-PROG-PL.NOM
'He is smoking with a/his friend.'

Although somewhat rare in Nez Perce texts, associatives can be made into objects in Nez Perce, as in (20) which, like (19), was obtained by direct elicitation. The associative case role is marked by the verbal suffix *-tiween* (here the allomorph *-twe*).

(20) *láwtiwaa-**na** **pée**-tuqi-**twe**-c-e*
friend-OBJ 3SUBJ.3OBJ-smoke-ASSOC-PROG-SG.NOM
'He is smoking with a friend.'

In Sahaptin texts associatives are perhaps most often treated as objects and marked by the verbal suffix *-twan*, cognate of the Nez Perce morpheme. Associative marking by the NP suffix *-in* and the verbal suffix *-twan* is contrasted in (21) and (22), both of which come from Warm Springs Sahaptin via direct elicitation. (Contrasted also is the associative function in (21) and the "obviative" ergative function in (22) of the NP suffix *-in*.)

(21) *pa-wyánawi-ya wínš-**in***
3PL.NOM-come-PST man-ASSOC
'He came with the man.'

(22) *pà-wyanawi-**twán**-a wínš-in*
3SUBJ.3OBJ-come-ASSOC-PST man-ERG
'The man came with him.'

The Sahaptian verbal suffix -**tiwén* also seems to have had a verbal origin. But the Nez Perce verb, which should have the form *tiwéen*, does not occur independently. The expected meaning survives in the verbal complex *tiwíikin* 'accompany, follow'.[8] The Sahaptin verb is *twánan* (with frozen 'thither' directional *-an*) 'follow'. The Klamath associative is marked by *dola* plus the objective form of the NP (with suffix *-s*), as in (23). The morpheme *dola* (with fused indicative suffix *-a*) is probably an old verb, the Klamath cognate of Sahaptian **tiwén* 'accompany'. The Klamath *l* is a common reflex (via diminutive sound symbolism) of an earlier *n*. But unlike in Sahaptian, Klamath *dola* never became suffixed to the main verb in the clause.

(23) *čoy hok 'at goWaskč-a č'asgaay'aak-**s** **dola***
and he now start off-IND little weasel-OBJ with
'And then now [he] started off *with* little weasel.' (Stern, 1951)

It is also possible that in the areally close Central Kalapuya language the verbal suffix *-dini*, which marks instrumental objects, is cognate. The construction is illustrated in (24).[9]

(24) *láw'mdé gidii-d-ts'máq-**dini** a-'wá'da'k heš an-k'wín*
now then ASP-DIR-hit-INSTR DET-stick this DET-racoon
And she hit this coon *with* the stick.' (Jacobs, 1945:II.11.1.3.)

5. COMPETITOR OBJECTS

Aoki (1970:100) cites a verbal suffix *-so'* and notes that it "indicates that an action is performed in competition with someone else." Aoki provides contrastive examples, here (25) and (26).

(25) *wistok-áhsa-s-a*
shoot-upward-PROG-SG.NOM
'I am shooting upward.'

(26) ***'a**-wstok-ahs-**ó'**-s-a*
1SUBJ.3OBJ-shoot-upward-competitor-PROG-SG.NOM
'I am shooting upward *in competition with him/her*.'

That the verb in (26) is transitive, in contrast to the verb in (25), is evident from the prefix *'a-* which codes a 1st or 2nd person subject acting on a 3rd person object. In (26) the verbal suffix *-ó'* marks the semantic role of the 3rd person object. It is not clear what the origin of this morpheme is, however. It is listed here in order to complete the description of transitivizing verbal suffixes which are descriptive of or code the semantic role of the object. Aoki provides only the one example of this verbal suffix, and I have not as yet run across any more in my data or in Phinney (1934) or in Aoki (1979). And so, based on just this one example, the underlying form of this morpheme could be either *-so'* or *-o'*. But if the form is actually *-o'*, then it would appear a possible cognate of the Klamath verbal suffix *-ee'* which Barker (1963:132) defines as 'in a game, competition, for fun', e.g.

(27) *'-otG-a*
long object-remove-IND
'Take a long object away from someone,' (Barker, 1963:24)

(28) *'-otG-**ee**'-a*
long object-remove-competitor-INDEF
'take a long object away from someone *in competition (as in a game)*.' (Barker, 1963:132)

6. CONCATENATED VERBS

When an NP object marker derives from a verb, presumably it originated in a serial verb construction with OV or VO word order. Patient object markers so derived usually begin as a verb *take*, *hold*, or *have*, and dative markers usually begin as the verb *give*. And verbs like *take* or *use* often come to function as instrument markers. See Heine and Reh (1984).

The Bantu language KinyaRwanda and also Indonesian have verbal suffixes which, much as in Sahaptian and Klamath, code the semantic role of objects.[10] A major difference, however, is that the object in those languages is marked by word order. Also, it is possible that in such VO languages as these (KinyaRwanda and Indonesian) where the semantic case role marker occurs verb final and the object NP follows immediately, these verbal suffixes may not have a verbal origin. An earlier *V Prep NP* may simply have been reanalyzed as *V-Suffix NP*.[11]

But what about free word order languages like Nez Perce where object NPs are already case marked? What kind of process would lead to verbal suffixes of verbal origin which disambiguate the semantic case role of morphologically marked object NPs?

In Nez Perce, at least, the subject and object (or agent and patient) in a transitive clause are never marked by word order. The same seems to be true in the related Sahaptin dialects and in Klamath. And so it seems reasonable not to reconstruct such a function for word order in Proto-Sahaptian. In Nez Perce verb and object form no tight syntactic unit, and thus 'give' did not develop into a dative marking adposition.[12] But how did it become a dative marking verbal *suffix*?

Although subject and object are never marked by word order in Nez Perce, the syntactic juxtaposition of verbs often functions to mark simultaneity in the events coded. This phenomenon is mentioned in (Aoki, 1978:31)

and briefly described in Rude (1985b:364–365). The two verbs suffix the same aspectual morphogy and express agreement with the same coreferential subject. No other word or particle intervenes between the verbs. This verbal concatenation in the perfective, progressive, and future is illustrated in (29) to (32). Transitive verbs as in (32) share the same coreferential object.

(29) *kaa hi-pe-wléeke'yk-e hi-pa-ckilíitoq-a 'iníit-kex*
and 3NOM-PL.NOM-run-PAST 3NOM-PL.NOM-return-PAST tepee-to
'And running they returned to the tepee,' (Rude, 1985:5:171–172)

(30) *kawó' qiláasx̣ hi-pe-ckiliin-e*
then otter 3NOM-PL.NOM-go home-PAST
hi-pe-hipst'úuyn-e
3NOM-PL.NOM-be satisfied-PAST
'Then the otters went home satisfied,' (Aoki, 1979:9:52)

(31) *wáaqo' 'éetx̣ew-c-e cikliin-'ipéecwi-s-e*
now be sad-PROG-SG.NOM go home-want-PROG-SG.NOM
'Now I am sad wanting to go home,' (Aoki, 1979:9:76)

(32) *'e-pe-tem'ik-úu-yu'*
1SUBJ.3OBJ-PL.NOM-go down-LOC-FUT
'a-pó-opci'yaw-ta-no'
1SUBJ.3OBJ-PL.NOM-kill-go-FUT
'Going down we will kill him,' (Aoki, 1979:17:51)

Thus the pattern currently exists which will account for the transformation of verbs into verbal suffixes which mark the semantic role of objects. The only structural modification has been the merger into a single clause. The process may have been accomplished before the development of the verbal person prefixes and the proliferation of aspectual markers in Sahaptian. The pronominal prefixes and wider variety of aspectual suffixes are absent in Klamath.

7. CONCLUSION

According to Givón (1984), languages which objectify nonpatient NPs and thus have a more purely syntactic object, are only "a subset of languages, so far only VO languages with a morphologically unmarked accusative." In the languages Givón looked at, dative shift is a topicalization involving a

leftward movement to the post verbal word order slot. Nez Perce is an example of a free word order language with marked object NP, but which nevertheless has developed the same kind of syntactic object.

The Sahaptian-Klamath process described in this paper shows how grammaticalization can proceed from pragmatic word order to affixation apart from the grammaticalization of word order for coding subject and object. The grammaticalization described here involves a structural change where two lexical units become one. There is a consequent phonological change — what was formerly two lexical units with two primary stresses have become one with one primary stress. And what began as a verb concatenation has given rise to clause merger: Two clauses have become one.

Along with these structural changes came the association of the concatenated verbs with transitivity. This could have been generalized from the high transitivity of the verb 'give'. The spread would have been from benefactive object to directive object, etc. And this association with transitivity has equalled a major functional change. What began as a pragmatic verb concatenation (which still codes the simultaneity of various events) has led to constructions for the promotion to object of various semantic roles. What has resulted is a grammatical construction with a discourse/pragmatic function — a construction for coding the importance of various nonagentive arguments. The case marking of object NPs is widespread in Penutian, and is assumed to have predated the grammaticalization process described in this paper. Distinguishing a patient from an agent is most critical when the patient is human. This explains the general tendency to limit morphological object marking to human nouns. It also suggests why such a system could be extended to marking objects of discourse/pragmatic importance (which thus might also sometimes include nonhuman or inanimate objects). The extension from human to discourse/pragmatic importance is to be expected. In Sahaptian and Klamath the morphological marking of objects of discourse/pragmatic importance has been further extended to include a wide variety of semantic case roles.

NOTES

1. The Nez Perce data found in this paper is from Phinney (1934), Aoki (1979), and my own field notes obtained on the Umatilla Indian Reservation near Pendleton, Oregon, in 1983–84 and supported in part by a grant from the National Science Foundation, BNS-831060. The Sahaptin data is from field work funded by NSF Research Grant No. BNS-

8711123, and research with the Klamath data of Stern (1951) was supported by a grant from the National Endowment for the Humanities, Grant # RO-21292–86, to Scott DeLancey. The terms *Sahaptin* and *Sahaptian* are confusing: Nez Perce and Sahaptin are two North American Indian languages which together comprise the Sahaptian language family.

2. In texts the vast majority of benefactive arguments are objectified. Although quite rare in texts, constructions where the benefactive is an oblique NP case marked with the suffix *-'ayn* do occur, as in the following example from Phinney (1934:18:7–8). The object of the verb in this example is the 'it' with sentential antecedent.

kawó' 'éetee-x wéc'u' 'e-kú-s-e
then surely-I stop 1/2SUBJ.3OBJ-do-PROG-SG.NOM

*'étke titóoqa-**'ayn** 'a-ko-s-áa-qa*
because people-BEN 1/2SUBJ.3OBJ-do-PROG-SG.NOM-PAST

'Then I am surely stopping it because I was doing it *for* the people.'

Such alternate constructions for the benefactive, expressed either as an oblique NP or as the syntactic object with the benefactive notion coded by a verbal affix, are areally common. Compare the following examples from Sliammon (a Salish language; data is from John Davis, personal communication).

t'utθ'u-t-as John kw qigaθ 'ə Mary
shoot-CAUS-3ERG John DET deer OBL Mary
'John shot the deer *for* Mary.'

t'utθ'-əm-t-as John Mary 'ə kw qigaθ
shoot-BEN-CAUS-3ERG John Mary OBL DET deer
'John shot Mary a deer.'

3. As in Aoki (1970:98). Rigsby (Forthcoming) found their occurrence in Sahaptin to be determined largely by dialect, although he deduces from Jacobs (1931) the possibility that Jacobs' Lower Snake informants used *-ni* to mark benefactives and *-y* genitives.

4. Compare, for example, the following sentences from Ashante Twi. In example (a) ***maa*** is the verb 'give' with its prototypical sense, but in the serialized constructions in (b) and (c) ***maa*** serves to mark dative and benefactive case roles.

(a) *Afua* **maa** *Kofi nhoma*
Afua gave Kofi book
'Afua *gave* Kofi a book'

(b) *Afua de nhoma* maa *Kofi*
Afua hold book give Kofi
'Afua gave the book *to* Kofi'

(c) *Afua yɛ abodoo* **maa** *Kofi*
Afua made cornbread gave Kofi
'Afua made cornbread *for* Kofi'

Perhaps even the Germanic preposition whose function includes the marking of datives has a verbal origin; the Indo-European verb 'give'. Indo-European **dō*-'give' (cf. Latin

dare, Greek *didōnai*, Russian *dat'*, Sanskrit *dadhāmi*, etc.) does not occur in Germanic. Germanic **giban* 'give' is cognate with Latin *habere* 'to hold, possess, have, handle'. The expected phonological reflex of PIE **dō*-'give' in Germanic would be **tō*- ([tu] in English), which happens to be the form of the preposition 'to'. Such a change in function could have arisen from serialized constructions like, 'He *gave* (<'hold, take') the book *to* (<'give') her. Possibility suggested by Robert Hetzron (personal communication).

5. See Aoki (1963), DeLancey et al. (1988) and Rude (1987) for evidence of this genetic relationship, and Barker (1963) for data from Klamath.

6. Probably the Sahaptian benefactive NP suffix — Nez Perce *-'ayn* and Sahaptin *-ay* — is also cognate to one or both 'give' formatives. In Nez Perce sound symbolism the alternation from *e* to *a* imparts a diminutive or augmentative sense.

7. In the orthography used here, *c* equals [ts] and e equals [æ].

8. There is also a verbal prefix *tiwii-* which means 'to lead (of warriors)', and *-tiwee* is a common noun formative: *láw* 'speaking', ***láwtiwaa*** 'friend'; *síiks* 'nest', ***síikstiwaa*** 'intimate friend'; *híik'ay* 'cup, dish', ***hik'áytiwaa*** 'brother-in-law'; *'ilúut* 'belly', ***'ilúutwee*** 'half-brother', etc.

9. That *-dini* is not a noun prefix in (24) is evident from the form of the object noun. The initial *a(n)-* of nouns always deletes after a prefix.

10. For a description of the KinyaRwanda system, see Kimenyi (1980), and for Indonesian see Chung (1978).

11. Indo-European is full of examples where prepositions attach to verbs, no matter what the word order. Peter Hook, a contributor to this volume, has pointed out (personal communication) examples (g) and (h) from Sanskrit:

(g) *ti-ṣṭa-nti* *upa grám-e*
RED-stand-3PL.NOM at town-LOC
'They stand/stay at the town.'

(h) ***upa**-ti-ṣṭa-nti* *grám-am*
at-RED-stand-3PL.NOM town-ACC
'They occupy the town.'

Compare also English *go* which takes oblique locative phrases (*go in the house*, *go out of the house*) with the verbs *enter* (with *en-* 'in') and *exit* (with *ex-* 'out of, from') which take direct object locatives (*enter the house*, *exit the house*).

12. The case marking of S and O NPs in Nez Perce is not directly derivable from verb serialization or word order position. NP case marking morphology is rather the result of analogical extensions from elsewhere in the grammar. In Chapter IV of Rude (1985a) it is argued that the Nez Perce ergative NP suffix *-nim* is the same as the verbal cislocative suffix *-nim*, and it is shown that the object NP marker *-ne* has the same form as an old locative suffix. The ergative NP suffix *-nim* never serves to topicalize the agent. Instead the case marking of ergative and objective NPs functions together to topicalize the object.

REFERENCES

Aoki, Haruo. 1963. "On Sahaptian-Klamath linguistic affiliations." *International Journal of American Linguistics* 29:107–112.

Aoki, Haruo. 1966. "Nez Perce vowel harmony and Proto-Sahaptian vowels." *Language* 42:759–767.
Aoki, Haruo. 1970. *Nez Perce Grammar*. Berkeley/Los Angeles: University of California Press. *University of California Publications in Linguistics* 62.
Aoki, Haruo. 1978. "Coyote and fox (Nez Perce)." In *Coyote Stories*, William Bright (ed.), 26–31. Chicago: University of Chicago Press. *International Journal of American Linguistics Native American Text Series* 1.
Aoki, Haruo. 1979. *Nez Perce Texts*. Berkeley/Los Angeles: University of California Press. *University of California Publications in Linguistics* 90.
Barker, M. A. R. 1963. *Klamath Dictionary*. Berkeley/Los Angeles: University of California Press. *University of California Publications in Linguistics* 31.
Bright, William. 1982. *Bibliography of the Languages of Native California*. Metuchen, NJ, and London: The Scarecrow Press, Inc. *Native American Bibliography Series*, 3.
Chung, Sandra. 1978. *Case Marking and Grammatical Relations in Polynesian*. Austin: University of Texas Press.
DeLancey, Scott, Carol Genetti and Noel Rude. 1988. "Correspondences in some Penutian languages." In *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, William Shipley (ed.), 195–224. Berlin: Mouton de Gruyter.
Givón, Talmy. 1984. "Direct object and dative shifting: Semantic and pragmatic case." In *Objects: Toward a Theory of Grammatical Relations*, Frans Plank (ed.), 151–182. New York: Academic Press.
Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
Jacobs, Melville. 1931. *A Sketch of Northern Sahaptin Grammar*. Seattle: University of Washington. University of Washington Publications in Anthropology 4:85–291.
Jacobs, Melville. 1945. *Kalapuya Texts*. Seattle: University of Washington. *University of Washington Publications in Anthropology* 11:1–394.
Kimenyi, Alexandre. 1980. *A Relational Grammar of KinyaRwanda*. Berkeley: University of California Press. *University of California Publications in Linguistics* 91.
Morrison, Henry. Warm Springs Sahaptin Dictionary. Manuscript.
Phinney, Archie. 1934. *Nez Perce Texts*. New York: Columbia University Press. *Columbia Contributions to Anthropology* 25.
Rigsby, Bruce. Forthcoming. *Sahaptin Grammar*. To appear in *Handbook of North American Indians* 17. Washington: Smithsonian Institution.
Rude, Noel. 1985. Studies in Nez Perce Grammar and Discourse. Unpublished PhD dissertation, University of Oregon.
Rude, Noel. 1986a. "Topicality, transitivity, and the direct object in Nez Perce." *International Journal of American Linguistics* 52:124–153.
Rude, Noel. 1986b. "Discourse-pragmatic context for genitive promotion in Nez Perce." *Studies in Language* 10:109–136.
Rude, Noel. 1987. "Some Sahaptian-Klamath grammatical correspondences." *Kansas Working Papers in Linguistics* 12:67–83.
Rude, Noel. 1988a. "Ergative, passive, and antipassive in Nez Perce: A discourse perspective." In *Passive and Voice*, Masayoshi Shibatani (ed.), 547–560. Amsterdam: John Benjamins.
Rude, Noel. 1988b. "Semantic and pragmatic objects in Klamath." In *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, William Shipley (ed.), 651–673. Berlin: Mouton de Gruyter.
Stern, Theodore. 1951. Unpublished Klamath Texts.

Grammaticalization of Postpositions and Word Order in Senufo Languages

Robert Carlson
University of Oregon and Summer Institute of Linguistics

1. INTRODUCTION

1.1. Word order in the Senufo languages

The Senufo languages are a group of closely related languages spoken in southeastern Mali, southwestern Burkina Faso, and northern Côte d'Ivoire, as well as one isolated pocket in northern Ghana. They are usually classified with the Gur family of Niger-Congo, though this has been called into question recently.[1] Typologically they share with Gur languages a suffixal noun class system, but in many other respects they diverge from central Gur languages. While the latter have a basic SVO word order, Senufo languages all have a "mixed" SOV word order: The direct object precedes the verb, but oblique objects, marked by postpositions, follow the verb, as do all but the most nominalized complement clauses. In addition, the subject is separated from the direct object by an auxiliary which carries the tense and most of the aspect and modality information of the clause. The label SOV is thus rather misleading, the reality being rather: [S AUX O V OBL]. It should also be pointed out that predicate nominals follow the copula/verb rather than precede it, and that genitives precede their head nouns.

The mixed SOV order of Senufo languages is actually an areal phenomenon. It is perhaps best known as the ordinary word order of the Mande family of Niger-Congo, which is geographically contiguous to the Senufo group. It is also the order of some dialects of the unrelated Songay language, which borders Mande on the north. It is an alternative to SVO order in some clause types in a few Gur languages spoken in western Burkina Faso (e.g. Toussian, Wara, both contiguous to Senufo languages), and in a number of

Kru languages, which are contiguous to Mande on the south (one isolate Kru language, Sɛmɛ, is contiguous to the Senufo language Sucite in Burkina Faso). It is clear that there has been mutual influence, but its direction and precise form remain obscure.

1.2. The origin of the S AUX O V order

Various attempts have been made to explain the [S AUX O V OBL] order as the modification of an earlier more orthodox order. Heine (1976; 1980) and Heine and Reh (1984) trace present day Mande order to the reanalysis of an earlier SVO order in which the direct object was a nominalized clause. The erstwhile main verb became the new auxiliary, and the nominalized verb was reanalyzed as the main verb. The placement of the direct object before the nominalized verb is explained as due to the normal order of genitives in Mande languages. It is assumed that the object was syntactically the genitive possessor of the nominalized verb, and as such preceded its head noun. Reanalysis of the nominalized verb as the main verb (concomitant with reanalysis of the main verb as auxiliary) led to the reanalysis of the genitive as the direct object. All constituents remained where they were in the original structure.

Marchese (1986) follows a similar argument for Kru languages, which have [S AUX O V] order in clause types with an auxiliary, but SVO when no auxiliary is present. Marchese argues that the former construction is a reanalysis of an earlier one in which the AUX was the main verb and the V was a nominalized verb. Unlike Heine and Reh do for Mande, however, she assumes that the OV ordering of the embedded nominalized clause reflects the earlier word order of all clauses (cf. also Hyman, 1975, Givón, 1975). Kru languages differ as to which arguments of the verb besides the direct object occur before rather than after it in constructions with auxiliaries. Marchese (chapter 5) assumes that this reflects a gradual movement of arguments to a position after the verb to bring these constructions into conformity with the current VO order of clauses without auxiliaries.[2]

In Carlson (1985) I put forward a different hypothesis for the Senufo languages: The order [S AUX O V] arose not as the reanalysis of an embedded structure ([S V [O V]]) , but from the grammaticalization of the initial verb in a [S V O V] serial construction.[3] Three factors were foremost in suggesting the serial hypothesis. One was the absence of any clear evidence of an embedded structure, such as nominal marking on the main verb. The second was that it required no word order change to explain present day word order.

The auxiliaries developed from verbs used intransitively, and the direct object was originally the direct object of the following verb just as it is in the modern languages. We will return to this point below. The third factor which rendered the serial hypothesis plausible was the occurrence in modern Senufo languages of just the sort of grammaticalization process hypothesized for the proto-language, namely the ongoing conversion of serial verbs into auxiliaries. The Senufo auxiliaries are in fact a good example of what Hopper (Volume I) calls "layering". Having once established an avenue of grammaticalization, speakers continue to exploit it in succeeding generations, with the result that strata accumulate, older auxiliaries coexisting with newer ones.

A crucial difference between the two hypotheses outlined above is the transitivity of the verb which developed into an auxiliary. If the original language was VO, then the auxiliary must have been transitive, with an embedded clause as direct object. If, on the other hand, the original language was OV, the auxiliary must have been intransitive, followed by one or more verbs (together with their arguments) in a serial construction. Given the flexible nature of transitivity in modern Senufo languages, where virtually any transitive verb may be used intransitively, and where there are very few verbs which are exclusively intransitive, the serial hypothesis must be refined to state that auxiliaries developed from verbs used intransitively rather than from intransitive verbs. The following examples from Supyire illustrate the ongoing process of auxiliarization. The first is an example of a developing tense marker, the second of a modality marker:[4]

(1) a. *U à pyi-a kàrè*
s/he PERF PAST-NF go
'S/he had gone.'

b. *U a jà-a kàrè*
s/he PERF be.able-NF go
'S/he was able to go.'

Both verbs may be used both transitively and intransitively, as the following examples show:

(2) a. U *a lì pyì*
s/he PERF it do
'S/he did it.'

b. *U à pyi bagé-e*
s/he PERF do/be house-in
'S/he was in the house.'

c. *U a ù jà*
s/he PERF s/he overcome.
'S/he overcame/defeated/beat up him/her.'

d. *U a jà*
s/he PERF be.able
'S/he was able.'

There is no evidence that the second verb in the examples in 1 above is now or ever was the embedded complement of the first verb. The first verb is simply used intransitively. The construction is an ordinary serial construction. In such constructions, transitive verbs may occur in either position, as shown by the following examples:

(3) a. *U à pa-a kù lwɔ́*
s/he PERF come-NF it take
'S/he came and took it.'

b. *U a kù bwɔn-a bò*
s/he PERF it hit-NF kill
'S/he hit and killed it.'

It seems clear then that the auxiliaries in Supyire developed from initial serial verbs used intransitively.

1.3. The place of postpositional phrases

If the [S AUX O V] order reflects an earlier [S V O V] order in Senufo languages, the placement of the oblique objects still remains to be explained. In a canonical OV ordering, these arguments ought of course to come before rather than after the verb. The aim of this paper is to show that an extension of the serial hypothesis can account for the placement of obliques in modern Senufo languages without the necessity of any change in word ordering. The proposal is that serial verb constructions of the type [S (O) V O V] were reanalyzed as [S (O) V O POST]. Of course in order to substantiate such a claim, verbal etymologies must be established for the postpositions in ques-

tion. Before proceeding to that task, however, several observations are in order.

There are two main sources of prepositions and postpositions: verbs (often in serial constructions) and head nouns of genitive constructions (cf. Heine and Noske, 1988). The latter type are well-attested in Mande (cf. Bird and Kendall, 1986) and Kru (cf. Marchese, 1986). Heine and Reh (1984) rightly point out that the postposing of such denominal adpositions points not to an earlier OV order, but simply to the present day (and presumably earlier) genitive-noun order. Senufo postpositions present a more complicated picture. In these languages, postpositions must be divided into simple and complex. The complex postpositions have an initial element which is indeed denominal, but this must in most cases be followed by one of the simple postpositions, in a structure rather like the English *on top of* or such compounds as *behind, beyond, beside*, where the *be-* is cognate with the preposition *by*. In other words, the noun did not give rise to a postposition, but only to the first element of one. Obliques must thus be marked with a simple postposition, either by itself or in composition with a nominal element.[5] The following examples of complex postpositions from Supyire are typical:

(4) a. *bagé ŋ̀kèrè nà* < *ŋ̀kèrè 'side'*
house.DEF side at
'beside the house'

b. *bagé fùnŋì-ì* < *funŋɔ* 'belly'
house.DEF inside-in
'inside the house'

c. *bagé ɲùŋì-ì* < *ɲuŋɔ* 'head'
house.DEF top-at
'on top of the house'

The hypothesis put forward here is that the simple postpositions of Senufo languages descend from verbs. These verbs were used transitively in a case marking function, that is, they added locative, dative, and benefactive arguments to the clause, one per verb, in a manner common to many serializing languages. It should be pointed out that such a use of serial verbs is already highly grammaticalized. The etymologies proposed in the following section show that most of these verbs were already highly abstract long before the proto-Senufo stage. The development of a category of postpositions thus represents simply one part of the over-all trajectory of these morphemes from less to more grammaticalized (Hopper, Volume I).

The data also make clear that the phenomenon of "layering", already noted in the auxiliaries, has also occurred in the category of postpositions. Some of the simple postpositions are shared by all the Senufo languages, and may thus be reconstructed for the proto-language. Others are less widespread, and obviously more recent. Having once established the route serial verb, postposition, speakers continue to exploit it creatively.[6]

The remainder of this paper is divided into three sections. Section 2 will present evidence of a verbal origin for most of the simple postpositions of Senufo languages. Section 3 will briefly examine comparable data from Mande languages. The final section will deal with problems presented by the serial hypothesis.

2. VERBAL ETYMOLOGIES FOR SENUFO POSTPOSITIONS

The task of establishing etymologies for Senufo languages is beset with the difficulty of the paucity of published materials for most of the languages concerned. No full-fledged grammars are available, and of the sketches which exist many suffer from various defects, such as not marking tone or not treating postpositions. This means that much of what is proposed here must be regarded as tentative and subject to revision. It is my belief, however, that enough data is available to make the serial hypothesis attractive in the Senufo case.

Information from fifteen Senufo languages is used in this paper. The sources range from relatively complete, e.g. Laughren (1973) for Cebari, to gleanings from wordlists, e.g. Jordan (1980) for Nafaara. Pending a better classification, I have divided up the thirteen languages which are geographically contiguous into Northern, Central, and Southern groups. Karaboro and Nafaara are isolated. The list of languages with the sources used is given in Table 1.

2.1. *na/la* 'at, on, to'

This postposition is found in all Senufo languages.[7] A typical exampl is the following from Supyire:

(5) *U a kàrè Sukwoo na*
s/he PERF go Sikasso to
'S/he has gone to Sikasso.'

Table 1 *Languages and Sources Cited*

LANGUAGE	SOURCES
NORTHERN	
MAMARA	Dembele (1983), Cauvin (1980)
NANƐRGHƐ	anon. (1983)
SUCITE	Garber (1987)
SUPYIRE	personal notes
SHƐNPIRE	Madou Kone (pers. comm.)
CENTRAL	
SENARI	Prost (1964)
SHENARA	Cissé (1986)
CEBAARA	Mills (1984, 1987), anon. (1974), Means (1966)
CEBARI	Laughren (1973)
SOUTHERN	
PILARA	Welch (ms)
TAGBANA	Clamens (1952)
JIMINI	Lejeune (ms)
FƆNDƆNDƆ	Boutin (1981)
OTHER	
KARABORO	Wichser (pers. comm.)
NAFAARA	Jordan (1980)

There is good evidence that *na* developed from a verb meaning 'be at'.[8] Although no Senufo language retains *na* as a main verb, in Tagbana[9] and Supyire it is used as the progressive auxiliary:

(6) a. *Wi na li* TAGBANA
he PROG eat
'He is eating.'

b. *U na lyí* SUPYIRE
s/he PROG eat.IMPFV
'S/he is eating.'

The Karaboro 'durative' auxiliary *naa* is probably related:

(7) *Nawee mpãy tər naa n tèr nàńgá Laboklo kuè ye*
people some strange DUR INCOMP go today Labo towards to
'Some strange people were going to Labo today.'

Note the similar dual function of *'nà* in Igbo (Welmers, 1973:312):

(8) a. *Ó nò nà Kánù*
s/he be at Kano
'S/he is at Kano.'

b. *Ó nà àbyá*
s/he PROG coming
'S/he is coming.'

Welmers suggests reconstructing a locative copula **na* 'be at'.

In Tyurama, a Gur language spoken in western Burkina Faso, and in geographical proximity to Karaboro, Senari, and Sucite, all three functions, verb/copula, progressive auxiliary, and postposition, are attested (Prost, 1964):

(9) a. *O na dugu na*
S/he be.at house at
'S/he is in the house.'

b. *Ba na*
they be.at
'They are here.'

c. *Me na me wu*
I be.at I eat
'I am eating.'

There is tonal evidence in Supyire (and in Sucite, cf. Garber 1987:298) that the postposition *na* is verb-like rather than noun-like. The tone of *na* is strong mid.[10] Strong mid nouns never accept the spreading of high or low tones from the previous word, whereas strong mid verbs and *na* do:

(10) a. *wùu bagé* < *'ba-* 'house', strong mid
us house.DEF
'our house'

b. *U a wùù ɲyà* < *ɲya* 'see', strong mid
s/he PERF us see
'S/he saw us.'

c. *wùù nà*
us on
'on, about, for us'

(11) a. *Zhyé bagé*
Zhye house.DEF
'Zhye's house'

b. U *à* *Zhyé* *ɲyá*
s/he PERF Zhye see
'S/he saw Zhye.'

c. *Zhyé ná*
Zhye on
'on, about, for Zhye'

2.2. *ni* 'in, at, to'

There is equally good evidence for a locative verb/copula **ni* in Niger-Congo. Welmers (1973:314) points out that the three functions of Yoruba *ní* (marking location, possession, and 'present action') may all be accounted for by assuming a copular/verbal origin with the meaning 'be at'. Wara, a Gur language geographically contiguous to Supyire, Senari, and Sucite, has a copula *ni* with both locative and equative uses (data from Prost 1968):

(12) a. *I* *nna* *ni*
he here be.at
'He is here.'

b. *I* *siyo tu* *ni*
he tree under be.at
'He is under the tree.'

c. *N* *su* *ni* *kyetumiyã*
my father be farmer
'My father is a farmer.'

The Dogon "presentative" (*-ɲ* in the Donno Sɔ dialect, *-i/-y/-ni* in the Tɔgɔ Kã dialect), which is also used as a focus marker, is arguably a reflex of the same copula (data from Tɔgɔ Kã, Prost, 1969b:33–35):

(13) a. *Bõ ma-ni*
father my-it.is
'That's my father.'

b. *Bé-y yɛrɛ*
they-it.is come
'It is they who have come.'

In Donno Sɔ *-ɲ* also has the function of marking datives and some accusatives (Prost, 1969a:32), and in Tɔgɔ Kã *ni* is a postposition meaning 'in' (Prost, 1969b:77):

(14) a. *Dì indɛ páy gɔ́-ɲ obo* DONNO SƆ
water man old the-to give
'Give water to the old man.'

b. *Dɛpɛ poŋo tɛrɛ ma ni yɛrɛ* TƆGƆ KÃ
baboon fonio pile my in came
'The baboon came on my pile of fonio.'

All Senufo languages for which I have data except Karaboro have *ni* as a postposition. At least one language, Jimini, seems to have retained a verbal/copular use:[11]

(15) a. *Wi ni go ni*
s/he be.at house in
'S/he is in the house.'

b. *Golo ni na mã*
chicken be.at me to
'I have a chicken.'

Traces of such a verbal usage are also found in Cebari:

(16) *Pí ní na* cf. *Pí pye na*
they be here they be.PAST here
'They are here.' 'They were here.'

Cebaara has similar data:

(17) *Pe i náa*
they be here
'They are here.'

Cebaara *i* derives from *ni* (still a possible pronunciation in some varieties). Unsurprisingly, the auxiliary for progressive in Cebari is *ni* and in Cebaara it is *i*/*ni*:

(18) a. *Wí ní ɲí lòzòwò* CEBARI
he PROG be.IMPFV hunter
'He is a hunter.'

b. *Pyàũ ní ɲí gbàgí ní*
child.DEF PROG be.IMPFV house.DEF in
'The child is in the house.'

c. *nɛ̌yì i wi fulolo ni syée* CEBAARA
cows.DEF PROG him drag PROG go.IMPFV
'The cows were dragging him.'

The tonal evidence in Supyire is identical to that illustrated for *na* above: *-i*/*-e* (from earlier *ni*) is strong mid tone, and it behaves in regard to tone spreading like a verb rather than like a noun. Compare the following with examples 10 and 11:[12]

(19) a. *wùù-è*
us-in
'in, to, for us'

b. *Zhyé-é*
Zhye-in
'in, to, for Zhye'

2.3. *ḿ*/*pan* [pã] 'to'

Most of the Senufo languages have a dative/benefactive and locative goal postposition with one of these two forms.[13] *pan* and *ma* (or *ma*/*maa* in some languages) are the perfective and imperfective forms, respectively, of the verb 'come'.[14] Most languages have grammaticalized the imperfective form. Cebaara is typical:

(20) a. *á wi lúrú ǹ pã kàtyóolì má*
and she return PERF come courtyard to
'and she returned home.'

b. *Wi ǹ gí kãa wi má*
he PERF it give him to
'He gave it to him.'

The etymology from 'come' is confirmed by data from Pilara and Nafaara, which have grammaticalized the perfective form:

(21) a. *Ki kã u pə̃* PILARA
it give him/her to
'Give it to him/her.'

b. *Wi pə̃ ga*
s/he come here
'S/he came here.'

(22) *Seewa u pan* [pã] NAFAARA
farm be.at him to
'He has a farm.'

Once again, the tonal facts in Supyire show that *-á* (from *má*) is verb-like rather than noun-like. High tone nouns do not accept the floating low tone which follows pronouns, whereas high tone verbs and *-á* do:

(23) a. *u fáágé* < *fáá-* 'stone'
s/he stone.DEF
'his/her stone'

b. *Ma u wìì* < *wíí* 'look at'
you s/he look.at
'Look at her/him.'

c. *Ku kan u-à*
it give s/he-to
'Give it to her/him.'

2.4. *tán* [tã̀n]/*táán* [tã́:]/*táánná* [tã́:na] 'beside'

Northern and a number of central Senufo languages have this postposition. There are two verbs which seem to be related: *taanna* 'to arrange in a line, to place side by side' (occurs in Cebaara, Shenara, and Supyire) and *taanna* 'to measure, to make equal' (Sucite *táánlá*, Shɛnpire *támáná*, Supyire *taanna*, weak mid tone).[15] The latter, with its high or weak mid tone, is perhaps to be preferred. However, it is likely that the two verbs are related historically, the meaning 'make equal, measure' deriving from placing two

things side by side in order to measure or equalize them. Examples of the postposition and related verbs from Supyire are:

(24) a. *U ɲyɛ bagé tààn*
s/he be house.DEF beside
'S/he is beside the house.'

b. *U à vɛ́ribíí taanna tabalíŋi na*
s/he PERF glasses.DEF line.up table.DEF on
'S/he lined up the glasses on the table.'

c. *U a bwɔ́nhigíí faanra-a tàànnà*
he PERF granaries.DEF build-NF make.equal
'He built the granaries to an equal height.'

In Supyire *táán* behaves tonally like *-á*, that is, like a high tone verb.

2.5. *fùn* 'without'

Several central Senufo languages have this postposition. It seems to be related to the verb *fùn* 'consider taboo, abstain or refrain from'. An example of the postposition in Cebaara is:

(25) *Wi i Senárì nyuu tɛ̀ɛ̀lɛ̀ fùn*
s/he PROG Senari speak fault without
'S/he speaks Senari without fault.'

An example of the verb *fùn* from Supyire:

(26) Sàànɔgɔbíí *màha sòòŋi fùn*
Saanogos.DEF HAB terrapin.DEF refrain.from
'The Saanogos do not eat the terrapin (because they consider it taboo).'

Supyire does not have the postposition *fùn*, but has instead borrowed *bàà* 'without' from Jula/Bambara *bali*, which, significantly, is both postposition and verb ('prevent from, forbid') in the source language (Sikasso Jula)– in standard Bambara it is not used as a postposition).

2.6 *kan* 'for'

Some southern Senufo languages have developed a benefactive marking from the verb *kan* 'give'. In these languages *má* is restricted to dative. The following examples are from Tagbana:

(27) a. *Golo syo na kã*
chicken buy me for
'Buy a chicken for me.'

b. *Ki yo kũdi kã*
it say chief for
'Say it for the chief.'

Jimini offers similar examples. It is unclear from the sources whether *kã* is a serial verb or a postposition in such sentences. The cognate in Karaboro, *kə̃*, however, is clearly a benefactive postposition, not a verb. Another verb *wãr* has replaced it for the meaning 'give':

(28) a. *Tá ń lɔ̀ʔɔ̀ kù ma kə̃*
I SUBJUNCT water draw you for
'I should draw some water for you.'

b. *Ù sro wãã rá ǹ di*
he mush give I SUBJUNCT eat
'He gave me mush and I ate.'

2.7. *kurugo/kunuŋɔ* 'along, through, because of, after'

This postposition is used in some northern and central Senufo languages. A possible cognate might be *kurugo* 'to rub against', in Shenara and Cebaara.[16] A nasalized variant appears in Nanɛrghɛ *kũnuŋɔ* 'hit, touch against', which in turn is probably related to the Supyire verb *'kunuŋɔ'* 'be close by, be beside', which has weak mid tone. The postposition in Supyire is tonally like a weak mid verb:

(29) a. *u cìge* < *ci-* 'tree', weak mid
s/he tree.DEF
'her/his tree'

b. *U bò < bo* 'kill', weak mid
s/he kill
'Kill him/her.'

c. *u kùrùgò*
s/he along
'in her/his pocket' or 'because of him/her'

An example of the verb *kunuŋɔ* from Supyire:

(30) *Yi a yì-yé kúnúŋɔ́*
they PERF they-REFL be.close.by
'They are (too) close to each other.

3. POSTPOSITIONS AND AUXILIARIES IN MANDE

It was noted above that many postpositions in Mande languages are indisputably nominal in origin. There are some, however, for which a verbal etymology is much more likely. Bird and Kendall (1986) point out that in several northern Mande languages certain forms are used as both auxiliaries and postpositions. These have in the past been treated as homophones, but the number of cases is too numerous to be accidental. An example is the Bambara form *ye*, which is used as (i) the affirmative transitive completive auxiliary, (ii) the equative copula, and (iii) an associative, dative/benefactive, and locative goal postposition. Examples of each of these uses (data from Bird and Kendall, 1986):

(31) a. *À ye nìn kɛ*
he AUX this do
'He did this.'

b. *Nìn ye Ali ye*
this be Ali ASSOC
'This is Ali.'

c. *Ali kɛ-ra dùgu-tigi ' ye*
Ali do-PAST village-chief the ASSOC
'Ali became the chief.'

d. *N bɛ nìn kɛ* Ali *ye*
I AUX *this do Ali* BEN/ASSOC
'I am doing this for/with Ali.'

In the closely related language Maninka (Kita dialect), the form *di* covers roughly the same territory as Bambara *ye*, viz. (i) recent completive auxiliary, (ii) associative postposition. In the more distantly related Soninke language, the form *dà* is both (i) transitive completive auxiliary and (ii) benefactive postposition. Also in Soninke the locative copula *wa* is identical to the locative postposition *wa* (though the latter is ordinarily converted to *ŋa* by a morphological process). Examples of these forms from Soninke (data from Bird and Kendall, 1986) are:

(32) a. *Ali wa saxa ǹ ŋa*
Ali be.at market the at
'Ali is at the market.'

b. *Ali dà nàni ǹ xami faaba dà*
Ali AUX cows the milk father BEN
'Ali milked the cows for his father.'

Bird and Kendall point out that often the auxiliaries in Mande languages cliticize onto the preceding subject noun phrase. This phonological observation, and the fact that both auxiliaries and postpositions follow noun phrases, leads them to hypothesize that auxiliaries in Mande languages somehow 'originated in the subject phrase' (p. 397), perhaps as some sort of case marking. This is in the light of the Senufo data a most unlikely explanation. The phonological encliticization of the auxiliaries onto the preceding subject phrase, which is widespread not only in Mande, but in Senufo and Kru as well, is entirely an accident of their position. If they had followed the verb, they would undoubtedly have developed into verbal suffixes.

A much more likely explanation of the 'homophony' between auxiliaries and postpositions is that they have a common verbal origin. It is undisputed that some of the auxiliaries have verbal origins (e.g. Bambara *bɛ* 'affirmative incompletive auxiliary' and 'be at'; *na* 'future auxiliary' and 'come'). Certainly the search for the etymologies of Bambara *ye* and Soninke *wa* and *di* should take verbs into consideration.[17]

If the above conjecture is correct, then some of the Mande postpositions may have arisen from a serial verb source as proposed for Senufo languages.

4. FROM SERIAL VERB TO POSTPOSITION

The development of adpositions from serial verbs is too well-known to require justification (cf. Lord, 1973; Givón, 1975; Heine and Reh, 1984; Heine

and Noske, 1988; Craig, Volume II). As an explanation for the origin of [S AUX O V OBL] word order in Senufo languages, however, the serial hypothesis raises at least three problems.

The first problem concerns transitivity. In their role of introducing participants into the clause, the case marking serial verbs were used transitively, or at the very least they were preceded by an otherwise unmarked object. However, some of the verbs proposed as cognates to the postpositions are ordinarily intransitive in modern Senufo languages. The verb *pan/má* 'come', for example, cannot normally be preceded by a locative object. In Supyire it can be used factitively, but in that case the direct object denotes something which is caused to come, such as rain, and the subject can only be God. Locative objects must follow:

(33) *U à pa kànhà nà*
S/he PERF come village to
'S/he came to the village.'

The serial hypothesis therefore requires that in proto-Senufo (or earlier) *pa* could be preceded by its locative object. This is precisely the order found in serials in Dogon (Prost, 1969b:65):

(34) *ɛŋə digɛ dyɛ Pɛlu* yɛrɛ
goats lead take Pel come
'Take the goats to Pel.'

Starting from a serial construction like that currently found in Dogon, the grammaticalization of *pa* (or *má*) as a postposition would be natural.[18]

As pointed out above, Senufo languages show a great deal of flexibility in the transitivity of many verbs. This is illustrated with the verb *kan* 'give' in Supyire. *kan* may take both a direct and an oblique dative object. When it appears as a serial verb, it usually takes only a dative object, but this may be either an oblique or a direct object:

(35) a. *U a kù kàn mìì-á*
s/he PERF it give me-to
'S/he gave it to me.'

b. *U a kù lwɔ́-à kan mìì-á*
s/he PERF it take-NF give me-to
'S/he took and gave it to me.'

c. *U a kù lwɔ́-á mìì kàn*
s/he PERF it take-NF me give
'S/he took and gave me it.'

The last example above illustrates the kind of construction which was the source of the Karaboro benefactive postposition *kɔ̃*.

A second problem faced by the serial hypothesis is the obligatory ordering of obliques after the verb. Put succinctly, an original OV word order predicts only the postposing of the adposition, not the post-verbal position of the entire adpositional phrase. Case marking serial verbs need not always follow the semantic main verb, as the following example of instrument marking from Supyire shows:

(36) *U à ŋwɔɔní tàha-a kyaàre kwɔ̀n*
s/he PERF knife.DEF use-NF meat.DEF cut
'S/he cut the meat with the knife.'

To explain the [V OBL] order, some explanation must be found for the obligatory placing of the grammaticalized serial in second position.

It is worthwhile noting in this connection that the ordering of serial verbs is generally not flexible. When both of the verbs are semantically main verbs, the ordering generally is iconic to temporal ordering of the events in the world. Thus 'take and come', 'hit and kill' follow the ordinary chronological order of the events concerned. Grammaticalized serial verbs similarly have a fixed position. The instrument marking verb in (36), for example, must precede the main verb. When *kan* is used to mark benefactive, as in (35c), it must come second. These orderings are obviously iconically motivated.

Since serial verbs marking dative/benefactive and locative goal would come second for temporal iconic reasons, it is not surprising that *má* obliques must follow the verb. The other locative postpositions may all be used with either static or dynamic senses. It is difficult to see how iconicity would affect the ordering of the former uses: 'she house be.in sleep' is not more or less iconic than 'she sleep house be.in'. However, in dynamic uses, where the locatives ordinarily encode goal, iconicity would lead to placement in second position: 'she go house be.in' is motivated whereas 'she house be.in go' is not. If grammaticalization took place first with motion verbs, the postverbal ordering would be most natural. It is of course also possible that [V OBL] ordering in neighboring languages played a role in the rigidification of word order in proto-Senufo.

The third and final problem to be addressed here is that of word order

in copular clauses. As noted at the beginning of this paper, predicate nominals ordinarily follow rather than precede the copula. Heine and Reh (1984) regard this as a remnant of earlier SVO order in Mande, whereas Williamson (1986) considers it an innovation.[19] In support of the latter position, it is certainly suggestive that Bambara requires the postposition *ye* following predicate nominals in equative clauses, in which the copula itself is *ye* (see 31b above).

In Section 2.1 and 2.2 it was shown that the locative verbs *na* and *ni* gave rise to progressive auxiliaries as well as to postpositions. In the postpositional use, it was claimed that they were used transitively, with their locative object preceding. It is well known that progressive constructions often arise from locative ones in which the main verb was originally nominalized or at least non-finite. If such a construction underlies the Senufo progressive, then of course the ordering [S AUX V] reflects the current copular ordering [S COP PN], and differs from the ordering claimed for the serial verb use presumed to underly the postpositions.

There is, however, another possibility. If other auxiliaries developed not from main verbs with embedded clauses, but from initial serial verbs used intransitively, the same possibility is open for *na* and *ni*. The following verb (and its arguments) never was the syntactic direct object of *na* or *ni*, but simply followed it in a serial construction similar to those in (1) above. Word order in copular clauses would thus be a later, and independant development, possibly itself influenced by the grammaticalization of the postverbal position for oblique objects.

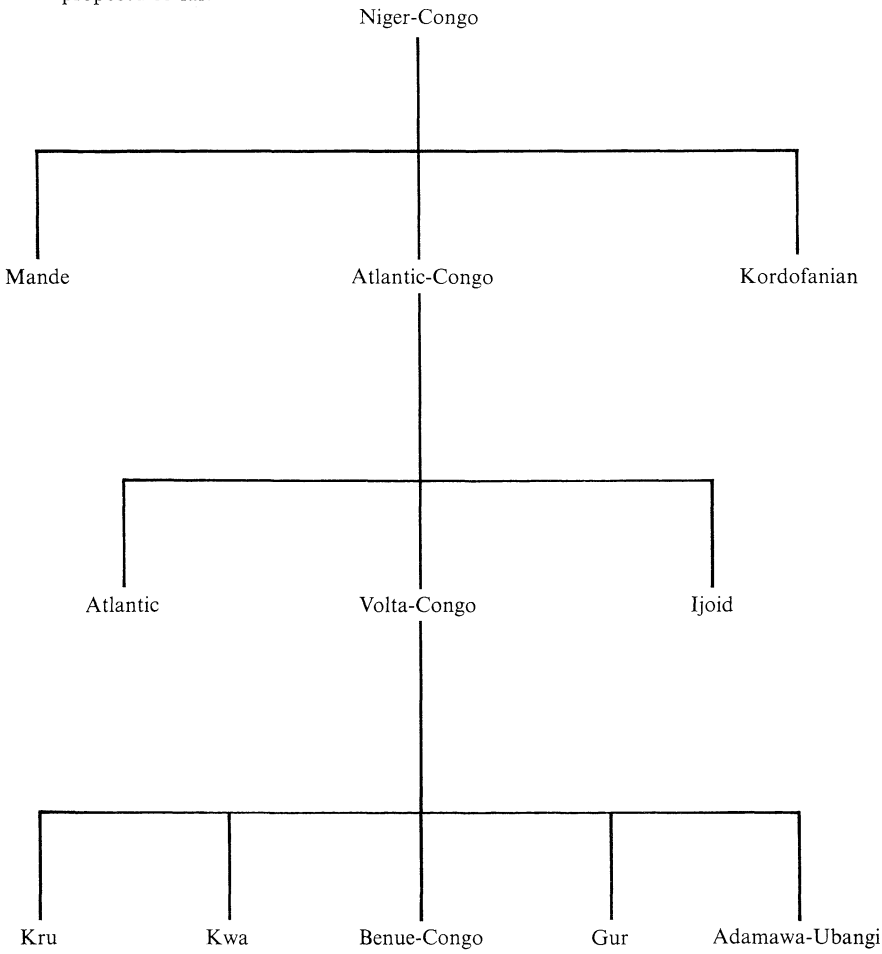
If these arguments are plausible, the serial hypothesis offers an explanation of the mixed SOV word order of Senufo languages. The original language was more strictly verb-final, but allowed serial verb constructions in which the direct objects of the verbs intervened. The evidence points in fact to word ordering somewhat like that found in present-day Dogon and Ijo.

ABBREVIATIONS

HAB	habitual auxiliary
INCOMP	incompletive
NF	non-final marker in serial constructions in Supyire, marks all but final verb
PN	predicate nominal

NOTES

1. The following classification of Niger-Congo (Bendor-Samuel, to appear) is the best proposed so far:



2. Claudi (1988) tries to show that Marchese is wrong in assuming an earlier OV order, proposing instead that the direct object in [S AUX O V] clauses was made the direct object of the auxiliary verb by a process she calls 'complement serialization', and that the nominalized V was 'added as a second complement of the auxiliary verb'. Apart from its *ad hoc* character, this proposal is unconvincing chiefly because it assumes attraction of an object which is semantically incompatible with the auxiliary verb until the latter has progressed so far on the road of grammaticalization and semantic bleaching as to make it unlikely that it would acquire a direct object, let alone two of them.

3. This is in basic accordance with the hypothesis advanced in Williamson (1986:7): 'Proto-Niger-Congo was a strict SOV language with verb serialization.' In some languages, auxiliaries developed from the second verb in a series, in other languages from the first verb. Williamson proposes that subsequently the verb 'became nominalized' after the auxiliary, though how this would happen is not at all clear.

4. The tones in the Supyire examples are surface rather than underlying tones. The great many morphotonemic processes in the language account for the variety of surface tones a given word may have.

5. Some Senufo languages have developed uses of denominal postpositions without the following simple postposition. In Supyire for example both *kantugo* 'behind' (originally a noun meaning 'back') and *kantugo yyéré* (*yéré* 'toward') are possible. Other languages such as Cebaara require the use of the simple postposition *má* with the cognate of *kantugo*.

6. It is perhaps worth emphasizing here that grammaticalization is something done by speakers, i.e. by agents using language in a purposeful way. It is not a mysterious epigenesis of some sort of semantic genotype which only needs the right environment to unfold. Speakers increasingly automate their speech because it suits their purposes to do so. The telicity of a grammaticalization 'trajectory' is due to the telicity of human behavior, not to some reified 'nature' attributed to lexemes or concepts.

7. Most of the languages have *na*, while Nanɛrghɛ and Sucite have *la*. There are frequent differences in nasalization between Senufo languages (e.g. Cebaara *pan* [*pã*], Supyire *pa* 'come'; Cebaara *can* [tsã], Supyire *ce* 'know'; Sucite *ne*, Supyire *le* 'put') of which this would be an example. According to the hypothesis in Bole-Richard (1982), there were probably two variants in the proto-language, **la* and **lã*.

8. There may have also been an element of motion in the original verb. In his list of 'preliminary' roots for Niger-Congo, Bole-Richard (1988) lists both *na* 'walk, go, come', and *na* 'be, stay'.

9. Clamens (1952:1426) was apparently the first to note that the postposition *na* and the 'present' auxiliary were 'the same' in Tagbana. This is a good example of what Hopper (Volume I) calls "divergence".

10. Supyire and Sucite have two mid tones, one 'weak', which perturbs in most environments, and the other 'strong', which is relatively stable, except that it accepts spreading tones under certain conditions. Garber (1987) presents evidence that the weak mid developed from a split of the proto high tone into an upper and a lower register.

11. It is quite possible that the various *ni* forms attested in Senufo languages are cognate with the copula *nye* or *nyɛ* 'be' which is found in most of them. Bole-Richard (1988) lists *nyi* as a variant of *ni* 'be'.

12. Given the flexible nature of transitivity in Senufo languages, it is not impossible that the verb 'put in' (Sucite *ne*, Cebaara *le*, Supyire *le* in some dialects, *ni* in others) is related to the postposition. A similar factitive use is possible with many verbs in the modern languages, as the following from Supyire demonstrate: *pèè* 'be big; make big', *wwɔ* 'be black; blacken', *pa* 'come; cause to come'.

13. Mamara has *mú* and Sucite has *-ũ*, apparently rounding the vowel following a labial. Supyire has *-á*, having dropped the initial [**m**] except in poetry.

14. In Carlson (1985) I suggested that Supyire *-á*, which is used almost exclusively to mark dative/benefactive rather than locative goal, was perhaps related to the verb 'give' in

Akan languages (e.g. *mã* in Anyi). With the evidence now available from other Senufo languages, the etymology given here is much more likely. The etymology from 'come' was first suggested to me by Betty Mills. The relation between *pan* and *má* is morphologically irregular. It is not clear whether this is a case of suppletion, or the sole relict of an earlier derivational process.

15. The final *-na* of both of these verbs may possibly be an old frequentative or 'plural' suffix, added when the direct object is plural and/or the basic action is repeated several times. Some pairs from Supyire are: *kùrù* 'fold, bend' and *kùrùnɔ́* 'coil up'; *láhá 'take off'* and *láhálá* 'peel'; *múgó* 'open' and *múgúló* 'open several objects'; *sanha* 'untie' and *sanhana* 'unroll'.

16. The postposition in these languages is high tone, whereas the verb is mid tone. The distinction between weak and strong mid does not exist in Shenara and Cebaara.

17. As a tentative suggestion, Bambara *ye* is perhaps related to the root **yi* 'come; go' listed in Bole-Richard (1988). There is a tantalizing similarity in Senufo: as shown above, the dative/benefactive postposition *má* comes from the verb *má* 'come'. In northern and central Senufo languages the perfect auxiliary is *mà*. The tone of course is wrong, but the similarity to the dual function of Bambara *ye* is striking.

18. This is also the order found in Kru languages in clauses with auxiliaries. Marchese (1986) points out that the locative objects of 'come' and 'go' in many Kru languages are syntactically like direct objects. See also Welmers' comments on other Niger-Congo languages (1973:454).

19. In Ijo, which together with Dogon is the most strictly verb-final of any group in Niger-Congo, the copula, when there is one, is ordinarily final. Some dialects are developing a medial copula, however, and Williamson suggests that a similar development may have occurred in Mande.

REFERENCES

Anon. 1974. *Tyebara Grammar Lessons*. Korhogo: Mission Baptiste.

Anon. 1983. *Lexique du Senufo de Ndorola*. Bobo Dioulasso: Mission Catholique.

Bailleul, C. 1982. *Dictionnaire Bambara-Français*. England: Avebury Publishing Company.

Bendor-Samuel, J. To appear. *Niger-Congo*. Dallas: Summer Institute of Linguistics.

Bird, C. and M. Kendall. 1986. "Postpositions and auxiliaries in northern Mande: Syntactic indeterminacy and linguistic analysis." *Anthropological Linguistics* 28(4):389-403.

Bole-Richard, R. 1982. "Hypothèse sur le développement des nasales en Niger-Congo." Paper presented at the 15th Congress of the West African Linguistics Society, Port Harcourt.

Bole-Richard, R. 1988. "Le fonds lexical commun Niger-Congo." Paper presented at the 18th Congress of the West African Linguistics Society, Niamey.

Boutin, P. 1983. "Éléments pour une systématique du Fɔndɔndɔ." Thèse de III cycle, Paris III.

Carlson, R. 1985. "Grammaticalized verbs in Supyire." In *Proceedings of the First Annual Meeting of the Pacific Linguistics Conference*, S. Delancey and R. Tomlin (eds). Eugene: University of Oregon.

Cauvin, J. 1980. *L'image, la langue et la pensée: l'exemple des proverbes (Mali)*. St. Augustin: Anthropos-Institut.
Cissé, N. 1986. *Lexique Tuvunro-Français*. Bamako: Direction Nationale d'Alphabétisation Fonctionelle et de Linguistique Appliquée.
Clamens, G. 1952. "Essai de grammaire sénoufo-tagwana." *Bulletin de l'Institut Francophone de l'Afrique Noire* 14(4):1402–1465.
Claudi, U. 1988. "The development of tense/aspect marking in Kru languages" (review of Marchese 1986). *Journal of African Languages and Linguistics* 10:53–77.
Dakubu, (ed.). 1980. *West African Data Sheets*.
Dembele, E. 1983. *Lexique Mamara-Francais*. Bamako: Direction Nationale d'Alphabétisation Fonctionelle et de Linguistique Appliquée.
Garber, A. 1987. A Tonal Analysis of Senufo: Sucite Dialect. Urbana: University of Illinois dissertation.
Givón, T. 1975. "Serial verbs and syntactic change: Niger-Congo." In *Word Order and Word Order Change*, C. Li (ed.), 47–112. Austin: University of Texas Press.
Heine, B. 1976. "A typology of African languages based on the order of meaningful elements." In *Kölner Beiträge zur Afrikanistik* 4. Berlin: Dietrich Reimer.
Heine, B. 1980. "Language typology and linguistic reconstruction: the Niger-Congo case." *Journal of African Languages and Linguistics*. Leiden, 2(2):95–112.
Heine, B. and Manuela Noske. 1988. "Adpositions in African languages. Paper presented at the Symposium on Grammaticalization, University of Oregon, May 12–15, 1988.
Heine, B. and M. Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
Hopper, Paul. Volume I. "On some principles of grammaticization."
Hyman, L. 1975. "On the change from SOV to SVO: Evidence from Niger-Congo." In *Word Order and Word Order Change*, C. Li (ed.), 113–147. Austin: University of Texas Press.
Jordan, D. 1980. "Nafaara list". *Dakubu* 1980.
Laughren, M. 1973. Analyse plérématique du tyebari (un dialecte sénufo). Thèse du Doctorat du 3e cycle, Nice.
Lejeune, n.d. "Essai de grammaire Dyimini." ms.
Lord, C. 1973. "Serial verbs in transition." *Studies in African Linguistics* 4(3):269–297.
Marchese, L. 1986. *Tense/aspect and the Development of Auxiliaries in Kru Languages*. Arlington: Summer Institute of Linguistics and University of Texas Press.
Means, M. (ed.). 1966. *Tyebara-English, English-Tyebara Dictionary*. Korhogo: Mission Baptiste.
Mills, E. 1984. *Senoufo Phonology, Discourse to Syllable*. Arlington: Summer Institute of Linguistics and University of Texas Press.
Mills, E. 1987. *Kàa joo! cours de senoufo (senanri)*. Abidjan: Université Nationale de Côte d'Ivoire.
Prost, A. 1964. *Senar. Contribution à l'étude des langues voltaiques*. Dakar: IFAN.
Prost, A. 1968. *Deux langues voltaiques en voie de disparition, le Wara et le Natioro*. Dakar: IFAN.
Prost, A. 1969a. *Les parlers Dogons: Tɔgɔ Kã*. Dakar: Université de Dakar.
Prost, A. 1969b. *Les parlers Dogons: Donno Sɔ*. Dakar: Université de Dakar.
Welch, n.d. "Notes on the grammar of Pilara." ms.
Welmers, W. 1973. *African Language Structures*. Berkeley: University of California Press.
Williamson, K. 1986. "Niger-Congo: SVO or SOV?" *Journal of West African Languages* 16(1):5–14.

SECTION C: SUBORDINATION

From Postposition to Subordinator in Newari

Carol Genetti
University of California, Santa Barbara

1. INTRODUCTION

In Tibeto-Burman languages it is common for case postpositions, markers of concrete spatial or social relationship between entities, to develop into adverbial subordinators[1], markers of temporal and logical relationships between propositions.[2] The syntactic mechanism by which such a development occurs is nominalization, which allows for verbs to be inflected with nominal morphology. The prevalence of this phenomenon in Tibeto-Burman is part of what led Konow to write in *The Linguistic Survey of India*:

> Another consequence of the nominal character of the Tibeto-Burman verb is that it can be used in connexion with postpositions like an ordinary noun. In this way the verbal noun is used to form various kinds of adverbial sentences. (1909:9)

The syncretism between postpositions and subordinators is illustrated in the following two sentences, taken from texts in the Dolakhali dialect of Newari:

(1) Dolakhali Newari
*cotan-**na*** *pol-ju*
spoon-INSTR strike-3SG:PAST
'He hit it with a spoon.'

(2) Dolakhali Newari
chẽ-ku *yer-**na*** *wā* *ām-e* *naku* *moŋ-an*
house-LOC come-when EMPH he-GEN cheek swell-PART
coŋ-gu
stay-3SG:PASTHAB.
'When he came to the house his cheek was swollen.'

In the first sentence, the ergative/instrumental postposition *na* marks the instrumental relation. In the second sentence, *na* is suffixed to the verb of a subordinate clause, and indicates a temporal relationship between the subordinate clause and the main clause.

In an earlier paper (Genetti, 1986), I presented a typological study of the syncretism between case postpositions and adverbial subordinating conjunctions in the Bodic branch of Tibeto-Burman.[3] All twenty-six languages of the study exhibited some syncretism between the two grammatical categories; the extent of the syncretism was too great to be attributable to chance, but implied a pattern of development whereby the function of postpositions was grammatically extended so that they developed into subordinators. The results of the study will be summarized below.

The phenomenon of syncretism between postpositions and clause linking morphology is not restricted to Bodic languages. Thurgood (1986) cites evidence of the same syncretism in the Lolo-Burmese branch of Tibeto-Burman. Austin (1981), along with a number of papers in Austin (1988), clearly establishes this syncretism in Australian languages. Craig (Volume II), Reesink (1988) and Lichtenberk (Volume I) give examples of the same syncretism in Chibchan, Papuan and Austronesian languages respectively.

The primary goal of the previous study was to offer a semantic explanation for a set of typological data. Clearly lacking in that paper was any detailed account of the mechanism by which such grammatical extension[4] occurred, and the syntactic readjustments that must accompany such a development. The current study is meant to supply this information. It is an examination of the relationship between postpositions and subordinators in the Kathmandu dialect of Newari, a Tibeto-Burman language of Nepal. I will show that in Newari, the development of postpositions into subordinators occurred repeatedly over the last several centuries, with each cycle of development following the same semantic patterns found in the typological study. I will also show that the morphosyntactic mechanism by which the development occurred was nominalization, followed by a reanalysis of originally nominal morphology as verbal morphology, via the reanalysis of unmarked deverbal nominals as erstwhile finite verbs.

1.1. The Typological Study Revisited

From the typological study it is possible to draw two important conclusions. The first is that the frequency of syncretism between case postpositions and clausal subordinators is great enough to imply a common trend of

grammatical extension. The second is that the extension follows regular patterns, such that postpositions with a given semantic value develop into a consistent set of subordinators, usually consisting of only one or two members. The patterns of extension are as follows:[5]

(3) Locative > if/although, when/while/after
Ablative > when/while/after, because, non-final[6]
Allative > purpose
Dative > purpose
Ergative/Instrumental > because, when, while

Each postposition in the list is followed by the subordinators that are most commonly marked by the postposition. The subordinators are listed in order of frequency. Thus it is most common to find ablatives functioning to mark temporal subordination, but there is also a less common pattern where ablatives mark causation, or mark non-final clauses in clause chains. The following examples illustrate each of the commonly attested patterns listed in (3).

LOC-IF Classical Tibetan (Bacot, 1948)

(4) *gyas* ***na*** *bsgyur*
right LOC turn
'Turn to the right.' (1948:63)

(5) *me yod* ***na*** *du-ba 'byung*
fire be if smoke become
'If there is fire, there is smoke.' (1948:64)

LOC-WHEN Bunan (Konow, 1909)

(6) *lag-****tog*** *la-sab bang-****dog*** *trad-pa re tsua*
hand-LOC ring feet-LOC shoe also put
'Also put rings on his hand and shoes on his feet.' (1909:477)

(7) *leb-cha-****astog****...*
arriving-when
'When he arrived...' (1909:477)

ABL-WHILE Vayu (Michailovsky, 1981)

(8) *ban bon laxtse a kem-nong-****na***
forest fly go her house-LOC-ABL
'She flew into the forest from her house.' (1981:252)

(9) *mithe-nong ko dinekal dzok-**na-na**...*
then TOP daily going-ABL-ABL
'Then while (he) was going each day...' (1981:261)

ABL-BECAUSE Gurung (Glover, 1974; Glover & Gurung, 1979)

(10) In the chart of Gurung postpositions (1974:70) is listed *seró* 'from (temporal)'

(11) *khiba mã:bá phódẽ ta-**seró**...*
old.man old.woman starving become-ABL
'Because the old man and woman starved...' (1974:76)

ABL-NON-FINAL Khambu (Rai) (Konow, 1909)

(12) "The suffix of the ablative is *ka*, which is usually preceded by one of the suffixes ***la*** or ***bi***; thus *pu-la-**ka*** 'from a father'; *tam-bi-**ka*** 'from her'." (1909:381)

(13) *tik janā chākara bra-tu-**ka** mā-wa ānā siku-bi*
one person servant call-ABL what.is saying asked
'He called a servant and asked "What is it?".' (1909:381)

LOC/ALL-PURPOSE Thulung (Allen, 1975)

(14) *ramli-ka rokomalung-**Da** kho sebDiu*
Ramli-ERG stone-LOC ax sharpened
'Ramli sharpened the ax on a rough stone.' (1975:149)

(15) *yo breb-**Da** ləksa*
salt buy-PURP go
'Go and buy some salt.' (1975:58)

DAT-PURPOSE Newari

(16) *jĩ-ĩ Raj-yā-**ta** biy-ā*
I-ERG Raj-GEN-DAT give-PAST
'I gave it to Raj.'

(17) *jĩ-ĩ kerā nyā-e-**ta** wan-ā*
I-ERG banana buy-INF-DAT go-PAST
'I went to buy some bananas.'

ERG/INSTR-BECAUSE Lhasa Tibetan (Goldstein, 1973)

(18) *stag mangpo yod pa-**s** nga-**s** gcig bsad pa-yin*
tiger many have NMNL-ERG I-ERG one kill PFV
'Because there are many tigers I killed one.' (1973:51)

ERG/INSTR-WHEN Thakali (Hari and Maibaum, 1970)

(19) *'nakyu-**ce** 'pohr-si yah-ci mu ro*
dog-ERG take-ANT go-PAST is PRTCL
'So the dog took her and went.' (1970:187)

(20) *taalwri pih-si pih-wa-**ce**...*
why say-ANT say-NONPAST-when
'When she asked why...' (1970:258)

The semantic framework which most readily accounts for this data is a localistic framework proposed by Diehl (1975). Diehl's approach to semantics differs from other localists (such as Gruber (1976) and Anderson (1971)) in a number of ways, but his primary innovation is the extension of localist theory into two dimensions. The first dimension contains the four deep semantic cases common to localistic theory: Theme, Location, Source and Goal. The second dimension is that of four semantic spheres or 'spaces' in which the four basic case relations function. The four semantic spheres are:

(21) Social: indicates location with respect to positions defined or interpreted as human beings.
Spatial: locates physical objects in space.
Temporal: locates events in time.
Logical: locates propositions with respect to each other.

These four spheres are inherently ordered based on an 'egodeictic hierarchy' which correlates with relative abstractness. The basic cases can thus be extended into increasingly abstract spheres to express slightly different semantic meanings, as represented in Table 1.

The patterns of grammatical extension listed above can thus be seen as a process by which the basic case relations are extended to more abstract domains, as they proceed from coding relations between arguments to relations between propositions.[7]

Table 1 *Distribution of Cases in Four Semantic Spheres (Based on Diehl 1975)*

	Location	*Source*	*Goal*
Social	associative	ergative	dative
Spatial	locative	ablative	allative
Temporal	when/while	since (abl.)	until
Logical	if	because	purpose

1.2. Newari

Newari is spoken by approximately 500,000 people, the majority of whom live in the Kathmandu Valley of Nepal. The precise position of Newari within the Tibeto-Burman family is unclear, although current data suggests that it may be related to the Bahing-Vayu branch (Genetti, 1987). Newari is one of a handful of Tibeto-Burman languages which has a written tradition. The oldest surviving documents written in Newari date back to the late fourteenth century, and there still exists a considerable amount of material dating from the fourteenth to the late nineteenth century. The language used in these documents is called Classical Newari, a term which does not refer to any one synchronic stage. The main accessible materials on Classical Newari are a grammar and dictionary by Hans Jørgensen (1936 and 1941 respectively), and two series of texts which Jørgensen edited and translated, one set being from the eighteenth century (Jørgensen, 1939), and one from the nineteenth century (Jørgensen, 1931). There is also a series of historical documents dating from the late fourteenth century, edited and translated by Vajracarya and Malla (1985). The Classical Newari data will be complemented by my own data on modern Newari of Kathmandu. This dialect is directly descended from the Classical Newari represented in the manuscripts.

The next section will examine these dialects in detail, focusing on the etymological sources of the subordinators found at different historical stages, and on the stages of development evidenced in the available textual materials. The following section will address the syntactic issues involved. The final section will discuss to what extent the phenomena described here are similar to prototypical examples of grammaticalization.

2. POSTPOSITIONS AND SUBORDINATORS IN CLASSICAL AND KATHMANDU NEWARI

2.1. Background

Before proceeding to a discussion of the etymologies and behavior of the Newari subordinators, I will present some background information on syntax and morphology. Newari is rigidly verb final, and the elements preceding the verb have a relatively free order. Like all Tibeto-Burman languages it is postpositional. Case marking follows an ergative pattern; patients can be unmarked or marked with the dative postposition.

The postpositions in Classical and Kathmandu Newari are given in Table 2.

Table 2 *Postpositions in Classical and Kathmandu Newari*

	CL. NEW	KATH. NEW
ERG	së-M, së-n[8]	nã[9]
INSTR	n, (na)[10]	nã
ABL	yã-ke-n	nã
LOC	s, (sa)	e
ASSOC	wo	(yã-)ke
DAT	(yã, s)-ta	(yã-)t0
ALL	twoM	e
GEN	yã, s[11]	yã

The first thing to note in Table 2 is that the Kathmandu Newari inventory of postpositions is much simpler than that of Classical Newari. This is due in part to the extension of the instrumental to the ergative and ablative cases[12], and the resultant loss of the old ergative and ablative markers.

Also note that we can see in this table the common Tibeto-Burman tendency to form complex postpositions out of concatenations of simple case elements (DeLancey, 1984). Thus the Kathmandu Newari associative and dative can be optionally preceded by the genitive (the conditioning factors for this use of the genitive are currently unknown). The final nasals in the Classical Newari ergative and ablative are most likely the remnants of an older concatenation with the instrumental morpheme. The first element of the ergative *së-M* may well be from an earlier concatenation of the locative

sa plus an old Tibeto-Burman ablative in the form of a front vowel (DeLancey 1983; Hargreaves, 1984).

Let us turn now to the Classical and Kathmandu Newari verb inflection. The system of finite inflection did not undergo much change from the Classical to modern dialects, and will not be discussed here in detail (see Hale, 1980; Genetti, 1988 for a discussion of the modern conjunct/disjunct system). Note however, that four of the finite verb forms in Classical Newari be used with a nominal function, and can be marked with postpositions to explicitly mark subordination (from Jørgensen, 1941:52; the abbreviations E,G,D,I,L, and S stand for ergative, genitive, dative, instrumental, locative and associative respectively):

(22)

Label	*Function*	*Derived KN Form*	*Suffixes*
A3	non-past/habit.	Non-Past Disjunct	G
A4	1st person irrealis	Non-Past Conjunct	G,D,I,L,S
A5	3rd person	Stative-Habitual	E,G,I
A6	1st/2nd person	Past Conjunct	G,I,S

When we compare the non-finite forms of Classical and Kathmandu Newari, we see that there have been some significant changes. Jørgensen identifies twelve non-finite forms in Classical Newari, as opposed to five in the modern dialect[13]. Of the twelve non-finite forms, five are formed on a 'short base' and function as infinitives and participles, and seven are formed on a single invariant 'long base'. Of these seven forms, four have in addition a suffix *ŋā* which comes between the long base and the final suffix:

(23) Long base forms, illustrated with the verb *kan-* 'tell':
kana-sā
kana-san-woM
kana-ŋā-seM
kana-ŋā-s
kana-ŋā-wo
kana-ŋā-n
kana-seM

Note that Jørgensen considers these forms all to be separate verbal inflections, whereas all the subordinators with finite verbs for the base he considers to be finite verbs suffixed by postpositions. This is probably due to the lack of virtual identity between the postpositions and some of the suffixes in this list (i.e. *sanwoM*).

In the modern Kathmandu dialect, there is no longer any trace of the long base. All subordinators are formed with inflected finite verbs; the only exception to this is the purposive *ta* which is suffixed to the infinitive, a form which cannot be derived from the long base. Thus in the development of the modern dialect, all forms with the long base were lost.

2.2. Subordinators

With this background information, let us look now at the form of subordinators in Classical and Kathmandu Newari. For the purposes of this paper I will focus on the forms found in the eighteenth, nineteenth and twentieth century texts. I also will limit my data primarily to those forms which appear to be derived from postpositions, and to mention only in passing those with nominal origin. I will divide the subordinators into classes by semantic content.

My methodology for studying the Classical Newari texts was to search for subordinate clauses first in the English translation, then to cross-reference my examples in the Newari texts. A better method would be to work from the Newari version directly, however at this point that project is not practical, since it would require a detailed analysis and gloss of all of the texts in question. Therefore the numbers given in the tables should not be considered absolute. The distributions in the Classical texts are insufficient to draw any firm conclusions, but they are still interesting in that they represent probable sequences of development.

Temporal. The subordinator most commonly used to explicitly signal a temporal relationship between clauses both in Classical Newari and the Kathmandu dialect is *belas* (in the modern dialect *bale*), which is formed from a noun meaning 'time'. There are some examples of other strategies in the Classical Newari texts; their distribution is given in the Table 3.[14]

Table 3 *Distribution of Temporal Forms*

	Batis.	*Vicit.*	*Kath.*
long base + syaM	5	0	0
long base + ŋā + va	5	0	0
stem + FI + naM	0	1	0

The two forms which occur with relative frequency in the eighteenth century texts are both built on the long base. The suffix of the first form

syaM is clearly the ergative *sëM* (the vowel /ë/ being a later reflex of /ya/). Thus this is an example of the ergative being used to code a temporal connection between clauses. The second form is built on the long base plus the synchronically meaningless element *ŋā*. The final suffix is transparently the associative postposition *va*. Both strategies for coding 'when' clauses are predicted by the localistic schema discussed above.

In the nineteenth century texts, neither of the above strategies are found. Instead, we find only one example of 'when' coded by a postposition; this is the ergative/instrumental *nã*, and it is suffixed to a verb with finite inflection. Note that this is the same semantic concatenation of the eighteenth century form with the long base plus *syaM*. In the nineteenth century data, the verb is finite, in keeping with the general trend away from using the long base and toward using finite inflection in these constructions. Also, the postposition is the newer form of the ergative, rather than the older, fading *syaM*. However, this particular concatenation did not gel into a regular pattern. This is not surprising since the subordinator *belas* is used in the large majority of cases in the late nineteenth century, and is the only form attested today; forms meaning 'when' constructed with the ergative postposition cannot be elicited from speakers of modern Kathmandu Newari.

Causal. Forms indicating a causal relationship between clauses are distributed through the texts as shown in Table 4.

Table 4 *Distribution of Causal Forms*

	Batis.	*Vicit.*	*Kath.*
stem + FI + naM	6	2	0
stem + a + M	2	0	0
stem + FI + n	1	0	0
stem + FI + guli + n	0	1	28

Although the first three forms in this table are listed separately, they are actually all instances of the same concatenation, finite inflection plus some form of the ergative/instrumental. The two examples with *-a-M* are both formed with the causative: *dayakaM* 'by letting' (lit. 'by causing to be'), *juyakaM* 'by making it be'. Jørgensen's grammar only lists the causative stem *-ka* as a form used in compounding, so its use in this construction is odd. These cases may well be collapsible with the following class.

The final form in the table is interesting because it clearly represents a newer concatenation than any of the other cases. *guli* is a nominalizer. As

will be discussed in more detail below, consistent overt marking of nominalization is a relatively new feature in Newari. Whereas none of the other postpositions underwent re-inflection following overt nominalization,[15] the ergative/instrumental was suffixed to newer nominalized forms. This form has now become frozen. In all other cases the nominalizer *guli* was reduced to *gu* by a regular process (this involved deletion of *Ci* sequences in second syllables following /u/, Malla, 1985:45). However, the form *gulin*, with the suffix *n*, did not undergo the sound change. Thus the modern form *gulĩ* is now frozen as a single unit. It is also interesting to note that Jørgensen (1941:73) lists another causal subordinator, formed by the long base plus *ŋā* plus the same ergative/instrumental *n*. This form was not found in the manuscripts, and so may have been an older form replaced by the forms with finite inflection. Note that it is formed with the same concatenated elements as the other forms. Thus we can see three different stages of 'because' using the same ergative/instrumental morpheme; first it is suffixed to the long base, then to the finite inflection, then to the finite inflection plus nominalizer.

Conditional. The primary subordinator indicating a conditional clause is *sā* throughout the Classical Newari texts and in the modern dialect. This morpheme is probably etymologically related to the locative *sa*, although the vowels differ in their degree of backness (possibly due to the laxing of the vowel in the locative to schwa, an allophone of the back vowel). Since the two forms are different phonologically in all of the available material, we can infer that they are synchronically separate morphemes. So although this form may be diachronically related to the locative, in our data it clearly only functions as a subordinator. The distribution of this and other 'if' forms in the texts is given in Table 5. The forms with *sā-M* all can be translated as 'even if/although', and consist of *sā* plus the contrastive particle *nã*. Although this form is not attested in the twentieth century texts, it can easily be elicited

Table 5 *Distribution of Conditional Forms*

	Batis.	*Vicit.*	*Kath.*
long base + sā	117	19	0
long base + sāM	2	6	0
long base + ŋā + s	2[16]	1	0
stem + FI + sā	0	6	40
stem + FI + sāM	0	1	0
stem + FI + ki	0	0	6

from Kathmandu Newari speakers. The two forms of the long base plus *ŋā* plus *s* can be considered as the long base plus the *sa* locative; assuming the above analysis of the etymological relationship of *sā* 'if' and *sa* LOC to be correct, then these are two separate examples of the same semantic concatenation being used to form a single subordinating function.

We see here an interesting pattern in the distribution of the forms with respect to the stem they suffix to. In the eighteenth century texts the forms are all suffixed to the long base. In the nineteenth century texts, the forms are suffixed to both the long base and the long base plus finite inflection. By the twentieth century, only the forms with finite inflection are found. In addition, there is a new form *ki*, apparently a borrowing of the Nepali disjunctive 'or', which has been recently introduced into the language and is suffixed onto the finite inflection, this being the only position of subordinators in Kathmandu Newari.

Concessive. The suffix meaning 'although' is *naM*, which can have as its source one of two homophonous suffixes, either the ergative/instrumental *naM*, or the contrastive *naM*.[17] These two suffixes may well be etymologically related (see Anderson, 1977:356 on the relationship between a topic marker and an ergative case marker in Chinookan). The two co-exist synchronically, the only difference between them being that the topicalizer does not reduce to nasalization of the preceding vowel, as does the ergative suffix. In Genetti (1986) I found no evidence of 'although' coming from a source morpheme (ergative/instrumental/ablative). On the other hand, the topicalization function seems natural considering the contrastive, counterfactual function of concessive clauses. Note that there are sources for 'although' involving emphatic morphemes in other Bodic languages (i.e. Tibetan *na-yang* 'if + EMPH'; compare English *even if*). In the collection of texts, this subordinator was found four times. Two examples were in the eighteenth century texts, and two in the twentieth century texts. All four were suffixed to verbs with finite inflection. All examples of 'although' in the nineteenth century texts took the form *sāM* 'even if', so were grouped with the 'if' cases above.

Purpose. The purpose clauses in all of the texts are marked by the dative *ta*, or by the genitive plus dative *yāta*. Both forms are found in Classical Newari, but only the short version is possible in the Kathmandu dialect. This again illustrates the relative productivity of the postposition suffixation process of forming subordinate clauses in Classical Newari as compared with the modern dialect.[18]

In all three sets of texts the purpose marker is suffixed to the verb inflected with the infinitive. In Newari, the infinitive is syncretic with a future

form (Jørgensen's form A4, Kathmandu Newari Non-Past Conjunct).[19] Since purpose clauses are only formed with verbs inflected by this one future category, we can infer that it is the infinitive function, as opposed to the finite agreement, that attracts the purpose marker.

The distribution of these forms in the texts is as follows. Two examples of *ta* were found in the eighteenth century texts. One example of *yāta* was found in the nineteenth century texts. But nineteen examples of *ta* were found in the twentieth century texts. The increase in distribution can be partially explained by another strategy for marking purpose clauses, which is to leave the verb of the purpose clause in the stem form. The use of this second strategy is very common in the Classical Newari texts, but is apparently now falling out of use.

Non-final. This category consists of markers of non-final clauses in clause chains, where there is no explicit marking of the semantic relationship between clauses. Clause chains normally consist of two to three clauses representing events in temporal sequence. The non-final verb is also non-finite, and takes its tense value from the fully inflected final verb. The following is an example from Kathmandu Newari:

(24) *manu-nã duku syān-āā wan-a*
man-ERG goat kill-NF go-PAST
'The man killed the goat and left.'

In Classical Newari, verbs of non-final clauses are marked by the associative, which suffixes onto one of the past tense forms (Jørgensen's A6):[20]

(25) *śaḍa ga-yā-va van-aM*
horse climb-A6-ASSOC go-PAST
'He mounted his horse and rode off.' (Batis., 23.26)

The Kathmandu Newari non-final marker *āā* developed directly from the Classical form *ā-va*. It is unclear if the synchronic status of the Kathmandu form is as a single unit, or as a sequence of a shorter participial vowel *ā* and a separate non-final marker *ā*.

The semantic relationship between the associative and the non-final marker is again predicted by Diehl's schema. The associative is essentially a Location (in that it puts two participants together in one place, cf. DeLancey, 1984); the non-final marker puts two or more events together at one point in time.

syaM. This form is not categorized by semantic content since it has a number of different uses. Hargreaves discusses the distribution of this form in detail and makes the following observations (1984:9):

a. syaM marks the manner in which the action is performed with respect to an adjacent verb.
b. *syaM* is a negative conditional marker when it occurs on a negated verb in a clause chain.
c. *syaM* marks the transition from direct quote to narrative when it occurs with direct quote complements, and when the speech act is made explicit in the narrative.

This subordinator is clearly related to the ergative *sëM*. Its use in these sentences can at least be partially attributed to the underlying semantic case of the ergative, which is Source. An example of this morpheme marking manner clauses is the following:

(26) *putrika cha-hma bo-syaM van-aM*
statuette one-CL fly-ERG go-PAST
'The one statuette flew away.' (Batis., 33.3)

A more accurate translation might be "The statuette went by flying', where *syaM* marks the manner or means by which the action of the main verb was performed. We can see here the relationship between manner and instrument; 'go by flying' is semantically analogous to 'go by boat', with a clear instrumental NP.

We can extend the relationship between manner and instrument to the more abstract relation of Source by noting that in examples such as (26), the action denoted by the manner clause is a direct cause (Source) of the action of the main verb; the going would not have taken place without flying, walking, or some other means of locomotion. This connection between manner, instrument and Source is more clear with examples with propositional NPs, as above, than examples with concrete nouns; it is not as clear that 'tongs' in a sentence like *He picked up the spaghetti with tongs* is the direct cause of the 'picking up'. However, the fact that the instrumental and ergative relations are commonly marked by the same morpheme in Tibeto-Burman indicates that there must be some semantic connection between them, as there is, for example, between dative and allative.

Constructions of the type illustrated in (26) tend to undergo reanalysis to a main verb followed by a directional or locational auxiliary. In this case,

the use of the ergative may be expanded from marking instruments to marking verb-auxiliary constructions. For example:

(27) *bhikṣā ma pho-syaM coŋ-a*
bhiksu not beg-syaM stay-PAST
'(The bhiksu) remained there ... without begging alms.'
(Batis. 146:10)

In this example the *syaM* clause denotes the manner, but not the cause of the staying. It is unclear to what extent the verb 'stay' is functioning here as a main verb or as an auxiliary.

The other uses of this morpheme in Classical Newari can also be seen as extensions of the basic Source meaning. An example of a negative conditional sentence is:

(28) *ji-va juddh ma yā-syaM amo bastu ma du*
I-ASSOC fight not do-syaM this thing not have
'Unless you fight me, you shall not eat these things.'
(Batis., 37:8)

The subordinate clause in this sentence codes the cause (Source) of the events of the second clause. In this case the cause would be the non-realization of a particular action, hence the negative conditional function.

The third use of *syaM* is to mark direct quotes when the speech act is elaborated upon in the main clause. *syaM* is again marking the manner (in this case the verbatim direct quote) of the speech act, as in the following example:

(29) [direct quote] *dhā-syaM bicar ya-ŋa-va*
say-syaM greet do-A6-ASSOC
'By saying this, he greeted her.' (Batis., 44:19)

This construction may have resulted in the extension of the function of the postposition to that of a complementizer.

An interesting fact about the *syaM* subordinate clauses is that the suffix is attached neither to the long base nor to an inflected finite verb, but to the stem vowel. There has been no change in this form over the last two centuries; in modern Kathmandu the suffix still cliticizes onto the stem vowel. There has, however, been a decrease in the use of this suffix. The first function has been primarily taken over by the short participle (the short vowel *ā*, apparently derived from the non-final marker), while the quotative complementizer is now *dhakāā* (which is also derived from a form with the non-final marker).

3. MORPHOLOGICAL REANALYSIS AND SYNTACTIC CHANGE

From the above description of the subordinators in Classical and Kathmandu Newari, we can make the following observations:

a. The majority of the subordinators were formed with postpositional elements.
b. The patterns of development found in Newari were the same found in the earlier study of Bodic languages.
c. There was a clear trend away from suffixing subordinators to the long base to suffixing them to finite verb morphology.
d. The change of the stem from the long base to the finite inflected verb appeared to happen at different times for different subordinators, implying independent development of the subordinators, as opposed to a single unitary process.

Now that the process of development from postposition to subordinator has been established for Newari, we can turn to the morpho-syntactic mechanisms by which the development occurred. I will argue that the process involved a reanalysis of originally nominal morphology as verbal morphology, via the reanalysis of unmarked nominal clauses as clauses with erstwhile finite verbs.

The strongest piece of evidence for the morphological reanalysis of nominal morphology to verbal, is the virtual identity of the majority of subordinators to postpositions in Classical Newari, and the lack of such identity in the modern Kathmandu dialect. Thus consider the Classical Newari case, represented in Table 6. The only subordinator which is not phonologically identical to a postposition is *sā* 'if', which has a low front vowel as opposed to the low back vowel of the locative. The rest of the subordinators have the same phonological form as the postpositions. Thus it is reasonable to propose that postpositions and subordinators were not distinct in Classical Newari. Compare this to the current Kathmandu Newari system, represented in Table 7. In Table 7 only two subordinators have the same shape as the postpositions. These are the purpose/dative *ta* and the although/emphatic *nã*. It could well be that speakers recognize the similarities between these postpositions and subordinators and consider the two functions to involve the same morpheme.

For the remaining forms, however, there is no longer any phonological connection between the subordinator and the postposition. Therefore an analysis which assumes that native speakers still retain some connection

Table 6 *Classical Newari Postpositions and Subordinators*

	Subordinator	*Postposition*	
Temporal	syaM	syaM	ERG
	va	naM, n, M	INSTR
	naM	s(a)	LOC
Causal	naM	ta	DAT
	M	yā	GEN
	n	va	ASSOC
Conditional	sā	naM	EMPH
	sā-M		
	s		
Concessive	naM		
Purpose	ta		
	yā-ta		
Non-final	va		
by, etc.	syaM		

Table 7 *Kathmandu Newarti Postpositions and Subordinators*

	Subordinator	*Postposition*	
Causal	gulĩ	nã	ERG
Conditional	sā	e	LOC
Concessive	nã	nã	EMPH
Purpose	ta	ta	DAT
Non-final	āā	wa	ASSOC
by, etc.	sẽ		

between the forms is unfounded. Rather, it is more plausible to state that at some point, these forms must have been reanalyzed as verbal morphology, separate from the postpositional forms.

Notice that distributional facts also favor this analysis. These forms (again excepting *ta* and *nã*) only occur suffixed to verb forms and do not occur as nominal morphology. Furthermore, these forms may occur on verbs following other verbal morphology, such as the enclitic particle *ni* 'yet':

(30) *chwĩkā ma-pu-u-ni-sā ma-bu-u*
sting not-burn-STAT-yet-if NEG-cook-STAT
'If it doesn't sting yet, it's not cooked.'

This indicates that the forms are indeed verbal, and not underlyingly nominal.

The syntactic mechanism which led to the reanalysis of nominal morphology as verbal concerns the degree to which nominalized clauses (and relative clauses, which are structurally similar to nominalized clauses) are explicitly marked. In Classical Newari examples of nominalized or relative clauses are found both with and without the nominalizers *gu*, *guli*, *pani*, and *hma*[21], and it is common for clauses with fully inflected finite verbs to function as nominals (the remaining Classical Newari examples are from Jørgensen, 1941):

(31) *thwa* ***bo-va*** *khaŋ-ā-wa*
this fly-A5 see-A6-ASSOC
'On seeing them fly.' (1941:65) (no nominalizer)

(32) ji-n *jāM swāmi-n doh* ***yā-k-guli*** *chuM ma khaŋ-ā*
I-ERG my husband-ERG crime do-A5-NMNL any NEG see-A6
'I don't see any crime that my husband has done.' (1941:66)

(33) ***na-va*** *kwathā*
eat-A5 room
'A room where one eats.' (1941:65) (no nominalizer)

(34) *meba sunā-naM* ***ma-twoŋ-gu*** *laMkha*
other who-ERG NEG-drink(A5)-NMNL water
'Water which others do not drink.' (1941:65)

In Kathmandu Newari on the other hand, every nominalized verb must be marked as such either by *gu*, *mha*, or *pĩ*:

(35) *thana wa-a-mha jimi kijā khaa*
here come-STAT-NMNL my brother is
'The one that came here is my younger brother.'

(36) *kerā na-i-gu ji-u*
banana eat-NPD-NMNL should-STAT
'It's good to eat bananas.'

(37) **thana wa-a jimi kijā khaa.*
**kerā na-i jiu.*

Sentence (35) illustrates a headless relative clause; the nominalizer agrees with the animacy and plurality of the head noun. (36) illustrates a typical construction with a nominalized clause. (37) shows that both constructions are ungrammatical without their respective nominalizers.

This distinction, then, is what gave rise to the process of reanalysis. In Classical Newari case particles suffixed to nominal finite verbs can express typical nominal case relations without creating subordinate clauses. Apparently this is possible whether or not the nominalized verb is explicitly suffixed by the nominalizer, although, at least for the examples given in Jørgensen's grammar, the nominalizer is much more frequent when bonafide postpositions are used:

(38) A5 + GEN (no nominalizer)
strii puruṣ ne-hmaM mo-ka-yā pāp
wife man two-CL live-A5-GEN sin
'The sin of the husband and wife being dead.' (1941:67)

(39) A6 + GEN (no nominalizer)
Citragriiba-yā pās Hiranyāk-an phe-ŋā-yā khaM
C. -GEN snare H.-GEN rend-A6-GEN talk
'The story of how Hirangyak rent Citragiiba's snare.' (1941:68)

(40) A6 + GEN (no nominalizer)
kumāl hŋer-ā-yā arth
boy laugh-A6-GEN meaning
'The meaning of the laughing of the boy.' (1941:68)

(41) A4 + GEN (no nominalizer)
mocā bu-ya-yā bedanā
child bear-A4-GEN pain
'The pain of childbearing.' (1941:64)

(42) A6 + NOM + GEN
syān-ā-hma-yā rā
kill-A5-NMNL-GEN flesh
'The flesh of one who's been killed.' (1941:67)

(43) A5 + NOM + ERG
saty ma dhā-o-pani-sen jakaM thathiMŋ ḅrtti yāyuo
truth not say-A5-NMNL-ERG only this.way speak do
'Only liars speak in this way.' (1941:65)

The first stage in the development of postpositions to subordinators is thus the suffixation of nominal postpositions to fully inflected finite nominal verbs, which lacked overt nominalizing suffixes. Originally the postpositions in these circumstances must have coded the same semantic meaning as they did when suffixed to nominal constituents. Thus by analogy with sentences like (45) sentences like (46) were formed:

(44) *thva deśa-s utpatā-naM khvayāva conaM*
this country-LOC disaster-INSTR cry stay
'(Somebody) was weeping on account of a public disaster.'
(Batis., 36:14)

(45) *āva chan daya-n jin rājy li kāya dhuna*
now you have-INSTR I kingdom back take finish
'Because you are here I have now won back my kingdom.'
(Batis., 151:4)

In these two sentences the semantic role of the instrumental marked NPs is equivalent; both are the causes of the action of the final verb. In (44) the postposition is clearly nominal morphologically, as it is suffixed to a noun. In (45), the morphological status of the instrumental morpheme is unclear. One can either analyze the construction as a nominal clause marked with a case suffix, or as a finite verb followed by a causal subordinator. Thus in the second stage the status of the morpheme becomes ambiguous. This ambiguity could, in turn, fuel the semantic extension of the postpositional meaning into more abstract domains.

The third step in the process was the reanalysis of ambiguous cases into verb plus verbal morphology. This was additionally motivated by the regularization of explicit marking of nominalization. Once a binary division between nominalized and non-nominalized clauses existed (depending on the presence or absence of one of the nominalizers *gu*, *guli*, *pani*, or *hma*), ambiguous cases had to be resolved in one way or the other. For the majority of the subordinators the conflict was resolved by reanalyzing the postpositions as subordinators. The instrumental, which marked causal clauses, was still treated as a postposition, and underwent an additional round of cliticization to the nominalizer *guli*.

This reanalysis of the new subordinators also precipitated a change from the use of long base forms to mark subordinate verbs to the use of finite inflected forms. It appears that as more and more postpositions were suffixed to finite forms via the process discussed above, and as these forms began to

function more like subordinators, the long base forms fell out of use. This can be seen most clearly in the development of *sā* 'if', which apparently already had status as a full subordinator suffixed onto the long base early on in Classical Newari. By the late nineteenth century, this morpheme began to be used on finite verb forms as well, and finally as this suffixation became prevalent the long base forms were lost entirely.

The fact that each morpheme has a slightly different history is not surprising in light of the fact that each had different functions, a different set of similar morphemes and constructions with which it competed, and different phonological factors which contributed to distinguishing it from the postposition. For example, the fact that the dative is always suffixed to infinitive verb forms to code the purpose function is a reflex of the affinity of purpose and future. As another example, take the case of the associative *va*. This postposition is found primarily suffixed onto form A6, a past tense which was used mainly with first and second person subjects, but also sometimes with third person subjects (Jørgensen, 1941:68). Jørgensen reports that in the earlier manuscripts, this form was also suffixed to A4, the non-past counterpart of A6 (Jørgensen, 1941:64). This may suggest that the function of the associative/non-final marker was being extended to other verb forms, in accordance with the general trend toward suffixation of finite forms. However, in this case the non-past form never became entrenched in the grammar, and consequently the non-final clauses marked by A6 plus the associative also became non-finite. At some point in this development, the suffix *va* assimilated to the front vowel of the stem-final *ā*, obscuring the associative origin of the morpheme.

4. CONCLUSION

In the preceding pages I have presented a historical study about a functional shift in which a group of morphemes which originally functioned as postpositions developed into verbal subordinators. I will conclude with a few comments on the implications of this study for the wider concepts discussed in these volumes, particularly the notion and definition of 'grammaticalization' itself.

As discussed in the editors' introduction to these volumes, there is debate among linguists as to exactly what types of historical change should be considered 'grammaticalization' and what types should not. While all seem to agree that the case of a lexical item taking on a grammatical function and becoming a bound morpheme of reduced phonetic substance is the prototypi-

cal case of grammaticalization, there is disagreement as to what criteria to use in order to include or exclude other phenomena. Lehmann (Volume II) proposes that the grammaticalization of a unit is "directed towards subjecting that unit more to the rules of grammar". A very different view is espoused by Hopper (1987, Volume I), who asserts that since all parts of grammar are at all times undergoing change, "all grammatical phenomena are to be thought of as 'under grammaticization".'

The Newari case is interesting because the morphemes which underwent the grammatical change did not start as lexical items but as bound postpositions with a clear grammatical status. Therefore this does not fall within the traditional definition of grammaticalization. It also does not clearly fit into the realm of grammaticalization as defined by Lehmann. It is not obvious by any means that postpositions are less subject to the rules of grammar than are subordinators, or that these morphemes became 'more grammatical' as a result of the category shift. However, a number of features characteristic of grammaticalization are found in the present study.

To start with, the development from postposition to subordinator appears to be unidirectional. I found no evidence of postpositions developing from subordinators in a typological study of this phenomenon in Bodic languages (Genetti, 1986), and while an increasing number of studies (Craig, Volume II, Lichtenberk, Volume I; Reesink, 1988; Thurgood, 1986) report on postpositional sources for subordinators, none of them mention any cases where the reverse is true. In addition, in the literature specifically concerned with the development of postpositions (Givón 1975; 1979; DeLancey, 1984; Carlson, Volume II; Lichtenberk, Volume I) the sources are typically verbal or nominal.

A second feature common to this study and clearer examples of grammaticalization is that of layering (Hopper, Volume I). Layering is the presence at one synchronic period of both the newer grammaticalized and the older, established methods of conveying a particular function. Layering can be seen for example in modern Newari, where purpose clauses can be coded either by the infinitive of a verb marked with a dative (*swa-e-ta wana* 'went to look') or by a simple stem form (*swa wana* 'went to look'). The latter strategy is more common in the older texts, while in the modern texts the form with the postposition occurs more frequently.

A third similarity between the data in this study and typical cases of grammaticalization concerns semantic change. The semantic changes that the morphemes underwent as they developed into subordinators are typical of changes found in grammaticalization. Traugott (1987:34–35) states three

semantic-pragmatic tendencies which characterize paths of semantic change; (cf. also Traugott and König, Volume I):

Tendency 1: Meanings based in the external described situation > meanings based in the internal (evaluative/perceptual/cognitive) situation.
Tendency 2: Meanings based in the external or internal described situation > meanings based in the textual situation.
Tendency 3: Meanings tend to become increasingly based in the speaker's subjective belief-state/attitude toward the situation.

All three of these tendencies are exemplified in the Newari case presented above. The development from spatial postpositions to temporal subordinators exemplifies Tendency 1. The use of postpositions to indicate relationships between textual units exemplifies the second tendency. And the development of markers of conditionals and concessives from spatial postpositions exemplifies Tendency 3.

In a metaphorical analysis of the data we also find familiar paths of semantic change. Heine et al. (Volume I) propose the following linear arrangement of conceptual categories:

PERSON > OBJECT > SPACE > TIME > PROCESS > QUALITY

Any of these categories can be conceptualized by any category to its right, resulting in what Heine et al. have termed 'categorial metaphors'. In Newari we find the metaphors TIME IS SPACE and PROCESS/EVENT IS OBJECT, both predicted by the framework proposed by Heine et al.. We also see in the Newari data the metaphor REASON IS INSTRUMENT, this being one of the metaphors proposed by Reesink (1988) for semantic change accompanying grammaticalization in Usan.

Thus we have seen that the Newari data presented above is similar to clearer cases of grammaticalization in a number of respects, especially in that the same types of semantic change are found in both cases. It is important that we consider studies of non-prototypical examples of grammatical development as we explore the vast field of grammatical change, and define our terminology.

ABBREVIATIONS

ANT	anterior
CL	classifier
NF	non-final

NMNL nominaliser
PASTHAB past habitual
PRTCL particle
Abbreviations such as A5, A6, etc. are labels for new forms in Classical Newari, as given in Jørgensen, 1941:55–56.

NOTES

1. There are a number of problems with the current binary division of clause combining into 'subordinate' clauses and 'coordinate' clauses (see e.g. Haiman and Thompson (1984) for discussion). For this paper I will use the term 'subordinate clause' to refer to hypotactic adverbial clauses and to non-final clauses in clause chains. 'Subordinator' refers to the affix (in these languages always suffixes) which explicitly marks the relationship of the subordinate clause to the matrix clause.

2. Funds for this study were provided by grant BNS-8313502-01 from the National Science Foundation, Scott DeLancey, PI. I am grateful to Scott DeLancey, Tom Givón, David Hargreaves, Bernd Heine and Elizabeth Traugott for valuable comments and suggestions.

3. The twenty-six languages of the study and the sources for the data are as follows:

Classical Tibetan:	Jaschke, 1881; 1954; Bacot, 1948
Balti:	Read, 1934
Purki:	Rangan, 1979
Ladakhi:	Francke, 1979; Koshal, 1979; 1982
Lahul:	de Roerich, 1933
Lhomi:	Vesalainen and Vesalainen, 1980
Lhoke:	Konow, 1909
Sherpa:	Schottelndreyer, 1973, 1980
Kagate:	Konow, 1909
Garhwal:	Konow, 1909
Jirel:	Maibaum and Strahm, 1973
Nyamkat:	Konow, 1909
Lhasa:	de Roerich and Phuntshok, 1978; Goldstein, 1974; DeLancey, p.c.
Danjongka:	Konow, 1909
Kanauri:	Joshi, 1909; Nigam and Neethivanan, 1971
Bunan:	Konow, 1909
Vayu:	Michailovsky, 1981
Chepang:	Caughley, 1982
Limbu:	Weidert and Subba, 1985
Thulung:	Allen, 1975
Khambu (Rai):	Konow, 1909
Khaling:	Toba, 1979
Sunwari:	Schulze and Bieri, 1973
Gurung:	Glover, 1970; 1974; 1979
Tamang:	Taylor, 1970
Thakali:	Hari and Maibaum, 1970

4. I have used the term "grammatical extension" as opposed to "grammaticalization" since this case differs from the prototypical pattern of grammaticalization in that the morphemes undergoing change already have a fully grammatical function, as opposed to a lexical origin. For more discussion of these points, see the concluding remarks (4.0).

5. This list does not reflect the problem of syncretic postpositions, the fact that ablative and ergative/instrumental, dative and locative, locative and allative, are often marked by the same morpheme. For discussion of this and other problems associated with the typological study, see Genetti (1986).

6. By the term "non-final" I am referring to markers of non-final clauses in clause chains. These clauses differ from the other subordinate clauses in that they are not explicitly marked as adverbial, i.e. of coding condition, concession, temporal sequencing, or causation. However, these clauses are included in this study since they are functionally and structurally related to subordinate clauses, and the non-final markers are commonly derived from postpositions.

7. The inherent ordering of the four spheres proposed by Diehl is SOCIAL-SPATIAL-TEMPORAL-LOGICAL. Evidence for this ordering and its relationship to the concrete/abstract continuum is given in Diehl (1975:101–104).

8. Jørgensen's [ë] represents "a variety of *e* (*ie*?) which cannot be defined more closely" (1941:10). See Hargreaves (1984) for a discussion of this and related forms.

9. In this paper I am using slightly different orthographical conventions than in my previous work. For the low back vowel I will now use *a* instead of *o*. For the low front vowel I will now use *ā* instead of *a*. I apologize for any confusion this may cause. I hope that it will be of benefit in the long run, since this conforms to standard practice in Nepal.

   For the Classical Newari data I will represent nasalization of a vowel by a following *M*; again this follows tradition.

10. The original form of this morpheme and the locative are unclear. Jørgensen lists both as a single consonant, without a following vowel. However, the distinction between a simple consonant and the consonant plus the vowel /a/ would not necessarily be marked in the Devanagari alphabet. Since in the modern dialects the ergative/instrumental morpheme is *nã*, it seems reasonable, to assume that the vowel was present in Classical Newari as well. However, evidence points to a locative form *s* without the vowel, since this morpheme has reduced to /e/ in the modern dialect. This is the result of a regular sound change by which word-final /s/ was reduced to /e/, e.g. Classical /nhas/ → Kathmandu /nhae/ 'nose'.

11. The Classical Newari genitive in *-s* is used on plural and animate nouns. It has survived only marginally in the modern dialects, being found in the declension of Dolakhali plural pronouns and in irregular inflected forms in the Kathmandu dialect.

12. The Classical Newari data suggest that at this stage *nã* was an instrumental marker that later supplanted the ergative and ablative morphemes. (For another case of development in this direction see Garrett (Forthcoming). However, due to the Tibeto-Burman tendency to continually recreate case postpositions from other postpositional elements (DeLancey, 1984) it is possible that the original meaning of this postposition was other than instrumental.

13. These are the Infinitive, Participle, Non-final form, Infinitive of Purpose and the form inflected with *se* (see below).

14. The tables listing the distribution of forms will have the following conventions: ***Batis.***

refers to the eighteenth century *Batïsaputrikākathā* (Jorgensen, 1939), *Vicit.* to the late nineteenth century *Vicitrakarnikāvadāoddhrta* (Jorgensen, 1931), and *Kath.* to my own collection of twentieth century texts collected from two informants. FI abbreviates 'finite inflection'.

Examples taken from the historical manuscripts will be cited as follows: (Batis. 30:16) refers to the 16th line on page 30 of the *Batïsaputrikākathā.*

15. Jørgensen (1941:53) states that nominalizers are found between finite verb stems and postpositions frequently in the younger manuscripts. I found no examples outside of *gulin* in my data.

16. One of these forms actually has the ending *LB-nā-s*, with a dental, as opposed to a velar /n/. I am currently treating this as an error in some stage of the writing, since there are no other *LB-nās* forms listed in Jørgensen or found in any of the other available texts. If my assumption is incorrect, this form requires explanation.

17. The precise discourse function of this particle has not yet been determined.

18. The extension of dative to purpose may be related to the use of *to* in purposive expressions such as *I'm watering the plants to help Aunt Elda.*

19. The syncretism between the infinitive and the Kathmandu non-past conjunct and that between the participle and the past conjunct are an interesting topic for students of Newari. Austin Hale pointed out to me (p.c.) that in almost all of the cases of the infinitive (the exception being *dhunk-* 'finish') the clauses marked with the infinitive are future with respect to the main verb. Thus the syncretism between the future and the infinitive. This still does not address the problem of the non-past conjunct being syncretic as opposed to the disjunct (for a description of these categories see Hale (1980)). My current view (though not necessarily my final analysis) is that the infinitive reflects an old future marker, while the participle is a phonological reduction of the non-final marker (historically Jørgensen's A6 plus the associative). The syncretisms between these and the conjunct forms are thus coincidental, and do not imply any functional connection between non-finite and conjunct.

20. Hargreaves (1984) considers this suffix, *syaM*, and other postpositions functioning as subordinators to be nominalizers. While I disagree with this part of his analysis (for reasons elaborated below), his arguments still stand.

21. The conditions determining the use of the nominalizers are not currently known.

REFERENCES

Allen, N.J. 1975. *Sketch of Thulung Grammar.* Ithaca, N.Y.: Cornell University China-Japan Program.

Anderson, John. 1971. *The Grammar of Case: Towards a Localistic Theory.* Cambridge: Cambridge University Press.

Anderson, Stephen R. 1977. "On mechanisms by which languages become ergative." In *Mechanisms of Syntactic Change*, Charles Li (ed.), 317–364. Austin: University of Texas Press.

Austin, Peter. 1981. "Switch-reference in Australia." *Language* 57(2):309–334.

Austin, Peter (ed.). 1988. *Complex Sentences*. Amsterdam/Philadelphia: John Benjamins. [*Typological Studies in Language* 15.]
Bacot, J. 1948. *Grammaire du Tibétain Littéraire*. Paris: Librairie d'Amérique et d'Orient.
Carlson, R. Volume II. "From verb to postposition in Senufo."
Caughley, Ross. 1982. *The Syntax and Morphology of the Verb in Chepang*. Canberra Australian National University [*Pacific Linguistics* B-84].
Craig, C. Volume II. "From verbs to postpositions and beyond in Rama."
DeLancey, Scott. 1984. "Etymological notes on Tibeto-Burman case particles." *Linguistics of the Tibeto-Burman Area* 8(1):59–77.
Diehl, Lon. 1975. "Space Case: Some principles and their implications concerning linear order in natural language." In *Work Papers of Summer Institute of Linguistics, University of North Dakota Session* 19:93–150. Huntington Beach, CA: Summer Institute of Linguistics.
Francke, A.H. 1979. *Ladakhi and Tibetan Grammar*. Delhi: Seema Publications. (Originally published 1900. "Sketch of Ladakhi Grammar." *Journal of the Asiatic Society of Bengal*, pt.1, extra no. 2, 63 pps.)
Garrett, Andrew. Forthcoming. "The origins of NP split ergativity." In *Proceedings of the Conference on the Theory and Practice of Historical Linguistics*, Barbara Need and Eric Schiller (eds). Chicago: University of Chicago Press [*University of Chicago Special Publications in Linguistics* 2].
Genetti, Carol. 1986. "The development of subordinators from postpositions in Bodic languages." In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.
Genetti, Carol. 1988. "A contrastive study of the Dolakhali and Kathmandu Newari dialects." In *Cahiers de Linguistique Asie Orientale* 27(2): 161–91.
Glover, W. 1970. "Gurung texts." In Hale and Pike 1970:1–31.
Glover, W. 1974. *Sememic and Grammatical Structures in Gurung*. Norman: Summer Institute of Linguistics.
Glover, W. and D.B. Gurung. 1979. *Conversational Gurung*. Canberra: Australian National University [*Pacific Linguistics* D-13].
Goldstein, Melvyn C. 1973. *Modern Literary Tibetan*. Urbana: University of Illinois [*Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics* 5].
Gruber, Jeffrey. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam: North Holland.
Haiman, John and Sandra Thompson. 1984. "'Subordination' in universal grammar." *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 510–523. Berkeley: Berkeley Linguistics Society.
Hale, Austin. 1973. *Clause, Sentence and Discourse Patterns in Selected Languages of Nepal*. Norman: Summer Institute of Linguistics.
Hale, Austin. 1980. "Person markers: Finite conjunct and disjunct verb forms in Newari." In *Papers in South-East Asian Linguistics* 7, Ronald L. Trail et al. (eds), 95–106. Canberra: Australian National University.
Hale, Austin and K. Pike, (eds). 1970. *Tone Systems of Tibeto-Burman Languages* 3, Texts I. Urbana: University of Illinois.
Hargreaves, David. 1984. "Case marking and nominalization in Classical Newari." Presented at the Seventeenth Annual International Conference on Sino-Tibetan Languages and Linguistics. Eugene, OR.
Hari, A. and A. Maibaum. 1970. "Thakali texts." In Hale and Pike 1970:165–306.

Heine, B., U. Claudi and F. Hünnemeyer. Volume I. "From cognition to grammar: Evidence from African languages."
Hodgson, B. 1857. "Comparative vocabularies of the broken tribes of Nepal." *Journal of the Asiatic Society of Bengal* 26:317–522.
Hodgson, B. 1858. "Comparative vocabularies of the broken tribes of Nepal." *Journal of the Asiatic Society of Bengal* 27:393–442.
Hopper, P. 1987. "Emergent grammar." In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.
Hopper, P. Volume I. "On some heuristic principles of grammaticization."
Jaschke, H.A. 1881. *Tibetan-English Dictionary*. London: Routledge and Kegan Paul Ltd.
Jaschke, H.A. 1954. *Tibetan Grammar*. New York: Frederick Ungar.
Jørgensen, Hans. 1931. *Vicitrakarṇikavadānoddhṛta*. London: The Royal Asiatic Society.
Jørgensen, Hans. 1936. *A Dictionary of the Classical Newari*. Copenhagen: Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser XXIII:1.
Jørgensen, Hans. 1939. *Batisputrikākathā: A Tale of Thirty-two Statuettes*. Copenhagen: Munksgaard.
Jørgensen, Hans. 1941. *A Grammar of the Classical Newari*. Copenhagen: Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser XXVII:1.
Joshi, P.T.R. 1909. "A Grammar and Dictionary of Kanawari." *Journal of the Asiatic Society of Bengal* 5.
Konow, Sten. 1909. *Tibeto-Burman Languages*. Calcutta: Motilal Banarsidass [*Linguistic Survey of India* 3].
Koshal, S. 1979. *Ladakhi Grammar*. Delhi: Moltilal Banarsidass.
Koshal, S. 1982. *Conversational Ladakhi*. Delhi: Motilal Banarsidass.
Lehmann, C. Volume II. "Grammaticalization and related changes in contemporary German."
Lichtenberk, F. Volume I. "On the gradualness of grammaticalization."
Maibaum, A. and E. Strahm. 1973. *Jirel Texts*. In Hale 1973:177–300.
Malla, Kamal P. 1985. *The Newari Language: A working outline*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo Gaikokugo Daigaku [*Monumenta Serindica* 14].
Michailovsky, B. 1981. *La Langue Hayu*. Paris: Université Paris III.
Nigam, R.C. and J. Neethivanan. 1971. *Survey of Kanauri in Himachal Pradesh*. Calcutta [*Language Monograph*, 3].
Rangan, K. 1979. *Purki Grammar*. Mysore: Central Institute of Indian Languages.
Read, A.F.C. 1934. *Balti Grammar*. London: Royal Asiatic Society.
de Roerich, George. 1933. "The Tibetan dialect of Lahul." *Journal of the Urusvati Himalayan Research Institute* 3:83–189.
de Roerich, George and T.T.L. Phutshok. 1978. *Textbook of Colloquial Tibetan*. New Delhi: Manjusri.
Reesink, G. 1988. "Discourse in space: The grammaticalization of metaphor in Papuan clause marking." Paper presented at the Symposium on Grammaticalization, Eugene, Oregon May 1988.
Schottelndreyer, B. 1980. "Person markers in Sherpa." *Papers in South-East Asian Linguistics*, R. Trail et al. (eds), 125–136. Canberra: Australian National University [*Pacific Linguistics* A-53].

Schottelndreyer, B. and H. Schottelndreyer. 1973. "Sherpa texts." In Hale 1973:53–176.
Schulze, Marlene and D. Bieri. 1973. "Chaining and spotlighting: Two types of paragraph boundaries in Sunwar." In Hale 1973:389–399.
Taylor, D. 1970. "Tamang texts." In Hale and Pike 1970:132–164.
Thurgood, Graham. 1986. "Lolo-Burmese subordinators from case postpositions: Several partial etymologies." In *Proceedings of the Second Annual Meeting of the Pacific Linguistics Conference*, Scott DeLancey and Russell S. Tomlin (eds), 449–454. Eugene: Dept. of Linguistics, University of Oregon.
Toba, S. 1979. *Khaling*. Tokyo: Institute for the Study of the Languages and Cultures of Asia and Africa, Tokyo Gaikokugo Daigaku [*Asian and African Grammatical Manual* 13d].
Traugott, Elizabeth. 1989. "On the rise of epistemic meaning in English: An example of subjectification in semantic change." *Language* 65:31–55.
Traugott, E. and E. König. Volume I. "The semantics-pragmatics of grammaticalization revisited."
Vesalainen, O. and M. Vesalainen. 1980. *Clause Patterns in Lhomi*. Canberra: Australian National University [*Pacific Linguistics* B-53].
Vajracarya, Dhanavajra and Kamal P. Malla. 1985. *The Gopālarājavaṃśāvali*. Wiesbaden: Franz Steiner Verlag [*Nepal Research Centre Publications* 9].
Weidert, A. and B. Subba. *Concise Limbu Grammar and Dictionary*. Amsterdam: Lobster Publications.

The Evolution of Dependent Clause Morpho-syntax in Biblical Hebrew

T. Givón
Linguistics Department
University of Oregon

1. INTRODUCTION*

This study tracks the development of the morpho-syntax of three subordinate-clause systems in Biblical Hebrew — relative clauses and verb complements and, to a somewhat lesser extent, adverbial clauses. One development — the transfer of the REL-clause subordinating morpheme to the V-complement paradigm, is the core of this investigation. This process is of particular typological-comparative interest, because the sharing of the same subordinating morphemes by REL-clauses and V-complements is widely attested in many languages, as in e.g. German, English, Spanish, French, Akkadian, Biblical Aramaic, Austronesian and no doubt many others.[1] The diachronic processes we track down here may thus have wider applicability.

Early Biblical Hebrew (henceforth EBH) also bears enough internal evidence for reconstructing several earlier diachronic processes involving the syntax of REL-clauses and ADV-clauses. First among those is the rise of the EBH REL-clause pattern, with the subordinator morpheme *'asher*, and the gradual displacement, by *'asher*, of an earlier pattern of *nominalized* REL-clauses (roughly of the Turkish/Northern Uto-Aztecan/Tibetan type).[2] This change is again of considerable typological interest, representing an instance of what may be called 'syntactic liberation', whereby erstwhile nominalized clauses gradually acquire finite-clause syntactic properties. Such changes have often resulted in a total re-evaluation of the syntactic markedness situation in a language, so that what used to be subordinate clause syntax became characteristic of the least marked, finite main clause.[3]

The history of REL-clause syntax in Biblical Hebrew is, further, inti-

mately intertwined with that of ADV-clauses. In the earliest stages of the BH dialect continuum, ADV-clauses preserve largely the archaic nominalized pattern. Throughout the BH continuum this pattern is gradually displaced by the finite REL-clause pattern. This displacement turns out to be intertwined with the transfer of REL-clause morphology into the V-complement paradigm, the main topic of this study. The three main changes thus form one complex, interactive diachronic drift, whose intricate course cannot be reduced to a simple one-cause — or even a simple one causal-chain — model.

In addition to helping us track down in considerable detail the nuts-and-bolts dynamics of the rise of dependent-clause morpho-syntax, the Biblical Hebrew data also offer us insight into the general process of analogical language change. Almost all creative-elaborative diachronic change in language, be it phonological, morpho-syntactic, semantic or discourse-pragmatic, is in principle *analogical*. That is, it involves the language user's recognition — conscious or subliminal — of *similarities* between two structural or functional contexts. Analogical creative change in morpho-syntax extends the applicability of a construction from its original (or 'older') functional context to a new context, one that is judged to be in some way 'similar'.

Even though analogy is the living core of diachronic linguistic change, its invocation in diachronic syntax has often been problematic, and it has on occasion been abused. In particular, proponents of so-called 'abstract syntax' have on occasion ascribed considerable feats of abstract analogical leaps to the human linguistic organism, in the process depleting the force of analogy-as-explanation to the point of utter vacuity.[4] The detailed tracking down of the complex series of gradual changes undertaken in this study reveals analogy to be alive and well. But it also suggests that analogy in diachronic syntax most likely operates across relatively *narrow* inferential gaps, where either the structural or functional similarity between two constructions or their functional contexts is relatively *concrete* and *transparent*. This is not to suggest that the human mind is incapable of far-out metaphoric leaps, complex multi-step inferences, or poetic flights of fancy, or — more relevant in this case — recognizing highly abstract similarities. Diachronic change, however, is the cumulative product of the *averaged-out* behavior of an entire *population* in a relatively concrete speech-processing domain. During such a process, the creative or abstract inferential fancy of the few exceptional minds tends to be *cancelled out*, while more prototypical similarity judgements tend to be transmitted.[5]

2. THE PRE-HISTORY OF BIBLICAL HEBREW RELATIVE CLAUSES

2.1. The use of the noun 'place' as REL-subordinator: A general typological perspective

In the typological-comparative study of subordinating morphology shared between REL-clauses and V-complements, one may distinguish between two main types. The first involves the shared use of *interrogative pronouns*, as in the Spanish or French *que*. It is not necessary to assume that this type of sharing arises through transmission from REL-clause to V-complement, since there are perfectly plausible analogical pathways leading directly from either WH-questions or REL-clauses to V-complements.[6]

The second type of subordinator-sharing involves the shared use of *demonstratives*, as in the English *that*, the German *das*, the Aramaic *di*, the Archaic Biblical Hebrew *zu* or the Akkadian *sha*. This feature is indeed most commonly traceable to a diachronic transfer of REl-clause morphology to V-complements.

The Biblical Hebrew subordinator *'asher* — contracted later into *she* — fits in precise detail neither category. It is etymologically most likely to have arisen from the noun *'athar* 'place'. This suggests that it had entered the REL-clause paradigm, at some phase of pre-Biblical Hebrew, as a *locative* REL-clause marker, and only later expanded its scope to mark REL-clauses of other case-role and adverbial categories.

The use of a *classifier generic noun* to mark the relativization of its semantically corresponding case-role is widely attested cross-linguistically. Thus, for example, in EBH itself the erstwhile noun **"d* 'time', which survives in many temporal and even locative uses in EBH,[7] is systematically used as the temporal subordinator 'until' in time-ADV-clauses, either in the nominalized or the finite pattern. Thus consider:[8]

(1) a. ***"ad*** *ha-yom ha-ze*
till the-day the-this
'till this day' (Kings II, 8.22)

b. ***"ad*** *shuv-xa* *'el-ha-'adama*
till return/INF-your to-the-soil
'till you return to the ground' (Genesis, 3.19)

c. *"**ad** 'asher* hishlix-am mi-pan-av
till REL throw/PERF/3MS-them from-face-his
'until he cast them away from his presence' (Kings II, 17.20)

The plausible earlier derivation of *"d* from a locative source is suggested from extant locative uses in EBH, as in:

(2) *va-yelxu "**ad**-"ir bet ha-Ba"al*
and-go/IMPF/3MP **until**-city/of house/of the-Baal
'and they went all the way **to** the city of the Baal's house'
(Kings II, 10.25)

The use of generic classifier nouns as REL-clause markers is intimately associated, at least at its diachronic point of origin, with *nominalized* clauses and genitival noun-compounding. In order to illuminate the naturalness of such a process, let us consider briefly the Tibetan REL-clause situation cited in DeLancey (1988). In Lhasa Tibetan, four REL-clause suffixal operators are used, with varying degrees of case-role specificity. The nominal lexical source of two of them — *mkhan* for the actor/subject case, from 'expert', and *sa* for the locative case, from 'earth', 'ground' — is relatively transparent, since they are found in recognizable noun compounds, as in:

(3) a. *šin-**mkhan*** 'carpenter'
wood-expert
b. *yod-**sa*** 'place of residence'
live-ground

In addition, the 'default' suffix *-yag* is used in Lhasa Tibetan for relativizing the patient and instrument cases (in the perfective aspect); and *-pa* is used for non-actors (in the perfective aspect). The original lexical meaning of *yag* is for the moment irretrievable. The suffix *pa* is widely attested as a *locative* suffix, as in *khang=pa* 'house'. The broad, less specific, highly 'bleached' grammatical meaning of both *yag* and *pa* suggests a more protracted history. To illustrate the use of all four suffixes in relativization, consider the following examples (Mazaudon, 1978):

(4) a. Actor REL-clause:
*stag gsod-**mkhan** mi*
tiger kill-AGENT man
'the man who killed the tiger'

b. Locative REL-clause:
*kho sdod-**sa**-'i khang=pa*
he/ABS live-PLACE-GEN house=PLACE
'the house where he lives'

c. Instrument REL-clause:
*kho-s stag gsod-**yag**-gi me=mda*
he-ERG tiger kill-INSTR-GEN gun
'the gun with which he killed the tiger'

d. Patient REL-clause:
*kho-s bsad-**pa**-'i stag*
he-ERG kill-PAST-GEN tiger
'the tiger that he killed'

The noun-compound, nominalized origin of Tibetan REL-clauses may be schematically illustrated by the suggestive paraphrases of (4a-d) above, respectively:

(5) a. the man, tiger-killing-*person*
b. the house, his living-*place*
c. the gun, his tiger-killing-*instrument*
d. the tiger, his killing-*object*

As a relativization strategy, the system in (5) is reasonably coherent: The semantic role of the 'missing' coreferent argument is coded inside the REL-clause by the generic classifier, which stands in a nominal-compound relation to the nominalized verb.

Other languages display either the footprint of this strategy or its early stages. Thus, for example, in Hewa (Papua-New Guinea, Sepik region) both place and time REL-clauses follow this pattern, as in (Volrath, 1985):

(6) a. Time REL-clause (adverbial clause):
yaú** an-a nö Atini-su m-i-m-e **yaú
sun I-SUBJ place Atini-to INDIC-go-REM-REAL **sun**
'**the day** (when) I went to Atini'

b. Locative REL-clause:
nö** an-a m-i-m-e **nö
place I-SUBJ INDIC-go-REM-REAL **place**
'**the place** (where) I went to'

The use of generic classifier head nouns, such as 'place', 'time', 'person' or 'thing', to classify the semantic case-role of arguments in complex clauses is not limited to REL-clauses of course. Thus, for example, in the history of negation in many languages, the NEG-intensifier is such an object noun. French is of course a convenient example, where nouns such as *personne* 'person', *rien* 'thing' (Lat. *rem*), or *pas* 'step' were first used as negative intensifiers for human, inanimate and locative objects, respectively. Of this cluster, the locative *pas* was eventually generalized as the unrestricted negative suffix.

The spread of the locative REL-clause marker into the general REL-clause subordinator paradigm can be seen in Krio, an English-based Creole. Here the current REL-clause marker *we* is transparently derived from the locative WH-particle 'where'. Thus consider:

(7) a. Locative REL-clause:
di haus ***we*** *di mami bin-liv na-in*
the house REL the woman ANT-live LOC-it
'the house where the woman lived'

b. Subject REL-clause:
di mami ***we*** *de-kam ia*
the woman REL PROG-come here
'the woman who is coming here'

c. Object REL-clause:
di tik ***we*** *di pikin bin-fes kat*
the stick REL the child ANT-take cut
'the stick that the child cut'

The very same development has also occured in Modern Greek (*pou* 'where?'), colloquial Southern German (*wo* 'where?') and Bambara (*min* 'where?'), with the current general REL-clause subordinator being etymologically the locative WH-element 'where?'.

From this typological perspective, the relativization pattern of Early Biblical Hebrew, and in particular the frequency distribution of the two main competing variants, can be readily interpreted as the *tail-end* of a protracted, profound typological change, whereby dependent clauses that were grammatically coded as nominalizations eventually adopted finite-clause syntax.

2.2. Relativization patterns in Early Biblical Hebrew

Even at the earliest attested stages of EBH, relativization with the subordinator *'asher'* is already the dominant pattern for most of the nominal case-roles. Nevertheless, a number of competing patterns are still attested. Most visible of those are:

(a) The participial-adjectival pattern in subject REL-clauses in the habitual/durative aspect.
(b) The receding use of the old nominal-infinitival pattern in the relativization of most external, adverbial case-roles.

In addition, other variant patterns are also found in EBH that are clearly archaic, and as such relevant to the reconstruction of gradual spread of the finite, *'asher*-marked pattern. In this section we will describe the competing patterns, reconstruct their likely prehistory, and quantify their text distribution. The latter will later serve as the base-line from which to track down the course of other changes, as well as the completion of some changes that had just begun by EBH times.

2.2.1. The finite 'asher-marked pattern

Let us consider first the most common pattern of subject and object relativization in the two main verbal aspects — perfect and the so-called 'imperfect'—[9] in EBH.

(9) a. Subject, perfect:
*ha-'adamah **'asher** patstah **'et-pihah***
the-ground REL open/PERF/3FS ACC-mouth/its
'the ground that opened its mouth' (Genesis, 4.11)

b. Subject, imperfect:
*ki 'im **'asher** yetse' mi-me''a-xa*
but if REL exit/IMPF/3MS from-gut-your
'except for (the one) that will come out of your guts' (Genesis, 15.4)

c. Accusative, perfect:
*mi-kol mlaxt-o **'asher** ''asah*
from-all work-his REL did/PERF/3MS
'from all the work that he had done' (Genesis, 2.2)

d. Accusative, imperfect:
ha-'arets ***'asher*** *'ar'eh-xa*
the-country REL show/IMPF/1MS-you
'the country that I will show you' (Genesis, 12.1)

e. Locative, imperfect:
ha-maqom ***'asher*** *tidrox kaf ragle-xa b-o*
the-place REL step/IMPF/3FS sole/of foot-you LOC-it
'the place where your foot will step' (Joshua, 1.3)

f. Abstract, perfect:
'et-ha-"ir ***'asher*** *dibarta*
ACC-the-town REL talk/PERF/2MS
'the town about which you spoke' (Genesis, 12.1)

In addition to these 'internal' clausal arguments, the *'asher*-marked pattern can also be found in EBH in the relativization of various oblique case-roles (some of which also display the infinitival-nominal variant pattern, see further below). Thus consider:

(10) a. Manner, perfect:
k-xol ***'asher*** *shama"nu le-Moshe*
like-all REL hear/PERF/1PL to-Moses
'just as we listened to Moses' (Joshua, 1.17)

b. Time, perfect:
*ka-****'asher*** *hikriv la-vo' mitsraym-a*
as-REL near/PERF/3MS to-come Egypt-DIR
'when he approached Egypt' (Genesis, 12.11)

c. Purpose, imperfect:
le-ma"an ***'asher*** *yetsaveh*
for-reason REL order/IMPF/3MS
'so that he may order' (Genesis, 18.19)

The *'asher*-marked relativization pattern can also be found with non-verbal predicates, where again competing patterns are also found. Thus consider:

(11) a. Locative-copular:
ha-basar 'asher b-o ruaḥ ḥayim
the-flesh REL in-it wind/of life
'the flesh that has in it the breath of life' (Genesis, 7.15)

b. Negative adjectival-copular:
*ha-behema **'asher** lo' ṭhorah*
the-beast REL NEG pure
'the beasts that are not pure' (Genesis, 7.2)

c. Locative-possessive copular:
*ha-behemah **'asher** 'it-o b-a-tevah*
the-beast REL with-him in-the-ark
'the animals that (were) with him in the ark' (Genesis, 8.1)

d. Possessive-copular:
*kol **'asher** la-hem*
all REL to-them
'all that they had' (Joshua, 2.14)

2.2.2. The participial relativization pattern

This pattern in EBH is limited to de-verbal modifiers that are, at least functionally, analogous to *subject* REL-clauses. Unlike their correspondents in the *'asher*-marked pattern that covers primarily the two main aspects, perfect and 'imperfect', participial relativization involves mostly the *durative* or *habitual* mode. The choice between the two patterns, for subject relativization, is thus completely determined by tense-aspect-modality.[10]

When the head noun is definite and marked with the definite prefix *ha-*, the modifying participial also takes the definite prefix. This pattern is characteristic of modifying adjectives, demonstratives and ordinals. Formally, participial modifiers in EBH display nominal-adjectival forms of the verb. The very same form is used in lexical *subject nominalizations*. Thus consider:

(12) a. *ve-ha-**'orev** qam mhera mi-mkqom-o*
and-the-**ambusher** rose/3MS quickly from-place-his
'and the ambusher(s) stood up quickly' (Joshua, 8.19)

The use of participials as relative-like noun modifiers may be illustrated with both definite and indefinite heads:

(12) b. Indefinite:
*kol 'esev **zore'a** **zera'***
all grass **seeder/of seed**
'all the grasses **that seed their seeds**'
'all the **seed-casting** grasses.' (Genesis, 1.29)

c. Definite, participle:
kol ha-basar ***ha-romes***
all the-flesh **the-crawling**
'all the animals **that crawl**'
'all the **crawling** animals' (Genesis, 7.21)

d. Definite, adjective:
ha-me'or-ot ***ha-gdol-im***
the-light/NOM-/PL **the-big-**PL
'the **big** lights' (Genesis, 1.16)

2.2.3. Lexical nominalizations

In addition to the subject-case nominal-participial pattern above, EBH furnishes a wealth of evidence for the prevalence, at some archaic stage, of a nominalized REL-clause pattern for the 'internal' clausal case-roles. Synchronically, the bulk of these are clearly *lexical* nominalizations — rather than the more *syntactic* infinitives, to be discussed further below. Some are action nominals with the nominalizing prefix *m-*, as in:

(14) *le-**mi**-mshelet* *ha-yom*
to-NOM-government/of the-day
'for the governing of the day' (Genesis, 1.16)

Others are object nominalizations without the prefix, as in:

(15) *ve-tishma''* *ha-'arets* ***'imrey*** *p-i*
and-hear/IRR/3FS the-earth sayings/of mouth-my
'and let the earth hear the sayings of my mouth' (Deut., 32.1)

While the *'asher*-marked pattern is numerically predominant for object REL-clauses, one may on occasion find a double-patterned hedge, where the lexical-nominal pattern is used first, then immediately recapitulated by its *'asher*-marked equivalent, almost as if clarification is in order. This pattern is very frequent, as in:

(16) *mi-shvu''at-ex* ***'asher*** *nishba''t-anu*
from-swear/NOM-your REL swear/PERF/2FS-us
'of the oath that you have sworn to us'
'from your swearing that you have sworn to us' (Joshua, 2.20)

The coding of subjects-agents as possessor — genitive — of the nominalized verb, as in (16), is of course the most typical sentence nominalization pattern cross-linguistically.

Locative nominalizations taking the lexical-nominalized pattern are abundant in EBH, all taking the nominalizing prefix *m-*.[11] They are invariably location nouns derived from location or motion verbs as in:

(17) *mo-rad* 'down-slope', 'descent', from **(y)rd* 'descend'
ma-'aleh 'up-slope', 'ascent', from **"l(h)* 'ascend'
ma-qom 'place', from **qom* 'stand'
mo-shav 'site', 'place of sitting', from **(y)shv* 'sit'

The same *m*-prefixed pattern is also abundant with agent, instrument, action or abstract lexical nominalizations, as in:

(18) *ma-"ader* 'hoe', from **"dr* 'to hoe'
m-oreh 'teacher', from **h-r'* 'cause-see', 'teach'
me-mshala 'government', from **mshl* 'govern'
me-rots 'race', from **rts* 'to run'

All these nominals are clearly lexical, and were not counted as REL-clauses in computing frequency distribution in this study. Nonetheless one must keep in mind that the distinction between lexical and syntactic nominalization — thus also between nominals and REL-clauses, respectively — is not always easy to draw.

2.2.4. Infinitival syntactic nominalizations

The clearly syntactic infinitival-nominal REL-clause pattern is most productive with oblique case-roles and ADV-clauses, where it is numerically the predominant pattern in EBH. As illustration, consider time-ADV clauses:

(19) a. *be-yom "asot YHWH*
LOC-day/of make/INF/of YHWH
'on the day that God made ...'
(lit.: 'on the day of God's making ...') (Genesis, 2.4)

b. *be-hibar'-am*
LOC-PASS/create/INF-their
'when they were created'
(lit.: 'at their being created') (Genesis, 2.4)

c. *"ad shuv-xa 'el ha-'adamah*
till return/INF-your to the-earth
'till your return to the earth'
(lit.: 'until your return to the earth') (Genesis, 3.19)

d. *'axarey holid-o 'et-Mahal'el*
after/of sire/INF-his ACC-Mahal'el
'after he sired Mahal'el'
(lit.: 'after of his siring Mahal'el') (Genesis 5.13)

e. *ki-shmo"a malxey ha-'emorim*
SUB-hear/INF/of kings/of the-Emorites
'when the Emorite kings heard'
(lit.: 'upon the Emorite kings' hearing of' (Joshua, 5.1)

The nominal-infinitival pattern is also found in purpose-ADV-clauses, where a competing pattern marked by *'asher* is just beginning its spread in EBH. Another competing purpose-clause pattern, also a spreading innovation, uses the *irrealis-imperfect* finite verb-form without *'asher*, often with the simple conjunction *wa*-preceding. Thus compare:

(20) a. Infinitival purpose clauses in (with a complex preposition):
le-ma"an haḥrim-am
to-reason/of destroy/INF-their
'In order to destroy them'
(lit.: 'for the purpose of their destruction') (Joshua, 11.20)

b. Infinitival purpose-ADV-clause (with the dative preposition):
va-yaniḥe-hu be-gan ha-"eden
and-put/IMPF/3MS-him in-garden/of the-eden
***le**-"avd-o u-**le**-shomr-o*
to-work/INF-it and-*to*-guard/NOM-it
'And he put him in the Garden of Eden to work it and to guard it...' (Genesis, 2.15)

c. Finite irrealis-imperfect purpose clauses:
na"ashe 'adam be-tsalme-nu u-vi-dmute-nu,
make/IMPF/1P man in-image/oh-our and-in-figure/of-our
v-yirdu bi-dgat ha-yam...
and-rule/IMPF/3MP LOC-fish/of the-sea
'Let us make Man in our own image, so that they shall rule over the ocean's fish...' (Genesis, 1.26)

d. Finite purpose clause with the subordinator **'asher-**:
hava nerdah ve-navla sham sfat-am,
EXH descend/IMPF/1P and-mix/IMPF/1P there lip-their
***'asher** lo' yishm"u 'ish 'et-sfat re"ey-hu*
REL NEG hear/IMPF/3MP man ACC-lip/of friend-his
'let's go down there and mix-up their language(s), so that they won't understand each other's language...' (Genesis, 11.7)

The infinitival-nominal relativization pattern is also found in manner adverbs of manner-ADV-clauses:

(20) a. Infinitival manner-ADV-clause:
***ki**-lḥox ha-shor 'et-yereq ha-sadeh*
as-chew/INF/of the-ox ACC-grass/of the-field
'the way the ox chews the grass of the field'
(lit.: 'like the ox's chewing the field grass') (Num., 22.4)

b. infinitival/nominal manner adverb:
*yerushah **k**-ma-ḥlkot-am*
inheritance **as**-NOM-divisions-their
'inheritance according to the way they divided'
(lit.: '...according to their divisions') (Joshua, 12.17)

2.3. The numerical distribution of the various relativization patterns in Early Biblical Hebrew

In this section we survey the frequency distribution of the three main types of REL-clauses in EBH — the *'asher*-marked, the participial, and the infinitival-nominal — in the various case-role categories. Purely lexical nominalizations, noted in section 2.2.3. above, were not counted. The texts counted are, separately, Genesis (chapters 1–20), and Joshua (chapters 1–20). The assumption made here concerning the temporal sequence of the BH dialect continuum will be discussed further below.

The distribution in the two books is rather similar, and may be summarized as follows:

(a) *Subject*: The *'asher*-marked form has taken over the perfect, imperfect/irrealis and copular-habitual modes; the participial form is used in the habitual mode;
(b) *Accusative*: The *'asher*-marked pattern is 100% entrenched;
(c) *Locative*: The *'asher*-marked pattern is 100% entrenched;

(d) *Manner*: The *'asher*-marked pattern is almost 100% entrenched;
(e) *Time & purpose clauses*: Only here does the lexical-nominal pattern survive massively.

Table 1 *The distribution of REL-clause types in the various case-roles and ADV-clause types in EBH (Genesis 1–20)*

	RELATIVIZATION TYPE							
	'asher (PERF/IMPF)		PARTICIPIAL		INF-NOMINAL		TOTAL	
CASE-ROLE	N	%	N	%	N	%	N	%
subject	9	32	19	68	/	/	28	100
accusative	32	100	na	na	/	/	32	100
locative	11	100	na	na	/	/	11	100
manner	9	100	na	na	/	/	9	100
time-ADV	2	5	na	na	41	95	43	100
purpose-ADV	1	33	na	na	2	67	3	100

Table 2 *The distribution of REL-clause types in the various case-roles and ADV-clause types in EBH (Joshua 1–20)*

	CONSTRUCTION TYPE							
	'asher (PERF/IMPF)		PARTICIPIAL		INF-NOMINAL		TOTAL	
case-role	N	%	N	%	N	%	N	%
subject perf/impf)	17	46	20	54	/	/	37	100
accusative	42	100	na	na	/	/	42	100
locative	12	100	na	na	/	/	12	100
manner	41	95	na	na	2	5	43	100
time-ADV	1	20	na	na	41	80	51	100
purpose-ADV	1	20	na	na	4	80	5	100

This distribution clearly separates the clause-internal case-roles from the adverbial ones. In the former, the archaic nominal pattern of relativization survives only lexically — except for the use of the participle-adjectival form for subject relativization in the durative-habitual verbal mode. This conspicuous survival is clearly motivated by the fact that by EBH time, the participle

has become the main *finite* form for the two imperfective aspects *in main clauses* as well.

If one assumes, as the etymology of *'asher* clearly entitles us to, that the new finite relativization pattern was initiated in locative case-role, then the spreading clearly proceeded through the object case-roles first, perhaps then onward to the subject case. At least at the beginning point of the BH dialect continuum, the new finite pattern had not yet penetrated deeply into the adverbial clause paradigm.

2.4. Because-ADV-clauses in Early Biblical Hebrew

One conspicuous ADV-clause type where only scant competition between relativization patterns is found, and where the penetration of the *'asher*-marked relativization pattern is least advanced in EBH, is the *because*-ADV-clause. All such clauses in EBH are finite, and the overwhelming majority are marked by the subordinator *ki-* — the very same one also used to mark indirect-quote verbal complements (and, incidentally, also an archaic 'if' subordinator for *conditional*-ADV-clauses; and in combination with other operators also part of the subordinator for *concessive* ADV-clauses; see extensive discussion further below). As an example of the *ki*-marked pattern of *because*-ADV-clauses, consider:

(20) Because-ADV-clauses subordinated by **ki-**:
va-yevarex 'elohim 'et-ha-yom ha-shvi"i...
and-bless/IMPF/3MS God ACC-the-day the-seventh
***ki**-b-o shavat mi-kol mla'xt-o...*
SUB-LOC-it rest/PERF/3MS from-all work/of-his
'and God blessed the seventh day ... *because* on that (day) he had rested from all his work...' (Genesis, 2.3)

The frequency distribution of the syntactic pattern of *because*-ADV-clauses in a large sample of EBH text is given in Table 3, below.

The importance of this pattern for our study of the syntax of verb-complements comes from the combination of two factors:

(a) By Late Biblical Hebrew, the REL-subordinator *'asher* has largely taken over the because-ADV-clause paradigm, partially displacing the old subordinator *ki*.
(b) The *because*-ADV-clause is one of the three main *analogical path-*

ways through which REL-clause morphology is transferred to verb complements.

Table 3 *Frequency distribution of types of because-ADV-clauses in EBH (Genesis 1–20)*

PATTERNS					
'asher-marked:		*ki*-marked		TOTAL	
N	%	N	%	N	%
/	/	38	100.0	38	100.0

3. COMPLEMENTS OF COGNITION-UTTERANCE VERBS IN EARLY BIBLICAL HEBREW

In this section we survey the syntax, and text-distribution of competing variants, of complements of cognition-utterance verbs[12] in EBH. That is, we characterize here the same beginning point on the BH dialect continuum as for the syntax of relativization, above. We will describe each pattern separately, then discuss their distribution.

3.1. Indirect-quote complements

Indirect-quote complements of cognition verbs, such as 'see', 'hear', 'understand', 'know', 'remember', 'forget' etc., are marked in EBH by either one of two subordinators, *ki* or *ve-hineh*. The first probably started, to judge from various types of comparative and internal evidence,[13] as some type of *locative* preposition. In EBH it functions as a *time-adverb* marker, *conditional* marker, the *comparative* 'as', 'like' marker, as well as others. The second is a compound of the conjunction *v-* 'and' and an archaic verb 'be' **hn*.[14] To contrast the two patterns:

(22) a. *va-yar'* *'elohim* *ki-ṭov*
and-see/IMPF/3MS God SUB good
'and God saw that it was good' (Genesis, 1.10)

b. *va-yar'* ***ve-hineh*** *ḥarvu* *pney* *ha-'adamah*
add-saw **and-be** dry/PERF/3MP faces/of the-earth
'and he saw that the surface of the earth had dried up'
(Genesis, 8.13)

The functional distribution of *ki* and *ve-hineh* is relatively transparent. The subordinator *ki* seems to mark highly predictable or *presupposed* information in the complement clause. In contrast, *ve-hineh* — often translated as 'and lo!' — marks surprising, non-presupposed information. This is to some extent indicated by the distribution of the two variants among various verbs in the EBH text. In Table 4 below, we summarize the frequency distribution of *ki-* and *ve-hineh*-marked complements of various main verbs in the Kings II text (chapters 1–20).

Table 4 *Frequency distribution of the two complementizers* ki *and* ve-hineh *with different main verbs in EBH text (Kings II, chs 1–20)*

	ki		*ve-hineh*		TOTAL	
verb	N	%	N	%	N	%
'know'	8	100.0	/	/	8	100.0
'hear'	4	100.0	/	/	4	100.0
'see'	4	57.1	3	42.97	7	100.0
nominalized ('sign')	1	100.0	/	/	1	100.0
TOTAL:	17	85.0	3	15.0	20	100.0

The only three instances of *ve-hineh* in the text involve the verb 'see'. In all three instances, *surprise discovery* is involved, either upon opening one's eyes, or upon entering the scene. The three examples are:

(23) a. *va-yifqaḥ 'adonay 'et-"eney ha-na"ar*
and-opened/IMPF/3MS God ACC-eyes/of the-boy
va-yar' ***ve-hineh*** *ha-har male' sus-im*
and-see/IMPF/3MS **and-be** the-mountain full horse-PL
'and God opened the boy's eyes and he saw **that** the mountain was swarming with horses' (Kings II, 6.17)

b. *va-yifqaḥ YHWH 'et-"eney-hem*
and-open/IMPF/3MS YHWH ACC-eyes-their
va-yir'd ***ve-hineh*** *b-tox Shomron*
and-see/IMPF/3MP **and-be** in-middle/of Shomron
'and God opened their eyes and they saw *that* (they were) inside Shomron' (Kings II, 6.20)

c. *va-tavo'* *'el-beyt* *ha-"am*
and-enter/IMPF/3FS to-house/of the-people
va-tere' ***ve-hiheh*** *ha-melex "omed* *"al-"amud*
and-see/IMPF/3FS **and-be** the-king standing/MS on-pillar
'and she entered the public hall and saw *that* the king was standing on top of a pillar' (Kings II, 11.13)

In all three instances, further, the complement verb (or implied verb) is in the *present-progressive* aspect.

In contrast, when 'see', 'know' or 'hear' take the complementizer *ki-*, the information in the complements is *presupposed*, the complement verb is often marked by the *perfect/anterior* aspect characteristic of presupposed clauses in EBH and elsewhere,[15] and the entire sentence is often embedded in a rhetorical question, the answer for which is assumed to be common knowledge. Thus consider:

(24) a. *va-yhi* *ki-shmoa"* *'elisha 'ish* *ha-'elohim*
and-be LOC-hear/INF/of Elisha man/of the-God
ki *qara"* *melex* *Yisra'el* *'et-begad-av*
SUB tear/PERF/3MS king-of Israel ACC-clothes-his
'and so when Elisha the man of God heard *that* the King of Israel *had torn* his clothes' (King II, 5.8)

b. *va-yo'mru* *'el-av ha-yada"ta*
and-say/IMPF/3MP to-him Q-know/PERF/3MS
ki *ha-yom 'elohim loqeaḥ* *'et-'adoney-xa...?"*
SUB the-day God taking/MS ACC-master-your
'and they told him: "Do you know *that* today God is taking away your master...?"' (King II, 2.5)

c. *ve-hu'* *'amar 'el-ha-zqen-im ha-r'item*
and-he said to-the-old-PL Q-see/PERF/2MS
ki *shalaḥ* *ben* *ha-meratseaḥ ha-ze...?*
SUB send/PERF/3MS son/of the-murderer the-this
'and he said to the elders: "Have you noticed that this son-of-a-murdered **had sent**...?"' (Kings II, 6.32)

The suggestion that *ki*-marked complements are presupposed is further supported by the fact that in EBH *ki-* is also the main marker of the (largely presupposed) *because*-ADV-clauses.

By Late Biblical Hebrew (henceforth LBH), the use of *ve-hineh* seems to have been discontinued in favor of *ki*, which may have thus spread to non-presupposed complements as well. This displacement seems to have been completed long before the invasion of *she-* into the V-complement paradigm, although the facts are not fully clear, except that *ki-* is clearly numerically predominant in the LBH text. In all likelihood then, *she-* replaced *ki-* alone.

3.2. Direct-quote complements

Only utterance verbs take direct-quote complementation in EBH, an almost universal distribution. Such complements appear either with no subordinator at all, or with two possible forms of the verb 'say' — one infinitival, the other finite and agreeing with the subject of the main verb.[16] As illustration of the three patterns, consider:

(25) a. Zero complementizer:
va-yo'mer *'elohim:* *'Yhi* *"or'*
and-say/IMPF/3MS God be/EXHO/3MS light
'and God said: "Let there be light"' (Genesis, 1.3)

b. Uninflected 'say':
va-yetsav *'elohim "al ha-'adam* ***le-'mor***:
and-order/IMPF/3MS God DAT the-man**to-say**
mi-kol *'asher "ets* *ha-gan* *to'xel...'*
from-all REL tree/of the-garden eat/IMPF/3MS
'and God commanded Adam thusly: "You may eat from whatever tree in the garden..."' (Genesis, 2.16)

c. Inflected 'say':
va-yevarxe-hum ***va-yo'mar***:
and-bless/IMPF/3MS-him **and-say/IMPF/3MS**
barux 'avraham...'
blessed Abraham
'and he blessed him saying: "Blessed be Abraham..."'
(Genesis, 14.19)

The *zero* complementizer in EBH is used predominantly when the main verb is itself the verb 'say' or 'tell'. The subordinator 'say' is most often used with less prototypical utterance verbs, as in (25b,c) above. The functional distribution of the two forms of 'say' remains to be fully determined.[17]

3.3. The accusative blend indirect-quote pattern

There exists a variant of the indirect-quote complementation pattern described above, where the complement is a *syntactic blend* between an accusative — direct nominal object — and a sentential complement, the latter marked by *ki* or *ve-hineh*. The presence of this blend pattern in EBH is noteworthy because it turns out to have figured prominently, as one of three analogical pathways, in the subsequent transfer of REL-clause morphology to verb complements. As illustrations consider:

(26) a. *va-yar' 'elohim* ***'et-ha-'or*** *ki ṭov*
and-see/IMPF/3MS God ACC-the-light SUB good
'and God saw that the light was good'
'and God saw the light, (and) that it was good' (Genesis, 1.4)

b. *va-yar' 'elohim* ***'et-kol 'asher "asa***
and-see/IMPF/3MS God ACC-all REL make/PERF/3MS
ve-hineh *ṭov me'od*
and-be good/ms very
'and God saw all that he had made and lo, it was very good'
(Genesis, 1.31)

The frequency distribution of the various V-complement forms in EBH is of no overriding interest at this point,[18] except for noting that the *'asher*-marked REL-clause pattern is totally absent in the EBH text of either Genesis (chapters 1–20) or Joshua (chapters 1–20). The relevant frequency distribution of the *ki/ve-hineh* vs. the *'asher/she* complementation patterns is given — and discussed in great detail — in section 5.

4. THE BIBLICAL HEBREW DIALECT CONTINUUM

4.1. Main variables used to identify continuum-points

In this section we survey the criteria used to identify successive points along the BH dialect continuum. Once these successive points are established, one can track down the transfer of REL-clause morphology to V-complements along the continuum. These dialect-continuum points were established by purely linguistic means, through tracking the progress — in terms of changes in the frequency distribution of competing variants — of four major dia-

chronic changes that span the BH dialect continuum. One of the changes is essentially phonological, two involve grammatical morphology, and one involves word-order. These four diachronic processes are:

(27) (a) The phonological contraction of the subordinator *'asher* into *she*.
(b) The transfer of *'asher/she* from REL-clauses to V-complements;
(c) The complex transformation of the tense-aspect-modality system;
(d) The drift from the predominant VSO order of EBH to the predominant SVO order at the end of the continuum.

With respect to all four variables, the next point on the continuum after BH, *Mishnaic Hebrew*, codified roughly around the 2nd Century AD, reveals the changes to be 100% complete. That is:

(a) All occurrences of the REL-subordinator in Mishnaic Hebrew (henceforth MH) are of the shortened, clitic form *she-* (rather than *'asher*);
(b) All indirect-quote complements use the subordinator *she-*, which has by then completely replaced the EBH subordinators *ki* and *ve-hineh*;
(c) The main line of the perfective, sequential, foreground, thematically-continuous narrative is coded by the EBH *perfect*, which thus takes over from the EBH '*imperfect*';
(d) The dominant main-clause word-order in *Song of Songs* and Mishnaic Hebrew is SVO, rather than the dominant VSO order of EBH.

These four changes are essentially complete by the end-point of the BH dialect continuum, the book *Song of Songs*. Examples illustrating the state of the four variables at the end-point of the continuum, in either *Song of Songs* or Mishnaic text, are as follows:

(28) a. The subordinator *she-* in REL-clause:
biqashti *'et-**she**-* *'ahava* *nafsh-i*
search/PERF/1S ACC-REL-love/PERF/3FS soul-my
'I was looking for the one that my soul loved'
(*Song of Songs*, 3.1)

b. The subordinator *she-* in V-complements:
*'amru l-o: limadta-nu **she**-'eyn mqablim*
say/PERF/3MP teach/PERF/2MS-us SUB-NEGreceive/3P
'they told him: "You have taught us that one does not receive..."' (*Zra'*im, *Braxot*, 2.7)

c. The *perfect* as the main narrative tense and

d. The use of the SVO word-order in foreground clauses:
dod-i yarad le-gan-o
lover-my descend/PERF/3MS to-garden-his
'my loved went down to his garden' (Song of Songs, 6.2)

Based on the four major diachronic processes outlined above, and on the frequency distribution of the competing variants in various book of the Old Testament, the following 6-point continuum can be established for tracking down the changes across the BH dialect continuum:

(29) Early Biblical Hebrew:
(i) Genesis
(as well as other major portions of the Five Books;
as well as probably the bulk of Joshua, Judges and Samuel)
(ii) Kings II
(probably also Kings I)
Late Biblical Hebrew:
(iii) Esther
(as well as Ezra, Nehemiah and Daniel, perhaps others)
(iv) Lamentations
(v) Ecclesiastes
(vi) Song of Songs

In establishing this continuum, we proceed without regard to the long scholarly tradition of dating the text by means other than such variables of linguistic change. In doing this, one leaves oneself open to a variety of possible criticisms from traditional Biblical scholarship. This venerable tradition has used historical dating methods that need not be quarreled with. Most commonly, it dates the time in which the texts were either codified or copied. But in many instances the language of the text may not reflect the spoken — or even written — language used at the time of codification, but rather harkens

back to some other time along the BH continuum. This is an extremely involved subject that will not be addressed here in any comprehensive way.

The fundamental assumptions underlying our choice of purely linguistic methodology is based, for better or worse, on over 100 years of linguistic study of the geographic and/or temporal distribution of linguistic variation and language change. They may be given as follows:

(30) 'If a number of variables, i.e. pairs of competing forms in phonology, morpho-syntax and word-order, line up the very same texts (or speech communities) in exactly the same order of gradually changing relative frequencies along either a geographic or temporal space, then only two coherent explanations exist for such remarkable convergence:
(i) These are contiguous points along a geographic dialect dialect continuum; or
(ii) These are contiguous points along a temporal diachronic change continuum'.[19]

To illustrate the graduality of the BH dialect continuum, and the degree to which information from separate linguistic variables indeed converges, and thus to justify treating the BH dialect continuum as a normal temporal continuum of diachronic change, consider the following frequency shift for three of the four main variables given in (27) above.[20]

As can be seen from comparing the three tables, the purely phonological change — contraction of *'asher* to *she-* — seems to lag behind the grammatical changes, essentially not appearing in any appreciable frequency until the Ecclesiastes point that is virtually the *midway* point along the continuum. The fact that in Ecclesiastes *'asher* and *she-* are essentially in free variation is well illustrated by their occurrence in adjacent clauses or even in the same clause. Thus consider:

(31) Free variation of *she-/'asher* in LBH:
a. *u-mah* ***she-haya*** *kvar* *huh,*
and-what REL-be/PERF/3MS already be/3MS
*va-**'asher** li-hyot* *kvar* *haya*
and-REL to-be/INF already be/PERF/3MS
'and whatever has happened already is, and what is to be has already happened' (Eccl., 3.15)

Table 5 *Frequency distribution of the* 'asher *and* she- *subordinator forms in REL-clauses and V-COMP-clauses in the BH/MH dialect continuum*

	REL-CLAUSES				V-COMPLEMENTS			
	'asher		*she-*		*'asher*		*she-*	
book	N	%	N	%	N	%	N	%
Genesis	93	100.0	/	/	/	/	/	/
Lamentations	9	75.00	3	25.0	/	/	/	/
Esther	100	100.0	/	/	7	100.0	/	/
Nehemiah	58	100.0	/	/	14	100.00	/	/
Ecclesiastes	69	52.6	62	47.4	11	47.8	12	52.2
Song of Songs	/[21]	/	25	100.0	/	/	6	100.0
Mishnaic H.	1	0.6	122	99.4	/	/	24	100.0

Table 6 *Frequency distribution of the VS and SV word-order in main clauses across the BH dialect continuum* [from Givón, 1977[22]]

	VS ORDER		SV ORDER		TOTAL	
book	N	%	N	%	N	%
Genesis	169	87.1	25	12.9	194	100.0
Kings II	174	76.8	53	23.2	227	100.0
Esther	99	73.3	36	26.7	135	100.0
Lamentations	36	50.0	36	50.0	72	100.0
Ecclesiastes	11	21.0	41	79.0	52	100.0
Song of Songs	2	8.0	26	92.0	28	100.0
Mishnaic	9	7.2	115	92.8	124	100.0

b. *ve-shabeaḥ 'ani 'et-ha-metim* ***she****-kvar* *metu*
and-praise I ACC-the-dead/MP REL-already die/PERF/3MP
min-ha-ḥayim ***'asher*** *hema ḥayim "adena*
from-the-live/MP REL they live/mp still
'and I praise the dead who have already died over the living who are still alive' (Eccl., 4.2)

Table 7 *Frequency distribution of the* perfect *and* imperfect *forms in the aspectual function of foreground, perfective, sequential, thematic- continuity across the BH dialect continuum* [From Givón, 1977]

	perfect		imperfect		TOTAL	
book	N	%	N	%	N	%
Genesis	/	/	379	100.0	379	100.0
Kings II	13	1.6	753	98.4	766	100.0
Esther	61	27.3	162	72.7	223	100.0
Lamentations	21	87.3	31	12.7	246	100.0
Ecclesiastes	97	96.0	4	4.0	101	100.0
Song of Songs	62	95.3	3	4.7	65	100.0

c. *tov* ***'asher*** *lo' tidor,*
good SUB NEG promise/IMPF/2MS
*mi-**she**-tidor* *ve-lo'* *teshalem*
from-SUB-promise/IMPF/2MS and-NEG pay/IMPF/2MS
'it is better that you don't promise, than that you promise and not pay back' (Eccl., 5.4)

The seeming lag in the change of our phonological variable may of course be due to the total independence of the variables from each other. It may also be altogether illusory, reflecting the well-documented conservatism of writing systems at the phonological level. Purely phonological contractions of this type are characteristic of spoken language, and are often not reflected in the writing system until much later. This particular contraction, preserving only the stressed second syllable of *'asher*, is so natural and predictable that it probably predates by far the point where it begins to be reflected in the written LBH text.[23]

4.2. Adverbial clauses and the BH dialect continuum

In addition to the three independent changes documented in Tables 5,6,7 above, another change also continues, toward virtual completion, along the BH dialect continuum. This is the continued extension of subordinator *'asher/she-* into the adverbial-clauses paradigm. Let us consider first purpose-ADV-clauses. In the first 17 chapters of the Mishnaic text of *Zra"im*,[24] 90% of

purpose-ADV-clauses were found to take the finite pattern with the REL-subordinator *she-* preceded by either one of three innovated prepositions. Only 10% — represented by a single case — displayed the infinitival-nominal. Thus consider:

(32) a. Finite purpose-clause with *kdey-she-*:
kdey ***she-lo'*** *yaku* *'ish* *be-re"ey-hu*
reason/of REL-not hit/IMPF/3MP man LOC-friend/of-his
'so that they may not hit each other' (*Zra"im, Pe'a*, 4.4.)

b. Finite purpose-clause with *"al-mnat-she-*:
"al mnat ***she-yelaqeṭ*** *bn-o* *'aḥar-av*
on reason/of REL-gather/IMPF/3MS son-his after-him
'in order that his son may gather after him' (*Zra"im, Pe'a*, 5.6)

c. Finite purpose-clause with *bishvil-she-*:
bi-shvil ***she-lo'*** *yo'vedu*
LOC-way/of REL-not get-lost/IMPF/3MP
'so that they do not get lost' (*Zra"im, Dmey*, 3.3)

d. Infinitival-nominal purpose-clause with *kdey*:
kdey *li-ktsor* *ve-li-shnot*
PREP to-mow/INF and-to-repeat/INF
'in order to mow repeatedly' (*Zra"im, Pe'a*, 3.6)

As we will suggest further below, the spreading of the *'asher*-marked pattern into the purpose-clause paradigm is important for understanding the spread of REL-clause morphology into the V-complement paradigm, since purpose-clauses may be a necessary intermediate step within one of the analogical pathways that transmitted REL-clause morphology to V-complements. The frequency distributions of the two competing variants of purpose-clause syntax along the BH dialect continuum is given in Table 8, below.

The systematic displacement of the old nominalized pattern by the finite *'asher*-marked pattern is also seen with time-ADV-clauses. As one may recall, the bulk of time-ADV-clauses in EBH displayed the old, infinitival-nominal relativization pattern (95% in the Genesis text and 80% in Joshua). In contrast, 88% of time-ADV-clauses in the Mishnaic text counted are marked with *she-*. The *Song of Songs* text, likewise, has no single instance of the old nominal/non-finite form. Rather, one finds there only *she*-marked finite forms of time-ADV-clauses, as in (33) below.

Table 8 *Frequency distribution of the variants of purpose-ADV-clauses across the BH dialect continuum*

	VARIANTS					
	NOMINAL-INF		*'asher/she*		TOTAL	
book	N	%	N	%	N	%
Genesis	2	67.0	1	33.0	3	100.0
Joshua	4	80.0	1	20.0	5	100.0
Esther	11	100.0	/	/	11	100.0
Nehemiah	1	100.0	/	/	1	100.0
Lament	/	/	/	/	/	/
Ecclesiastes	4	50.0	4	50.0	8	100.0
Song of Songs	/	/	/	/	/	/
Mishnaic H.	1	10.0	9	90.0	10	100.0

(33) *"ad* ***she****-matsa'ti* *'et-****she****-'ahava* *nafsh-i*
till REL-find/PERF/1S ACC-REL-love/PERF/3FS soul-my
'until I found the one whom my soul loved' (Song of Songs, 3.4)

The relative frequency of the two competing variant patterns of time-ADV-clauses along the BH dialect continuum is given in Table 9, below.

By Mishnaic time or even earlier, further, the finite *she*-marked pattern had also become the dominant type of concessive-ADV-clause, as in:

(34) *'af "al pi* ***she****-'eyn ha-baqar yaxol la-"avor*
also on mouth REL-NEGthe-cattle can/3MS to-pas/INF
'*even though* the cattle cannot pass' (*Zra"im*, *Pe'a*, 2.2)

Finally, one may recall (cf. Table 3) that because-ADV-clauses in EBH display no penetration of the REL-marker *'asher*, only the *ki*-marked finite pattern. By Mishnaic times — and in fact in LBH — because-clauses have begun to display the penetration of the *'asher/she* subordinator. To illustrate this pattern from MH, consider:

(35) a. Because-ADV-clause, with a preposition:
mi-pney ***she****-'efshar*
from-face/of REL-possible/MS
'because it is possible' (*Zra"im*, *Pe'a*, 5.3)

Table 9 *Frequency distribution of the variants of time-ADV-clauses across the BH dialect continuum*

	VARIANTS					
	NOMINAL-INF		*'asher/she*		TOTAL	
book	N	%	N	%	N	%
Genesis	41	95.0	2	5.0	43	100.0
Joshua	41	80.0	10	20.0	51	100.0
Esther	13	81.2	3	18.8	16	100.0
Lamentations	2	50.0	2	50.0	4	100.0
Song of Songs	/	/	9	100.0	9	100.0
Mishnaic H.	7	12.0	47	88.0	54	100.0

b. Because-ADV-clause, without a preposition:
***she**-ken derex bney mlaxim*
REL-thus way/of sons/of kings
'because this is the way of princes' (*Zra"im, Braxot*, 1.7)

In fact, the pattern is already visible earlier on in LBH, as in:

(36) a. Because-ADV-clause, with a preposition:
itke-dat mah la-"asot ba-malka Vashti
as-law/of what INF-do with-queen Vashti
*"al-**'asher** lo' "asta 'et-ma'mar ha-melex*
LOC-REL NEG do/PERF/3FS ACC-order/of the-king
'according to what law should one deal with Queen Vashti **because** she didn't comply with the King's order...?'
(Esther, 1.15)

b. Because-ADV-clause, without a preposition:
la-nu bo'shet ha-panim ... ve-la-'avotey-nu,
to-us shame/of the-face and-to-fathers-our
'we are the guilty/shameful ones...
***'asher** ḥṭa"nu le-xa*
REL sin/PERF/1P LOC-you
because we had sinned against you' (Daniel, 9.8)

c. Because-ADV-clause, without a preposition:
madu"a lo' yer"u pan-ay,
why NEG sad/IMPF/3PS face-my
'why shouldn't my face look sad,
***'asher** ha-"ir beyt qivrot 'avot-ay ḥareva*
REL the-city house/of tombs/of fathers-my ruined
since the city of my father's tombs is ruined' (Nehemiah, 2.3)

d. Because-ADV-clause, without a preposition,
in free variation with *ki-*:
***ki** mi yode"a mah ṭov la-'adam ba-ḥayim,*
SUB who knows/PARTIC/SM what good for-man in-life
'**because** who knows what is good for a man in his life,
mispar ymey ḥayey hevl-o...,
number/of days/of life/of vanity-his
the days of his vainful life are numbered...,
***'asher** mi yagid la-'adam mah yihyeh*
SUB who tell/IMPF/3MS to-man what be/IMPF/3MS
because who can tell a man what will be (after he is gone)'
(Eccles., 6.12)

In fact, sporadic examples of *'asher*-marked because-ADV-clauses can be found in EBH, at least as early as Joshua, as in:

(37) *ki-yish'alun bney-xem maḥar le-'mor*
if/when-ask/IMPF/3MP sons-your tomorrow to-say
'so if your children will ask you in the future:
mah ha-'avan-im ha-'elej la-xem,
what the-stone-PL the-these to-you
"What are these stones to you?"
va-'amartem la-hem
and-say/PERF/2MP to-them
then you shall tell them:
***'asher** nixretu mey ha-Yarden...*
REL split/3MP water/of the-Jordan
"Because the Jordan's water were divided..."' (Joshua, 4.7)

The frequency distribution of the *ki* and *'asher/she*-marked variants of because-ADV-clauses at several points along the BH dialect continuum is given in Table 10, below.

Table 10 *Frequency distribution of the variants of because-ADV-clauses across the BH dialect continuum*

	VARIANTS					
	ki-		*'asher/she*		TOTAL	
book	N	%	N	%	N	%
Genesis	38	100.0	/	/	38	100.0
Esther	20	90.9	2	9.1	22	100.0
Nehemiah	22	91.6	2	8.4	24	100.0
Lamentations	15	88.2	2	11.8	17	100.0
Ecclesiastes	57	89.0	7	11.0	64	100.0
Song of Songs	5	71.4	2	28.6	7	100.0
Mishnaic H.	/	/	35	100.0	35	100.0

In sum, the continuing spread of the *'asher*-marked pattern across the BH dialect continuum, a process that began in pre-Biblical Hebrew, clearly suggests that the dialect-continuum we assume, along which we track the transfer of REL-clause morphoplogy to V-comp-clauses, is also evident in other types of diachronic syntactic change in Biblical Hebrew.

Finally, the further spreading of the *'asher*-marked pattern into both the purpose-clause and because-ADV-clause paradigms is of course important for a reason that has more direct bearing on our study — namely that these ADV-clauses may figure directly in the line of transmission of REL-clause morphology to V-complements.

5. THE DYNAMICS OF TRANSFER OF REL-CLAUSE MORPHOLOGY INTO THE V-COMPLEMENT PARADIGM IN BIBLICAL HEBREW

5.1. The shift of frequency distributions

We come now to the core of this study, the tracking down of the transfer of REL-clause morphology into the V-complement paradigm. In pursuing this task, we will first document the gradual spreading of the REL-clause subordinator *'asher/she-* into the V-complement paradigm. The most telling figures are those involving the relative frequency of the two competing forms of indirect-quote complementizer for cognition-utterance verbs — the older

ki/ve-hineh and the expanding *'asher/she-*. The figures are given in Table 11, below.

Table 11 *Frequency distribution of the* 'asher/she- *subordinator and the* ki/ve-hineh *subordinator in V-complements in the BH/MH dialect continuum*

	'asher/she-		*ki/vehineh*[21]		TOTAL	
book	N	%	N	%	N	%
Genesis	/	/	24	100.0	24	100.0
Kings II	/	/	20	100.0	20	100.0
Lamentations	1	12.5	7	87.5	8	100.0
Esther	7	50.0	7	50.0	14	100.0
Nehemiah	14	66.6	7	33.4	21	100.0
Eccl	23	65.7	12	34.3	35	100.0
Song of Songs	24	100.0	/	/	24	100.0

As can be seen, the gradual shift in frequency distribution finds its terminal point in the *Song of Songs* and its half-way point in Esther. Ecclesiastes, the half-way point for many of the other variables studied, is at the 65% completion point here. In the next section we attempt to outline the mechanism via which this protracted process of analogical extension must have occurred.

5.2. Analogical pathways

5.2.1. Preamble

There is strong evidence that the transfer of REL-clause morphology to V-complements during the BH dialect progression involved not only one but *three* analogical pathways, together responsible for the penetration of the REL-subordinator at *three distinct points* of the of the V-complement paradigm. Each of these putative pathways involves an intermediate syntactic construction that is a *syntactic blend*:

(a) The accusative-propositional blend
(b) The purpose-subjunctive blend
(c) The because-factive blend

We call these intermediates 'blends' for cogent reasons: In each case, an *intermediate syntactic pattern* exists that may be interpreted as coding either

one context — relative clause — or the other — verb complement. In other words, the two variant contexts are — semantically or pragmatically — so close that the syntactic 'blend' structure may aptly receive *either* interpretation.

Putative syntactic blends of this type are created, presumably, via analogic extension of one form to a new functional context that is *sufficiently similar* to the form's original context. Presumably, considerations of *relative coding density* — i.e. functional load — also play a role in such extension. That is, the old functional context in such instances tends to be relatively 'over-coded', while the new one is 'under-coded'.

One may consider our 'syntactic blends' — or 'intermediates' — *analogical bridges* that span otherwise unbridgeable gaps between functionally-dissimilar points along a change continuum. In this case, the initial functional gap between the prototype REL-clause and the prototype V-complement is by itself substantial. The significance of the existence of these intermediates is important both in this case, as well as for the theory of analogical change at large, and we will return to it in section 6.

5.2.2. The accusative-propositional blend

The accusative blend is an intermediate construction that capitalizes on the systematic polysemy of many perception-cognition verbs, whereby such verbs can take either a *nominal direct object* or a *proposition* as their complement. Verbs most commonly in this group are:[25]

'see', 'hear', 'know', 'remember', 'forget',
'believe', 'guess', 'understand'

The accusative blend involves cases where *both* complements are present together in the very same sentence. One finds numerous cases of this type all over the EBH texts, as well as at most other points along the BH dialect continuum. Indeed, such blends should be viewed as a normal by-product of the polysemy of the verbs in question. That is, they represent normal spoken-language *parataxis*, whereby the speaker first chooses to consider the important *nominal topic* as the object of the verb, then — perhaps as an 'afterthought' and with an intonation break — adds the proposition in which that nominal is the topical participant. Schematically, this may be illustrated with:

(38) a. Nominal object:
He saw *the woman*

b. Propositional object:
He saw that *the woman* was there
c. The paratactic accusative blend:
He saw *the woman*, that *she* was there

The only difference between the use of such blends in EBH and their use in LBH is in the subordinators: In EBH the subordinator of the propositional complement is either *ki* or *ve-hineh*, while in LBH it is *'asher/she-*.

As an illustration of the accusative-blend pattern from various points along the BH dialect continuum, consider:

(39) a. *va-yar' 'elohim **'et-ha-'or** **ki** ṭov*
and-see/IMPF/3MS God ACC-the-light SUB good
'and God saw that the light was good'
'and God saw the light, (and) that it was good' (Genesis, 1.4)

b. *va-yar' 'elohim **'et-kol** **'asher** "asa*
and-see/IMPF/3MS God ACC-all REL make/PERF/3MS
***ve-hineh** ṭov me'od*
and-be good/MS very
'and God saw all that he had made and lo, it was very good'
(Genesis, 1.31)

c. *le-har'ot ha-"amim ve-ha-sarim **'et-yofy-a***
to-show the-peoples and-the-officials ACC-beauty-her
***ki** ṭovat mar'eh hi'*
SUB good/3F/of visage she
'to show the people and the court officials **her beauty that** she is good looking' (Esther, 1.11)

d. *ro'im 'atem **'et-ha-ra"a** **'asher** **'anaḥnu** **b-ah***
seeing/3P you ACC-the-evil REL we in-it
***'asher** Yerushalayim ḥareva*
SUB Jerusalem ruined/FS
'you see **the misfortune we are in that** Jerusalem is destroyed...'
(Nehemiah, 2.17)

e. *'al tir'd-**ni** **she-**'ani shḥarḥoret*
NEG see/IMPF/3MP-me REL-I dark/FS
'don't see **me that** I am dark-skinned' (Song of Songs, 1.6)

Another intermediate form, perhaps the next stage in re-analyzing the double-complemented accusative-blend, involves nominal objects that are

largely *semantically depleted*, such as the noun 'thing'/'saying'. As illustration of this, consider two examples, one from EBH, the other from LBH:

(40) a. *diber* YHWH ***'et-ha-davar ha-ze*** *'el-Moshe*
say/PERF/3MS YHWH ACC-the-thing the-this to-Moses
'asher *halax* *Yisra'el*
SUB go/PERF/3MS Israel
'God said (this thing) to Moses that Israel had gone...'
(Joshua, 14.10)

b. *zxar-na'* ***'et ha-davar*** ***'asher*** *tsivita*
remember/IMP-IMP ACC **the-saying** SUB order/PERF/2MS
'et-Moshe "avde-xa ***le-'mor...***
ACC-Moses slave-your ***to-say***
'remember (the thing) that you have ordered your servant Moses:...' (Nehemiah, 1.8)

Next, one also finds a more advanced intermediate syntactic form, where the two complements begin to be integrated into a single — presumably *syntacticized* — form. In such an early integrated form the accusative *'et-* does not mark a nominal any more, but rather is directly followed by the subordinator *'asher*, which in turn then marks the propositional complement. One finds such examples fairly early in EBH, with many examples already as early as Joshua. Thus consider:

(41) a. *hugad* *le-"avde-xa* ***'et-'asher*** *tsiva*
told/PERF/3MS to-slave-your ACC-SUB order/PERF/3MS
YHWH 'elohey-xa 'et-Moshe la-tet la-xem...
YHWH god/of-you ACC-Moses to-give to-you
'your humble servant has been told that YHWH your God had ordered Moses to give you...' (Joshua, 9.24)

b. *shama"nu* ***'et-'asher*** hovish YHWH
hear/PERF/1P ACC-SUB dry/PERF/3MS YHWH
'we've heard that God had dried up

'et-mey *yam* *Suf*
ACC-water/of sea/of Suf
'we've heard that God had dried up the waters of the Red Sea'
(Joshua, 2.10)

c. *zxar-na'* ***'et-'asher*** *hithalaxti* *le-faney-xa*
remember/IMP-EXH ACC-SUB walk/PERF/1S LOC-face-your
'please remember **that** I have walked in front of you'
(Kings II, 20.3)

One also finds this pattern in the LBH text, as in:

(42) *va-yesaper* *la-hem Haman* ***'et-kvod*** ***''oshr-o***
and-tell/IMPF/3MS to-them Haman ACC-weight/of wealth-his
*ve-**'et-'asher*** *gidl-o* ha-melex
and-ACC-SUB increase/PERF/3MS-him the-king
*ve-**'et-'asher*** *nis'-o* *''al ha-sar-im*
and-ACC-SUB raise/PERF/3MS-him on the-official-PL
'and Haman told them (of) the extent of his wealth and *that* the king had made him big and *that* he raised him over the court officials...' (Esther, 5.11)

The process of analogical change can be now given as a three-step progression, schematically:

(43) a. Paratactic accusative-blend:
He saw ACC-*the woman*, SUB-she was there
b. First syntactic intermediate:
He saw ACC-*the thing* SUB-she was there
He saw ACC-*the woman* SUB-she was there
c. Second syntactic intermediate:
He saw ACC-SUB she was there
d. Final simplification:
He saw SUB she was there

Variant (43a), the initial step in the sequence, is attested in the earliest stages of EBH, and there only with the subordinator *ki* or *ve-hineh*. Very clearly, the transfer of the REL-subordinator to V-complements did not depend on this variant. Variants (43b,c), on the other hand, are only attested with *'asher/she-*. The analogy between REL-clause and V-complement must have been created at stage (43b) — for a relatively transparent reason: This must have been the stage where the intonation break, characteristic of the parataxis in (43a), was removed, thereby leaving a syntactic-semantic configuration of a verb-complement that is still *sufficiently similar* to that of the prototypical restrictive REL-clause in terms of three major conditions:

(a) A nominal (here marked as ACC-object),
(b) directly followed by a proposition
(c) in which one argument is co-referential to that nominal

It is at this point that analogy must have exercised its option, so to speak, interjecting the REL-subordinator *'asher/she* into the slot previously occupied, exclusively, by either *ki* or *ve-hineh.*

5.2.3. The purpose-subjunctive blend

The second likely channel for analogical extension of the REL-subordinator into the V-complement paradigm in BH involves two contiguous points on a semantic continuum:

(a) Purpose clauses
(b) Subjunctive exhortative complements

Both are *irrealis, future-projecting, modal* complements of main verbal clauses. In many languages indeed the grammar accords these two similar treatment.[26]

As one may recall, the REL-subordinator *'asher/she-* by the end of the BH dialect continuum — and certainly by the Mishnaic period — has largely replaced the earlier nominal-infinitival form of purpose-ADV-clauses with the *'asher*-marked finite pattern.[27] In addition to the strong semantic-structural affinity between purpose clauses and subjunctive complements, a crucial ingredient of this analogical pathway is the *systematic polysemy* of another group of complement-taking verbs. These are verbs that may take either *subjunctive* or *indicative* complements. As such, they constitute two contiguous points along the V-complementation continuum.[28] Verbs most conspicuously in this group are:

Manipulative: 'tell'/'say', 'order','ask', 'suggest',
Facilitative: 'permit', 'allow','let'
Desiderative: 'want'/'wish', 'hope'

To illustrate this contiguity with schematic English examples, consider:

(42) (i) Desiderative verbs:
a. She wanted *him* to leave (manipulative)
b. She wished *that* he *would* leave (subjunctive)
c. She hoped *that* he *would* leave (subjunctive)
d. She hoped *that* Joe was there (indicative)

(ii) Manipulative verbs:
a. She told *him* to leave (manipulative)
b. She told him *that* he *should* leave (subjunctive)
c. She told him *that* John left (indicative)
d. She ordered *him* to leave (manipulative)
e. She ordered *that* he leave (subjunctive)
f. She asked *him* to leave (manipulative)
g. She asked *him* if he *could* leave (subjunctive)
h. She asked *that* he leave (subjunctive)
i. She suggested *that* he leave (subjunctive)
j. She suggested *that* he had already left (indicative)

(iii) Facilitative verbs:
a. She let *him* leave (manipulative)
b. She allowed/permitted *him* to leave (manipulative)
c. She allowed *that* it *could* be done (subjunctive)
d. She permitted *that* it be done (subjunctive)
e. She allowed *that* he was drunk (indicative; colloquial)

The precursors of this analogical change in EBH are relatively subtle, involving primarily the use of *'asher/she-* either in the *subjunctive of uncertainty*, or in *purpose clauses*. Thus consider:

(45) a. Subjunctive of uncertainty:
va-yo'mer 'el-bney Yisra'el le-'mor
and-say/IMPF/3MS to-sons/of Israel to-say
***'asher** yish'alun bney-xem maḥar...*
SUB ask/IMPF/3MP sons-your tomorrow
'and he told the children of Israel thusly: **if** your children ask you in the future...' (Joshua, 4.21)

b. Purpose-ADV-clause:
ha-yom ha-ze 'aḥel gadel-xa
the-day the-this start/IMPF/1S raise/INF-you
be-"eyney kol Yisra'el,
in-eyes/of all Israel
***'asher** yed"u ki ka-'asher hayiti...*
SUB know/IMPF/3MS SUB like-REL be/PERF/1S
'this very day I will begin to raise you in the eyes of all Israel, **so that** they know that the way I was...' (Joshua, 3.7)

c. Purpose-ADV-clause:
ve-hayta nivlat "izevel ke-domen
and-be/PERF/3FS corpse/of Jezebel like-trash
"al-pney ha-sadeh, ***'asher*** *lo' yo'mru zo'a 'izevel*
on-face/of the-field SUB NEG say/IMPF/3MP this/FS Jezebel
'and Jezebel's corps shall be like trash on the fields, **so that** people would not say: "This is Jezebel"' (Kings II, 9.37)

But even in EBH one already finds instances where this usage had spread to complements of manipulative verbs. Thus consider:

(46) *va-'ashbi"a-xa b-YHWH 'elohey ha-shamayim...*
and-swear/IMPF/1S-you in-YHWH God/of the heavens
'asher *lo' tiqaḥ 'isha li-vn-i mi-vnot*
SUB NEG take/IMPF/2MS wife for-son-my from-daughters/of
'and I enjoin you in the name of YHWH God of the heavens... that you should not take a wife for my son from the daughters of (the Canaanites)...' (Genesis, 24.3)

And this subjunctive V-complement usage is indeed massive by LBH, as in:

(47) a. *ṭov* ***'asher*** *tidor mi-****she****-tidor*
good SUB promise/IMPF/2MS from-SUB-promise/IMPF/2MS
ve-lo' teshalem
and-NEG repay/IMPF/2MS
'it is better that you promise than you promise and not pay' (Eccles., 1.17)

b. *va-'omra* ***'asher*** *lo' yifteḥu-m*
and-say/IMPF/1S SUB NEG open/IMPF/3MP-them
'and I said that they should not open them'(Nehemiah, 13.19)

c. *va-'omra l-a-nevi'im* ***'asher*** *yihyu miṭṭaharim*
and-say/IMPF/1S to-the-prophets SUB be/IMPF/3MS cleansing/MP
'and I told the prophets that they should be cleansing themselves' (Nehemiah, 13.22)

d. *va-yikkatev... 'asher lo' tavo' Vashti*
and-written/IMPF/3MS SUB NEG come/IMPF/3FS Vashti
'and it was written that Vashti should never come' (Esther, 1.19)

e. *ki* *Mordexay tsiva* *"aley-ha*
because Mordechai order/PERF/3MS on-her
'asher *lo'* *tagid*
SUB NEG tell/IMPF/3FS
'because Mordechai ordered her that she should not tell'
(Esther, 2.11)

f. *va-yevaqesh* *mi-sar* *ha-sarisim*
and-ask/IMPF/3MS from-chief/of the-eunuchs
'asher *lo'* *yigga"el*
SUB NEG defiled/IMPF/3MS
'and he asked the Chief of Eunuchs that he not be defiled'
(Daniel, 1.9)

g. *va-yo'mar* *ha-tirshḥa'* *la-hem*
and-say/IMPF/3MS the-minister to-them
'asher *lo'* *yo'xlu*
SUB NEG eat/IMPF/3MP
'and the minister (Nehemiah) told them that they should not eat' (Ezra, 2.63)

h. *'im "al-ha-melex ṭov*
if on-the-king good
ve-'im yiṭṭav *"avde-xa* *le-faney-xa*
and-if good/IMPF/3MS slave-your LOC-front-your
'asher *tishleḥe-ni* *'el-Yehuda*
SUB send/IMPF/2MS to-Judea
'if the King deems this good and if your humble servant is pleasing to you would you please send me to Judea'
(Nehemiah, 2.5)

In sum, the purpose-subjunctive analogical pathway transfers REL-clause morphology to V-complement along the following semantic continuum:

(a) Purpose clauses
(b) Subjunctive complements of manipulative, desiderative or facilitative verbs
(c) Indicative complements of the same verbs

2.3.5. The because-factive blend

The third likely analogical pathway by which REL-clause morphology could have been transferred to V-complements involves the well-known pragmatic properties of some *factive/presuppositional* predicates, whereby the state or event coded in their propositional complement is also the *cause* of the state/event coded by the main clause. Predicates in this group are typically:

> 'be happy', 'be sad', 'be good', 'be bad' 'resent', 'appreciate', 'like', 'regret', 'be fortunate', 'be unfortunate', 'be nice'

Given the well-known *factivity* ('presuppositionality') of because-ADV-clauses, the interpretation of these factive predicates, when followed by an indicative complement, is semantically very similar to their interpretation when followed by a because-ADV-clause. To illustrate this schematically, consider:

(48) a. He's happy *that* she's here (indicative)
b. He's happy *because* she's here (because-ADV)
c. It's nice *that* she's here (indicative)
d. It's nice *because* she's here (because-ADV)
e. It's terrible *that* you hurt yourself (indicative)
f. It's terrible *because* you hurt yourself (because-ADV)
g. She resented *that* they ignored her (indicative)
h. She resented it *because* they ignored her (because-ADV)

Now, earlier above we noted the slow replacement of the because-ADV-clause subordinator *ki* of EBH with the REL-subordinator *'asher/she-*. That replacement is far from complete by the end-point of the BH continuum: Only 28% of because-clauses in *Song of Songs* are marked by *she-*. (The shift is 100% complete in the Mishnaic dialect). But be the frequencies as they may, one begins to find, all across the LBH continuum, blends whereby *she-* — or its predecessor *ki* — is used with factive complements of the type described in (48) above. As illustrations, consider:

(49) a. *ve-sane'ai 'ani 'et-kol "amal-i she-'ani "amel*
and-hate/PERF/1S I ACC-all work-my REL-I doing/MS
***she-**'aniḥe-no le-'adam she-yiḥye 'axar-ay*
REL-leave/IMPF/1S-it to-man REL-live/IMPF/3MS after-me
'and I hated all the toil that I toiled **that/because** I will leave (it) to some man who would come after me...' (Eccles., 2.18)

b. *mah dod-ex mi-dod* ***she****-kaxa hishba"t-anu*
what lover-your from-lover SUB-thus swear/PERF/2FS-us
'what is so special about your lover **that/because** you have thusly sworn us (to secrecy)?' (Song of Songs, 5.11)

c. *kol'oyv-ay... sasu* ***ki*** *'atah "asita*
all enemies-my rejoice/PERF/3MP SUB you do/PERF/2MS
'all my enemies...rejoiced **that/because** you did...
(Lament., 1.21)

d. *shma" 'elohey-nu* ***ki*** *hayinu buza*
hear/IMPER god-our SUB be/PERF/1P shame
'hear oh God **that/for** we have become despicable'
(Nehemiah, 3.36)

e. *va-yismexu* ***ki*** *ha-'elohim simx-am*
and-rejoice/IMPF/3MP SUB the-God gladden/PERF/3MS-them
'and they were happy **that/because** God has gladdened them'
(Nehemiah, 12.43)

One interesting version of this blend occurs in Ecclesiastes, spanning two adjacent clauses, one clearly a V-complement subordinated by *she-*, the other interpretable either as V-complement or a because-ADV-clause and subordinated by *ki-*:

(50) *yada"ti* ***she****-gam zeh hu' ra"yon ruaḥ,*
know/PERF/1S SUB-also this it notion/of wind
ki *be-rov ḥaxma rov ka"as*
SUB in-increase/of knowledge increase/of anger
'I knew that this too was total nonsense, **that/because** with more knowledge comes more aggravation' (Eccles., 1.18)

The because-factive analogical pathway may be rendered schematically as follows:

(a) The REL-subordinator *she-* slowly invades because-ADV-clause paradigm
(b) Factive-cause presuppositional predicates thus display *she*-subordinated complements that are syntactic blends. That is, they can be interpreted as either indicative ('that') or cause-ADV ('because')
(c) The new morphology is eventually transmitted to the complements of other factive verbs in the cognition-utterance paradigm.

6. DISCUSSION

6.1. Multiple causation in grammatical evolution

Much additional distributional and time-course data need to be gathered and analyzed in order to fully understand the complex interaction between the many variables that contributed, directly or indirectly, to the grammatical evolution of dependent clauses in Biblical Hebrew. The results of this study, suggestive as they may be, are thus tentative and preliminary. Among the things the study does demonstrate, perhaps even vividly, is how our old reductionist habits, of seeking single causes to complex linguistic events, single explanations to linguistic phenomena, or — in diachrony — single pathways for linguistic change, are inadequate and often misleading. The data surveyed here suggest that, contrary to old reductionist propensities, diachronic grammatical change is often complex, multi-causal and interactive.

6.2. The scope of analogy

It is often tempting, when one sees only the beginning and end points of a protracted, complex diachronic process, to jump to conclusions about the proper context within which analogy may exercise its option as the main venue for stretching extant linguistic structures to perform novel functions. The second general point suggested by this study is how one must resist the facile invocation of analogy over global spans. The analogical extension of code-function relations in language is most likely a down-to-earth, highly concrete process. What at first glance may seem to be a global, abstract analogy, in this case along the so-called *X-bar* lines, often turns upon more detailed examination to involve much smaller, more concrete — and above all psychologically more plausible — intermediate steps. The three 'blends' invoked here are precisely such phenomena — down-to-earth media through which analogical intuition flows relatively unimpeded — given the impetus of functional need.

6.4. The because-factive analogical pathway: Second time around?

The degree to which the so-called because-factive blend indeed served as analogical pathway in the transfer of REL-clause morphology to the V-complement paradigm in EBH remains to be more fully assessed. What seems to reinforce the case for this blend is the real possibility that the very same

analogical transfer of ADV-clause morphology to V-complements may have already occurred once before in pre-Biblical Hebrew. The subordinator *ki* in EBH is both the exclusive marker of because-ADV-clauses and the exclusive marker of factive verb-complements. Further, *ki* must have already had a protracted history by the Genesis stage of EBH. On comparative grounds alone, one must assume that *ki* started as a *locative* preposition, a usage that still survives in the Amharic bleached locative *k-*.[29] In EBH one still finds *ki* as a noun-marking preposition, in the manner-comparative case-role as well as in manner-ADV-clauses. Thus consider:

(51) a. Manner nominal case:
na"aseh *'adam* ***ki-dmutey-nu***
make/IMPF/1MP man **like**-image-our
'let's make a man in our image' (Genesis, 1.26)

b. Manner-ADV-clause:
va-ya"as *Noaḥ* ***k-xol*** *'asher tsivah*
and-do/IMPF/3MS Noah **as**-all REL order/PERF/3MS
'and Noah did exactly *as* (God) had ordered (him)'
(Genesis, 6.22)

Next, the use of *ki-* to mark *time-ADV-clauses* is well entrenched by the Genesis point of EBH, both with the older, nominal-infinitive pattern and with the finite one. Thus consider:

(52) a. Time-ADV, nominal pattern:
va-yhi ***k-hotsi'-am*** *'otam ha-xuts-a*
and-be **when**-let-out/INF-their them the-out-DIR
'and lo, **when** he let them out' (Genesis, 19.17)

b. Time-ADV, finite pattern:
va-yhi ***ki-heḥel*** *ha-'adam la-rov*
and-be **when**-begin/PERF/3MS the-man to-multiply
'and then **when** people started multiplying' (Genesis, 6.1)

On more general grounds, the well-entrenched temporal use of *ki* tends to further reinforce the probability of its locative origin (cf. Traugott, 1975). The case is further strengthened by the existence of a parallel time-ADV pattern, i.e. the preposition *b-*. This preposition is still primarily a locative marker in EBH ('in', 'at', 'on', 'with'). Thus consider:

(53) a. ***b***-*yom* *"asot* *YHWH 'elohim 'erets ve-shamayim*
LOC-day/of do/INF/of YHWH God land and-sky
'**on** the day when God made the earth and the sky'
(Genesis, 2.4)

b. *k-a-davar* *ha-ze* *tdabrun* *'el-"eysav*
like-the-word the-this talk/IMPF/2MP to-Esau
be-*mots'a-xem* *'ot-o*
LOC-find/INF-your ACC-him
'in this manner talk to Esau *when* you find him'
(Genesis, 32.20)

Next, the use of *ki* to mark *conditional-ADV-clauses* is still attested in EBH, although it is fast receding and in competition with the newer 'if' particle *'im*. Such irrealis-conditional ADV-clauses tend to, universally, shade semantically into the irrealis- 'when' ADV-clauses, as in:

(54) ***ki*** *ta"avod* *'et-ha-'adamah, lo'* *tosef*
if/when work/IMPF/2MS ACC-the-land NEG continue/IMF/3FS
la-tet koḥ-a *le-xa*
to-give/INF vigor-her to-you
'**if/when** you cultivate the land, it will not continue to give its vigor to you' (Genesis, 4.12)

One finds suggestive traces of the conditional use of *ki* in *concessive*-marking combinations, as in:

(55) *'af-**ki*** *'amar* *'elohim lo'* *to'xlu*
also-**if** say/PERF/3MS God NEG eat/IMPF/3MP
'**even though** God said: 'You shall not eat', (Genesis, 3.1)

An analogous combination, where the new 'if' particle *'im* is used to mark a *temporal*-ADV-clause, again illustrates the semantic shading between 'if' and irrealis-'when' in EBH. Thus consider:

(56) *lo'* *'oxal* *'ad '**im** dibarti* *dvar-ay*
NEG eat/IMPF/1S till **if** speak/PERF/1S words-my
'I shall not eat until I've spoken my piece' (Genesis, 24.33)

The shading from irrealis-'when' to 'if', i.e. from *temporal* to *conditional*, is again a widely accepted universal semantic process (see e.g. Traugott, 1985; 1986; Traugott and Wiegand, 1981; Givón, 1979b, ch. 8; *inter alia*). One may

thus suggest that we had in the history of Biblical Hebrew *two* separate instances — in pre-Biblical Hebrew with *ki*, and later across the BH continuum with *'asher/she* — of the very same diachronic-semantic extension, along a considerable portion of the following universal continuum:

locative > temporal > conditional > causative > factive

A continuum of roughly the same sort has of course been suggested elsewhere (see Heine *et al*, This volume).

6.3. Contact as an alternative explanation

It has always been tempting in historical linguistics, for reasons that are far from obvious, to cast around for contact explanations of grammatical change. The reasoning is usually this:

(a) Other languages have a similar phenomenon;
(b) There may have been some contact;
(c) Therefore the phenomenon must have been borrowed.

While contact-induced diffusion of grammatical traits is a well-known phenomenon, its scope is rigidly circumscribed.[30] My own inclination has always been to defer, at least initially, to two general methodological injunctions:

(a) Explain externally, i.e. by contact, only what has no reasonable internal explanation; and
(b) Explain by contact only changes that are counter-normative, i.e. go against the more common diachronic drift.

With this in mind, two hypotheses have been suggested to account for the two main changes investigated here. The first involves the possibility that the form *she-* was borrowed from the Akkadian REL-subordinator *sha-*. There are only two instances of *sha-* in EBH, both in a single archaic passage, the Song of Deborah:

(57) *'ad* ***sha****-qamti Dvorah,* ***sha****-qamti 'em be-Yisra'el*
till REL-rise/PERF/1S Deborah REL-rise/PERF/1S mother in-Israel
'Until I rose, Deborah, I rose, an Israelite mother' (Judges, 5.7)

There are several reasons why the passage in (57) is rather thin evidence for the Akkadian *sha-* being the precursor of the LBH *she-*:

(a) As we have seen above, there is no other instance of the clitic form of *she-* in EBH. It appears much later in LBH.

(b) The Song of Deborah text is archaic in form; it does not fit in the colloquial mold of the bulk of the EBH text; its exact linguistic lineage is yet to be identified.

(c) The contraction of *'asher* into *she-* is an extremely natural, automatic consequence of rapid speech delivery, whereby in grammaticalizing a bi-syllabic morpheme, only the stressed — in this case the second — syllable is retained and simplified. To invoke contact to explain such a garden-variety natural internal change is, to say the least, training a large gun on a rather small fly.

(d) The process of contraction into *she-* may have been, for all we know, already completed at the spoken level by EBH times. The fact that it has no printed attestations except in this one aberrant passage may be simply due to the conservatism of the written genre. And the unpredictable form *sha-* may simply represent a double scribal slip — colloquialism and mis-transmission of the vowel values.

(e) Finally, Rosen (1959) notes that in closely related Semitic dialects, Phoenician and Canaanite, a similar REL-subordinator is found, attested in the consonantal writing most often as *'sh* but occasionally as *'shr*.

Let us turn now to the second contact hypothesis. In Biblical Aramaic, one finds the REL-subordinator *di-* — an erstwhile demonstrative — also used in V-complements. Since Aramaic has largely taken over as the colloquial medium by LBH times, it may be argued that the transfer of *she-* to V-complements was prompted — or at the very least accelerated — by contact with Aramaic. Since three specific internal pathways have been identified, and since the transfer of REL-clause morphology to V-complements is highly universal — as in German, English, Spanish, French, etc. — the invocation of contact to explain the change is again somewhat unmotivated. Although of course contact may accelerate the rate of an internally-motivated change.

Further, Rosen (1959) also notes that there are in fact a small number of examples in Biblical Hebrew where the Aramaic pattern — using the feminine-singular demonstrative *zu* 'this' — is attested as REL-clause subordinator. Thus consider:

(58) a. *"am* ***zu*** *ga'alta*
people DEM save/PERF/2MS
'the people that you've saved'
'this people that you've saved' (Exodus, 15.13)

b. be-reshet ***zu*** *ṭamnu* *nilkda* *ragl-am*
in-net DEM hide/PERF/3MP caught/PERF/3FS foot-their
'their foot was snared in the net they had hidden'
(Psalms, 9.16)

c. *bi-mzimot-ay* ***zu*** *ḥashavti*
in-plots-my DEM plan/PERF/1S
'in the plots that I have devised' (Psalms, 10.2)

d. *ve-'ore-xa* *be-derex* ***zu*** *telex*
and-show/IMPF/1S-you in-path DEM walk/IMPF/2MS
'and I will instruct you in the path that you shall follow'
(Psalms, 32.8)

There are a number of reasons why this pattern does not lend itself well to being an Aramaic source of the LBH *she-*:

(a) The pattern's existence in EBH (Exodus), or in the extremely conservative Psalms language (which harkens back to LBH in all other respects), makes it unlikely that it was borrowed from Aramaic, which was not the dominant spoken language during EBH times.

(b) The pattern is highly natural, and probably represents the same process of grammaticalization — from paratactic to condensed-syntactic — by which German and English demonstratives became REL-clause markers. It involves the fronting — by Y-movement — of stressed demonstratives in a paratactic configuration, with later syntacticization. Schematically this process may be described as:[31]

(59) a. Paratactic Y-movement pattern:
The woman, *that one* I saw, left

b. Condensed syntactic pattern:
The woman *that* I saw left

Rather than being an Aramaic borrowing, the *zu*-marked relativization pattern may simply be an aborted natural development in EBH syntax.

6.5. The Biblical Hebrew dialect continuum

The most sensitive ramifications of this study, and the ones I am most reluctant to close the book on, concern the status of the Biblical Hebrew

dialect continuum. We deal here with traditions and texts that may have been first transmitted orally, later transcribed, surely also edited, and later on codified. One may assume relative stability of transmission only from the point of codification, scribal slips aside. But the codified text seldom reflects the colloquial — or even the written — linguistic usage prevailing at the time of codification. Nor does the written language necessarily reflect the spoken. A few examples illustrate this vividly:

(a) Our tracking of the contraction from *'asher* to *she-* suggests that it is *under-represented* in the EBH texts.
(b) The grammar of many later-codified books, such as Psalms, Job, the Chronicles and others, reflects — in all variables studied here — the conservative EBH dialect of Genesis. These late books were most likely written by native speakers of Aramaic whose scholarly Hebrew style harkened back to the archaic EBH model.
(c) The Dead Sea scrolls reflect the same pattern, written during the late period, when Aramaic had taken over as the spoken medium, and for that reason conforming in their grammar to the Genesis dialect, rather than the end-point of either Song-of-Songs or Mishnaic Hebrew.
(d) Some texts clearly reflect more colloquial usage. Song of Songs is probably the best example of this.
(e) Some texts are stylistically mixed. In Lamentations, for example, relatively late features are intermixed with formulaic incantations and clear archaeisms.

For all these reasons, it was considered most prudent to order the BH dialect-continuum points by purely linguistic means.

Given all the reasons why *any* relative chronology should be taken with a certain amount of caution, the results of this study (and of its predecessor, Givón, 1977) nonetheless confirm the general validity of the approach pursued here. The BH dialect continuum as ordered here behaves like a live natural diachronic continuum. A great number of phonological, morphological and syntactic changes progress through this continuum, with the points largely retaining their relative ordering from one diachronic process to the next. The most common discrepancy among various changes is in their *rate* — some progress faster than others. But Ecclesiastes is the 50% point in almost all of the processes tracked, *Song of Songs* is the 100% point, and Mishnaic Hebrew is essentially identical to the *Song of Songs* dialect. Whether our relative

ordering of books along a diachronic continuum may contribute to Biblical scholarship in its wider perspective remains to be seen.

6.6. Syntactic change and secondary grammaticalization

The rise of great many morpho-syntactic patterns can only be understood as a process of *secondary grammaticalization*. For example, *past tense* morphemes seldom arise directly, but rather as reanalysis of either the *perfect* or *perfective* aspects.[32] Similarly, *present* tense usually arises through the reanalysis of the *durative*, or even the *perfect* aspect.[33] And the *future* quite often arises as a reanalysis of *irrealis*. *Nominative* case-markers seldom arises directly, but most often as reanalysis of either the *genitive* (Kurylowicz, 1964), or the *ergative* (Givón, 1981; Saxena, 1988).

The same phenomenon of secondary grammaticalization is also found in the evolution of syntactic constructions. For example, the *passive* often arises as reanalysis of the *reflexive*. What is suggested in this paper is that existing, earlier-grammaticalized morpho-syntax can give rise, via secondary grammaticalization, to other morpho-syntactic patterns. Some of these secondary pathways may turn out to be highly regular. The pathways most explicitly suggested by the Biblical Hebrew data are:

(60) NOMINAL CLAUSES → RELATIVE CLAUSES → V-COMPLEMENTS
↓
ADV-CLAUSES

ACKNOWLEDGEMENTS

* An early precursor of this paper benefitted from comments and suggestions from Robert Hetzron, Stanislaw Segert and Joe Malone. I am also indebted to Bernd Heine and Elizabeth Traugott for many helpful suggestions and comments.

NOTES

1. Shared morphosyntax between REL-clauses and V-complements is not confined to the subordinator morpheme. Many languages nominalize both types of structures, with the result that both construction types share much of their morpho-syntactic structure. Traces of this may indeed be found in Early Biblical Hebrew.

2. For details on nominalized REL-clauses and the typology of REL-clause formation, see Givón (1990, ch. 15).

3. A discussion of this process in Japanese may be found in Akiba (1978). The transformation (or 'surfacing') of nominalized, subordinate clause structures into current Japanese main clauses also resulted in a predictable concommitant — the reinterpretation of the *genitive* morpheme as *nominative*. A similar transformation may have occurred in early Indo-European, if Kurylowicz (1964) was right about the genitive origin of the old IE nominative. For extensive discussion of such typological changes, see Givón (1979a; 1988).

4. In Givón (1971), where the early precursor of this discussion appeared (see also Givón, 1975), I hypothesized, somewhat tongue-in-cheek, that someone might want to use the *X-bar* convention to "explain by analogy" how VP and NP structures get to resemble each other (in that case in terms of word-order typology). Sure enough, Lightfoot's (1975; 1976) subsequent work indulged liberally in precisely such 'explanation'. Lehmann's (1973) 'structural principle' was in essence founded upon the same notion of abstract analogical explanation, as was Vennemann's (1973) or Hawkins' (1979) 'explanations' of word-order universals.

5. The exact analogue of this can be found in biological evolution. Many favorable genetic mutations, as well as creative extensions of adaptive behavior patterns, tend to cancel out or dilute within the central large genetic pool. Far-out innovations can usually survive — to the point of *speciation* — only in isolated population pockets. For discussion of gapped speciation, see Mayr (1959), as well as Givón (1989, ch. 10).

6. A direct analogical pathway from WH-questions to V-complements exists, at least in principle, in that cognition-utterance verbs such as '(not) know', 'ask', 'tell'/'say', 'forget' etc. can take both indicative and WH-Q complements. Whether any actual case of direct transfer of WH-Q morphology to V-complements can be documented remains to be shown.

7. Some of the related etyma are *"idan* 'epoch', *"od* 'more, further, still/yet', *y-'ad* 'target', *"ada-ying* 'already', *mo-'ed* 'date', 'holiday' **y-"d* 'meeting', ('*gathering place'), **y-"ad* 'target', 'goal', *"ed'* 'witness', 'testimony', 'place-marker' (thus by inference 'place').

8. The transcription used here marks all consonantal distinctions, but follows Israeli pronunciation in ignoring geminate consonants. Vowel values are also rendered by Israeli pronunciation. While this is not, strictly speaking, the most traditionalist transcription, it makes the text more readily available to those conversant in current Hebrew. Since we do not deal with phonology here, loss of information is deemed negligible.

9. The 'imperfect' of Biblical Hebrew is a gross misnomer, since it is a historical collapse of two erstwhile distinct prefixal paradigms: (a) the foreground, perfective, sequential, main narrative aspect; and (b) the irrealis/jussive form. For an extensive discussion of the tense-aspect-modality in Biblical Hebrew, see Givón (1977). Most REL-clauses of the *'asher* pattern display the perfect aspect, for excellent semantic-pragmatic reasons. The 'imperfect' examples here are taken largely from the irrealis paradigm, either future/conditional or habitual.

10. In EBH, the participial is either *lexical-nominal*, *habitual*, or — at least incipiently — also the *durative* aspect. The latter usage has come to full bloom in Modern Hebrew. As the finiteness of this ex-nominal aspect increased, it began to diverge syntactically from prototypical adjectives, eventually adopting the finite-verbal relativization pattern with *'asher/she-*. Thus compare:

Adjectival: *ha-ish ha-gadol*
the-man the-big
'the big man'
Participial: *ha-ish* ***she***-*yoshev kan*
the-man REL-sitting/MS here
'the man who is sitting here'

This change has clear precursors in EBH, although they are numerically insignificant. Thus:

kol ***'asher*** *romes*
all REL crawl/NOM/MS
'all (the creatures) that crawl' (Genesis, 7.8)

11. The nominalizer *m-* in Hebrew is another multi-functional operator that is not restricted to a particular role, but may have arisen in a limited case context and then spread to many other environments. At the moment it is not clear if such a scenario can be reconstructed.

12. For the cross-language typology of verb complements see Givón (1980; 1990, ch. 13).

13. In Amharic, *k-* is either 'from' or a bleached general locative, reinforced semantically by later ex-nominal post-positions.

14. Inflected forms of **hn* survive in BH, as in *hine-ni* 'here I am', *hin-xa* 'here you are', *hin-o* 'here he is', *hin-a* 'here she is', etc. It is presumably cognate to the interjection *ken* 'yes' and the Arabic **kn* 'be', perhaps also to the Ethiopic **nw* (*näbbärä*) 'be'.

15. Perfect-anterior aspects have a strong statistical association with presupposed subordinate clauses, such as REL-clauses, ADV-clauses and factive V-complements. See Givón (1977) for a description of this association in Early Biblical Hebrew, and Givón (1984, ch. 8) for a cross-linguistic perspective.

16. The use of 'say' as a subordinator for utterance verb complements, direct as well as indirect, is widespread cross-linguistically. For a recent survey, see Saksena (1988).

17. On general grounds, most likely the inflected form of 'say' may denote stronger epistemic commitment on the part of the speaker to the information in the complement, while the uninflected form may denote weaker commitment. (In Bemba (Givón 1972) the same morphological contrast is used to code direct vs. indirect quote, respectively.) This is obviously a conjecture at this point.

18. In the first 34 chapters of Genesis, 16% of direct-quote complements were subordinated by either form of 'say'; the rest were *zero*-marked. In the first 20 chapters of Joshua, 32% were marked by 'say', the rest *zero*-marked. These distributions are most likely determined by the frequency of lexical verbs of utterance. The frequency distribution of the competing indirect-quote complementizers *ki* and *ve-hineh* is presumably also determined by the relevant functional considerations.

19. This assumption is in essence a more elaborate recapitulation of the Labovian dictum that today's synchronic variants — distributed in geographic space — are tomorrow's diachronic stages distributed in temporal space. This dictum closely parallels the relation, in biology, between the range of genetic variation within a population, and the range of evolutionary options open to it.

20. The fourth variable, to be illustrated further below, is of course the main variable under study here.
21. The only token of *'asher* in *Song of Songs* occurs in the first line which gives the *title* of the book, and is thus not counted as part of the text.
22. The chapters counted in Givón (1977) are: Genesis (35); Kings II (20); and the entire books of Lamentations, Esther, Ecclesiastes, Song of Songs. The Mishnah counts in this table represent the 9 chapters of *Masexet Braxot* (*Seder Zra'im*). All counts of Mishnaic Hebrew presented elsewhere in this study are of the first 17 chapters of *Zra'im* (*Braxot* and *Pe'a*). Counts from the LBH book of Nehemiah represent the entire book.
23. Sporadic occurrences of *she-* begin to appear in Jonah (1 occurrence as REL-subordinator, none in V-comp), Chronicles (1 occurrence as REL-subordinator, 1 as V-comp), Ezra (1 occurrence as REL-subordinator, none as V-comp).
24. A facsimile of the Kaufmann Mishnah edited by Dr. G. Beer, Prof. of Theology, Heidelberg University (original ms. in the Budapest library; facsimile made in Jerusalem in 1968).
25. See Givón (1984, ch. 4) for a cross-linguistic perspective on this systematic polysemy.
26. See Givón (1980; 1990; ch. 13) for the cross-language typology of verb complements.
27. Here again Ecclesiastes represents the 50% point along the continuum.
28. See Givón (1980; 1990; ch. 13) for the subjunctive-indicative continuum in complements of manipulative and cognition-utterance verbs.
29. See Hetzron (1970) for an exhaustive description of the Amharic case-marking system.
30. See Thomason and Kaufman (1988) for a discussion of the contact transmission of grammatical traits. While the authors hedge at various points in advocating diffusion of grammar by contact, they eventually point out that the strongest cases where such 'diffusion' has been demonstrated involve *substratum* interaction, where pidginization — i.e. language switch and imperfect second language learning — was involved in the early stages. This argument was made earlier in Givón (1973/1978). Recent extensive examples of the dynamics of such *substratum-pidginization* may be seen in Keesing (1988) and Li (1985). Neither the putative contact of BH with Akkadian, nor the shift from Hebrew to Aramaic during the later LBH period represents such a mode of contact, during which one expects a great number of grammatical traits to be 'transmitted'.
31. Theo Vennemann (in personal communication). See also some discussion in Givón (1979b, ch. 5,6) and Givón (1990, ch. 15).
32. The same has also been shown for child language (Antinucci and Miller, 1976).

REFERENCES

Akiba, Katsue. 1978. Topics in Japanese Diachronic Syntax. Ph.D. Dissertation, UCLA.

Antinucci, Francisco and Ruth Miller. 1976. "How children talk about what happened." *Journal of Child Language* 3:167–89.

DeLancey, Scott. 1988. "Relativization and nominalization in Tibetan and Newari." University of Oregon, Eugene, MS.

Givón, T. 1971. "Historical syntax and synchronic morphology: An archaeologist's field trip." *Chicago Linguistic Society* 7:394–415.
Givón, T. 1972. *Studies in ChiBemba and Bantu Grammar. Studies in African Linguistics* Supplement 3.
Givón, T. 1973/1978. "Prolegomena to any sane creology." In *Readings in Creole Studies*, Ian Hancock, Edgar Polomé, Morris Goodman, and Bernd Heine (eds), 3–35. Ghent: E. Story-Scientia.
Givón, T. 1975. "Serial verbs and syntactic change: Niger-Congo." In *Word Order and Word Order Change,* Charles N. Li (ed.), 47–112. Austin: University of Texas Press.
Givón, T. 1977. "The drift from VSO to SVO in biblical Hebrew: The pragmatics of tense-aspect." In *Mechanisms of Syntactic Change*, Charles N. Li (ed.), 181–254. Austin: University of Texas Press.
Givón, T. 1979a. "The drift away from ergativity: Diachronic potentials in Sherpa." *Folia Linguistica Historica* 1:41–60.
Givón, T. 1979b. *On Understanding Grammar*. New York: Academic Press.
Givón, T. 1980. "The binding hierarchy and the typology of complements." *Studies in Language* 4:333–77.
Givón, T. 1981. "On the development of the numeral 'One' as an indefinite marker." *Folia Linguistica Historica* 2:35–53.
Givón, T. 1984. *Syntax: A Functional-Typological Introduction*, Vol. I. Amsterdam: John Benjamins.
Givón, T. 1988. "The evolution of the object and genitive case-marking in Uto-Aztecan: Internal reconstruction in Ute." *Foncción* 3.
Givón, T. 1989. *Mind, Code and Context: Essays in Pragmatics*. Hillsdale, N.J.: Erlbaum Associates.
Givón, T. 1990. *Syntax: A Functional-Typological Introduction*. Vol. II. Amsterdam: John Benjamins.
Hawkins, John. 1979. "Implicational universals as predictors of word order change." *Language* 55:618–48.
Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. Volume I. "From Cognition to Grammar: Evidence from African Languages."
Hetzron, Robert. 1970. "Toward an Amharic case-grammar." *Studies in African Linguistics* 3:301–54.
Keesing, Roger. 1988. *Melanesian Pidgin and the Oceanic Substratum*. Stanford: Stanford University Press.
Kuryłowicz, Jerzy. 1964. *The Inflectional Categories of Indo-European*. Heidelberg: Winter Verlag.
Lehmann, Winfred. 1973. "A structural principle of language and its implications." *Language* 49:47–66.
Li, Charles N. 1985. "Contact-induced semantic change and innovation." In *Historical Semantics, Historical Word Formation*, Jacek Fisiak (ed.), 325–37. Berlin: Mouton de Gruyter.
Lightfoot, David. 1975. "Diachronic syntax: Extraposition and deep-structure reanalysis." In *Proceedings of the Fifth North-Eastern Linguistics Society*, Ellen Kaisse and Jorge Hankamer (eds). Cambridge, Mass.: Harvard University Press.
Lightfoot, David. 1976. "Syntactic change and the autonomy thesis." Paper read at the Symposium on Diachronic Syntax, University of California, Santa Barbara, May 1976.

Mayr, E. 1959. *Evolution and the Diversity of Life*. Cambridge, Mass.: Harvard University Press.
Mazaudon, M. 1978. "La formation des propositions relatives en tibétain." *Bulletin de la Société de Linguistique de Paris* 31.
Mithun, Marianne. 1986. "Evidential diachrony in Northern Iroquoian." In *Evidentiality: the Linguistic Encoding of Epistemology*, Wallace Chafe and Johanna Nichols (eds), 89–112. Norwood, N.J.: Ablex.
Rosen, H. 1959. "Zur Vorgeschichten des Relativsatzes im Nordwestsemitischen." *Archiv Orientálni* 27.
Saxena, Anju. 1988. "The grammaticalization of the verb 'Say' as a clausal subordinator." University of Oregon, Eugene, MS.
Thomason, Sarah and Terence Kaufman. 1988. *Language Contact, Creolization and General Linguistics*. Berkeley: University of California Press.
Traugott, Elizabeth C. 1975. "Spatial expressions of tense and temporal sequencing: A contribution to the study of semantic fields." *Semiotica* 15:207–30.
Traugott, Elizabeth C. 1985. "Conditional markers." In *Iconicity in Syntax*, John Haiman (ed.), 289–307. Amsterdam: John Benjamins.
Traugott, Elizabeth C. 1986. "On the origins of 'And' and 'But' connectives in English." *Studies in Language* 10:137–50.
Traugott, Elizabeth C. and Nancy Wiegand. 1981. "On the renewal of discourse markers: The case of causal connectives." Paper read at the Philological Association of the Pacific Coast, Stanford University, November 1981.
Vennemann, Theo. 1973. "Explanation in syntax." In *Syntax and Semantics* Vol. II, John Kimball (ed.), 1–50. New York: Academic Press.
Volrath, P. 1985. "Relative clauses in Hewa." Summer Institute of Linguistics, Papua New Guinea, MS.

SECTION D: MODALITY

A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English

Sandra A. Thompson and Anthony Mulac
University of California, Santa Barbara

Everybody has to believe something. I believe I'll have another beer.
—Bumper sticker in Santa Barbara, CA

1. INTRODUCTION AND BACKGROUND[1]

The focus of this paper is the interplay between the phenomenon known in English grammar as '*that*-deletion', illustrated in (1) and (2), and the grammaticization of 'epistemic phrases' (EPs), illustrated in (2) and (3):

(1) I think *that* we're definitely moving towards being more technological. (10.16)[2]

(2) I think *0* exercise is really beneficial, to anybody. (1.31)

(3) It's just your point of view you know what you like to do in your spare time *I think*. (41.77)

Examples such as (1)–(3) have generally been thought of in terms of a process of '*that*-deletion', i.e. an alternation between constructions like (1) and (2), with and without *that*. The purpose of this paper, however, is to suggest that what has been thought of as '*that*-deletion' is better understood as an alternation between constructions like (1), in which *I* and *think* are main subject and verb, with *that* introducing a complement clause, and constructions like (2) and (3) in which *I think* is an epistemic phrase, expressing the degree of speaker commitment (Palmer, 1986:51, Traugott, 1987), functioning roughly as an epistemic adverb such as *maybe* with respect to the clause it is associated with.

Further, we will argue that this alternation can be viewed as a clear example of grammaticization. That is, unlike much work of the 1960's and 70's on the syntax of English, we will not take cases such as (1) and (2) to illustrate an 'optional' process of deleting *that*, but will take cases such as (2) and (3) as grammaticized versions of cases such as (1).

To support our point we will consider the role of discourse frequency in the grammaticization process. Specifically, we will show that the discourse frequency of patterns like (2) is directly related to the frequency of patterns like (3). That is, analysis of ordinary spoken conversation shows that forms such as (1) alternate with forms such as (2) and (3) in a subtle and statistically significant way.

We suggest that an appreciation of this relationship provides a deeper understanding of the process of grammaticization and its reliance on recurrent discourse patterns.

The relationship between frequency of tokens in text and the emergence of grammar has been pointed out in several places in the last decade or so, starting with Givón (1979, 1984), and including the important contributions of DuBois (1985, 1987), Durie (1988), Faltz (1987), and Hopper (1987, 1988, to appear). We see our study as a contribution to this body of writing, and as confirmation of the principle that 'recurrent patterns in discourse tokens exert pressure on linguistic types' (DuBois, 1985:359–360).

McDavid (1964), Elsness (1984), and Underhill (1988) have drawn attention to the sorts of factors that seem to be involved in the use vs. non-use of *that* in written English.[3] Our purpose in this paper, however, is to show that there are certain facts that can only be discovered if we look at conversational language; in particular, it is only with conversational data that we can find clues to the grammaticization of EPs.

Emonds (1969, 1976), Hooper (1975), and Hooper and Thompson (1973) were among the first to relate the use of *that* to the relationship between the 'main' verb and the 'complement'. Emonds refers to constructions such as (3) in terms of a rule which he called 'Complement Preposing', where the erstwhile 'complement' becomes the 'main clause'.

Underhill (1988), using articles from major newspapers, pursues this line of reasoning, making a strong case in favor of the following hypothesis (also suggested in Noonan 1985:86):

> The deletion of *that* causes the lower clause to cease behaving like a self-contained embedded clause.

Underhill's study shows the following:

a. *That* is deleted when the subject of the complement clause is the topic of the discourse; *that* is retained when the subject of the main clause is the topic.
b. *That* is deleted when the writer makes or endorses the assertion of the complement clause; *that* is retained when the writer does not necessarily endorse the assertion, but attributes it to someone else.

Underhill is thus claiming that these discourse facts can be accounted for in terms of his hypothesis: Where the embedded clause loses much of its embeddedness, its subject, rather than the main clause subject, tends to be the topic of the discourse, and its content, rather than that of the main clause, tends to be what the writer is endorsing.

We would suggest that this line of reasoning can be taken one step further: These epistemic verbs together with their subjects behave very much like unitary epistemic morphemes in other languages, to the point of being 'transportable' to positions other than that which they could occupy if they were only functioning to introduce a complement, as seen in (3) above.

2. METHODOLOGY

The transcripts analyzed for this study come from 116 8-minute recorded conversations between university students. For the purposes of a separate series of studies, each student was randomly paired with two individuals he or she did not know well, creating a mixed-sex and a same-sex dyad. In this way, 58 male/female, 29 male/male, and 29 female/female dyads were formed. They were asked to discuss at least one of several topics provided them, such as "What are the qualities of a good roommate?", and "What do you do to relieve school tension?"

The recordings were transcribed orthographically and verified by trained observers. A computerized word count showed that the nearly 15 1/2 hours of conversation resulted in more than 240,000 words spoken.

We first identified all instances of what we called the 'target' construction: All occurrences of 'main' subject and verb, as in (1) and (2), which could occur with *that*, whether or not the *that* was present.

The first author analyzed the transcripts for instances of the target construction. The second author provided a reliability check by independently coding 20% of the transcripts. The point-by-point percent of agreement between the two of us was 97%. In all, 1287 instances of the target form

were identified and coded for selected parameters, including whether or not a *that* occurred. We also counted and coded separately all epistemic parentheticals, as in (3). Thus, although we are claiming that the grammaticization process goes from examples like (1), with *that*, to examples of EPs like (2) and (3), the presentation of our findings will contrast 'target' constructions like (1) and (2) with EPARs like (3).

Because of their function as pragmatic expressions[4] we did not count occurrences of *you know* or *I mean*, as in:

(4) An' I'll get everything accomplished and this and that and I may even do good. *You know* but it's just all the work that you realize that you have to do before Monday morning at 8 o'clock. (14.375)

(5) B: Food's terrible, huh?
A: Well, it's not actually that bad. It all depends on your, on your taste. *I mean*, how well, what you're used to *I mean* you're used to really fine cuisine at home you're not going to be able to tolerate dorm food. (15.31)

3. TERMINOLOGY

For the remainder of this paper, we will refrain from using the term '*that*-deletion' because it connotes an inappropriate processual metaphor. We will use the terms 'main' and 'complement' clauses only for utterances such as (1), which contains a complementizer *that*, and when we need to informally distinguish between the two subjects or the two verbs in such utterances as (1) or (2), since our central point is that the blurring of the distinction between 'main' and 'complement' clause (as in (2)) is precisely what is involved in creating the conditions which are giving rise to EPs. When we use the terms 'subject' and 'verb', we will be referring to the 'main' clause subject and verb, not the 'complement' clause subject and verb. As mentioned above, we will use the term 'target' to refer to utterances of the types illustrated in both (1) and (2), with or without *that*.

The term 'verb type' will occasionally be used in implicit opposition to 'verb token'.

The term 'epistemic phrase' (EP) will be used to refer to the combinations of subject + verb, either in contexts such as that of (2), i.e. without *that* and before a clause that could be analyzed as its complement, or in contexts such as that of (3), i.e. in some position other than before a clause that could be

considered its complement. Thus, *I think* will be considered a complement-introducing main clause in (1) above, but an EP in (2) and (3). All instances of utterances of the form of (2) will be considered to exemplify EP's; that is, we are not distinguishing between epistemic and non-epistemic uses here, but between combinations of subject + main verb (as in (1)) and epistemic phrases (as in (2) and (3)).

Finally, the term 'epistemic parenthetical' (EPAR) will refer only to cases such as (3), i.e. to the epistemic phrase consisting of a subject and a verb which appears in some position other than before a clause that could be considered its complement.

4. HYPOTHESIS

The central hypothesis of this paper can be stated as follows:

> EPs in English (as in (2)) are grammaticized forms of subjects and verbs introducing complement clauses (as in (1)). The use of EPs as EPARs (as in (3)) is evidence of this grammaticization.

The evidence supporting this hypothesis is of two types:

1. Frequency: There is a direct relationship between the frequencies of subjects and verbs occurring in target clause EPs without *that* and the frequencies of subjects and verbs occurring as EPARs. That is, those subjects and verbs occurring most frequently without *that* are precisely those which occur most frequently as EPARs.

2. Semantics: The meanings of the verbs most frequently used as EPs are those associated with belief as a mode of knowing (see Chafe, 1986:263, 266).

In other words, the evidence suggests that the most frequent subjects and verbs occurring with what syntacticians have considered to be '*that*-less' 'complements', as in (2), have in fact been reanalyzed by speakers as epistemic phrases, which have a degree of freedom not possible for subject-verb combinations; in particular they are 'free' to occur in other positions, just as other epistemic phrases, such as epistemic adverbs, do in English.

It is clear that the term 'grammaticization' is being used here in a sense somewhat extended from that first posited by Meillet (1912:13): *'Attribution du caractère grammatical à un mot jadis autonome'*. That is, whereas Meillet

focussed on grammaticization of bound morphemes from previously free forms, our use of the term is more in line with other research, e.g. Bybee and Pagliuca (1985, 1987), Li and Thompson (1974), Marchese (1979), and many papers in these volumes, which are concerned with what Hopper (p.c.) has called "the co-opting by languages into their grammars of lexical elements." That is, this research broadly involves category shifts from more 'lexical' to more 'grammatical', including such processes as the bleaching and narrowing of free verbs to free markers of mood and free prepositions. The only difference between the case we are examining here and those discussed in the earlier literature is that we are here dealing with the grammaticization of a phrase rather than a single lexical item. In every other respect, we suggest, they are entirely analogous.

According to Langacker's (1977) treatment of reanalysis, the phenomenon we are concerned with can be considered as a type of 'semantic reformulation', specifically 'constructional reformulation' (p. 85), that is, a reanalysis involving a shift in semantic and syntactic categories affecting an entire clause or sentence. In our case, this shift involves the periphrastic verb-subject combination *I think* becoming a single element behaving as a member of the grammatical category of adverb.

This grammaticization process thus represents a kind of category shift, from subject-verb phrase to EP, the latter having distributional properties not possible with the former. As mentioned above, these epistemic phrases can be thought of as similar to epistemic particles in other languages. For example, Wallace Chafe (p.c.) notes that the Seneca particle *iiwi* is always glossed as 'I think' by speakers who are bilingual in Seneca and English. And Robert Kirsner (p.c.) reports that the Afrikaans *glo* 'possibly' appears to have followed precisely the pathway we are positing for parenthetical *I think*: *glo* is etymologically derivable from *glo'k*, itself from the parenthetical *glo ek* (literally) 'believe I'.

As DuBois (1988), however, points out, restricting our understanding of grammaticization to a move from the lexicon into the grammar may cause us to miss the fact that grammaticization always involves the reanalysis of one kind of pattern (in our case, a frequent collocation of subject and complement-introducing verb) as another kind of pattern (in our case, an epistemic phrase with no syntactic complement-introducing properties). As DuBois argues, grammaticization can usefully be regarded as a reanalysis of a pattern in the domain of language use as a pattern in the domain of language structure.

We will return to the issue of what EP's reveal about the nature of

grammaticization after examining the evidence that supports our hypothesis that EPs are grammaticized versions of complement-introducing subject-verb combinations.

5. RESULTS

In his discussion of aspectogenesis in Indo-Aryan, Hook (Volume II) provides a striking demonstration that an increase in a item's text frequency is an important concomitant of its grammaticization. In a similar way, we wish to show that out of a range of possible forms, only a small subset occur with great frequency in the data, and these are the very forms that show up grammaticized as EP's.

5.1 Verbs

Tables 1 and 2 show that *think* and *guess* account for the majority (65%) of the target verb types, and the vast majority (85%) of the EPARs.

Table 1 *The most frequent 'main' verbs out of 44 main verb types*

think	683	(53%)
guess	150	(12%)
Other	454	(35%)
Total	1287	(100%)

Table 2 *The most frequent verbs as Epistemic Parentheticals (EPs)*

think	76	(47%)
guess	62	(38%)
Other	21	(15%)
Total	159	(100%)

We note that there are 44 different target verb types altogether in our data; it is striking that *think* and *guess* account for 65% of all the target verb tokens, while the other 42 verbs are spread over the remaining 35%.

On the other hand, there are only 18 verb types occurring as EPARs; here the same two verbs, *think* and *guess*, account for 85% of the total EPAR

tokens. Furthermore, it is clear that *think* and *guess* are in a class by themselves: No other EPAR verb is used more than 3 times in our data base. This illustrates one of Hopper's grammaticization principles (Volume I), SPECIALIZATION:

> Within a functional domain, at one stage a variety of forms with different semantic nuances may be possible; as grammaticization takes place, this variety of formal choices narrows and the smaller number of forms selected assume more general grammatical meanings.

That is, as the function of a grammaticized element narrows, so does the variety of forms.

Now we observe that *think* and *guess* also occur as target verbs without *that* significantly more often than the rest of the verbs taken together, as shown in Table 3.

Table 3 *Occurrence of* that *with* think *vs.* guess *vs. all Other verbs*

	-that	*+that*	*Total*
think	622 (91%)	61 (9%)	683 (100%)
guess	148 (99%)	2 (1%)	150 (100%)
Other	342 (75%)	112 (25%)	456 (100%)
Chi-square (df = 2, $N = 1287$) = 79.24, $p < .0001$			

Table 3 shows that 91% of all tokens of target *think* occur without *that*, and 99% of all tokens of target *guess* occur without *that*. For all other verbs, only 76% of target tokens occur without *that*.[6]

Comparing Table 1 with Table 3, then, we see that the two most frequent target verbs in the data also occur significantly more often without *that* than the other verbs do.

Thus, as far as verbs are concerned, the frequency conditions for grammaticization are met: Each of the two most frequently occurring target verbs, *think* and *guess*, also occurs without *that* proportionately more frequently than the aggregate of other verbs. And it is just those verbs that occur most frequently as EPARs.[7]

5.2 Subjects

Tables 4 and 5 show the percentages of subjects occurring with target verbs and in EPARs.

Table 4 *The most frequent subjects with 'main' verbs*

I	1065	(83%)
you 2p	61	(5%)
Other	161	(12%)
Total	1287	(100%)

Table 5 *The most frequent subjects in Epistemic Parentheticals (EPs)*

I	154	(95%)
you 2p	6	(4%)
Other	2	(1%)
Total	162	(100%)

The proportions are similar to, but even more striking than, those found for verbs. *I* and second person *you* account for the vast majority (88%) of the target verb subject tokens, and for the overwhelming majority (99%) of the EPAR tokens.[8]

Moreoever, in both cases *I* is more frequent than *you*, which is in turn more frequent than all others.

As with verbs, we can now show that these same two subjects, *I* and *you*, occur more frequently with target verbs without *that* than do other subjects; see Table 6.

Table 6 *Occurrence of* that *with* I *vs.* you *vs. Other subject*

	− that		*+ that*		*Total*	
I	955	(90%)	110	(10%)	1065	(100%)
you 2p	55	(91%)	6	(9%)	61	(100%)
Other	102	(64%)	59	(36%)	161	(100%)
Chi-square ($df = 2$, $N = 1287$) = 83.22, $p < .0001$						

Table 6 shows that *I* and *you* occur in target clauses without *that* more frequently than is the case for other subjects. And, just as with verbs, *I* and *you* are each significantly more likely ($p < .0001$) to occur in target clauses without *that* than is any other subject.

Our findings for verbs have been confirmed with respect to subjects, then, and we see that the frequency conditions for grammaticization are again

met: The two most frequently occurring target subjects, *I* and *you*, also occur without *that* more frequently than do all other target subjects. And it is just those subjects that occur most frequently in EPARs.

Furthermore, these results support our hypothesis in a much stronger way. It is no accident that the most frequently occurring epistemic subjects are *I* and *you*: It has often been noted that markers of evidentiality and epistemicity are skewed towards first person singular declaratives and second person questions.

For example, Woodbury (1986:192), in a discussion of evidentiality in Sherpa, points out that: "The first person vs. nonfirst person distinction is widespread in Sherpa, but the term 'first person' is something of a misnomer. *In the interrogative all so-called first person phenomena are associated with second person.*" (Emphasis added [ST & AM]). In his discussion of subjectivity in language, Benveniste (1971: 228–9) notes that certain verbs do not have the property of "permanence of meaning" when the person of the subject is changed: In particular, verbs of "mental operations", such as *believe*, when used with *I*, "convert into a subjective utterance the fact asserted impersonally."

Similarly, in his discussion of 'parenthetical verbs' in English, Urmson (1963:222) notes: Part of what I design to show is how differently these verbs [i.e. what he calls "parenthetical" verbs, e.g. *think* and *believe* - ST & AM] are used in the first person present and in other persons and tenses. Neither Benveniste nor Urmson considered interrogative *you* with the verbs in question, but our data suggest that it forms a class with *I* in the present tense, and that the two together must be distinguished from all other persons and tenses.

All of the EPARs with *you* are questions, and 55 out of 67 (82%) of the target clause *you*'s are in questions. In the examples below, using intonation as the criterion[9], (6) illustrates an EPAR question, (7) a target clause question, and (8), just for comparison, a target clause *you* which is not a question:

(6) What's the point of that *do you think*? (110.45)

(7) So what *do you think* you're going to major in now that you're down here? (108.133)

(8) (talking about dormitory regulations on alcohol)
A: ... they probably shouldn't have any rules at all but
B: [(laughs) *you think* it should be wide open.]
A: [If they're legally responsible] (95.274)

Finally, just as we saw that verb semantics interacts with frequency for verbs in the grammaticization of EPs, we note that the discourse difference between first and second person reference interacts with frequency for subjects. Table 5 showed that *I* outranks *you* in EPAR's 83% to 5%, but Table 6 shows that *I* and *you* are almost identical in percentage of tokens without *that* (91% vs. 90%). Again, if frequency were all that were involved, this would be an unexpected result. However, the figures make sense if we consider that *I* and *you* are roughly equally likely to be used with verbs introducing complements in other roles besides indicating epistemicity. For example, in 'I told her that we were going swimming', epistemicity is not at issue. These non-epistemic uses are more likely to occur with *that*, which predicts that *I* and *you* should have roughly the same percentage of tokens with *that*, and that is just what we find.

5.3 *I think*

Further support for our claim can be found in the behavior of the phrase *I think*. Since *I* is far and away the most frequent target subject in the data and *think* is far and away the most frequent target verb, and since they both occur 90% of the time without *that*, it is natural to ask whether the target phrase *I think* is significantly more likely to occur without *that* than all other combinations; Table 7 shows that it is.

Table 7 *Occurrence of* that *with* I think *vs. all Other combinations*

	− that	*+ that*	*Total*
I think	599 (92%)	54 (8%)	653 (100%)
Other	513 (81%)	121 (19%)	634 (100%)
Chi-square (df = 1, $N = 1287$) = 32.03, $p < .0001$			

6. ARE EPS AN EXAMPLE OF GRAMMATICIZATION?

EPs provide an intriguing illustration of grammaticization, viewed in the broad sense discussed at the beginning of this paper. Similar to comparable examples discussed in Heine and Reh (1984), the case of EPs illustrates the type of grammaticization in which a governing or head element is reanalyzed as a governed or dependent element.

Nevertheless, it is not a "textbook case" study in grammatization. First, as mentioned earlier, *I think* is a phrase rather than a single lexical item. Second, it is not clear what the grammatical status of EP's is. Our position is that although EP's do not form an obvious lexical category, they can be argued to comprise a grammatical sub-category of adverbs; in this sense they can be viewed as similar to auxiliaries, which also form a grammatical category resulting from the grammaticization of lexical elements.

Another problem with the case we are examining here is that we are describing a process which is largely attested only synchronically; currently available evidence on the history of the relevant constructions does not allow us to say anything definitive about whether the synchronic alternation between 'main clause subject + verb' and EPs has a diachronic parallel (Traugott, p.c.). However, again, what we consider significant about this phenomenon is that it exemplifies the process of something more 'lexical' becoming something more 'grammatical'.

An interesting question is whether the relationship between target subjects and verbs and epistemic phrases might be best described as a process of lexicalization rather than of grammaticization. Evidence in favor of this hypothesis includes the fact that the input to the EP is a syntagm consisting of subject + verb, rather than a single lexical item. Thus an EP such as *I think* might be seen as analogous to a word such as *today* (< OE *to* + *daege*). However, the fact that EP's are still available for ordinary negation and questioning ('It's cute, don't you think?') provides strong evidence against such a hypothesis.[10]

As support for our position that EP's are instances of grammaticization in spite of their non-canonical nature, we note that, as described in the Introduction to this volume, grammaticization is a theory of "the fixed and the less fixed in language, ... of the organization of categories and coding", and that "it foregrounds the tension between relatively unconstrained lexical expression and more constrained morphosyntactic coding." The process of EP-creation in English, where a more lexical phrase becomes used as a distinct category with a more restricted meaning and different privileges of occurrence, is precisely this type of process.

As further support for our hypothesis, we can show that the five principles of grammaticization discussed by Hopper (Volume I), apply to the process of EP formation.

Hopper's first principle is DIVERGENCE, the tendency for the lexical source to change according to predictions based on its behavior as a lexical item. In our case, DIVERGENCE provides an account of the treatment by

speakers of lexical *I think* as an ordinary subject-verb combination; in the progressive, for example the auxiliary *am* appears, as in 'I am thinking that it's time to go'.

Related to this first principle is the second, LAYERING, the coexistence of several layers of a set of related grammatical phenomena such that the older forms live on in the language alongside the newer grammaticized versions. This principle nicely describes the fact that both subject-verb phrases and EPs co-exist in contemporary English.

Hopper's third principle, SPECIALIZATION, refers to the singling out of just a few, or perhaps one, of a cluster of forms for grammaticization. As shown above, EP formation illustrates this principle as well: Out of 44 different verb types found in subject-verb combinations, only 18 show up as EP's.

Hopper's fourth principle, PERSISTENCE, refers to the retention of vestiges of the earlier lexical meaning in the grammaticized form. That this principle is at work in the formation of EPs can be seen clearly in the differential meanings of the EPs *I think* and *I guess*: *I think* is a stronger assertion of belief than *I guess*. This is traceable to the difference between *think* and *guess* as verbs: *guess* implies an assertion based on little or no evidence, and hence less commitment to a proposition than *think* does, as seen in such locutions as "*Guess* how many people came to the open house."

The fifth principle articulated by Hopper is DE-CATEGORIALIZATION, which states that grammaticization tends to involve shifts away from the major lexical categories of Noun and Verb towards secondary categories such as Participle, Preposition, etc. As discussed just above, decategorialization is very much in evidence in the evolution of EPs: The lexical category of the erstwhile combination of canonical Noun + Verb can best be characterized as the secondary category 'Adverb'.

According to Hopper's five principles, EPs, then, appear to be straightforward instances of grammaticization, suggesting that the range of processes legitimately covered by this term may be somewhat broader than has been recognized.

Once again, however, we want to stress the point made in DuBois (1988) that grammaticization involves not just the reanalysis of lexical material as grammatical material, but also the reanalysis of a discourse pattern as a structural pattern. We have just such a reanalysis in the creation of a class of EP's clearly modelled on the most frequent collocations of epistemic target verbs and their subjects.

7. CONCLUSIONS

We have examined the evidence in favor of our claim that EPs in English are grammaticized forms of subjects and verbs introducing complement clauses. As evidence for this claim, we showed that the most frequent target clause subjects and verbs are just those which are most frequently found as EPAR expressions, and we suggested that the frequent occurrence of *I* (in declaratives) and *you* (in questions) without *that* in target constructions has led to their re-interpretation as epistemic phrases with verbs expressing belief, such as *think* and *guess* (cf. Chafe, 1986). As epistemic phrases, then, these combinations are free to float to various positions in the clause to which they are providing testimony, as other epistemic particles in English do, such as *maybe*. Again, we found strong correlations between the frequency of the forms found in these EPAR's and the frequency of those found in target epistemic phrases without *that*.

In this paper, we have considered grammaticization from the point of view of the behavior of EPs, a non-prototypical instance of grammaticization. Yet the process described has all the properties of less controversial instances of grammaticization.

With this study, then, we hope to have extended the understanding of the term 'grammaticization' to include phenomena such as these, and to have underscored the intimate relationship between discourse frequency patterns and the emergence of grammar.

NOTES

1. We would like to thank the following people for helpful comments on earlier versions of this paper: Edith Bavin, Dwight Bolinger, Joan Bybee, Wallace Chafe, John DuBois, Suzanne Fleischman, Julie Gerhardt, David Hargreaves, Bernd Heine, Paul Hopper, C. Douglas Johnson, Tsuyoshi Ono, William Pagliuca, Gillian Ramchand, Gisela Redeker, Nathan Salmon, Arthur Schwartz, Dan Slobin, Timothy Shopen, and R. McMillan Thompson. We are particularly grateful to Elizabeth Traugott for her careful and thoughtful comments, which have contributed greatly to the clarity and substance of this paper. None of these people necessarily agrees with the interpretation we've given to their input.

2. The numbers in parentheses indicate dyad number and line number in our data base.

3. Pilot studies confirm that some of these factors are at work in spoken English as well. Space limitations prevent us from presenting a number of interesting findings not directly related to the issue of grammaticization; for more discussion, see Thompson and Mulac (to appear).

4. For discussions of such discourse particles in conversation, see Goldberg, 1980; Ostman, 1981; Schourup, 1985; Erman, 1986; and Schiffrin, 1987.
5. Gramley (n.d.) relates this freedom to an iconic principle according to which verbs of 'propositional attitude' (= our epistemic verbs) are 'semantically' the most distant from their complements. Gramley does not relate this point to Haiman's (1983, 1985) discussions of iconicity, but it is clearly an instance of Haiman's general observations on iconicity and language.
6. Not only is the three-way comparison significant at the level of $p < .001$, but chi-square pair-wise comparisons also show that *think* and *guess* are each significantly more likely (at the level of $p < .001$) to occur without *that* than is the complement set of all other verbs.
7. Dwight Bolinger and Gillian Sankoff have pointed out to us that the nature of our data base, conversations in which speakers were expressing opinions, may have influenced the proportions of verbs used. One might have expected more occurrences of *bet*, *figure*, or *suppose* in less restricted conversation, for example. We acknowledge this possibility, but note that our hypothesis is strongly confirmed by the frequency correlations, independent of verb types, described in this section.
8. Our coding distinguishes second person *you* from generic *you*. From here on, whenever *you* is mentioned, we mean second person *you*; generic *you* is included in the 'other' category.
9. We follow the standard transcription practice of indicating rising intonation with a question mark and falling intonation with a period.
10. Intriguingly, though, Keesing (1988:101) notes that Tok Pisin, Bislama, and Solomons Pidgin all have an epistemic monomorphemic adverb [atiŋ] 'maybe' from English *I think*.

REFERENCES

Benveniste, Emile. 1971. *Problems in General Linguistics*. Coral Gables, FL: University of Miami Press. Translated by M. E. Meek.

Bybee, Joan L. and William Pagliuca. 1985. "Cross-linguistic comparison and the development of grammatical meaning." In *Historical Semantics*, Jacek Fisiak (ed.), 59–83. Berlin: Mouton de Gruyter.

Bybee, Joan L. and William Pagliuca 1987. "The evolution of future meaning." In *Papers from the 7th Conference on Historical Linguistics,* Anna G. Ramat, Onofrio Carruba, and Giuliana Bernini (eds), 109–122. Amsterdam: John Benjamins.

Chafe, Wallace, 1986. "Evidentiality in English conversation and academic writing." In *Evidentiality: the Linguistic Coding of Epistemology,* Wallace Chafe and Johanna Nichols (eds), 261–272. Norwood, NJ: Ablex. *Advances in Discourse Processes 20.*

DuBois, John. 1985. "Competing motivations." In *Iconicity in Syntax,* John Haiman (ed.), 343–365. Amsterdam: John Benjamins.

Dubois, John. 1987. "The discourse basis of ergativity." *Language* 63(4):805–855.

Dubois, John. 1988 "Discourse as pattern model for Grammar. The possessor = ergator affiliation." Paper presented at the Symposium on Grammaticalization, Eugene, Oregon, May 1988.

Durie, Mark. 1988. "Preferred argument structure in an active language." *Lingua* 74:1–25.
Elsness, Johan. 1984. "*That* or zero? A look at the choice of object clause connective in a corpus of American English." *English Studies* 65: 519–533.
Emonds, Joseph. 1973. "Parenthetical Clauses." In *Proceedings of the Ninth Annual Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago.
Emonds, Joseph. 1969. *Root and Structure-preserving Transformations*. Bloomington: Indiana University Linguistics Club.
Emonds, Joseph. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.
Erman, Britt. 1986. "Some pragmatic expressions in English conversation." In *English in Speech and Writing*, Gunnel Tottie and Ingegerd Backlund (eds). Stockholm: Almqvist and Wiksell International. Studia Anglistica Upsaliensia 60.
Faltz, Leonard M. 1987. "On inference and semantic change." To appear in *Studies in Language*.
Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
Givón, Talmy. 1984. *Syntax*. Vol I. Amsterdam: John Benjamins.
Goldberg, J. 1980. Discourse Particles: An Analysis of the Role of *y'know, I mean, well,* and *actually* in Conversation. Cambridge University: Ph.D. dissertation.
Gramley, Stephan. n.d. "The infinitive forms of English as verb complements." ms. University of Bielefeld.
Haiman, John. 1983. "Iconic and economic motivation." *Language* 59:781–819.
Haiman, John. 1985. *Natural Syntax*. Cambridge: Cambridge University Press.
Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
Hook, Peter E. Volume II. "The emergence of perfective aspect in Indo-Aryan languages."
Hooper, Joan B. 1975. "On assertive predicates." In *Syntax and Semantics* 4. 91–124. New York: Academic Press.
Hooper, Joan B. and Sandra A. Thompson. 1973. "On the applicability of root transformations." *Linguistic Inquiry* 4:465–497.
Hopper, Paul. 1987. "Emergent grammar." In *Papers from the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 139–157. Berkeley: University of California Linguistics Department.
Hopper, Paul. To appear. "Principles of grammaticization: Toward a diachronic typology." In *Proceedings of the IREX Conference on Comparative Linguistics*.
Hopper, Paul. Volume I. "On some principles of grammaticization".
Keesing, Roger. 1988. *Melanesian Pidgin and the Oceanic Substrate*. Stanford: Stanford University Press.
Langacker, Ronald W. 1977. "Syntactic reanalysis." In *Mechanisms of Syntactic Change*, Charles N. Li. (ed.), 57–139. Austin: University of Texas Press.
Li, Charles N. and Sandra A. Thompson. 1974. "Co-verbs in Mandarin Chinese: Verbs or prepositions?" *Journal of Chinese Linguistics* 2(3):257–278.
McDavid, Virginia. 1964. "The alternation of 'that' and zero in noun clauses." *American Speech* 39: 102–113.
Marchese, Lynell. 1979. Tense-aspect and the Development of Auxiliaries in Kru Dialects. UCLA: Ph.D. dissertation.
Meillet, A. 1912. "L'évolution des formes grammaticales." *Linguistique historique et linguistique générale*, Tome I, 130–148. Paris: Klincksieck.

Noonan, Michael. 1985. "Complementation." In *Language Typology and Syntactic Description*, Vol. II (Complex Constructions), Timothy Shopen, (ed.), 42–140. Cambridge: Cambridge University Press.

Ostman, J.-O. 1981. *You Know: A Discourse-Functional Approach*. Amsterdam: John Benjamins. *Pragmatics and Beyond* 2.7.

Palmer, F.R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schourup, Lawrence C. 1985. *Common Discourse Particles in English Conversation*. New York: Garland Publishing.

Thompson, Sandra A. and Anthony Mulac. To appear. "The discourse conditions for the use of complementizer *that* in conversational English."

Traugott, Elizabeth C. 1989. "On the rise of epistemic meanings in English: An example of subjectification in language change." *Language* 65: 31–35.

Underhill, Robert. 1988. "The Discourse Conditions for *That*-deletion." San Diego State University.

Urmson, J.O. 1963. "Parenthetical verbs." In *Philosophy and Ordinary Language*, Charles E. Caton (ed.), 220–240. Urbana: University of Illinois Press.

Woodbury, Anthony C. 1986. "Interactions of tense and evidentiality: A study of Sherpa and English." In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe and Johanna Nichols (eds), 188–202. Norwood, NJ: Ablex. *Advances in Discourse Processes* Vol. 20.

The Grammaticization of the German Modal Particles

Werner Abraham
University of Groningen, The Netherlands

1. INTRODUCTION

This paper investigates "modal particles" (henceforth MPs) in German and their grammatical status. It will be argued that MPs in German are partly the result of a process of grammaticization of lexemes which originally had a different grammatical status. Where they are not, i.e. where MPs can be shown to have always existed they nevertheless share all of the distributional characteristics of the entire category. In this respect, MPs in German form a homogeneous grammatical class of words.

Although there exists an extensive literature in German on MPs (certainly since the seminal book by Weydt published in 1969), it is striking that this is virtually a non-topic in the English-language linguistic tradition, perhaps because English has no grammatical category akin to the German modal particle. The same is true for all Romance languages, as contrastive research has shown. By contrast, there is an abundance of MPs in other Germanic languages such as Dutch, Frisian, and the Mainland Scandinavian languages.

In the ensuing discussion it will be shown that (a) MPs in German are correlative with distinct syntactic, semantic, and pragmatic characteristics, and that (b) these characteristics developed diachronically. It will be argued that MPs in German represent a case of grammaticization. Additional support for this conclusion will be adduced from the fact that the other European languages for which the same abundance of typical MPs is attested share the syntactic, semantic, and pragmatic characteristics which have been observed for German. This adds a certain, albeit restricted, typological status to our findings.

2. PHENOMENA AND QUESTIONS

The discussion and the conclusions in this article are based on the following linguistic phenomena: There is a limited number of lexemes in German each of which is ambiguous with respect to grammatical category. Under one classification, they are adverbs (both as sentential connectors, such as *aber* "but", *schon* "already", and as adjuncts, such as *halt* "to be sure", *wohl* "well"), scalar particles (see, e.g. *nur* "only", *auch* "also"), adjectives (see, e.g. *eigentlich* "real" or *eben* "flat"), or interjections (see *ja* "yes"). However, in another usage, that of particles, their grammatical status is undetermined. Along with this categorial division between the two functions of one single phonetic form, there is a division of meaning between the two homonymic forms. Under non-particle (= full) categorial status, the lexeme has a full referential (=adjectival or adjunct) or explicit syncategorial (= conjunctional, scalar particle, or interjection) meaning; see the illustrations above. Under the vague, categorially undetermined status, however, the meaning of the unit in question is elusive, and can at best be coined "modal" in the broad sense of the term. A "modal particle" use of such a single phonetic form is thus syntactically and semantically distinguished from its non-particle and non-modal use.

A principal claim behind this discussion is that for each pair of distinct functions under one single phonetic form, a relation obtains such that the modal particle function can be derived from its counterpart with full categorial status either synchronically or diachronically (see Abraham, 1989b for a discussion). The present paper will not deal with this issue. However, it is remarkable that two meaning-related linguistic forms exist side by side, where one of these forms (the MP) shows systematic bleaching of the semantic content of the other form (the lexical adjunct, adjective, or scalar particle), and where at the same time novel, syntactically distinct distributions constrain the occurrence of the second form. It is this coincidence of semantic bleaching and new syntactic constraints that suggests crucial affinity with the concept of grammaticization.

Categorially, i.e. with respect to distributional properties, MPs come closest to adverbials as will be shown in some detail presently. What distinguishes MPs from adverbials, however, is the fact that they cannot move as freely in the sentence as do members of the class of adverbials. In contrast to normal adverbials, they cannot be paraphrased by an adverbial sentence. Furthermore, they share with conjunctions the property of obtaining meaning syncategorially, i.e. only in collocation with very specific semantic and/or

syntactic units, as illustrated in Table I (following, with some modifications and additions, Helbig and Kötz (1981: 29)).[1] See Table I on p. 334.

The empty boxes in Table I show that MPs are syncategorial in the sense that they are constrained in co-occurrence possibilities with respect to sentential types and illocutive force. Furthermore, their linear position is restricted to what has been called the "middle field" (henceforth MF) in German and Dutch (roughly, the linear stretch between the "verbal brackets"; for a more detailed definition see Table III as well as (9)–(11) in section 3 below), and they usually cannot carry stress. Also, their scope is the whole proposition rather than the relation between subject and predicate, as is often the case with adverbials. They are thus akin to speaker-oriented units such as epistemics and deontics, yet much more specific than these: They make visible, by way of some specific illocutive force, what the speaker (user of the modal particle) deems adequate to fully disambiguate a certain prior or following text or discourse portion. In a way, MPs can most adequately be described as functioning as fillers of what has been left enthymemic in the text or discourse portion in question. In other words, MPs signal gaps in the argumental configuration of a prior text portion and partly indicate how a missing argument is to be reconstructed by the hearer or reader.

The most striking characteristic of MPs is the fact that there are invariably two forms of the same phonetic shape, viz. one representing the MP, with at most some fluid lexical meaning component, and more often, however, just some equally fluid and apparently subjective illocutive force; and a second form of the category adverb, adjective, scalar particle, or conjunction. Consider the formal pairs in (2) below, which are either polysemous (P) or homonymic (H).[2] See Table II on p. 335.

While the MPs do not provide the lexical meaning that their fully lexical counterparts possess, the illocutive force of the MPs, although at variance sometimes between two or more illocutive types, is nevertheless discrete enough to allow the (Fregean) composition of illocutive force (for attempts to systematize the inferential forms in the semantic description of MPs see Abraham 1989b). See the following examples.

(1) Kommt er *denn*?
Comes he then?
"Will he come after all?" ... DOUBT
(the speaker sees reason to believe that the person referred to will not come)

Table I *(The English equivalents, in the leftmost column, refer to the non-modal lexeme. The values of sentential type and illocutive force characterize the counterpart lexeme as an MP.)*

Sentential type	Declarative		Yes-No question		Word question			Imperative
Speaker's illocutive intention	Assertative force	Emphatic force	Interrogative force	Emphatic force	Interrogative force	Emphatic force	With implied answers	Directive
aber "but"		+		+		+	(rhetor. question)	+
auch "also"	+	−	+	−	−		+	+
bloß "only"	+		+	+	+	+	+	+
denn "because"			+		+		+	
doch "though"	+	+		+	+	+		+
eben "as I said"	+	+						+
eigentlich "to be precise"	+		+		+		+	
etwa "perhaps"			+					
halt "as I said"	+	+						+
ja "yes; as I said"	+	+		+				+
(ein)mal "once"	+							+
nur "only; but"	+		+	+	+	+	+	+
schon "already"	+						+	+
vielleicht "perhaps"			+	+	+			
wohl "well"	+		+		+			

Table II *The column "lex(ical) mean(ing)" (3rd from right) refers to the item outside the MP-use; "meaning relation" (P vs. H) indicates whether or not some polysemous meaning association (core or periphery) is drawn between the two representatives under one identical phonetic form; "sync(ategorial) mean(ing)" (2nd column from the right) classifies those MPs which have* grammatical *morphemes as counterparts (i.e. conjunctions, interjections, scalar adverbs).*

Adverb	Verbal (form)	Adverbial conjunction	Coordinating conjunction	Scalar particle (ad-adjective)	Adjective	(Assertive) interjection	MP	Lex. mean	Sync. mean	Meaning relation: polysemy (P) vs. Homonymy (H)
			aber				aber	−	+	P
		auch		auch			auch	−	+	P
				bloß			bloß	+	−	P
			denn				denn	−	+	H
					eben		eben	+	−	H
		eigentlich					eigentlich	+	−	P
	halt						halt	−	+	H
						ja	ja	−	+	H
(ein)mal							(ein)mal	+	−	P
		nur		nur			nur	−	+	H
		vielleicht					vielleicht	+	−	P
		wohl					wohl	+	−	P
		zudem					zudem	+	−	P

(2) Kommt er *denn wohl*?
Comes he then well?
DOUBT + PERSONAL INVOLVEMENT OF THE SPEAKER

(3) Kommt er *denn wohl auch*?
Comes he then well also?
DOUBT + PERSONAL SPEAKER INVOLVEMENT + WISH OF THE SPEAKER THAT THE PROPOSITIONAL CONCEPT COME TRUE

What we have seen so far suggests that the MPs in German are perhaps exemplary for what has come to be known as grammaticization in the development of languages. To make this claim conclusive we shall have to show, first, that the data of MPs in German comply with the input and output characteristics of some process of grammaticization, and, second, that the constraints (distributional, semantic, etc.) characterizing contemporaneous MPs in fact triggered the grammaticizing process of the MPs. In other words, we shall first check whether MP-lexemes satisfy properties of what are generally described in the literature to be grammaticizing processes and their results, and, second, whether conditioning distributional factors can be found in the history of German leading to the new particle category. Due to the absence of morphophonological properties, in the narrow sense, characterizing the output category of MPs, particular attention will be devoted to other syntactic distributional properties. It will be seen that, among the conditioning factors, the syntactic, topological notion of the "MF" plays a crucial role in the emergence of MPs. It will also be shown that the rise of some of the MP uses coincides with the rise of this sentential MF (somewhat reminiscent of the more traditional term "verbal bracketing") in the period of Early New High German (1450–1650) and that older MPs (i.e. lexemes which existed as MPs in pre-MF periods) immediately complied with the linear restrictions which triggered the bleaching process characteristic of the younger MPs. From what we can judge today, there has always been a closed paradigm with respect to the distributional properties of the category of MPs in German, irrespective of their status as old or young particles.

The question may arise whether the conditioning syntactic factor of the MF is an idiosyncratic development of German. Note that this structural notion has been connected with the traditional term of verbal bracketing; the latter phenomenon, however, has not been observed for any other language but German as far as I know. In addition, the structural property of the MF hinges upon two further structural characteristics, viz. the strict verb-second

property (as realized in the German main clause) and, simultaneously, the verb-last property (realized in the independent as well as the dependent German clause). Both properties have been observed to hold only for the continental West-Germanic languages (German, Dutch, Frisian). This, then, combined with the association with verbal bracketing, may feed the suspicion that the MF is a phenomenon idiosyncratic to German and its closest relatives. We shall see, however, that there must be some deeper, underlying property to the syntactic phenomenon of the MF. The fact that also the Mainland Scandinavian languages, which share with German, Dutch, and Frisian a considerable number of MPs, also have a structure akin to the German MF corroborates this view. We shall hypothesize in a concluding section of this paper that the MF, or the structural equivalent in the Mainland Scandinavian languages, is the prominent place for rhematic information to be carried by the sentence. This finding not only gives the MF condition of the grammaticization of MPs a somewhat weightier typological status, but it also makes it conceivable that the structural place for rhematic material may be the syntactic factor fundamental to the development of MPs thereby allowing a typological analysis of the MP-phenomenon to extend to languages which display a comparable abundance of MPs, but lack a structural MF proper. This, however, will not be pursued in detail in this paper (see for more details Abraham 1988b).[3]

Does all this suggest an instance of grammaticization? The concept of grammaticization is often spoken of in reference to a linguistic unit which, in more or less unmodified form, moves from a lexical function (lexical morpheme), to a grammatical function (grammatical morpheme) thereby shedding (some of its) lexical properties and taking on grammatical properties. According to Meillet (1912:13), Kuryłowicz (1965:52), Heine and Reh (1984:15), and Lehmann (1985; Volume II), the process of grammaticization should involve at least one of the following phenomena:

(a) increase of the morphosyntactic range of a morpheme
(b) shift from lexical to grammatical (or from less grammatical to more grammatical) status (e.g. from a derivative formant to an inflectional one)
(c) loss of semantic complexity
(d) loss of pragmatic significance
(e) loss of syntactic freedom
(f) loss of phonetic substance
(g) increased constrainedness by grammatical rules

At first sight, MPs would appear to instantiate the following characteristics of the list above: (b), (c), and (e). Table II above is indicative of some of the properties of grammatical conversion in the following sense: MPs, although they are not grammaticized down to mere grammatical morphemes, are nevertheless subject to severe constraints of syntactic mobility (more so than full lexicals), or they do not possess as explicit a grammatical or semantic function as conjunctions or scalar particles. The widest generalization thus appears to be the following: The information provided by MPs (either lexical or syntactic-functional) is bleached in comparison to that of their correlative lexemes.

Characteristics that are not met so obviously by MPs are the following:

(d) loss of pragmatic significance: On the contrary, we will demonstrate that, at least for the grammaticization of MPs in German, there is a considerable increase in the pragmatic force of the particle lexemes concerned. Quite obviously, also, this is a characteristic interrelated in many respects with the loss of semantic and syntactic properties in (c) and (e) above.

(f) loss of phonetic substance: which probably holds only for purely grammatical markers; and

(g) increased constraint by grammatical rules (unless this property is taken to imply, with respect to our specific range of phenomena, that the grammatical elements concerned are highly constrained by their co-occurrence with specific sentential types [see Table I above] and that their occurrence is restricted to the structural MF [see (6)–(9) below]).

3. SYNTACTIC CONSTRAINTS FOR MPS

3.1. Structural positions and linearity

The central observation in section 2 was that for every MP there is a lexical homonym in the function of an adverb, adjective, or a scalar particle. In this section the linear restrictions for MPs in German will be demonstrated and, in particular, it will be shown to what extent these restrictions differ from those for the lexical or syncategorial homonyms. MPs will be seen to have linear properties that are clearly different from those of adverbial sentence connectors; what will remain to be looked at is the distinction between

adverbials and MPs. Consequently, our discussion of the syntax of the MPs will be oriented towards the linear constraints holding for adverbials, as in (6). It will be seen that MPs can occur in any of the MF positions reserved structurally for adverb(ial)s except for the MF position directly in front of the verb-last position (which is usually reserved for the direct object or some preopsitional phrase). What distinguishes MPs crucially from adverb(ial)s, however, is that they disqualify for positions to the left or to the right of the MF boundaries (traditionally, the verbal brackets). This leaves us with two tasks. The first is to define as accurately as possible the linear and structural implications of the notion of the MF. The second is to show to what extent adverb(ial)s occur under lesser linear restrictions and what other distributional distinctions, beyond the linear constraints, hold between MPs and adverb(ial)s, and suggest possibilities for their explanation. We shall see that other such distributional distinctions between MPs and adverb(ial)s can be found in the nature of their restrictions on their respective co-occurrence with syntactic types of sentences. See Table I above for a rough overview of the latter constraints. Since sentential type is interconnected with illocutive force, this distributional feature will be reserved for section 5 below.

Let us first take a look at the positional constraints holding for adverb(ial)s (as in (6)) and MPs (as in (7)).
(Adverbials in positions (9)(α)-(δ) are italicized.)

(6) gestern hatten *in der Mittagspause*(α) die Kollegen *im Umkleideraum*(ß) dem Lehrling *völlig unbemerkt*(Γ) zehn schmierige, verfaulte Kiwifrüchte *mit grimmiger Schadenfreude*(δ) in die Aktentasche gestopft.
yesterday had *in the lunchbreak* the colleagues *in the dressing room* the apprentice *completely unnoticed* ten dirty rotten kiwis *with grim malice* into his bag stuffed
"Yesterday during lunch time, the colleagues of the apprentice maliciously put ten rotten kiwis into his bag in the locker"

(6) illustrates all positions (viz. (α) through (δ)) that can be occupied by adverb(ial)s. MPs qualify for identical positions. See, for example, the distribution of *nur/bloß* in (7a-d) ((α)–(δ) signalling identical positions in examples (6) above and (7) below).

(7) a. Wie konnte ich *nur/bloß*(α),(ß) den Schlüssel verlieren?
how could I PART:*only* the key lose
"How in heaven's name could I lose the key?"
b. Wie konnte ich den Schlüssel *nur/bloß*(δ) verlieren?

c. Wie *nur/bloß*(Γ) konnte ich den Schlüssel verlieren?
d. *Nur/Bloß* — wie konnte ich den Schlüssel verlieren?

Note that (7c) reflects topicalization under *wh*-movement. In other words, the MP *bloß/nur* has not been topicalized in its own right, but only as a structural sister to the *wh*-word *wie*. (7d) illustrates *nur/bloß* in a position that is left-dislocated beyond the sentential structure (notice the pause indicated by the dash). This distributional variation can be traced furthermore in different sentence types or illocutive types. See (8) (ACC... accusative; PART...(modal or scalar) particle).

(8) a. Gib mir *doch* den Löffel! ... imperative
give me PART *the spoon* (ACC)
b. Hättest du mir *doch* gefolgt damals! ... wish
had you me PART *obeyed then*
c. Wer kommt *denn* heute abend? ... *wh*-interrogative
who comes PART this evening
d. Aber kann *denn* jemand aufpassen auf die Tiere?
but can PART *someone take care of the animals*
... *yes/no*-interrogative
e. Zumal sie *doch* dem Lehrling *immerhin* die
PART *they* PART *the*(DAT) *apprentice* PART *the*
Aktentasche völlig versaut haben. ... assertion
bag all messed up have

This yields the structural topology ("landing site structure", or "field structure") in Table III of the MP in the German sentence.[4] According to this first, provisional, topological framework the MF is the linear and structural domain between V-second ("COMP" or "V-2" below) and V-final ("V-FIN"). This is what traditionally has been called the "verbal bracket" in German.

Table III (p. 341) presents all possible positions of MPs in both independent and dependent sentence structures. Note that the particles take the positions of adverbials as sketched in (10a–e). In hierarchical structuring this will yield the subsequent trees (with adverbials as daughters of Chomsky-adjoined nodes [encircled]; such adjoined nodes are introduced to distinguish lexically governed [roughly, valency-governed] arguments from free adjuncts); MPs are italicized. See (9)–(11) on p. 342.

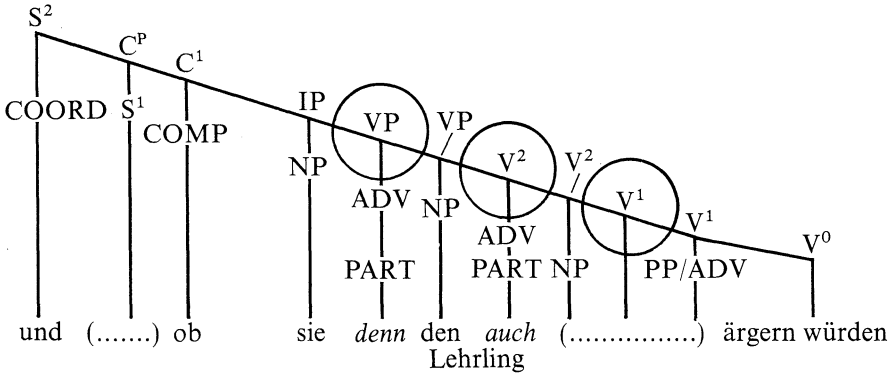
Table III *COORD = coordinating adverbials; IF = initial field; MF = middle field; FF = final field; V-2 = verb-second; V-FIN = verb-final position.*

	COORD	Topic(=IF)	COMP(V-2)	MF	V-FIN	FF
(a)	und/oder	S	ob	sie den Lehrling *auch*	ärgern würden	heute mittag
(b)	und	es reut ihn	–	den Löffel *doch*	abzuschlecken	–
(c)	–	–	da	es *ja eben doch*	schneit	–
(d)	–	Er fragte	wann	es *denn eigentlich*	stattfinde	–
(e)	–	der Teufel	der	ihm *übrigens*	den Garaus machte	–
(f)	–	umso ungeduldiger	je länger	er *halt/eben*	fragt	–
(g)	Aber	–	kann	jemand *auch*	aufpassen	auf die Tiere
(h)	Und	–	streiken	die Arbeiter *eigentlich auch*	–	wenn nur 75DM zu holen sind?
(i)	–	–	Hättest	du mir *doch*	gefolgt	damals!
(j)	–	–	Gib	mir *doch* den Löffel!	–	–
(k)	–	–	Hat	er *auch*	aufgegessen?	–
(l)	Aber	dem Greis, dem	will	*doch* niemand	wehtun!	–
(m)	Denn	daß es schneit, damit	haben	wir *eigentlich* nicht	gerechnet.	–
(n)	Weil	–	es schneit	*ja*!	–	–
(o)	–	–	Wie steht	es *denn*?	–	–
(p)	–	–	–	*Doch*!	–	–

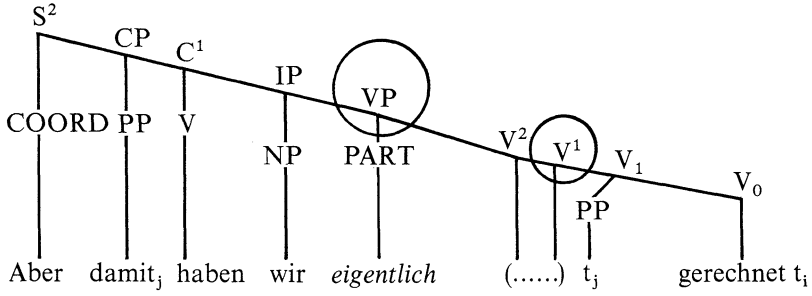
Glosses and/or translations for Table IIIa–p above (A... Accusative case, D...Dative case, MP...MPs):

(a) and/or (matrix clause) whether they the(A) apprentice MP irritate would this afternoon
(b) and he regrets the(A) spoon MP to lick
(c) because it MP-MP-MP snows
(d) He asked when it modal MP-MP took place
(e) the devil who him(D) MP the deathblow dealt
(f) ...the more impatient the longer he MP/MP keeps asking
(g) But can someone MP take care of the animals?
(h) And strike the workers MP-MP if there are only 75 Marks to gain
(i) (I wish) had you me(D) MP obeyed then!
(j) Give me(D) MP the spoon!
(k) has he MP eaten up everything?
(l) But the(D) old man, this(D) wishes MP nobody to hurt.
(m) Since that it snows, that have we MP not counted on
(n) Since it snows MP!
(o) What is the score MP?
(p) MP ("I did after all")

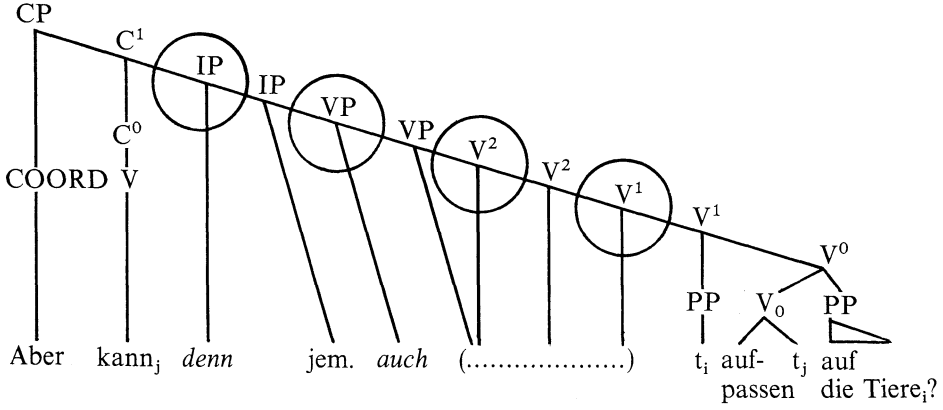
(9) *dependent sentence*: see Table III,(a-f), above.



(10) *matrix sentence*: see Table III, (l-o) above.



(11) *wh-sentence*: see Table III, (g-k) above



The range of linear variants is shared by adverb(ial)s, prepositional phrases, and MPs. However, the final field is strictly excluded as a landing site for MPs, other than for adverbials. Likewise, the initial field is excluded for MPs, except for a small, closed class of what has also come to be called "discourse particles", and only if they are under contrastive stress, i.e. in a highly marked occurrence. See Abraham (1989b) for some discussion of the distinct distributional properties of MPs and discourse particles. We will not be concerned with the latter type of particle in the remainder of this article.

We have now delineated the linear restrictions holding for MPs and adverb(ial)s. What distinguishes the two categories is that the latter, but not the former, can occupy positions in the forefield and the final field. This property of adverb(ial)s is interconnected with the faculty to carry contrastive stress, a faculty that the MPs lack completely.

In what follows we shall briefly consider a semantic difference between MPs and adverb(ial)s that appears to be necessary in order for us to be able to distinguish particles and adverbials beyond the linear constraints in the MF. Consider (12a,b). Note that the inverted sequences in (12b) and (12d), with permutations of the two adjectival adverbs (in b) and the prepositional adverbial groups (in d), are ungrammatical in distinction to their uninverted sequences in (12a,c).

(12) a. daß ich die Kühe *leider maschinell* gemolken habe
that — sorry to say — I have machine-milked the cows
b. *daß ich die Kühe *maschinell leider* gemolken habe
that I the cows mechanically sorry-to-say milked have
c. ich die Kühe *zum Unglück mit der Maschine* gemolken habe
I the cows to (my) distress with the machine milked
d. *ich die Kühe *mit der Maschine zum Unglück* gemolken habe
I the cows with the machine to (my) distress milked

What can we conclude from this distribution? *Leider* as well as *zum Unglück* can no doubt be taken as constituent, or sentential, adverbials. Consequently, they will be seen as operators extending scope. Now note that the scope readings enforced by (12b) and (12d) do not appear to be very felicitous in that the operator *leider* splits up what is conveyed in the sentence as semantically inseparable information ([[*maschinell*/*mit der Maschine*]$_{ADV}$ [*melken*]$_{V}$] "milk mechanically". *Leider* can extend scope only over the undivided verbal constituent (i.e. verb + adverb(ial)). We thus see that what is placed in the somewhat syntactically more neutral MF can nevertheless be subject to linearity constraints with semantic effects.

We shall now show that the same type of linear constraint with semantic effects holds for MPs. Let us illustrate this with the two examples, (13) and (14), below. (MPs have been underscored; angled brackets indicate alternative positions.)

(13) Er fragte, wann ⟨*denn eigentlich*⟩ es ⟨*denn eigentlich*⟩ stattfände.
"He asked when *precisely* it would take place."

(14) Aber kann ⟨*auch*⟩ jemand ⟨*auch*⟩ auf die Tiere ⟨*auch*⟩ aufpassen/
aufpassen *⟨*auch*⟩ auf die Tiere?
"But can someone *really* take care of the animals?"

What is new in (13) and (14), by comparison to (12a–d), is that the different scope relations have different readings in terms of rhematic and thematic material. What scope inversion of the MP-elements appears to achieve is to partition the utterance into a thematic part and a rhematic one, the latter always occurring to the right of the MP-element.

3.2. Results and implications

The distributional survey in Table III and in (12) renders the following picture:

A. With the exception of discourse-MPs (which are clearly different in distributional terms; see the brief distinguishing remarks above in connection with (11)), MPs proper (which form the bulk of MPs) occur only in the MF (MF). More specifically, they cannot carry contrastive (semantic) stress any more than they can occupy any of those positions in the initial field (IF) or the final field (FF) that their non-MP homonyms may occupy. For example, while the coordinating adverbials *aber* and *doch* occur under C^2 in the structural tree in (11), their MP-homonyms are excluded from this position. See the one-word clause in Table III, (p), specifically, where one has to assume an elliptical sentence such as the one bracketed in (15):

(15) *(Und sie bewegt sich) d o c h*
"Yet (, it moves)" or "(It moves) all the same"

with *und sie bewegt sich* taking the positions under the nodes C^2, C^1, and $COMP^0$, respectively. Compare again Table III as well as

(9)–(11) for what the structural positions, C^2, C^1, and $COMP^0$ mean in terms of lexical material that can occupy these positions.

B. MPs and scalar particles (as well as other constituent-modifying adverbials) share some linear variants in the MF. What they do not share, however, are movement variants that characterize fully lexical elements in the German sentence such as extractions to the left of the left bracket, C^1, or to the right of the right bracket, V^0. See again Table III above.

C. MPs cannot carry main stress, whereas adverb(ial)s (ADVs) as well as scalar particles (SCs) can. This distinction appears to follow from the fact that ADVs as well as SCs can "float" to "extreme" clausal positions, i.e. occupy positions in the IF and FF. Discourse MPs, in the strict sense of the term described under A. above, share some properties with SCs and ADVs. See the following examples — the discourse particles with contrastive stress in (17a, b) occupy positions in the final field and in the forefield, respectively, both of which are excluded for MPs. (*e* signals the base position of a category emptied after movement.)

(16) a. Er hat *ja* *nur* ein großes Maul.
he has MP SC a large beak
b. Maul_i hat er *ja nur* ein großes $[e_i]_N$
c. (?)*Nur* ein großes Maul_i hat er *ja* $[e_i]_{NP}$

(17) a. Eingroßes Maul hat er *ja* *d o c h* !
a large beak has he MP DISCOURSE-MP
"He has a large beak all the same."
b. *D o c h* hat er ein großes Maul!
"Nevertheless, he has a large beak."

As (16a–c) illustrate, the SC *nur* moves along with (part of its) constituent. The MP, *ja*, however is restricted to one position in the MF.

D. In distinction to MPs, scalar particle operators bind locally under what appear to be well-defined syntactic conditions (anaphoric binding; see Fanselow, 1988). Suffice it to say at this point that no such domain as the MF plays the syntactic constraining role for scalar particles, much in distinction to MPs.

E. MP-operators extend scope relations in a local, syntactic sense; see the discussion in connection with (12)–(14) above. They bind non-

lexical elements or constituents, namely rhematic (vs. thematic) and, within the rhematic material, epistemic and deontic functors. See Jacobs (1989).

F. The NPs carrying normal focal (rhematic) stress in a sentence cannot be reordered with respect to MPs. In other words, V^0-adjoining (adjoining to the verbal predicate, i.e. the structural position between the verb, V^0, and its direct object, V^1) is excluded. Also, any pre-IP position is excluded for the MP. See the distributional characteristics in Table III above and the structural description in (9)–(11).[5]

4. TYPOLOGICAL CORRELATION: MP-LANGUAGES AND MIDDLE FIELD-LANGUAGES

The observations in section 3 lead us to the following questions:

1. What is it that distinguishes so prominently the MF from the peripheral sentential positions so as to host exclusively MPs? Which are the general properties that distinctly mark this topological area? What is it about the MPs that matches these distinct properties?
2. What are the constraints that hold for the occurrence of MPs with respect to sentential types and/or illocutionary types? One of the ways to arrive at an answer to this question may be to see what the properties of sentential and illocutionary types are that constrain the collocation of two or more MPs, since one expects intersections of the constraints for the single MPs to hold.

Question 2 relates to 1 in the sense that it extends the general characteristic beyond a syntactic generalization. Findings about an interplay between syntactic and other restrictions for MPs may help us to find a point of departure for the description of MPs in those languages which have no MF in the sense that German has one.[6]

In this section we will consider only the questions in 1, i.e. the previous claim that the MF constraint is not an idiosyncrasy of German. Question 2 will be addressed later (sections 5 and 6).

The topological MF is the result of two very striking structural characteristics, namely the very strict word order properties of both V-second and V-last. This implies that the sentence structure hinges upon two very distinctly organised poles: A verb-node in, and only in, the second clause-structural position, and another verb-node in sentence-final position. Either position

can be filled only by a very limited range of grammatical categories (verbal predicates, copulas, verbal prefixes, and clitics). Every other position, and most prominently the linear range between these two polar positions, shows nearly unbounded access, and freedom of sequentiality, to any grammatical category, constrained only by non-syntactic properties which lend themselves to a structural organization in terms of discourse properties (see Abraham, 1986b; 1989a for a discussion).

More concretely, the corollary is that sketched in (21)–(23). (INFL = inflectional node (Chomsky, 1981); IF = initial field in the sentence topology, FF = final field, MF = middle field)

(18)	*German, Dutch, Frisian*	*English, Romanian*
	V-2 & V-last: MF;	V-2/V-3: void of a MF;
	i.e. MF only if both verb-second and verb-final structurally anchored	"open ended" linear structure: no verbal boundary right of verb second/third; thus, no MF

structural design:

(see Table III and (9)–(12) for illustrations and hierarchical corrolaries of the following linear structures)

(a) "structural bracketing", i.e. middle field (MF):
$\{^{IF-IF}\}[_{V-2}INFL]\{_{MF-MF}\}[_{V-last}V]]$
$(\{_{FF-FF}\})$

(b) "open ended", i.e. no MF:
$\{_{1_}\ (_{2_})\}[_{3}V]\{___\}$

Examples: GERMAN: cf. (a) above

(c) $\{Er\}_{IF}$ $[hat]_{V-2}$ $\{ja$ den Mann $wohl\}_{MF}$ $[gesehen]_{V-last}$ $\{heute\}_{FF}$
he has PART the man PART seen today

ENGLISH: cf. (b) above

(d) $\{_1$ He $(_2$ hardly)$\}$ $[_3$ has seen] {the man last night}

Now note the following typological correlation: The three West Germanic MF-languages (not, however, the rest of the Germanic languages and the Romance languages) typically share a wide class of MPs, with all the characteristics sketched above. As has been shown (Abraham, 1988) and will be briefly illustrated below, this correlation can be extended to also include the (Mainland) Scandinavian languages.

In contrast to the initial field (IF) and the final field (FF), the middle field (MF) is open for hierarchical structuring and linear order in terms of the following features: ±definite; ±generic (in terms of nominal number); pronouns, anaphors, and expletive elements (such as *it*, *there*; German *es*) in contrast to full, referential nouns; furthermore, to linear movement, and accent-shift dependent on discourse function (rhema vs. thema). See Abraham (1989a) for details.

I hypothesize that the typological correlation regarding the assumption that the syntactic MF has a particular potential to express discourse functional elements in German is as follows: MPs develop from lexical categories in those languages (German, Dutch, West-Frisian) which have developed a structurally defined syntactic MF; languages without such a MF extending between the structurally marked positions V-2 and V-final do not possess MPs in any sense comparable. This hypothesis appears to hold for Italian and French which are both non-MF languages and non-MP languages (see Helling, 1983) as well as for English, for the same reasons (cf. Panet, 1981; Abraham, 1981).

However, the hypothesis based on the correlation between the occurrence of the well-defined MF (see Table III as well as (9)–(11) above) seems to be on shaky ground when we look at the Scandinavian languages. Scandinavian languages are rich in MPs, but they have no MF. If the linear restrictions that we postulated for the co-occurrence and, as we shall see later, also for the diachronic development of the MP-category holds for German, but not, however, for other languages with a similar abundance of MPs, this could hardly be called an "explanatory", typological generalization. It is therefore tempting to test the hypothesized correlation between MF and the linear restrictions of MPs by looking at the Mainland Scandinavian languages. Danish, Norwegian and Swedish, which do not possess a structural V-final, resemble English and the Romance languages in this respect (compare (18b) above), but they differ from them in that they have a rich lexicon of MPs, sharing a number of semantic and illocutive assets of their German counterparts. The question thus is: Does the typological correlation fail with respect to Mainland Scandinavian?

I have shown elsewhere in detail (Abraham, 1988b) that the Mainland Scandinavian languages indeed restrict the occurrence of MPs to what is called the *nexus field*, which is comparable to the German MF in that it shows a form of verbal bracketing. In other words, although none of the Scandinavian languages has a verbal bracket to the right along the lines of the rigid German V-last, there is nevertheless a structural demarcation right-bounding

the range of linear positioning of MPs. See the following examples in (19) and (20). Compare the structural bracketing with that in (18c,d) above.

SWEDISH (the structural subscripts make use of Diderichsen's (1962) terminology: FF = forefield, identical with the German IF (= initial field above); NF = nexus field; CF = content field).

(19) {Vem}$_{FF}$ [hade]$_{V-2}$ {*troligen*}$_{NF}$ {[köpt]$_{V}$ boken}$_{CF}$
who had PART bought book-the
"Who has presumably bought the book?"

NORWEGIAN:

(20a) {Jeg}$_{FF}$ [er]$_{V-2}$ {*nok sikkert*}$_{NF}$ {veldig trett}$_{CF}$
I am PART PART very tired
"Quite surely, I am very tired."

(b) {A du}$_{FF}$ [tor]$_{V-2}$ {*na vel*}$_{NF}${komme utnar...}
oh you dare PART come out to ...
"O, you really dare come out to...after all?"

What associates Mainland Scandinavian basic word order with that of English and French is that verb-governed objects as well as prepositional objects and adverb(ial)s are exclusively in CF, i.e., after the last unit of the predicate (*köpt*; *komme ut*; that is, to the right of [$_{3}$V] in (18b) above). What differs is that the nexus field separates the finite predicate component (usually filled by some auxiliary verb) from the infinite element (such as the past participle). The domain between these two verbal units is the only place for MPs (and the negative particle), just as in German.[7] The nexus and the MFs are thus similar to the extent that they involve verbal bracketing, but they are distinct, among others, to the extent that the occurrence of non-particle, referentially full adverb(ial)s in Mainland Scandinavian is restricted to the content field after the second, infinite verbal component, which is to the right of the right verbal bracket. In other words, just like German (and Dutch and West Frisian), the Mainland Scandinavian languages also have a number of homonymic forms in two structurally and linearly distinct positions: full lexical adjuncts and modal particles. Note further that the MF-like nexus field in Mainland Scandinavian is the only position for sentential negation. This adds a clause-semantic component to the similarity to MF-languages in that the rhematic sentential negation in German (and Dutch and West Frisian) can only occur in the MF. The continental West Germanic languages and the Mainland Scandinavian languages thus share an abundance of modal particles

in unique linear and structural positions which can be subsumed under the general property of an MF-like characterization. Furthermore, both the MF and the nexus field are the locus of prominently rhematic as well as illocutive material.

5. PRAGMATIC CHARACTERISTICS OF MPS: FROM LEXICAL OR FUNCTIONAL BLEACHING TO ILLOCUTIVE FORCE

As regards the content transfer from the original lexicals to MPs, all MPs (possibly, except for *doch*; see below) have developed under bleaching of their semantic and functional content. The semantic bleaching was accompanied by a conspicuous increase in pragmatic function.

This pragmatic function is of a two-fold nature:

1. Illocutive or meta-communicative: see the assignments of illocutive force below (cf. discussions in Lütten, 1977; Franck, 1980; Gornik-Gerhardt, 1982; Weydt et al., 1983; Wolski, 1986; Shibles, 1988; to mention but a few).
2. Rheme-focussing (cf. the discussion in Hentschel (1986: 231ff., 277f.).

The present section addresses the first of the two prominent pragmatic properties. It will be shown what the lexical and/or functional input components to the grammaticizing process in diachrony were and which steps in the transition from the original, homonymic lexeme to the eventual modal output function can be reconstructed both in terms of semantic scope and in terms of constraints of illocutive force. Table IV below contains a list of most of the MPs in their original referential and in their syncategorial functions. It illuminates their original status as one-place sentential operators, or functors (in the sense of logical semantics) over clausal material or NP, designated A (see column (3)); the shift from a one-place functor to a two-place functor for A and B (as in (4)); and, finally, to a temporal or logical relator between the linguistic complex units, A and B. The last column, (5), represents the step immediately preceding the emergence of the MP-elements with illocutive or meta-communicative force. Note that it has been argued above that the "quasi-logical" step (as in ((5)) is the indispensible predecessor of the new modal function of MPs.

As can be observed, only the last three items in Table IV derive from a source without referential, lexical meaning; they represent syncategorial mean-

ings. Among these, *doch* can be traced back as a modal connector from its earliest appearance in the Indo-Germanic area (see Hentschel, 1986). As regards the formal transfer for many a particle or connector, the historical development implies a functional change from a monadic functor to a dyadic (from one-place to two-place) functor; sometimes this involves a change in the identity of arguments. A most fitting example is *auch*: according to the etymological literature, it started as an imperative predicator of one single argument, meaning something like Latin *auge*! "increase! grow!". This appears to have been followed by a scalar particle function such as *again*(A) *with respect to some former mention of A*, and, in due course, by disjunction of two appearances of A, *adding up to*(B, A). From this derived, in further course, CAUSE(B,A) and/or ADVERS(B,A) as (parts of) clausal connectors. This, in the last instance, yielded a modalised function as an independent (non-focus dependent) particle in discourse.

The correlation between the structural properties in Table IV and the typological distribution of MPs suggests very strongly that there is some conditioned relation holding between the metacommunicative, or illocutive, force of MPs, and the sketched structural properties. This ties in with the general observation that properties of textual coherence in the MF languages can be expressed by illocutive indicators (such as MPs) to a considerable extent, while the non-MF languages fail to require this stylistic property in a crucial way. I will come back to this point later.

Let us now address the questions posed in section 4 concerning what restrictions for MPs hold in terms of sentential type and illocutive force. The studies of Helbig and Kötz (1984: 29) and, in the same vein, of Hentschel (1986: 199–203) can be exploited to yield the following combinatorial constraints with respect to formal types of sentences and illocutionary types. I restrict myself to a few examples. (I — illocutive intention of the speaker; S — sentence; ADHORT — adhortation; EXCL — exclamation; QU — question; ASS — assertion; WHQU — word question; *yes/no*-QU — *yes/no*-question.)

(21) a *aber* as a MP: only with the I of an exclamation (I-EXCL) and an adhortation (I-ADHORT) irrespective of S, thus not with the intention of either a question (I-QU) or an evidential assertion (I-ASS)

b *auch*: if with S-ASS or S-QU then not with I-EXCL; if with S-QU then not with I-QU

c *bloß/nur*: occur everywhere except in S-ASS with I-ASS

Table IV

(1) Syncat-egorial meaning only	(2) Lexical item in NHG	(3) Original lexical meaning + valency (one-place functor) (A...usually sentence or NP)	(4) Derived relational (two-place functor) (A,B...senten-ces or NP)	(5) Derived quasi-logical (epistemic/ deontic)
	aber	again(A)	POST(B,A)	OPPOSITION(B,A)
	auch	increase(A) (IMPERATIVE)	AGAIN(A) > (A,A) > > ADD(B,A)	CAUS (B,A) ADVERS (B,A)
	bloß	naked(A)	void of (A,B)	only(B)
	eben	equal(ly high)(A)		CAUS(B,A)
	(so)gar	ready/complete/fully prepared (A) (cooked)		highly valued with respect to(A) ILLOC
	eigentlich	characteristic of (A) > > precisely(A)		
	etwa	somewhere(A)		possibly(A)
	halt	biassed toward(A)		very(A) > CAUS (B,A)

	ja	yes/in this way/ truly/surely(A)	CAUS(B,A) or: *ja*(A) B(*ja*)
	einmal	once(A)	
	schon	in a complete way(A)> >as is fitting(A)> >faultless(A)	*already*(TEMPORAL, SCALAR) ADVERS (not-A,A)
	vielleicht	"viel + leicht"(A) much-easy	PROBABILITAS(A)
	wohl	"willingly(A) (<wollen)	VOLUNTAS(A) >ASSUMPTIO(A)
denn (from temporal *dann*)			FOLLOWING (B,A)> >CAUS (B,A)
doch			ADVERS (B,A)
nur			cf. MHG *ni waere*(B,A) = NHG *es sei denn*(B,A) = *egal was sonst sein könnte*; viz. French *ne...que*)> > *allein*(A)

d *denn*: only in S-QU as long as they are I-QU or implicit answers
e *doch*: everywhere except in *yes-no*-QU with I-QU
f *eben/halt*: not in S-QU irrespective of which illocutive intention is to be expressed
g *ja*: not in S-QU with I-QU
h *schon*: excluded with I-QU or I-EXCL
i *vielleicht*: not with I-ASS or in S-WhQU with I-QU; however, occurs in any S as long as with I-EXCL as well as in S-*yes/no*-QU with I-QU
j *wohl*: in S-ASS with I-ASS as well as in S-QU with I-QU; not, however, with I-EXCL or I-ADHORT

I will not compute at this point the combinatory gaps following from the restrictions spelled out in (21). Suffice it to list in (22) classes of MPs the members of which can collocate in the linear order of (i)–(v) (following Helbig and Kötz, 1984:41f.).

(22)

i	ii	iii	iv	v
denn	aber	d o c h	auch	nur
doch	eben	schon	mal	bloß
eigentlich	halt			
etwa	vielleicht			
ja	wohl			

See the following examples (only MPs considered):

(23) a. Er hat *denn doch auch* nur wenig
he has MP MP MP only little
verloren (*nur* "only" = SP)
lost
b. Das hat *doch wohl* schon Geniecharakter (*schon* = scalar)
this has MP MP already properties of a genius
c. Ich kann *aber eben auch* bloß 24 Stunden im Tag arbeiten
I can MP MP MP only 24 hours a day work
d. Sie haben *eigentlich doch wohl* schon alles getan
they have MP MP MP already all done

It is not at all clear why other sequential orders, like those in (24), should be ruled out, as in fact they are (see Thurmair 1989 for an account). (24) and (23) are lexically identical except for the fact that the linear order of the MPs is altered.

(24) a. *Er hat *doch auch denn/doch denn auch* nur wenig
he has MP MP MP MP MP MP/SP SP little
verloren.
lost

b. *Das hat *wohl doch* schon Geniecharakter.
this has MP MP SP genial property

c. *Ich kann *eben aber auch/auch eben aber* bloß 24[h]
I can MP MP MP ADV MP MP SP 24 hrs
bleiben.
stay

d. ??Sie haben *doch wohl eigentlich* schon
they have MP MP ADV SP
alles getan.
everything done

e. *Sie haben *wohl eigentlich doch* schon
they have MP ADV MP SP
alles getan.
everything done

It is beyond doubt that the constraining properties ruling out the sequential orders in (24) and predicting those in (23) cannot be a mere intersection of the properties in (21) since this would not rule out the linear orders in (24). Despite the lack of distributional evidence I do not wish to assume *a priori* that the controlling force is one outside of syntax. Given the uncontroversial distribution of MPs in terms of sentential types (cf. (21a–j)) translatable into distinct syntactic structures, as well as in terms of illocutive types, all one can say at this moment is that these constraints possibly result from a conspiracy of syntax and the illocutive component (see Abraham 1991).

6. PRAGMATIC CHARACTERISTICS OF MPS: THEIR SCOPE AND THE DISCOURSE FUNCTION OF RHEME.

This section takes up the second of the two pragmatic characteristics of MPs indicated in section 4, rhematic force. As we have seen in sections 2 and 3, MPs generally occupy one out of several MF positions. As a rule, they are unstressed. Though this is not without exceptions, there are only a few MP-lexemes that may appear under sentential, contrastive or semantic stress: *schon*, *ja*, *doch*, and *wohl*. Due to the linear variance and restriction of accent

they do not seem to have scope over any grammatical sentential relation. Nevertheless it is often claimed that the MP has in its focus the relational category of the discourse rheme within a sentence (see for example Hentschel, 1986:231ff.; 277). See the following examples which illustrate this:

(25) Was$_i$ hat er ⟨*denn*⟩ dem Koch ⟨*denn*⟩ [e$_i$] getan?
What has he then the cook-DAT then done?

(26) Was$_i$ *denn* hat er dem Koch [e$_i$] getan?
What then has he the cook-DAT done

(27) *Dem Koch hat er was *denn* getan?

(28) Wem ⟨*denn*⟩ hat er ⟨*denn*⟩ etwas getan?

(29) *Dem Koch was ⟨*denn*⟩ hat er ⟨*denn*⟩ getan?

(30) Hat er ⟨*denn*⟩ dem Koch ⟨*denn*⟩ etwas ⟨**denn*⟩ getan?

(31) Hat ⟨*denn*⟩ der Mann ⟨*denn*⟩ dem Koch ⟨*denn*⟩ einen Hieb
Has then the man then the cook-DAT then a blow
⟨**denn*⟩ versetzt?
then given?

(25)–(31) show that the MP *denn* extends scope over the entire verbal phrase VP, or at least one constituent within the VP (cf. (26), (28)), and eventually down to the [V° + AUX] (cf. (32)).

(32) Wem hat er das Paket/ die Äpfel *denn* gegeben?
To-whom has he the parcel/the apples then given

(33) Wem hat er ein Pakket/apples then given?
To-whom has he a parcel/apples then given?

(34) Was hat er ⟨*denn*⟩ Freunden ⟨*denn*⟩ gewöhnlich ⟨*denn*⟩
What has he then friends-DAT then normally then
anvertraut?
confided to?

(32) and (33) confirm the constraints of the rheme-carrying element: Its linear position may be such that definite NPs are not in its scope (see (32)); however, indefinites will invariably have to be within its scope. If indefinites are rheme-prominent in morphological terms (as has been discussed in detail in Abraham 1989a), it follows that MPs like *denn* are rheme-sensitive. However, it can

easily be shown that a stricter notion of rheme-government cannot be based on phenomena of surface structure. In (25)–(31) which show movement under topicalization in question sentences or in assertive sentences, the source position of the moved element has been marked empty [e_i]. This is clear evidence that the scope range of MPs is of a deeper structural nature and cannot be defined in terms of surface ordering (Abraham 1991). While the rheme orientation of MPs was noticed before (Hentschel, 1986: 231ff.), the structural definition of rheme (in terms of the theory of movement and trace) has not previously been considered (see, for this issue, Abraham, 1988a; 1991).

7. DIACHRONIC EVIDENCE

This is what our discussion has yielded so far: We have seen that originally full lexical and/or functional meanings of present day MPs in German have been replaced by pragmatic meanings of an illocutive and of a rhematic type. Using general information derivable from historical grammars and etymological dictionaries, we have reconstructed intermediate steps showing that the bleaching process was accompanied by a rise of scope-relating force (see Table IV in section 5). As far as we can estimate, this rise of scope relations coincided roughly with the emergence of illocutive force of the elements in question. It must be emphasized that none of the illocutive meanings of any of the MPs is random or contingent upon context, but, rather, independent and very systematic. See section 5 (particularly (21)). What we do not know, however, is whether our claim about the MF as a strict synchronic constraint is also a conditioning factor in the diachronic development of MPs in German. Only if we could show that the development of the syntactic MF triggered the rise of the pragmatic properties characteristic of the MPs in German would we be able to speak of grammaticization in the concrete sense of the term.

The following diachronic section is intended to yield answers to the following questions: Can the rise of the MP functions be observed to coincide with the emergence of the syntactic MF constraint? Can any of the scope relations reconstructed in Table IV above be confirmed by looking more closely at the historical material? Finally, what was the syntactic behavior of those MPs which existed as such before the emergence of MF, judging from the earliest historical evidence?

In order to get closer to answers to these questions, we need longitudinal

observations of particle lexemes, notably from Old High German (OHG; dating from roughly 750 to 1050 A.D.), which had not yet developed a systemic MF, via Middle High German (MHG; 1050–1450), at the end of which the MF began to emerge, to New High German (NHG; from 1450 onwards), where the MF is fully established. Since OHG had developed some statistical restrictions on word order, we shall also look briefly at the stage prior to OHG, i.e. to word-order-free Gothic.[8]

In the ensuing historical survey, I shall focus upon four particles of German. They were chosen for the longitudinal, diachronic investigation for two practical reasons: First, the etyma in question already appear in Gothic, albeit in somewhat different functions; and, second, one can rely on existing diachronic work done on these particles (above all by Hentschel, 1986). The MPs are, in alphabetical order: NHG *doch*, *eben*, *halt*, and *ja*.

7.1. The NHG MP *doch*

Consider the following example:

(35) *Denk doch*! "Do think! Use your brain!"

Doch can be rendered in English in a number of ways: "after all", "as you already know for a fact", "if only", "isn't that so?", "just", "of course", "really", "thus", "yes" (Shibles, 1988:4). In the example above, it can express the following attitudinal types: Impatience, annoyance, disapproval, indignation, blame and anger. In the function of a sentence-connecting adverbial, it simply expresses contradiction or opposition, as in the following sentence:

(36) Er ist klein. *Doch* kann er hoch springen/*doch* er kann hoch springen.
"He is little. Nevertheless, he can jump very high."

The adversative meaning, both in the function of MP (without stress) and as a sentence connecting adverbial (cf. English *but*, German *aber*), can be traced back to Gothic and is also present in OHG and MHG. However, in the function of a discourse particle with heavy stress, meaning positive assertion and presupposing the refutation of an utterance with contradicting force, *doch* cannot be traced until in the eigtheenth century (Hentschel, 1986: 46). See the following NHG example:

(37) A: Er ist klein; er kann die Torlatte nicht erreichen
He is small; he can not reach the full height of the goal.
B: D o c h ! Er erreicht die Latte; er kann gut springen.
Yes, he can! He can reach the goal. He can jump well!

Heavily stressed *d o c h* in (37) is what has been called a *discourse particle* in that it occurs only in explicit discourse between two discourse partners. Other than MPs proper, it can carry heavy, refutatory stress.

7.1.1. Gothic

Gothic *thau* (meaning roughly "(al)though; all the same; nevertheless"), the source of NHG *doch*, appears in functions that have a clear metacommunicative reading. It resembles NHG *doch* in its use as an MP as well as an adversative adverbial (sentential connector).

It is noteworthy that other Gothic MPs such as *-a* (interrogative particle; cf. German *denn?*) or *-(u)h* "and", which never had homonymic referential counterparts, disappeared in the subsequent development (Hentschel, 1986:74). Gothic *thau*, however, in its presumed MP-function, has been retained next to its other coordinating, logical uses. If this observation should prove to be generalizable then the following can be assumed: The bleached illocutive meaning of the MP is supported, in some unknown fashion, by the lexical, or syncategorial, meaning of the homonym. This assumption will be tested in our further investigation.

The fact that Gothic *thau* occurred in apparent MP-function (i.e. with some illocutive force besides the sentence-connecting function) despite the fact that Gothic had not developed any MF may seem to disqualify our main formal property as a condition for grammaticization in the first place. It is to be noted, however, that the MP-status of Gothic *thau* is unique in this respect. No other NHG MP has a source as MP in Gothic or OHG. The earliest traces of MP-like illocutive force are traced in MHG, which is the period when verbal bracketing, and consequently the MF, emerged. For Gothic, however, no such linear restrictions have been observed. The tentative conclusion this permits is the following: MPs, and, more generally, rhematic material associate prominently with linearly unconstrained clausal domains *and*, preferably, with domains in which they can signal rhematic status without interference by other sentential functions, notably by the prototypical rhematic category, the verb. V-2, however, is such an interfering rhematic position. We can therefore say that Gothic, with its free word order, and also to an

extent OHG had a means to make prominent rhematic status of MP *thau*. As Luraghi (1988) has observed, other Indo-European free word order languages such as Old Greek, Hittite, and Latin shared MP-vocabularies comparable to that of modern German.

Naturally, all this can show is that Gothic had no structural reason for *not* having MPs from the very start. The correlation of the development of the MF and the emergence of illocutive MPs within, and never outside of, the MF remains more or less untouched as a formal condition for the process of desemanticization. This is what we have adduced as evidence for classifying the development of MPs as an instance of grammaticization in some broad sense.

7.1.2. Old high German thoh

Whereas it is not easy to presuppose configurational word order in Gothic we shall cast a closer look at the linear ordering of lexical elements in OHG and MHG, assuming that the development of verbal bracketing plays a crucial role in the elicitation of the modal meaning. For all practical purposes, linear distinctions of some statistical order must have played a meaning — and function — distributing role already in OHG. As Tomaselli (1989: chapter 2) has observed, word order in OHG was V-2 in independent clauses in the majority of cases investigated (= "statistical V-2"), whereas V-2 virtually never occurred in dependent clauses. V-last, to be sure, had not emerged yet in any statistical order beyond chance. Consequently, there was no MF in OHG.

In sentence-initial position OHG *thoh* appears to have meant something like NHG "also", Latin "*verum tamen*" (translated together with OHG *uuâr* (Hentschel, 1986: 88; the example is taken from Tatian).

(38) (TATIAN 95.3)
OHG: notdurft ist thaz quemen assuicha, *thoh* uuidaro uue manne thuruh then quimit assuich!
LATIN: necesse est enim ut veniant scandala, *verumtamen* vae homini per quem scandalum venit!
ENGLISH: it is compelling that annoyances will be aroused, but, *alas*, woe to the man through whom this annoyance arises

The use of OHG *thoh* in Otfrid is considerable. Hentschel (1986: 93) cites more than 200 occurrences, of which two thirds are adverbials or

conjunctions, that is, in the strong syncategorial function meaning "although, despite".

For what was more of a modal (i.e. less logical) reading, a position between subject and direct object in the embedded clause seemed to be instrumental (Hentschel, 1986: 90; Otfrid). See the following example:

(39) (OTFRID 79.5)
OHG: Inti gisuuo iru, thaz só uuaz só thú bitis só gibu ih thir, *thoh thoh* halftanod mines rihhes.
LATIN: Et iuravit illi: quia quidquid petieris dabo tibi, licet *dimidium* regni mei.
NHG (gloss): Und schwur ihr, daß so was so du bittest, so gebe ich dir *doch doch* die Hälfte meines Reiches.
ENGLISH: And swore to her that what you beg for I will give to you, *be it even* half of my empire.

The NHG MP-equivalent is conjectured to be "und sei es selbst/sogar" (= "be it even") which is certainly a fitting rendering.

A great number of occurrences both in Tatian and in Otfrid yield evidence that in non-initial sentential position OHG *thoh* is a very versatile modal element. See the following examples (Hentschel, 1986:90):

(40) (TATIAN 82.1)
OHG: ...inti batun inan thaz si *thoh* tradon sinis qiuuates ruortin.
LATIN: Et rogabant eum, ut *vel* fimbriam vestimenti eius tangerent.
NHG (gloss): Und baten ihn, daß sie *doch* einen Faden seines Gewandes berührten.
ENGLISH: And begged him that they might touch *no more than* his clothes.

(41) (HELIAND 2846–2846; from Hentschel 1986: 96)
OLD SAXON: ...(that sie ... mêr ni habdin) biûtan brôd fiƀi an ûsaru ferid endi fiscos tuêne, huat mag that *thoh* thesaru menigi?
NHG (gloss): (daß sie mehr nicht hatten) als Gerstenbrot, fünf auf unserem Weg, und der Fische zwei. Was vermag das *doch* dieser Menge?
ENGLISH: (that they did not have more) than barley-bread, five on our journey, and of fish two. What, *for all practical purposes*, does that mean for such a lot of people?

(42) (HELIAND 919–920; from Hentschel 1986:96)
OLD SAXON: ... ef thu nu ni bist that barn godes bist thu than *thoh* Helias, the hêr an êrdagu ...
NHG (gloss): ... wenn du nicht bist das Kind Gottes, bist du dann *doch* Helias, der hier auf Erden ...
ENGLISH: ... if you are not God's son, then *at least* you are the Saviour who here on earth ...

In both occurrences the logical "despite" for *thoh* is not a possible reading. Hentschel (1986: 90; 96) suggests a translation with the scalar particle "only". Equally possible, however, is also a modal reading in the sense of NHG *vielleicht* ("possibly") or *etwa* ("may be"). The same holds for the Heliand reading: The context in the first Heliand example refers to the food that the people at lake Genezareth had: Just five loaves of bread and two pieces of fish. Again, the adverbial reading of "despite" makes no sense. The sense of the NHG MP *doch*, however, with its bleached, thinned-out, adversative meaning, fits perfectly. Finally, see the second Heliand reading: Either NHG *doch* in the MP function, or *aber jedenfalls*, an equally modal pair, renders a sensible reading.

It is to be noted that in none of these examples — nor in any other of those investigated and documented by Hentschel, 1986 — is there reason to assume that OHG *thoh* really has the function of a MP as in NHG. This is for the simple reason that the occurrences of non-initial position are restricted to embedded *thaz*-sentences. Illocutive readings, however, cannot obtain in embedded clauses of this type; see the clause restrictions to which the MPs are subject, described in section 2. Since a full logical meaning is to be excluded in the contexts concerned, then what remains are very unspecific, thinly adversative, loosely context-establishing elements. For example, *thoh* in the following three examples can be rendered by three distinct translations into NHG (following Hentschel, 1986: 97):

(43) (OTFRID IV.1,11–13)
OHG: ... joh sie thar lérta filu frám io gidago fora thiu thaz sie irkantin *thoh* bi thiu, thaz er was druhtin heilant.
NHG (gloss): ... und lehrte sie dort immerzu, jeden Tag aufs neue, damit sie erkannten *doch* (="jedenfalls") daran, daß er war der Herr Heiland.
ENGLISH: ... and taught them there on and on, every day again, so that they recognized *in any case* that he was the Saviour.

(44) (TATIAN III.1,34)
OHG: (súntar si imu munto) theiz iaman *thoh* ni wunto.
NHG (gloss): (sondern sie schützt es) damit es jemand *doch* (= "nur nicht; nicht etwa") nicht verwundet.
ENGLISH: (but she protects him) so that *not* anybody *ever* could hurt him.

(45) (OTFRID III. 9,9–10)
OHG: sie wúnsgtun, muasin rúan *thoh* sinan tradon einan in sînen giwâtin
NHG (gloss): sie wünschten, sie dürften berühren *doch* (= "nur; zumindest") seinen Faden einen an seinem Gewande
ENGLISH: they wished that they might touch *at least* one thread of his clothes

This appears to support our thesis that MPs develop their thinned-out readings in two stages: First, where the logical meanings cannot be supported contextually; and second, where the co-occurring sentential type is not in contradiction with the original non-modal reading. In other words, our examples permit the conclusion that OHG *thoh/doch*, which was not linearly bound to particular positions, first entered into contexts which did not themselves represent the logical adversative relation of the logical sentence connector *thoh/doch*. Second, the generalization of *thoh/doch* was further carried through to the extent that it entered into sentential types beyond those that the original logical connector was restricted to (i.e. in the case of *thoh/doch*, assertions). Such a new type of sentence for OHG *thoh/doch* could be an imperative. Thus, we assume that, first, conditions of co-occurrence (semantic adversativity; assertion) were loosened. Since *thoh/doch*, in the original environments, was retained in its original function and meaning, this prompted new (modal) meanings under the new identifiable conditions. Such meanings, to be sure, were still associated with those of the original logical connector. What this situation in OHG was waiting for was some more rigid, formal, and thus grammaticizing, condition. This was to come through the linear constraint of occurrence in the newly developing MF. The MF-emergence guaranteed the overt identifiability of the specific semantics of MPs in distinction to sentential connectors (in the function of adverbials). Provided that this is generalizable then this would give support to the view that the development of MPs and their illocutive, or meta-communicative, semantics is one of the exclusion of non-sensical distributions on the basis of the original, logical or lexical,

meaning under distinct conditions of co-occurrence and linear position. We shall try to elucidate and corroborate this assumption below.

7.1.3. MHG doch

The range of occurrences and distributions of MHG *doch* appears to be fully congruent to that of NHG. MHG *doch* occurs in assertive clauses with interrogative syntax. See (46) below, which presents a dependent interrogative clause with the syntax of an independent interrogative structure, or (47), which is a conditional with a question syntax. The examples are taken from Hentschel (1986:110 ff.):

(46) (IWEIN 89)
MHG: sage durch got, wer weiz ez *doch*
NHG: Sage bei Gott, wer weiß es *doch*!
ENGLISH: Say,for God's sake, who knows it *after all*!

(47) (DES MINNESANGS FRÜHLING: 109, 18–HEINRICH VON RUGGE)
MHG: Het ich von dirre sumerzît
doch zwêne tage und eine guote naht
mit ir ze redenne âne nît (Hentschel 1986:112)
NHG: Hätt' ich von dieser Sommerzeit *doch* zwei Tage und eine gute Nacht, mit ihr zu reden ohne Groll.
ENGLISH: If *only* I would have of this summertime two days and one good night to speak with her in a good spirit.

MHG *doch* occurs further in generalised relative clauses, exclamation clauses, relative clauses, and matrix assertions. For the last type, see (48) below.

(48) (MINNESANGS FRÜHLING 113,9)
HG: Ich mac von vröiden ertoben âne strît: swâ waer ein wált beidiu lanc unde wît, dâ möhte man mich *doch* springende sehen
NHG: Ich kann ohne Kampf vor Freude rasen: wo immer ein noch so langer oder breiter Wald wäre, da könnte man mich *doch glatt* vor Freude herumhüpfen sehen
ENGLISH: I can be enraged by joy, without fighting: where ever there is a forest whatever long or deep, I could be seen *really* jumping with joy.

The linear distribution is very explicitly the MF, which had slowly developed (although by no means made indispensible until early NHG) during MHG between the two verbal brackets (left bracket *möhte* and right bracket *sehen* in (46) above; for this syntactic innovation in MHG see Wells, 1985: 256 ff.). One item deserves particular mention: See *doch* in wishes such as in (47) (Minnesangs Frühling 109). Such wishes do not have a dependent sentential function as in OHG, but have received independent status. It is this step — clausal independence in distinction to OHG dependence — that facilitates the modal reading in MHG, since only independent clauses are able to carry the modal load of a speech act.

7.2. The MP NHG *eben* and its diachronic development

Eben, in its full lexical status, is an adjective meaning "even", "just", "precisely", "similar", "level". In Grimm's *Deutsches Wörterbuch* (1862:6) the equivalent is "complanatus" = "eingeebnet, flach; flattened". There is no occurrence of the simple adjective in OHG or in Gothic, whereas it is not infrequent in compounds; see NHG *ebenfalls*, *ebenso*, *ebenmäßig*, *ebenda*.

The accepted conjecture is that the local-topographical, adjectival meaning receives a reanalysis of temporal deixis, in the sense of an anaphoric backward orientation: It renders something equal to something that was uttered shortly before; see Table 4 in section 5 above. It is from this temporal anaphor that, finally, the logical, communicative anaphor in the sense of causal confirmation with respect to what had been uttered just before develops.

In MHG, however, there are not yet occurrences of the modal meaning or even the temporal meaning (Lexer *Mittelhochdeutsches Wörterbuch* 1872, I: 499). After the adjectival "*paris*; *aequus*" meaning, the next form that appears to have occurred is the adverbial "accurate; *proprie*" (Grimm *DW* 1862:9). Occasionally, however, Hentschel mentions (1986: 105) that there is an MHG adverbial (Nibelungenlied 565: 4) meaning "momentary". This temporal meaning appears to be the necessary step for a further development of the meta-communicative, causal anaphor, i.e. the reference to an argumental status between the speech act utterance and a prior utterance. The earliest sources for occurrences of this reading come from the 18th century.

The development of the MP *eben* is thus as follows: (1) topographical evenness; (2) the addition of the meaning of temporal similarity next to the original, topographical evenness in MHG; and (3) a causal relation between two utterances, in the late 18th century.

7.3. The MP NHG *halt*

Halt can be traced back to an old comparative form meaning "prior to", temporally as well as modally, see (25). This diadic relator in OHG was continued in MHG meaning "*vielmehr*", "*viel eher*; (much) rather". "*A halt B*", in its modal modification, is thus equivalent to "*B eher wahr als A*" or "*B more probably/much rather true than A*". The original comparative morphology, still extant in Gothic *haldis* "potius; much rather", was probably not understood as comparative any more in OHG. In any case, in MHG the generalised relative clause connectors *swie(z) halt* or *swaz halt* (or *ob halt*) were no longer dyadic functors, but served as complementizers to sentences. Although one of the two original propositional functors (*swiez*, *swaz*) was lost in the later development *halt* remains as a component in an operator with scope over the whole proposition of an embedded (generalised) relative clause. As long as the comparative meaning was still felt to hold, this form of syntactic reanalysis (from a dyadic to a monadic relator) is accompanied by the following semantic reanalysis: "much rather S" (S = embedded clause). From this the reading "despite/although S'" (S' = the matrix sentence) developed. This is in line with what Hentschel conjectured (i.e. "wenn/was/ob auch; although, whatever", see Hentschel, 1986: 100f.). Cf. also Paul, Moser & Schröbler (1969: 177). It goes without saying that these three steps in the development of the MP *halt* were not possible until the original linear restriction of the comparative adverbial (including the generalizing relative complementizer) was lifted so that *halt* could move freely in the MF developing during MHG, cf. the following example from Hentschel (1986: 103f.):

(49) [$_S$ er sî *halt* [$_{S'}$ swer er sî]]
"may he then be whoever he may"

which presupposes a prior

(50) [$_S$ er sî [$_{S'}$ swer *halt* er sî]]
(same meaning as in (49))

with a transport of *halt* from the matrix clause in (49) into the relative pronoun *swer halt* of the dependent clause in (50). The step taken in (49) presupposes the complete emancipation of the particle lexeme *halt* from any lexical meaning restriction. *Halt,* in a way, moved from the embedded clause into the embedding clause much in the way that a negator can move into the matrix clause from the embedded clause (*Neg*-transport). See (51):

(52) [I believe [he is *not* there]] → [I do *not* believe [he is there]]

It is not totally unimaginable that once the meaning of "ja auch; eben (auch)" was established, the particle came to be influenced by the lexical homonym meaning "stop". This would make plausible the meaning paraphrase often found in the literature: "invariability; known; obviousness".

7.4. The MP NHG *ja*

Together with *doch, ja* is the most frequent lexeme in the German language. Likewise, it can be traced back in its use even to Gothic, prominently as an interjection, that is in vocative function. (Brugmann, 1911: 328; 333; Grimm, 1877: 2187.) Other occurrences, however, admit the interpretation of Gothic *ja(i)* as a MP, i.e. in the sense of a meta-commentary.

The same picture emerges for *ia* in OHG (cf. in Tatian and Otfrid) as for a) Latin *nonne*, in rhetorical questions, b) in negated sentences for Latin *numquid*, and c) in positively asserted sentences as an adverbial directly before the finite verb translating Latin *verumtamen (dico vobis)* (see Hentschel, 1986: 86). To all appearances, the two Latin particles *nonne* and *numquid* are rendered by the same form *ia*, but in different functions in OHG. The equivalent to *nonne* is OHG *ja* in an adverbial function meaning "*fürwahr, wahrlich*; truly, indeed". See the following example following Hentschel's (1986:84) material:

(53) (OTFRID III.23, 33)
ja sint, quad er, bî nôti zuelif dagoziti?
yes are, spoke he, exactly twelve day times
"It is alas not exactly twelve hours?"

Beyond Hentschel's reading (1986: 84 "sind es nicht genau zwölf Stunden?") it is also possible to assume that the OHG *ia* (for Latin *nonne*) renders the illocutive force of the Latin text, since *nonne* in the interrogative clause presupposes a positive assertive answer. What Otfrid translates is the expected, presupposed positive answer, and not, however, the presupposing question that is articulated in Latin. We see thus that the correct pragmatic equivalence between Latin and OHG entails a change in sentence type, namely from a rhetorical question premodelled in Latin to an OHG answer, which was only presupposed by the rhetorical question in Latin. (The question mark in (53), naturally, is not Otfrid's punctuation.)

Numquid, on the other hand, to all appearances entails a different linear order. See the following example:

(54) (OTFRID IV.12, 19–20; from Hentschel, 1986: 84)
... *ia* íh iz, drúhtin, nibin
ia iz herzá mini rurait noh súlih walo fuariti!/?
"In fact, master, it is not me. Quite truly, my heart not contemplates such evil"

Note that this is not an adverbial position, but, rather, a position in the far left forefield of the sentence. This position does not enforce inversion of subject and predicate, but rather the verb final position. This gives it the linear order of an embedded sentence rendering *ia* in the function of an embedding conjunction.

The third type of occurrence is truly that of a MP. See again Otfrid:

(54) (OTFRID I.2, 1; from Hentschel, 1986: 85)
uuola drohtîn mîn, *ia* bin ih scalc thîn.
"Very well, my master, I am MP your knight"

This modal reading is not so much enforced by the linear position (which is sentence-initial), but rather by the absence of textual support either for a reading of *ia* in the function of an adverbial or of a conjunction.

The picture is replicated in MHG. See the following example from *Minnesangs Frühling:*

(55) (MINNESANGS FRÜHLING 202, 3; from Hentschel, 1986: 109)
des enwil ih nimmer wîbe
mêr getrûwen einen tac
waz red ih? jâ sint si guot.
"Therefore I would never trust a woman one day long. What do I say? They are *truly* good".

Ja is again to be found only in sentence-initial position. This appears to hold generally in MHG, in contrast to the MF position in NHG which developed only later (Wells, 1985: 256 f.; Hentschel, 1986: 110). What (55) shows is that *ja* expresses contradiction to an explicit negative statement, but the exact meaning equivalence and functional type can be derived most clearly from the following example in the Nibelungenlied:

(56) (NIBELUNGENLIED 176,4; from Hentschel, 1986: 108, fn. 59)
jane wart den Sachsen geriten schedelicher nie
yes-no was to the Saxons ridden more damaging never

Jane can not be anything but the adverbial *ja* "surely" extending scope over the sentence negator *ne*, thus rendering the following sentential sense: "Certainly there was never a war worse for the Saxons." We can only assume that the MP with the somewhat thinned-out lexical meaning derived from this emphatic, explicitly refuting, usage. *Ja* as an adverbial in initial sentential position always has an anaphoric, or deictic, function, in that it takes up a whole proposition (or at least part of it, e.g. just the rheme of this proposition) referring to it by way of a modal appreciation.

7.5. Summary

We have looked at four different particles. Their development in the earlier, well documented stages of German has rendered a non-compact picture in the sense that a common line of development of MPs cannot be traced through time. These are the main points to be registered. They are to be compared with the rough reconstructions in Table IV in section 5 above.

1. *Properties of the early lexical representatives*: There are lexemes such as *doch* appearing in modal function and in all adverbial positions from the earliest documented evidence onwards (as early as Gothic). The original, non-modal lexeme can be of varied types: A truly lexical adjectival as in the case of *eben*; a comparative particle of somewhat diminished lexical content as in the case of *halt*; a clear vocative interjection and adverbial such as in *ja*; and the double function of a MP as well as a logical, sentence-connecting adverbial in the case of *doch*.

   This is in line with what we have found to hold synchronically. The fact that in two cases (cf. *eben* and *halt*) no early MP function could be registered shows that for a number of lexemes the developmental status of MPs is later than that of the adverbial modal.

2. *Syntactic reanalysis*: A clear case of syntactic reanalysis could be found with *halt*: What was originally part of a relative definite pronoun, i.e. bound in a compound relative pronoun (e.g. *swiez halt*), came to be a complementizer particle as a first developmental step, sliding into an adverbial position in its own functional status in a later step. Note that this is only possible if we assume that

halt always carried some full lexical meaning along with these steps of syntactic reanalysis.

This last observation is again in support with our crucial claim that the original (lexical or functional) meaning determined the development under syntactic reanalysis and semantic change of the modal meaning. It is to be noted that the steps of distributional reanalyses in the case of the four lexemes were not identical. This was hardly to be expected given their different functions and meanings. Yet, in each case, the following three conditions can be found to play a role: (1) "emancipation" from the logical, or referential, contextual co-options that the original sentential connectors required; (2) extension of the co-occurrence with specific sentential types; and, most prominently, (3) the formal and distinguishing constraint of the MF-position.

3. *Graduality of semantic development*: Unless there is a logical/modal meaning from the very start, the development seems to presuppose the following sequentiality: Topological meaning > temporal sequence > logical relation. The development of *eben* is typical.
4. *Text deixis*: What appears to be instrumental in the development of MPs is the text-anaphoric, or text-deictic, function as in the case of the epistemic, vocative *ja*. Such a text-deictic anaphor takes up a prior proposition, or part of it, to extend to an epistemic, metacommunicative comment on it. It is this commenting, referring, and anaphoric function that yields, in all probability, the function of the MP with its metacommunicative force.
5. *The MF-condition*: It is common textbook knowledge that verbal bracketing (and, consequently, the structural MF) had not yet developed categorically in Middle High German. As can be also observed in our examples from (46) onwards, the finite predicate can also occur in V-second in dependent clauses (besides V-3 — as in English! — V-final, and, in fact, any other position in the sentence). Likewise, the V-second position of the finite predicate is not yet fixed in the independent clause. Nevertheless, there is a clear propensity toward V-2 in the MHG independent clause (as it is in the OHG indenpendent clause). Note that left-peripheral position, V-2, alone is not a sufficient syntactic condition for the emergence of MPs. For example, Old English also had V-2, much like modern German, but unlike Modern English, but it had no MPs. The crucial factor in this undecided phase of the German

linear syntax is that where MP suspects occur in the MHG clause they avoid topicalized positions in the sentence, which are in principle open for adverb(ial)s. See, for example, *dêswâr* "truly", *zewâre* "truly", *(ih) waen(e)* "presumably", and *sicherlîchen* "certainly", all of which are possible in topicalized and even left-dislocated positions in the MHG clause (Seidl & Schophaus, 1979: 156f.). On the other hand, *halt*, *(vil) lîhte* "perhaps", *et* "as claimed before" (cf. German *eben*), clearly in non-referential use, hold clausal positions in other parts of the sentence. This is what our examples from (46) onwards confirm beyond doubt. In other words, nothing in our material nor in the literature contradicts the claim that MP uses had a strong propensity for a sentence-central distribution even in a phase when the MF was not fully developed yet. This also appears to explain why old MPs prove to comply fully with the newly arising MF distribution in Early New High German (see Wells 1985: 256ff.).

8. CONCLUSIONS: TOWARD A THEORY OF GRAMMATICIZATION

As has been noted in the opening section, the grammaticization of MPs in the West and North Germanic areas (excluding, most prominently, English) sheds new light on the range of this process to the extent that it deviates partly from the grammaticizing properties usually observed and summarized in section 2 above. This is how and where they differ from other phenomena under grammaticization: First, their lexical forms are not reduced phonologically; second, the rise of particle functions does not entail total semantic bleaching of the morpheme or the reduction to purely grammatical morphemes; and third, in relation to the former points, their pragmatic and discourse functional ranges extend considerably.

We have found that MPs in German as well as other West Germanic and North Germanic languages derive from what are invariably homonymic lexical elements, through a loss of semantic complexity. This is in line with one of the prominent generalizations observed for grammaticization. However, this semantic bleaching is coupled with a rise of pragmatic complexity, a fact contradicting property (d) on the list of accompanying characteristics of grammaticization (see again section 2). We have argued that the rise of MPs is a specific type of grammaticization to the extent that it is not foresee-

able that MPs will ever come to be purely grammatical morphemes. Thus, loss of pragmatic significance as an identifying condition of grammaticization is rated as secondary and dispensible in this context.

Despite this partial deviance from the list of identifying properties of grammaticization (according to Meillet 1912 and Kuryłowicz 1965; see section 2 above) there can be no doubt that, given the rigid syntactic control of the development of MPs as well as the loss of their original meanings and functions, which are still manifest in the homonymic counterparts, the development of the MPs in German (and probably in other languages as well) is to be considered a case of grammaticization in an extended sense.

The specific characteristics of the grammaticizing process of MPs in German and their earlier stages allow us to relate them to other theoretical questions that have been raised in the literature and in the introduction to this volume.

8.1. Grammaticization vs. language change

Grammaticization is no doubt a specific form of language change. For the ease of exposition, let us list its most prominent and indispensible characteristics in a slightly new form (cf. the parameters of grammatic(al)ization in Lehmann, 1982): (1) *syntagmatic shift* from lexical to grammatical (or from less grammatical to more grammatical) status; in Lehmann's terms, then, the MP loses its full linear freedom, although it does not reach the severely restricted grammatical status of a clitic or an affix[9]; (2) loss of semantic complexity; in particular, loss of referential (adverbial and adjectival) meaning in favour of illocutive force; this implies reduced *paradigmatic* variability; (3) increased constrainedness by grammatical rules (such as co-occurrence restrictions with respect to sentential types); this means reduced *syntagmatic* variability; and (4), the emergence of a *formal paradigm* (in terms of specific linear constraints), of which the arising form is a member. If the emergence of MPs is to be an instance of grammaticization as distinct from the more general language change it will in any case have to satisfy these necessary conditions.

The most crucial aspect of the grammaticization of MPs is the linear MF-constraint. This restriction also meets the paradigm condition: From the longitudinal, diachronic discussion it has emerged that, despite the fact that MPs departed from different grammatical categories, different sentential functions, and different semantic meanings, all MP-elements abandoned their distinct and different syntactic properties and aligned with the single MF-

restriction. In this way, they constitute a unique and easily identifiable formal paradigm.

Last but not least, it is to be noted that the paradigmatic effect is reinforced by the fact that the original, pre-MP lexemes have always retained their full meanings and syntactic functions. In other words, the identification of the specific functions and meanings of MPs is facilitated by virtue of the distinct formal properties that characterize the MP-paradigm in a unique way. The grammaticization of MPs may thus be a highly specific type, which may not be found to agree in every respect with other types of grammaticization. But it is grammaticization for the very reason that it is paradigm-forming on formal grounds.

The latter property also sets the grammaticization of MPs off from pure semantic change, which is restricted to semantics; i.e. semantic change is not constrained by formal, paradigmatic properties (see Hock, 1986, ch. 8 for a discussion).

8.2. Towards a typology of stages of grammaticization

What has been found to hold for the emergence of MPs in German is in line with other findings based on independent evidence. Thus, where a process of abstraction from a lexical source was involved the notion of locality, or topology, preceded the related temporal notion, which in turn gave rise to logical relations (see the case of *eben* or *denn*; likewise, but not discussed here, the case of colloquial and dialectal *nachher* [nocha]). The specific MP-development, however, involves one more step, namely that of illocutive force or discourse-commentary. This, then, induces the following undirectional cline:

(57) LOCALISTIC > TEMPORAL > LOGICAL > ILLOCUTIVE/ DISCOURSE FUNCTIONAL

A very similar generalization has been observed to hold for the development of epistemic meaning of modal verbs in the historical development of English (Traugott, 1989:8; see also Traugott and König (Volume I)). This observation made about wholly independent phenomena lends extra credence to the thesis in (57). Referring to a terminology introduced by Bybee (1987) we can say that the grammaticization of MPs does not yield *grams* (= grammatical morphemes), but *pragmagrams* (= pragmatic morphemes).

It is to be noted that another observation made in Traugott and König (Volume I), namely that grammaticization in *its early stages* involves an

increase in pragmatic meaning, bears directly on the emergence of MPs. If the assumption made by Traugott and König proved to be an empirically valid generalization, the emergence of MPs in German could be seen as a process of grammaticization still in its prime — which could also be seen as plausible by force of the fact that the strong syntactic correlation, the emergence of the verbal bracket (and the MF), was not completed until the 18th century (Wells, 1985: 256ff.). Yet, whether MPs will develop further, what their destiny will be, and under which circumstances such a process should be seen to be finalized can only be a matter of pure speculation.

Naturally, the illocutive, or commentative, characteristic of MPs is derivable directly from the context-functional use of conjunctions as source (and partner) lexemes (cf. the case of *denn* and *aber*) as well as indirectly from the lexical adverbials whose propositional reference was found to extend to implied or presupposed contexts (see, most notably, the cases of *ja* and *halt* in the diachronic discussion). Clearly, the hybrid status of any element between a categorially bleached MP and an adverb or conjunction in its full categorial standing, as well as the meaning relation between these corresponding lexemes, must contribute to the special function of MPs. It remains to be seen, however, whether the particular text-anaphoric, or text-deictic, steps in the diachrony of German MPs are generalizable for a theory of grammaticization.

8.3. Motivation and teleology of grammaticization

It is interesting to see that this point of inception can be loaded with extra-linguistic teleological weight so that it attains characteristics of a motivational starting point. Krivonossov (1986; 1987) has noted that MPs occur with significantly higher frequency either in those contextually established syllogisms that share *negated* modalities (i.e. syllogisms of the "Celarent" type or the "Camestres" type), or with those syllogisms where the second premise *follows* the conclusion. The latter serial order collocates with a use of MPs that is roughly 4 times as high as with that syllogistic order, i.e. where both premises *precede* the conclusion. Contextually established syllogisms with *positive* assertions occur with significantly less frequency. This is enough to make the following assumptions in terms of motivation and teleology:

(1) MPs are prominent signals for enthymemic (i.e., syllogistically incomplete) text portions, or for what the speaker thinks can be

enthymemic for the addressee (see Abraham, 1989b for reconstructions of the complete chains of inference induced by different MPs).

(2) "languages" — or probably more precisely, language cultures as well as styles of behavior — differ typically with respect to their propensity of avoiding, or at least of making visible, the existence of enthymemes (incomplete inferential steps) in texts thereby allowing the addressee to have a sense of being brought abreast of all of the information that the speaker possesses. In fulfilling this function, the use of MPs creates transparency, initiates trust, and it helps to clarify positions both socially and with respect to the flow of information in the discourse texts.[10]

It is to be noted that the function of indicating, or filling in, enthymemes across text-sections has fundamental rhematic status for the discourse. This ties in with what we have observed in section 6 above. While this in a way poses a plausible motivation and teleology for the development of MPs in German it must remain open as to what that means for a general theory of grammaticization.

NOTES

1. Other classifications yield somewhat different pictures: See Kemme (1979: 18–20) and Hentschel (1986: 198–203). These differences mirror the elusiveness, semantic more than syntactic, of the phenomenon under discussion, which may perhaps be taken to signal an ongoing process of which the terminal status has not yet been reached. See section 8.1. for such a speculative argument.

2. However, it is often not easy for native speakers of German to decide whether such a form is polysemous or homonymic. It is to be noted that the distinction between a polysemous and a homonymic lexical relation will have to be based purely on synchronic linguistic intuition. For example, the adversative adverb, or adversative adverbial conjunction *aber* ("but") retains an adversative meaning, albeit weak, in all of its uses as an MP. Likewise, *bloß* ("naked") implies, in all of its modal uses, a faint "stripped of everything else" sense operating upon the proposition in which it is used. This is, by way of example, what is meant by the relation of polysemy. Homonymy, on the other hand, pertains in the case of *halt* ("stop") or *eben* ("flat"), since none of the modal uses can be inferred from the meaning of the full lexical unit. (Abraham 1989b). The very fact that there is always a word pair, i.e. a modal particle *and* a full lexical element of identical phonetic form, is crucial to the assumption that grammaticization is at play in the first place.

3. See, for example, Luraghi (1988) for a discussion of particle elements in Old Greek, Hittite, and Latin.

4. See Höhle (1986) and especially Scherpenisse (1986) for a synthesis and comparison of

assumptions, both traditional and modern, about the basic order of the clausal constituents in German.

5. It would be attractive to adopt Fanselow's restriction for NPs in the German sentence structure, namely that NPs under focus weight cannot move to a pre-IP position due to ECP-constraints (Fanselow, 1988: 341–346), and to extend this constraint to any, i.e. also non-NP, focus-carrying element (see Abraham, 1986b for an argument supporting this categorial extension). But it is easy to see where the pitfalls on this path of reasoning lie. First, modal particles never occupy argument positions, which, however, has always been held to be a canonical requirement for establishing anaphorical binding conditions for movements to the pre-IP field. And, second, modal particles share certain strictly local linear restrictions with SPs and ADVs; see B. above. It seems more plausible to assume that what modal particles share with SPs are local properties of scope extension, except that the import of the scope distinction is of a discourse-functional sort in the case of modal particles, whereas in the case of SPs it is restricted to grammatical categories. Chances are that any final account will probably be of a complex nature. Cf. section 6 below for an extension of the scope property of modal particles and the discourse function *rheme* that the modal particle focusses on.

6. If the facts from the old Indo-European languages can be trusted, the correlation would also include word-order free and V-1 & V-last languages (Vedic and Hittite being V-last, Old Greek and Latin being word-order free; see Luraghi (1988).

7. This is not entirely true, however, since positions after the content field (CF) can be occupied as well, although the restrictions on this trait are not entirely clear. Particle meanings in this position are invariably modal, yet thinner than those in the nexus field (Askedal, 1986:86ff.; Nyholm, 1986).

8. Luraghi (1988) has shown that particles can indeed occur in very early stages of Indo-European such as Hittite, Old Indic, Old Greek, and Latin. A step this far back into the history entails some disadvantages, however, as we cannot be certain about the syntactic structure of some of these languages. Hittite, for all that is known to us, is an SOV-language. Greek, to be sure, has a totally free word order, notably as far as the finite verb is concerned. Latin, finally, is at best *statistically* V-initial. What we would want to find, however, is something like the verbal bracketing that exists in German. This was our crucial correlation in sections 3 and 4 above. But we have no clue to any syntactic property of these languages that might replicate the particular syntactic structure of German (and Dutch and West Frisian) or the Mainland Scandinavian languages. This leaves us with the possibility that while our correlation between verbal bracketing (implying the existence of a middle field with special discourse-based properties) is correct it may not be a necessary condition. What may be the crucial common denominator between the middle field and totally non-configurational (word-order free) languages is the deeper, non-syntactic property of "configurating", or structuring, the sentence in terms of discourse properties such as *theme* and *rheme*.

9. In distinction to Standard German, MPs in German dialects do in fact develop reduced forms of MPs or reduce them even to clitics. See the following examples:
   (i) Der hat *fei* gar ka Ahnung (Bavarian)
   der hat MP gar keine Ahnung
   "He has not the faintest (idea)!"

fei corresponds to NHG *vielleicht* reduced in MP-function. The epistemic adverb, however, can never occur in the reduced form; see (ii).

(ii) *Vielleicht*/**Fei* hat er gar ka Ahnung?
vielleicht hat er gar keine Ahnung?
"Maybe he has no suspicion at all?"

A somewhat similar, but functionally different particle form, reduced from the same NHG *vielleicht*, exists in Carinthian.

(iii) Der hat *leicht* a Ahnung!
Der hat MP eine Ahnung!
"He has not the faintest!"

Again, if the element in question is topicalized as an epistemic adverb it must occur in its unreduced form.

(iv) *Vielleicht*/**Leicht* hat er gar keine Ahnung
"Maybe he has no suspicion at all"

Another such reduced and clitisized form is to be found in Viennese (and, in fact, in all of Austrian-Bavarian).

(v) Was hat-a-*n* gmacht? Was hat-er-MP gemacht?
"What did he do then?"

Note the unreduced forms in different functions.

(vi) Was hat er *d e n n* gmacht?

(vii) Was hat er dan gmacht?
"What else did he do?"

Not all dialectal MPs develop such reduced or clitic forms, however.

10. It is perhaps fitting to point out that language-cultures may differ with respect to the amount of enthymemes occurring in texts. In other words, there may be languages that avoid inferentially incomplete texts in the first place, thereby foregoing the need to fill in modal particles or the like. Or else there may be languages that embody little need to fill the reader, or addressee, in on presuppositions shared by the writer or speaker, while other languages have more rigid requirements in this area. Note that this probably characterizes the degree of cultivation, or concern for the purity, of a language rather than the language system. But this may eventually have an effect upon the system of a language, such as when one language develops modal particles, while others do not. As far as I can see, little, if anything, is known in terms of these perspectives.

REFERENCES

Abraham, Werner. 1981. "Partikeln und Konjunktionen — Versuch einer kontrastiven Typologie Deutsch-Niederländisch." In *Partikeln und Deutschunterricht*, Harald Weydt (ed.), 168–188. Heidelberg: Julius Groos.

Abraham, Werner. 1985. "Wortstellung im Mittelfeld im Deutschen." In *Erklärende Syntax des Deutschen*, Werner Abraham (ed.), 27–52. Tübingen: Gunther Narr, Studien zur deutschen Grammatik 25.

Abraham, Werner. 1986a. "Die Bedeutungsgenese von Modalpartikeln: die bedeutungs-

konstituierenden Variablen 'Kontrastdomäne' und 'Kontext'." *Groninger Arbeiten zur linguistischen Germanistik (GAGL)* 27:1–44. Groningen: Dept. of German.
Abraham, Werner. 1986b. "Word order in the middle field of the German sentence." In *Topic, Focus, and Configurationality*, Werner Abraham and Sjaak de Mey (eds), 5–38. Amsterdam: John Benjamins (= Linguistik Aktuell 4).
Abraham, Werner. 1988a. Thema-Rhema-Gliederung und Thema-Rhema-Empirie. Paper read at the Annual Meeting of the Societas Linguistica Europaea, July 1988, at Freiburg/Breisgau.
Abraham, Werner. 1988b. "Vorbemerkungen zur Syntax der Modalpartikeln." *Linguistische Berichte* 118:443–465.
Abraham, Werner. 1989a. "Ergative Subjekte, die Partitivlösung und die DP/NP-Frage." *Groninger Arbeiten zur germanistischen Linguistik (GAGL)* 29:161–189. Groningen: Dept. of German.
Abraham, Werner. 1989b. "Discourse particles in German: How does their illocutive force come about? Between a maximalistic and a minimalistic position." In *Discourse Particles*, Werner Abraham, (ed.). Amsterdam: John Benjamins.
Abraham, W. 1991. Review of Thurmair 1989. *Linguistics.*
Altmann, Hans. 1976. *Die Gradpartikeln im Deutschen: Untersuchungen zu ihrer Syntax, Semantik und Pragmatik.* Tübingen: Max Niemeyer, Linguistische Arbeiten 33.
Altmann, Hans. 1978. *Gradpartikel — Probleme.* Tübingen: Gunther Narr, Studien zur deutschen Grammatik 8.
Askedal, Jan Ole. 1986. "Topologisk feltanalyse koderingssystemer og pragmatiske funksjoner." *Nydanske Studier & Almen Kommunikationsteori* 16/17:18–55.
Askedal, Jan Ole. 1987. "Überlegungen zur sprachtypologischen Beschreibung norwegischer Partikelstrukturen." *NOR-Skrift: Arbeitsskrift for nordisk språk og litteratur* 56:79–113.
Bayer, Josef. 1985. "Adjazenz und Kettenbildung: Bemerkungen zur Syntax der Gradpartikeln." Unpublished paper, Max Planck Institute, Nimway.
Brugmann, Karl. 1911. *Vergleichende Laut- und Stammbildung und Flexionslehre nebst Lehre vom Gebrauch der Wortformen der indogermanischen Sprachen.* Straßburg. (= K. Brugmann and B. Delbrück 1893–1916. *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*, vol. 2, 2. Abteilung.)
Bybee, Joan L. 1987. "The grammaticization of tense and aspect across languages." Plenary lecture at the Annual Winter Meeting of the *Linguistic Society of America*, San Francisco, December 27th, 1987.
Bybee, Joan L. and William Pagliuca. 1987. "The evolution of future meaning." In *Papers from the VIIth International Conference on Historical Linguistics*, Anna G. Ramat, O. Carruba, and G. Bernini (eds), 109–122. Amsterdam: John Benjamins.
Bybee, Joan L., William Pagliuca and Revere D. Perkins. Volume I. "Back to the future."
Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.
Diderichsen, Paul. 1962. *Elementaer Dansk grammatik.* Copenhagen: Gyldendal.
Fanselow, Gisbert. 1988. "Word syntax and semantic principles." In *Yearbook of Morphology*, Geert Booij and Jan van Maarle (eds), 95–122. Dordrecht: Foris.
Franck, Dora. 1980. *Grammatik und Konversation: Stilistische Pragmatik des Dialogs und die Bedeutung der deutschen Modalpartikeln.* Königstein/Ts.: Skriptor Verlag.
Geeraerts, Dirk. 1985. "Cognitive restrictions on the structure of semantic change."

In *Historical Semantics: Historical Word-Formation*, J. Fisiak (ed.), 127–154. Berlin: Mouton de Gruyter.
Geeraerts, Dirk. 1988a. "Where does prototypicality come from?" In *Topics in Cognitive Linguistics*, B. Rudzka-Ostyn (ed.), 207–230. Amsterdam: John Benjamins.
Geeraerts, Dirk. 1988b. "Cognitive grammar and the history of lexical semantics." In *Topics in Cognitive Linguistics*, B. Rudszka-Ostyn (ed.), 647–678. Amsterdam: John Benjamins.
Gornik-Gerhardt, Herbert. 1982. *Zu den Funktionen der Modalpartikeln 'schon' und einiger ihrer Substituentia*. Tübingen: G. Narr.
Grimm, Jakob and Wilhelm. 1854–1954. *Deutsches Wörterbuch*. Leipzig. (1860: part 2; 1862: part 3; 1877: part 4)
Heine, Bernd and M. Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: H. Buske.
Helbig, Gerhard and W. Kötz. 1981. *Die Partikeln*. Leipzig: VEB Verlag Enzyklopädie.
Helling, C. 1983. "Deutsche Modalpartikeln und ihre italienischen Entsprechungen." In *Parallela*, M. Dardano, W.U. Dressler, and G. Held (eds), 376–384. Tübingen: Gunther Narr.
Hentschel, Elke. 1986. *Funktion und Geschichte deutscher Partikeln: ja, doch, halt, eben*. Tübingen: Max Niemeyer, Reihe Germanistische Linguistik 63.
Hock, Hans Henrich. 1986. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter, Trends in Linguistics. Studies and Monographs 34.
Höhle, Till N. 1986. "Der Begriff 'Mittelfeld': Anmerkungen über die Theorie der topologischen Felder." In *Kontroversen, alte und neue: Akten des VII. Internationalen Germanisten-Kongresses Göttingen 1985*, Volume 3, Albert Schöne (ed.), 329–341. Tübingen: Max Niemeyer.
Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen*. Tübingen: Max Niemeyer, Linguistische Arbeiten 138.
Jacobs, Joachim. 1989. "On the semantics of modal particles." In *Discourse Particles*, Werner Abraham (ed.). Amsterdam: John Benjamins.
Kemme, Hans M. 1979. *'ja', 'denn', 'doch' usw. Die Modalpartikeln im Deutschen*. Goetheinstitut München.
Krivonossov, Alexander T. 1986. "Zur logischen Struktur des Textes." *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 4: 415–435.
Krivonossov, Alexander T. 1987. "Modale Partikeln und logische Schlüsse in der natürlichen Sprache." Program Booklet of the International Conference on Linguistik Particles, Berlin (West), p. 61.
Kuryłowicz, Jerzy. 1965. "The evolution of grammatical categories." In *Esquisses Linguistiques* II, 38–54. München: Wilhelm Fink.
Labov, William. 1963. "The social motivation of a sound change." *Word* 19:273–309.
Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
Lehmann, Christian. 1982. "Thoughts on Grammaticalization. A Programmatic Sketch." *AKUP* 48, vol. 1. Cologne: Institut für Sprachwissenschaft der Universität zu Köln.
Lehmann, Christian. 1985. "Grammaticalization: Synchronic variation and diachronic change." *Lingua e Stile* 20(3): 303–318.
Lehmann, Christian. Volume II. "Grammaticalization and related changes in contemporary German."
Lexer, Matthias. 1869–1872. *Mittelhochdeutsches Wörterbuch*. Leipzig: S. Hirzel.

Luraghi, Silvia. 1988. "Clitic placement in Indo-European languages." Paper read at the Annual Meeting of the Societas Linguistica Europaea, July 14, 1988, at Freiburg/Breisgau.
Lütten, J. 1977. *Untersuchungen zur Leistung der Partikeln in der gesprochenen deutschen Sprache*. Göppingen: W. Kümmerle.
Meillet, Antoine. 1912. "L'évolution des formes grammaticales." *Revista de Scienza 12*. (Reprinted in Meillet, 1948, *Linguistique Historique et Linguistique Générale*, 130–148. Paris.)
Nyholm, Leif. 1986. "Svansdubbleringen den västnyländska." *Xenia Huldeniana: En vänskrift tillägnad Lars Hulden på hans 60-arsdag* 5.(2):205–217. Helsingfors.
Panet, Edith. 1981. "Partikeln im Unterricht — Erfahrungen mit englischen Studenten." In *Partikeln und Deutschunterricht*, Harald Weydt (ed.), 101–110. Heidelberg: Julius Groos.
Paul, Hermann, Hugo Moser, and Ingeborg Schröble. 1969. *Mittelhochdeutsche Grammatik*. Tübingen: Max Niemeyer.
Platzack, Christer. 1986. "Diderichsens positionsschema och generativtransformationsgrammatik." Nydanske Studier & Almen Kommunikationsteori 16/17: 161–170.
Scherpenisse, Wim. 1985. "Die Satzstrukturen des Deutschen und Niederländischen im Rahmen der GB-Theorie: Eine Reaktion auf Marga Reis' Doppelkopfkritik." In *Erklärende Syntax des Deutschen*, Werner Abraham (ed.), 311–322, Tübingen: Gunther Narr.
Scherpenisse, Wim. 1986. *The Connection Between Base Structure and Linearization: Restrictions in German and Dutch*. Frankfurt am Main: Peter Lang.
Schwartz, D. 1986. *Gefühle erkennen und positiv beeinflussen*. Landsberg/Lech.
Seidl, Karl O and R. Schophaus. 1979. *Einführung in das Mittelhochdeutsche*. Wiesbaden: Athenaion, Studienbücher zur Linguistik und Literaturwissenschaft 8.
Shibles, William. 1988. "An analysis of German emotive particles and interjections." Unpublished paper University of Wisconsin, Whitewater.
Thurmair, M. 1989. *Modalpartikeln und ihre Kombinationen*. Tübingen.
Tomaselli, Alessandra. 1989. La sintassi del verbo finito nelle lingue germaniche. Un problema di tipologia nel quadro teoretico della grammatica generativa. Ph.D. Thesis, University of Pavia, Italy. Unpublished.
Traugott, Elizabeth Closs. 1989. "On the rise of epistemic meanings in English: An example of subjectification in semantic change." *Language* 65:31–55.
Traugott, Elizabeth Closs, and Ekkehard König. Volume I. "The semantics-pragmatics of grammaticalization revisited."
Wells, C.J. 1985. *German: A Linguistic History to 1945*. Oxford: Clarendon Press.
Wesemann, Monika. 1981. "'Das ist doch kein Problem.' Zu den dänischen Entsprechungen der deutschen Abtönungspartikel *doch*." In *Partikeln und Deutschunterricht*, Harald Weydt, 238–248. Heidelberg: Julius Groos.
Weydt, Harald. 1969. *Abtönungspartikel*. Bad Homburg: Gehlen.
Weydt, Harald (ed.). 1981. *Partikeln im Deutschunterricht: Abtönungspartikeln für Lerner des Deutschen*. Heidelberg: Julius Groos.
Weydt, Harald, Theo Harden, Elke Hentschel and D. Rösler. 1983. *Kleine deutsche Partikellehre*. Stuttgart: W. Kohlhammer.
Wolski, Werner. 1986. *Partikellexikographie*. Tübingen: Max Niemeyer.

SECTION E: SPECTRA OF GRAMMATICALIZATION PATTERNS

Areal and Universal Dimensions of Grammatization in Lahu

James A. Matisoff
Department of Linguistics
University of California, Berkeley

1. INTRODUCTION

The term *grammaticalization*, despite its heptasyllabic cacophony (the more concise *grammaticization*, or even *grammatization* would be preferable) represents one of the most important phenomena in diachronic linguistics.

Grammat(ic[al])ization is inherently a diachronic concept. It refers to a historical semantic process whereby a "root-morpheme" with a full *lexical* meaning assumes a more abstract *functorial* or "grammatical" meaning. Such processes may take centuries to complete (as Bréal suggests: *see below*), but this does not seem necessarily to be the case. We can see grammatizational phenomena occurring rapidly before our eyes in any language we choose.

Traditional Chinese grammar distinguishes between *shí-cí* "full/solid/real/true words" and *xū-cí* "empty words", so that a good Chinese translation of *grammaticalization* is *xū-huà* ("empty-izing, emptification, evacuation"). This metaphor has much to recommend it, since a hallmark of many grammaticized morphemes is that they are more conspicuous *in absentia* than they are *in praesentia*. Despite their "emptiness", their absence can immediately stamp one as a non-native speaker, or at best, as a child. (Try saying, e.g. **I going soon*, or **I read it in book*.)

The pioneer semanticist (and reputedly the inventor of the word *sémantique*, Michel Bréal, conceived of grammatization in terms of *specialization* rather than evacuation:

> We can see from this...example in what the Law of Specialization consists. Among all words of a certain kind, distinguished by a certain grammatical imprint, there is always one which is little by little drawn apart from its

> fellows. It becomes the pre-eminent exponent of the grammatical conception of which it bears the stamp. But at the same time it loses its individual value, and is no more than a grammatical instrument, one of the wheels of the phrase. When the French say *un temps plus long, une journée plus courte*, the word *plus* serves to determine the adjective by which it is followed; but by itself it has no more existence from a semantic point of view than the inflection *-ior*. In the same way we can guess the reason why the Law of Specialization demands the help of centuries before it can operate. Words have too much individual significance to adapt themselves at once to playing the part of auxiliary. It is only a long use of words in various connections that slowly prepares the mind to remove from them their superfluous value. (Bréal 1897; trans. 1964, p. 15)

A more vivid characterization of the process is Givón's term *bleaching*, which nicely captures the partial effacement of a morpheme's semantic features, the stripping away of some of its precise content so it can be used in an abstracter, grammatical-hardware-like way. The key notion here is *abstractness* (etymologically "drawing away"), a hard concept to define, since it seems to be almost a semantic primitive. One crude way of approaching it is in terms of what I have called *juxtapository productivity* (Matisoff, 1969), an index of a morpheme's combinatory possibilities in collocations — the fewer or more general the semantic features of a morpheme, the less likely they are to conflict with those of others. A full verb meaning '*send someone on an errand*' will combine with fewer other lexemes than a bleached verb that means *causative*.

Grammatization may also be viewed as a subtype of *metaphor* (etymologically "carrying beyond"), our most general term for a meaning shift, or *glissement sémantique*. Grammatization is a metaphorical shift toward the abstract. Here G. Stern's distinction (Bréal would say *répartition*) between the etymologically identical terms *metaphor* and *transfer* is useful: "metaphor" being defined as an originally conscious or voluntary shift in a word's meaning because of some perceived similarity, while "transfer" is a meaning shift so natural that it happens unconsciously. Needless to say, the ordinary speaker is synchronically unaware of the semantic history of the functors of his language — such metaphors — (better, such transfers) are "dead".

Human beings are preeminently creatures of habit, partly because of laziness, but partly because of the rational need to save effort. The scholar who eats the same lunch every day frees his mind from the time-consuming necessity of choosing something different each time. If a new collocation of morphemes proves useful (if it satisfies a *Sprachnot*), it can be stored in the

brain as a unit, much like creating a "macro" key on a computer. Grammatized usages of a word may seem strange at first, but they can quickly begin to seem normal.

It could be claimed that all grammatization involves *analogy* — as indeed grammatical change in general seems to. Yet not all grammatical changes are instances of grammatization. One can think of many instances of ongoing grammatical change, where the younger generation accepts constructions which their elders find abhorrent — these are all analogical, though not necessarily grammatizational: e.g. *bored of* (instead of *bored with*), by analogy with *tired of*, *sick of*; *so fun* (instead of *such fun*), by analogy with, e.g. *so amusing*, *so funny*; *between she and I* (instead of *between her and me*), by analogy with, e.g. *She and I are going*); *I could care less* (instead of *I couldn't care less*), perhaps by analogy with *I should worry?!*; even the infamous new "double copula" construction (*What I think is, is that he's an idiot*), perhaps by "analogical conflation" of a topicalization (*What I think is — he's an idiot*) and an extraposition (*It's that he's an idiot*).

It should also be pointed out that a morpheme may lose semantic features without becoming grammatized. The English expression *white shirt* was borrowed into Japanese as *waishatsu* in the Meiji period to mean 'formal Western white shirt', but was soon bleached (how appropriate!) to mean 'Western shirt in general', leading to such historically oxymoronic collocations as *aoi waishatsu* (lit. *"blue white-shirt"*). Similarly, the Lahu applied the expression *Kâlâ-phu* (lit. "white Indian") to the first "Europeans" they met, and this term has stuck; when they later encountered American blacks they unhesitatingly called them *Kâlâ-phu-nâ?* (lit. "black white-Indian"). These are instances of *generalization without grammatization*. One could also call this *idiomatization* (see the discussion of Lahu "idiomatic verb concatenations", 4.31 below).

What is really going on in such cases is a *total* loss of morphemic identity, or an obliteration of a morpheme-boundary. This goes far beyond what happens in grammatization, where the bleached item at least retains its morphemic autonomy. Examples of morphemic obliteration may be cited at will. When phonological reduction has not been too extreme we may talk of *morphanization* (Matisoff, 1973; 1978a), i.e. becoming an "orphan morph" (e.g. the *cran-* in *cranberry* or the *-tril* in *nostril*). More spectacular instances include the swallowing up of *eye* in *window* (< *wind* + *eye*) and *daisy* (< *day's* + *eye*), and my all-time favorite, the *-k* in *ink* (ult. < Greek *en-kaustos* 'burnt into', with the second element < *kaiein* 'burn').

1.1. The Southeast Asian linguistic area

The Southeast Asian linguistic area, defined broadly to include China south of the Yangtze, NE India, and the Himalayan region, as well as peninsular and insular SE Asia, is one of the great linguistic areas of the world, with five major language families (Sino-Tibetan, Tai-Kadai, Hmong-Mien (Miao-Yao), Austroasiatic (Mon-Khmer), and Austronesian (Malayo-Polynesian) flourishing in symbiotic profusion. Despite their genetic diversity, long periods of intimate contact have largely homogenized all aspects of the phonology, grammar, and semantics of these languages, so that it makes eminent good sense to consider phenomena like grammatization in the widest possible areal context.

Although many SE Asian languages have more morphology than is generally realized, it is true that they are overwhelmingly analytic and (except for Austronesian) monosyllabic/tonal or sesquisyllabic ("syllable-and-a-half") in structure. This is especially true of the languages that are genetically related to or culturally dominated by Chinese. In the latter category of "Sinospheric" languages are the unrelated Tai-Kadai and Hmong-Mien families, as well as Vietnamese (genetically Mon-Khmer but under intense influence from Chinese). The more than 250 Tibeto-Burman (TB) languages are demonstrably related to Chinese, but are different from it typologically, being (except for the Karenic branch) SOV — undoubtedly the original Sino-Tibetan word order.

Lahu, a language with half a million speakers on which I have been working since 1965, is typical of the Lolo-Burmese branch of Tibeto-Burman — the branch most firmly planted in the Sinosphere. (The TB languages of India and the Himalayas have more morphology, are less tonal, and are more under the cultural/historical influence of India than of China.) In this paper we will explore issues in Lahu grammatization that are of primary importance in the SE Asian linguistic area in general. The Sinospheric languages rely mostly on independent words rather than on affixes to express grammatical categories — particles, auxiliary verbs, classifiers, adverbs, etc. We will first consider the process whereby root-morphemes become grammatized into particles; then the dynamics of verb-bleaching in complex verbal nuclei; finally some phenomena which illustrate the truism that grammatical/semantic categories are continua rather than discrete, neat compartments.

All the Lahu examples are from my own fieldwork. The Thai and Mandarin examples are from standard sources, as is most of the Vietnamese, Khmer, and Hmong data (see References). The remainder of the latter, as

well as the examples from Yao, are from my field-methods classes conducted at Berkeley.

2. "PARTICULIZATION" IN LAHU

Particles are morphemes with abstract grammatical functions that cannot constitute the head of a phrase. Lahu, like other SE Asian languages, has dozens of them. They are conveniently divided into 3 subtypes: (a) *Noun-particles* (Pn) may only occur after nouns. Unlike true nouns, Pn's may not be quantified or classified. (b) *Verb-particles* (Pv) only occur after verbs. Unlike true verbs, Pv's may not be negated. (c) *Unrestricted particles* (Pu) may occur after either nouns or verbs. The Pu's may be further subclassified according to whether they occur only in non-final clauses (*non-final unrestricted particles* or Punf's), only in final clauses (final unrestricted particles or Puf's), or in both final and non-final clauses (*universal unrestricted particles* or Puniv's). Note that in Lahu, as in the other SOV languages of the Tibeto-Burman family, all particles are postpositional.

Particles convey such essential grammatical information as case relationships, verbal aspect, inter-clausal propositional relations, evidentiality, the speaker's emotional attitude, degrees of politeness, etc. Although as a group particles are obviously more abstract (or "empty") than either nouns or verbs, within the particle class itself there is a hierarchy of abstractness. As one might expect, the "unrestricted" particles are the most abstract of all — their meanings being so general as to be congruent both with nominality and verbality.

Sentence-final emotive particles are particularly richly developed in tone languages, since mere intonation is less salient when every syllable has a distinctive "melody" of its own.

Despite their centrality to the grammar of SE Asian languages, particles are generally open classes. Not only are particles readily borrowable from other languages, e.g. Lahu *šē* (Pv) 'attitude of regret' < Tai (cf. Siamese *sĭa*), they are also continually developing from fully lexical morphemes — i.e. particulization is an ongoing process. It is not always easy to say, in fact, where a root-morpheme ends and a particle begins. "In some cases the borderline between particles and members of other word classes is not too clear." (Thompson, 1965: 258; below 4.4.)

Particles (along with other high-frequency abstract morphemes like adverbs and numerals) are often marked by special phonological features. In Lahu they are likely to lose their initial consonants: *thà?* (Pn) 'accusative' is

often pronounced *à?* (below 2.11); *mâ* (Adv) 'negative' becomes simply *â*; *lê* (Punf) 'topicalizer' varies with *è*; *tā* or *tá* 'perfective; durative' (below 2.33, 4.42) frequently becomes *ā* or *á*, etc. Even if they manage to retain their initial consonant, particles may well lose their vowel quality through destressing: The Burmese negative adverb has been reduced to *mə-*, though it historically had the full vowel *-a* (cf. Lahu *mâ*); the Burmese numeral *tɪ?* 'one' (Written Burmese *tac* < PTB **g-tyik*) is usually pronounced with a schwa (*tə-*). In extreme cases, a particle may get so reduced that it fuses with the preceding root-morpheme (cf. Lahu *qay* 'go' < **ka* 'go' + *e* (Pv) 'motion away from deictic center'; below 4.51). The limit of phonological slightness is reached when a particle comes to lack any segmental phoneme at all — it is convenient to treat an abrupt intonational contour on a Lahu verb as a verb-particle, "imperative glottal stop", e.g. *qay-ʔ* 'Go!' (Matisoff, 1973:353–4).

2.1. In the noun-phrase: A noun becomes a noun-particle

2.1.1. thàʔ *'upper surface' > OBJECT MARKER (direct or indirect); efficacy/agentivity depressant*

This word has a good etymology that can be traced all the way back to Proto-Tibeto-Burman (PTB) **l-tak* ⨯ **g-tak* 'ascend; above' (see Matisoff, 1972a: #42; Benedict 1972: 52, 110, 123).

As an independent noun prefixed by *ɔ̀-*, or a constituent in noun-compounds, Lahu *thàʔ* means 'upper surface; top part':

ɔ̀-thàʔ-phɔ̂ ⨯ *ɔ̀-thàʔ-pá* 'the outer side; top surface'
khi-tɔ-qɔ-thàʔ 'instep; upper side of foot'
lìʔ ɔ̀-thàʔ 'on the book; the top surface of the book'

(1) *šɛ̂-šɪ̄ chi mì-châ* **ɔ̀-thàʔ** *bu tɛ a*
sand DEM floor surface pile put PTCL
'Pile the sand up on the floor!'

As a noun-particle (often reduced to *hàʔ* or even *à?*) (see above), this word has developed into an object-marker (Matisoff, 1973: 155–8). In this function it is used rather sparingly, only where clarity demands or emphasis is required. When both direct and indirect objects are present, *thàʔ* will follow the indirect object. This is because indirect objects are typically human (i.e. potential agents, high on the scale of efficacy), so that an explicit non-agent marker is sometimes required to exclude such an interpretation:

(2) *lì?* *chi* *ŋà* ***thà?*** *pî* *tā* *ve* *yò*
book DEM 1SG *** give PTCL PTCL PTCL
'(Someone) has given me that book.'

Without *thà?*, (1) could be interpreted as "I have given (someone) that book." This Pn may also be used in a temporal sense, marking a sort of "accusative of time":

(3) *ŋà-hi* *khɔ̄* *dɔ̂?* *ve* *qhɔ̀?* câ ve ha-pa ***thà?*** *yò*
1PL top hit REL year eat GEN month *** PTCL
'We play with tops in the month that we celebrate New Year's.'

thà? may also mark embedded sentential objects:

(4) *yɔ̂* *qɔ̀?* la ve ***thà?*** *ŋà* *dɔ̂-lɔ* *ve* *yò*
3SG return PTCL PTCL *** I hope PTCL PTCL
'I hope that he comes back.'

A clue to the semantic development of this etymon is found in the archaic Red Lahu ritual language, where one finds a compound noun-particle *ú-thà?* with a general locative or accusative meaning. The first syllable here is undoubtedly to be identifed with *ú-* ≈ *ó-* 'head' (cf. *ú-gɛ̂* 'pillow', *ó-qō 'head')*:

(5) *chɔ-yâ-vâ-yâ* ***ú-thà?*** *ɡɔ* *mâ* *ɡa*
people *** affect NEG reach
'May (these misfortunes) not be visited upon the people!'
(cf. Eng. 'upon the heads of the people')

(6) *mû-ni* *cɔ̀* *na* ***ú-thà?,*** *lɔ̀-pa* *cɔ̀* *na* *ú-thà?*
sun be above *** moon be above ***
'in the place above the sun, in the place above the moon'

It is easy to see how the basic meaning 'top surface' could have been bleached into the abstract idea of 'impingement upon; being topped by the action of the verb.'

2.1.2. lo *'road/way' > locative particle*

Unlike the previous example, where the fully lexical meaning of the etymon continues to coexist alongside its grammaticized one, the original meaning of the locative Pn *lo* can only be deduced from comparative data.

There is a well-attested PTB etymon **lam* 'road; way' (STC #87) (cf.

Written Tibetan (WT) *lam*, Written Burmese (WB) *lâm*), which has not survived with this meaning in Lahu. (The Lahu word for 'road' is now *yà?-qɔ*, perhaps related to the verb *yà?* 'descend'.) Instead, Lahu has developed a locative Pn *lo*, which I believe to be a variant of the 'road' root. *-o* is the regular Lahu reflex of **-am* (cf. WB *lam*, Lahu *lò* 'cord (measurement)'); the Lahu mid-tone, unmarked in my transcription, points to Proto-Lolo-Burmese (PLB) Tone *3, while the WB cognate reflects Tone *2, but this is no problem. Tonal developments are frequently irregular with functors, and Tone *3 is demonstrably secondary in any case.

As a Pn, *lo* is somewhat more formal than the other locative particles *ɔ̄* and *kà?*. By itself *lo* does not specify direction of motion, or even motion vs. rest; the interpretation depends on the following verb, or the sentence as a whole:

(7) há-qō ***lo*** *mi chè ve*
cave *** sit PROG PTCL
'He's sitting in the cave.' ("ESSIVE")

(8) *há-qō* ***lo*** *lò?* e ò
*** enter PTCL PTCL
'He has already gone into the cave.' ("ADESSIVE")

(9) *há-qō* ***lo*** ***tɔ̄?*** *e ò*
*** emerge
'He has already come out of the cave.' ("ABESSIVE")

Convincing comparative evidence for this grammatization comes from an obscure, moribund language of E. Nepal, Hayu (Michailovsky 1989), which has a cognate noun *lom* 'way; road', as in *lom he* (N + Pn) *'sur le chemin; on the way'* and *ek din mʊ lom* ("one + day + GEN + road") *'un jour de chemin;* a day's journey'. In Hayu also this noun seems already to have developed into a noun-particle, as in (10):

(10) *ang lom na ju!*
1SG *** EMPH descend
'Come down my way!'

Sequences like *ang lom* have changed their internal modificational structure, passing from a MODIFIER + HEAD (where the pronoun is genitivally attributed to the following noun: *my way*) to a HEAD + MODIFIER (where the bleached noun is the satellite of the preceding pronoun: *first person locatively conceived*).

2.1.3. 'object/possession' → genitive marker

Thai, Vietnamese, and Cambodian have all developed genitive markers out of a noun meaning "thing; object":

a. Thai *khɔ̌ɔŋ*
As noun:
(11a) *paj sýy* ***khɔ̌ɔŋ***
vV Vh OBJ
'go buy things'

As particle:
(11b) *mia* ***khɔ̌ɔŋ*** *phŏm*
wife GEN 1SG
'my wife'

b. Vietnamese ***của***
As noun:
(12a) ***của*** *công*
Nh public
'public funds; state property'

As particle:
(12b) *cái va-li* ***của*** *ông Nam*
CLF valise GEN Mr.
'Mr. Nam's suitcase'

c. Khmer ***rəbɒh***
As noun:
(13a) ***rəbɒh*** *nuh kee haw thaa kmaw-day*
thing DEM 3SG call QUOT pencil
'That thing is called a pencil.'

As particle:
(13b) *puəq-maaq touc* ***rəbɒh*** *kñom pii neəq nih*
friend little GEN 1SG two CLF DEM
'these two little friends of mine'

2.1.4. 'female proprietary spirit' > feminine agentive nominalizer

The original or basic meaning of Lahu *(ɔ-)šɛ̄-ma* seems to have been 'female proprietary spirit', e.g. the guardian spirit of a stream. The syllable

šɛ̄ is probably from the same etymon as *qhâ²-šɛ* 'village headman'. *šɛ* (V) 'lead' (< PLB **siŋ*^{1} ⧸ **siŋ*^{2} or **sin*^{1} ⧸ **sin*^{2}; cf. WB *sañ*, *²əsañ* 'owner: proprietor'); *-ma* is a feminine suffix.

From there the meaning branched out in several directions. In composition with the morpheme *yè* 'house' (*yè-šɛ̄-ma*) it developed two separate meanings: (a) the more "literal" one, *lady of the house*, and (b) the more idiomatic one 'house lizard', a harmless, ubiquitous creature that lives on the ceilings and walls of every house in Thailand. They are called *cîŋcòk* in Thai, and, like our *ladybugs*, are regarded with some affection.[1] From (b) the meaning developed into 'bug; critter' in general (e.g. *pû-sɛ̄-ma* 'silkworm'), and very recently the missionaries have encouraged a neologistic development into 'microbe, germ' (probably a calque on English, e.g. a *flu bug*). All this is certainly not grammatization, but merely *glissement sémantique*, metaphor, idiomatization.

On the other hand, the notion 'female proprietary spirit' developed more mundanely into 'female owner in general', and (in the direction of inalienable possession) into 'the female body':

(14a) *yɔ̂ ɔ̀-šɛ̄-ma dà² jâ*
3SG FEM.body good very
'She's got a great body.'

From there it was only a small further step to two thoroughly grammatized usages:

(a) *feminine reflexive*

(14b) *yɔ̂ **ɔ̀-šɛ̄-ma** kán chi qha pə-è ɡa te ve yò*
3SG *** work this all must do PTCL PTCL
'She has to do this work all by herself.'[2]

(b) *feminine agentive nominalizer*

(15) *yâ ɡa ga pɔ pî **šɛ̄-ma*** 'midwife'
(i.e. "she who must help give birth")[3]

This is an interesting example of the same lexeme developing both highly concrete and highly abstract senses.

2.2. Verb > noun-particle: 'Copula' > pluralizer

Here is a case that can only be "teased out" as the result of much comparative/historical labor. In "*God and the Sino-Tibetan copula*" (1985), I

traced the phonological and semantic evolution of a PST etymon **s-ray* ⨯ **s-way*, which originally seems to have had a highly abstract copular meaning. The variant in *-w-* gave rise to the most important and complex Lahu particle *ve*, which now does duty as a *nominalizer* (as such it occurs in the citation-form of verbs, like Eng. *to*) as well as a subordinator (*relativizer*, *genitivizer*) (see Matisoff, 1972b). The variant in **s-r-*, on the other hand, developed into a pluralizer in various TB languages — maximizing either the subject or the object of the verb:

> Lakher *hrai* 'plural affix for denoting company, etc.'; Mikir *háy* 'plural of non-animate nouns'; Boro *sray* 'to V everything' (e.g. *za-sray* 'eat everything', *zankri-sray* 'shake everything', *ran-sray* 'distribute everything'.

In Lahu, the descendant of this variant of the etymon is the morpheme *-hi*, used as a pluralizer with the personal pronouns and some proper nouns or nouns referring to people: *ŋà-hi* 'we', *nɔ̀-hi* 'you (pl.)', *yɔ̂-hi* 'they', *Cà-lɔ̂-hi* *'Cà-lɔ̂* and his friends; *Cà-lɔ̂* and his group'.[4]

2.3. In the verb-phrase: From verb to verb-particle

This sort of development is highly typical of SEA languages in general. The pan-areal criteria for distinguishing full verbs from verb-particles are *negatability* and *occurrability* as *the only morpheme in a VP*. The presumed development is from full verb to verb-particle. I have found no instances of the development proceeding in the opposite direction, i.e. "degrammatization" of particles to full verbs.

A neat example of the difference is the behavior of the Thai morpheme *ʔàat*, which behaves sometimes as a *full modal verb* meaning 'be able to; capable of' (in which case it may be directly preceded by the negative adverb *mâj*), and sometimes as a modal particle meaning 'likely to; imminent possibility', in which role it may not be directly negated:

(16a) *mâj ʔàat cə paj*
NEG Vmod PTCL Vh
'isn't able to go'

(16b) *ʔàat cə mâj paj*
PTCL PTCL NEG V
'probably won't go'

(In the terminology of Noss 1964, a modal verb is one which can precede another V in the same verbal nucleus. A modal particle also precedes the verb it modifies.)

Similarly, in Chrau (a Mon-Khmer language of the Bahnaric group, spoken in Vietnam), where "auxiliary verbs are distinguished from preverbal particles in that the auxiliaries are negatable" (Thomas, 1971: 144–8), the word *co·nh* may function either as a full, negatable "intentional verb" meaning 'want to' (17), or as a non-negatable preverbal particle meaning 'almost; about to' (18):

(17) *anh* ***co·nh*** *saq*
1SG Vint Vh
'I want to go'

(17a) *anh* ***êq co·nh*** *saq*
NEG
'I don't want to go.'

(18) *anh* ***co·nh*** *chu·t*
PTCL V
'I'm about to die'

But not:

*(18a) *anh êq co·nh chu·t]*
NEG PTCL V

The same semantic connection between imminence and volition is also illustrated by the range of meanings of the Lahu particle *gâ* (see 2.31 below and note 5), and the Burmese particle *hcĩ*, as well as by the Yao verb/verb-particle *ʔɔ̌ɔy* — though we are here dealing with a meaning constellation that is more universal than areal, as witness colloquial English expressions like *"Looks like it wants to rain, doesn't it*?"

2.3.1. 'think/desire' > desiderative particle

In modern Lahu, the important morpheme *gâ* combines freely with verbs to form desiderative expressions:

qay gâ 'want to go', *šĩ gâ* 'want to know', *g̈ì gâ* 'want to laugh'

Yet *gâ* is not itself a verb, since it cannot be preceded by the negative adverb *mâ* (*mâ qay gâ* 'not want to go', NOT **qay mâ gâ*), nor can it constitute a

verbal nucleus all by itself. To translate an English sentence where 'want' is the main verb (e.g. "I want mangoes"), a Lahu verb must be supplied for the *gâ* to hitch onto:

(19a) *ŋà* *má-mo-šī hɔ̂ʔ gâ* (lit. "I want to get mangoes."[5])
1SG mango get PTCL

(19b) *ŋà má-mo-šī câ gâ* (lit. "I want to eat mangoes.")
eat PTCL

Other occurrences of *gâ* give clues as to what its original non-grammatized meaning must have been. Besides its use as a desiderative particle, it also occurs as a bound morpheme in "*elaborate expressions*" (an elaborate expression is a 4-syllable collocation formed with elegant intent, where either the 1st and 3rd, or the 2nd and 4th syllables are identical), where it appears as the "couplet" of the free verb *dɔ̂* 'think'. An *elaborate couplet* consists of the non-repeating syllables in an elaborate expression, usually synonymous, antonymous, or otherwise correlative in meaning, thus:

dɔ̂-phu-gâ-phu 'pure thoughts' ("think-white-*gâ*-white")
dɔ̂-yè-gâ-yè 'be strong-minded; have equanimity' ("think-firm-*gâ*-firm")
dɔ̂-ša-gâ-ša 'be serene; have peace of mind' ("think-easy-*gâ*-easy")

Certain expressions in the archaic ritual language clearly indicate that *gâ* must once have had an independent verbal meaning like "think of conatively; desire":

gâ-pi-gâ-šɛ̂ʔ "desire-bundle" (ritual object offered in the temple, consisting of a bundle of sticks tipped with cottonwool to simulate flowers).

2.3.2. 'come' > cisative particle and marker of non-3rd person beneficiary

Sometimes the grammatization of a Lahu root-morpheme is accompanied by a morphophonemic change — notably a change in tone (which may be the synchronic reflex of a now-vanished segmental affix). The full verb 'come' is in the low-falling tone, *là* (e.g. *mû-yè là ve* 'It's raining' ("rain comes")). This verb has been grammatized into two semantically related particles, each of which is under its own distinctive tone: *la* (mid-tone (unmarked in the transcription)) and *lâ* (high-falling tone). [6,7]

The mid-toned particle *la* indicates cisativity, either in the "literal" sense of motion toward the deictic center, or figuratively, in the sense of "almost coming to pass; nearly V", e.g. *mə̀ʔ la* (V + Pv) 'blow in this direction; blow hither', *pò la* 'come flying', *cà la* 'come and eat', *ši la* 'be close to death', *cɔ̀ la* 'come to possess; get a hold of'.[8]

The falling-toned particle *lâ* has an even more abstract meaning. It is a benefactive particle indicating that the verbal action is for the benefit of (or impinges upon) a *non-3rd* person, e.g.:

chɔ lâ (V + Pv) 'chop for me/us/you'
phê lâ (V + Pv) 'release me/us/you'
phə̀ʔ-mā lâ (V + Pv) 'confess to me/us/you'

In this usage, *lâ* forms a tight paradigmatic opposition with a grammatized verb *pî* 'give', which indicates that the verbal action impinges on a *3rd* person, e.g.:

chɔ pî (Vh + Vv) 'chop for him/her/them'
phê pî (Vh + Vv) 'release him/her/them'
phə̀ʔ-mā pî (Vh + Vv) 'confess to him/her/them'

See Figure 1.

The highly grammatized nature of the opposition between *lâ* and *pî* is noteworthy, despite their different grammatical statuses (*lâ* is a Pv, tonally differentiated from its parent verb, while the *pî* is homophonous with its parent verb, so must be regarded as a post-head versatile verb ("Vv"; below 3.1). (For a brief discussion of the causative function of *pî*, see below 3.2.3.)

2.3.3. 'place/put down' > perfective, durative, permanent action

The Lahu verb *tā* 'place on; put on; set on; stand sthg. upright' has impeccable Tibeto-Burman credentials[9], reflecting PLB **ʔta*^{2} (cf. WB *thâ*) and ultimately PTB **s-ta* (cf. WT *sta* ~ *stad*):

(20) *pɛ̂-hɔ tā ve*
candle ***
'Set up a candle.'

(21) *ni-ma tê šī tí tā ve*
heart one CLF only *** PTCL
'concentrate one's efforts; be united in spirit with others'

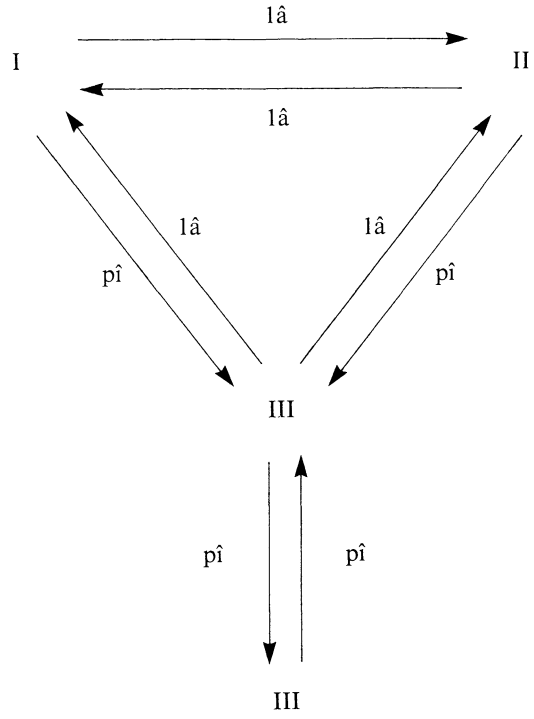


Figure 1: *The Lahu benefactive system*

This etymon has been bleached into one of the most important Lahu Pv's, in which function it indicates that the verbal action is stable, quasi-permanent, likely to last (*durative*); or that it is over and done with (*perfective*):[10]

(22) *yɔ̂-hi á-qhɔ tí chɛ̂ **tā** ve*
3PL home only stay *** PTCL
'They just stayed at home.'

(23) *cɔ̂ **tā** ve và²-ó-qō*
boil *** GEN pig-head
'a pig's head that has been boiled'

The particle *tā* does not collocate with adjectival verbs, apparently because of pernicious redundancy (adjectives *already* refer to quasi-permanent qualities or states). This fact provides one of the best criteria for distinguishing

the class of verbal adjectives from the other verbs in the language (adjectives are a subclass of verbs by virtue of their negatability (above 2.3)):

ORDINARY VERB		*ADJECTIVAL VERB*	
mâ qay	'not go'	*mâ dàʔ*	'not good'
qay tā	'had gone; went for good'	**dàʔ tā*	(nonsense)

This morpheme is straddled precariously on the interface between versatile verbhood and particulitude. (See below 4.2.1.1, 4.4)

2.4. Verb becomes a clause-particle (complementizer)

Often a verbal morpheme comes to govern a whole clause, in the manner of a subordinating conjunction. Of the many examples which could be given, we confine ourselves here to a couple of typical developments. (We will return to this topic under "Causatives", below 3.23.)

2.4.1. 'say' > quotative/topicalizer/conditional

2.4.1.1. Thai wâa *'say; quotative complementizer'*

The Thai word *wâa* still retains its full verbal meaning 'say' in certain contexts:

(24) ***wâa*** *phŏm duu-thùuk nán, mâj ciŋ ləəj*
*** 1SG despise DEM NEG true EMPH
'To say that I look down on them is simply not true.'

At the end of a non-final clause containing a verb of utterance or of cognition,[11] before the intonation break, this morpheme appears in grammaticized form as the introducer of the *quotative* clause — i.e. the clause that specifies what has been said or thought about:

(25) *phôm kô jaŋ mâj nêɛ-caj* ***wâa,****cə*
1SG CONJ still NEG sure *** PTCL
paj dâj rўy mâj
go able or not
'I'm still not sure whether I'll be able to go or not.'

Among the large number of verbs that may take a quotative *wâa-* complement are:

(a) UTTERANCE:
phûut 'speak; say', *bɔ̀ɔk* 'tell', *rîak* 'call', *sănjaa* 'promise', *tamnì?-tì?-tian* 'blame; criticize', etc.

(b) COGNITION:
rúu/sâap 'know', *khít* 'think', *rúu-sỳk* 'feel' *nêɛ-caj* 'be sure', *dâjjin*, 'hear', *wâŋ* 'hope', etc.

Sometimes there is a slight difference of emphasis, according to whether the *wâa* is taken in its more verbal, postpausal sense (26a), or its more grammaticized, prepausal sense (26b) (see Noss 1964, p. 200):

(26a) *khăw phûut səmăə, **wâa** ?aahăan mâj ?ərɔ̀j*
3SG speak always *** food NEG tasty
'They're always talking (about it, saying among other things) that the food is no good.'

(26b) *khăw phûut səmăə **wâa**, ?aahăan mâj ?ərɔ̀j*
'They always say that the food is no good.'

2.4.1.2. Khmer thaa *'say; quotative complementizer'*

As a main verb Khmer *thaa* means 'say' (27). As a quotative complementizer (28, 29) it behaves exactly like its Thai counterpart, *wâa*:

a. As main verb:
(27) *look **thaa** məc*
2SG *** how
'What did you say?'

b. As quotative complementizer:
(28) *kñom kɨt **thaa** look qayuq prəhael məphɨy-pram*
1SG think *** 2SG age about twenty five
'I think that you're about 25 years old.'

(29) *rəbɒh nih kee haw **thaa** qwəy*
thing this 3SG call *** what
'What do they call this?' ("This thing they call saying what?")

(See (13a), above.) This Khmer expression *haw thaa qwəy* is exactly analogous to Thai *rîak wâa ?əraj*, also meaning literally "call-say-what".

2.4.1.3. Lahu qôˀ *'say'* > qo *'if',* qôˀ qo *("if one says") 'topicalizer.'*

The Lahu verb *qôˀ* 'say; tell; call by a name' is routinely used to mark off a preceding clause as quoted material:

(30) *yɔ̀ kàˀ mâ qay gâ* ***qôˀ*** *ve*
3SG PTCL NEG go want *** PTCL
'He says he doesn't want to go either.'

The etymology of this verb is not certain[12], but it seems definitely to be related to a mid-toned particle *qo 'topicalizer; if, when'*:

(31) *phɨ̂ chi tê g̈ɨ̀* ***qo*** *à-thòˀ-ma te tù le*
dog DEM one CLF *** what do PTCL Q
'As for this bunch of dogs, what'll we do (with them)?'

(32) *nɔ̀ ô-ve câ* ***qo,*** *nà tù ve yò*
2SG DEM eat *** sick PTCL PTCL PTCL
'If you eat that, you'll get sick.'

(33) *ha-pa chi tê šɪ̄ pə-ò* ***qo***
month DEM one CLF end ***
'when this month is over'

Speakers seem to retain an instinctive feel for the relationship between *qôˀ and qo*, since they are sometimes used together to form a more emphatic topicalizer, lit. "if one speaks of CLAUSE":

(34) *lìˀ hê ve* ***qôˀ qo*** *ɔ̄, ɔ̀-yân g̈a*
book study PTCL *** PTCL time must
pə̀ mâ ve yò
spend much PTCL PTCL
'When it comes to studying now, you've got to spend a lot of time at it.'

2.4.1.4. Tamang ³pi-sam *and Mandarin* de huà

The development of topicalizers or conditionals from verbs meaning 'say' is documentable in many other languages of East and Southeast Asia: In Tamang (a TB language of Nepal), the sequence of morphemes *³pi + sam* (lit. SAY + IF "if one says") behaves as a unitary conditional marker, exactly like Lahu *qôˀ qo*:

(35) ^{1}*ni-ci* ^{3}*pi-sam*
's'il est parti',

"..où le verbe dire, ^{3}*pi*, tend à jouer le rôle d'une particule, d'un monème grammatical" (Mazaudon, 1973:50).

In Mandarin, the noun *huà* 'speech' may occur as the head of a relative clause (marked by the preceding subordinating particle *de*), with the whole clause then functioning as the protasis of a conditional sentence:

(36) *tā yě bù kěn lái de huà,*
3SG PTCL NEG willing come ***
wǒmen zěnme bàn ne?
1PL how do PTCL
'If he's not willing to come either, what shall we do?'

(Literally the first clause means something like "(as for) the talk of his also not being willing to come".)

Yet here too we seem to be dealing with a universal grammatizational tendency, as witness the English use of *say* to introduce hypothetical clauses: "*Say you give me two coconuts...*"

2.4.2. 'come' > purpose marker

The Chinese verb *lái* 'come' plays many special grammatical roles in the language. As an auxiliary after verbs of motion, it functions like Lahu *la*, to show that the verbal action is carried out toward the deictic center (Mand. *fēi lái* 'come flying', Lahu *pò la* 'id.') (see above 2.3.2).

As a coda to quantified expressions *lái* conveys an approximative notion: *wǔ shí lái suì* 'about 50 years old'; *liǎng mí lái gāo* 'about 2 meters high' (lit. "50 COME years", "2 meters COME tall"; cf. Eng. *give or take 50 years*).

So abstract may this verb become that it is frequently used as a pro-verb (much like English *do*), substituting for verbs of more specific meaning:

(37) *nǐ xiē-xie, ràng wǒ* ***lái*** *ba*
2SG rest let 1SG *** PTCL
'You rest — let me do (it)' (e.g. dig, pull, cook, etc.).

(38) *wǒ zìjǐ* ***lái*** *ba*
1SG self *** PTCL
(a) 'I'll do (it) myself.' (b) 'I'll help myself (to food).'

(39) *zài* ***lái*** *yí-diǎn ba*
again *** a little PTCL
'Would you like to have ("come") some more (to eat)?'

(40) *zài* ***lái*** *yí ge*
again *** one CLF
'Encore!' ("Come again (with) one!")

Finally, in its use as a subordinating conjunction, *lái* is not only abstractified but is grammaticized as well, serving to introduce a purpose-clause, or a clause that signifies intended future action:

(41) *nǐ néng yòng shénme fāngfǎ* ***lái*** *bāngzhù*
2SG can use what method *** help
ta ne?
3SG PTCL
'How are you going to help him?'

(42) *qǐng nǐ* ***lái*** *gěi dàjiā dú bào*
beg 2SG *** for everybody read paper
'Please (come on and) read the paper to everybody.'[13]

3. VERBLEACHING: FROM CONCRETE VERB TO ABSTRACTER VERB

"Verbleaching" looks as if it should be a German word, but it is actually only a jocular haplology for *verb-bleaching*, the semantic process by which a full verb undergoes abstractification, at the same time as it expands its syntactic co-occurrence possibilities, or "syntactic versatility". When this process is far enough advanced, the bleached morpheme can become a satellite to an indefinite number of heads — an auxiliary, helping, or "versatile" verb.

3.1. Verb concatenation in Lahu

Even by Lolo-Burmese standards, Lahu is remarkable for the apparent ease with which up to five verbs may be strung together or "concatenated" by simple juxtaposition to form complex verbal nuclei. This is a subject I had to begin to tackle almost a quarter of a century ago (see Matisoff, 1969; 1973: 199–265), and it has acquired an ever fresher interest in the context of the

general theory of grammatization. Lahu verb concatenation, with its minimum of morphology and its intricate idiosyncracies, operates according to an interplay of syntactic and semantic factors that resists any rigid formalization. It is in fact a "discourse-driven" aspect of the grammar.

One of the verbs in each concatenation is the verb-head or "Vh". The others (the "juxtapositorily productive", or "versatile" verbs) stand in a semantically subordinate relationship to the head, occurring either all to the head's left, or all to its right, or, often, flanking the head on both sides.

Lahu thus accomplishes in a uniform surface way what languages like English need many different structures to do. See Figure 2 (reproduced from Matisoff, 1973: 201), where the verb *qɔ̂* 'hoe' appears in simple juxtaposition with 12 versatile verbs, while the English translations use 6 different types of constructions.

complementary infinitives	*qɔ̂ ša*	'easy to hoe'	*ga qɔ̂*	'help to hoe'
-ing complements	*qɔ̂ kì*	'busy hoeing'	*tà qɔ̂*	'start hoeing'
modal auxiliaries	*qɔ̂ cɔ́*	'should hoe'	*ɡa qɔ̂*	'must hoe'
adverbs	*qɔ̂ bà*	'hoe away'	*qɔ̀ʔ qɔ̂*	'hoe again'
prepositional phrases	*qɔ̂ pî*	'hoe for smn'	*phôʔ qɔ̂*	'hoe in a group'
subordinate clauses	*qɔ̂ ni*	'hoe and see'	*ca qɔ̂*	'go and hoe'

Figure 2: *The gamut of English subordinating devices vs. Lahu juxtaposition*

The verbs in a true concatenation (there are several other types of multiverbal constructions in Lahu — see 3.13 below) form a single verbal idea, and are all deemed to belong to the same clause. They function as a semantic/syntactic unit.

How my ideas on this subject fell on deaf ears in the late 60's/early 70's! For it flew in the face of a cherished dogma of the generative panjandrums of the day — that each individual verb in a multiverbal "surface" expression must necessarily represent a separate underlying sentence. I felt obliged to defend my monoclausal analysis against the philistines:

> It might well be objected at this point that what is actually involved is a nesting of embedded sentences, with each verb in the concatenation deriving ultimately from a separate underlying sentence. For a variety of reasons[14] I find that approach cumbersome for the present purpose. Yet note that it does not really affect the substance of the argument at all whether one operates with embedded sentences or simply embedded verb strings. In either case we are still faced with the problem of stating the complex restrictions on the concatenative process, the nature of the hierarchical relationship among the concatenated entities, etc. (Matisoff, 1973:215)

And how would it account for semantic contrasts between full vs. bleached meanings of the lexeme, according to whether one or more underlying clauses were involved? For sometimes there is a real ambiguity, which proves the correctness of the unitary VP interpretation in one reading. Cf. the Lahu pair of sentences:

(43a) *yɔ̀ yɛ̀ te chɛ̂ ve*
3SG house make PROG
Vh Vv
'He is building a house.'

(43b) *yɔ̀ yɛ̀ te (lɛ) chɛ̂ ve*
3SG house make live
Vh Vh
'He built a house and is living (in it).'

(For the criterion of *lɛ*-insertability in "fortuitous concatenations" like (43b), see below 3.131.)

The non-head verbs in a concatenation get bleached semantically — occasionally in ways that may be idiosyncratic to Lahu, but more often in accordance with well-attested areal (or even universal) tendencies.

Although the grammatical structure of Lahu verb concatenations depends crucially on the semantic features of the individual lexemes involved, brute syntactic constraints also play a role. Though most versatile (= grammatized) verbs follow their Vh's, about a dozen always precede. We call these two syntactic classes *post-head versatile verbs* (Vv's) and *pre-head versatile verbs* (vV's), respectively. It is impossible to give a general semantic characterization of the vV's as opposed to the Vv's. (One such hypothetical dichotomy would be between "prerequisite" (vV) and "result" (Vv), but this does not hold water.) This arbitrariness makes it impossible mechanically to assign constituent structure to 3-verb "fore-and-aft" concatenations, which contain both a vV and a Vv in addition to the Vh (see Figure 9, below).

Concatenations containing 2 verbs may thus have either of two hierarchical structures. Either the first verb modifies the second (this is a *binary pre-head concatenation,* symbolized as vV + Vh), or the second modifies the first (*binary post-head concatenations*, symbolized as Vh + Vv).

3.1.1 Subclasses of post-head versatile verbs.

With only one major exception (see the discussion of *ɡ̈a* 'get', 3.2.2 below) the classes of vV's and Vv's are disjunct. As just mentioned, the set

of pre-head versatiles differs significantly from the post-head set in terms of number of members: There are only about a dozen vV's, and they seem to constitute a closed class, while there are several dozen Vv's, including a number of apparently recent loanwords. The two classes are also strikingly different in their internal ordering properties. The vV's have few mutual exclusions, and do not fall into semantically coherent subsets. The Vv's do have such exclusions, making it feasible to divide them into 4 subclasses (i.e. *juxtacapitals*, *medials*, *caudals*, and *variables*) on the basis of both semantic and syntactic criteria. (See Figure 3.)

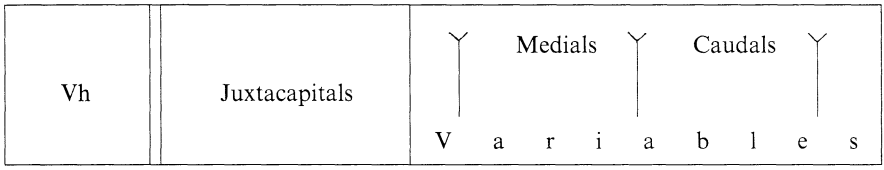


Figure 3: *Classes of Lahu Post-head Versatile Verbs*

3.1.1.1 The juxtacapital Vv's

The *juxtacapitals*, as their name implies, always occur right next to the head. They number about twelve, and have concrete meanings related to modes of motion or directionality. This is a relatively highly grammaticized group, and the members are mutually exclusive. They are as closely welded to the Vh as are such English "particles" as *out* or *away*. (See the examples in Figure 4.)

VERB	MEANING AS Vh	MEANING AS Vv	EXAMPLES	
kə	'put into, insert'	'Vh into'	*mâ? kə*	'blow into'
			cà? kə	'push into'
bà	'throw; discard'	'Vh away; Vh so sthg. is displaced; Vh in an irrevocable way'	*tú bà*	'burn away'
			qhè? bà	'chip away'
yà?	'descend'	'Vh downward'	*pu yà?*	'roll down'
			pɔ̂? yà?	'jump down'

Figure 4: Some juxtacapital Vv's

3.1.1.2 The caudal Vv's.

The caudals occur at or near the tail end of concatenations, and have highly abstract meanings, many connected with potentiality. (See Figure 5.)

VERB	MEANING AS Vh	MEANING AS Vv	EXAMPLES	
phɛ̀ʔ	'be; be the case'	'able to Vh'	*ti phɛ̀ʔ*	'able to plant'
			qay phɛ̀ʔ	'may go'
gà	'reach; arrive at'	'manage to Vh'	*tu gà*	'manage to stand'
			tɔ̂ʔ gà	'manage to emerge'
cɔ̂	'be right; correct'	'ought to Vh'	*thɔ̂ cɔ̂*	'ought to touch'
			yù cɔ̂	'should take'

Figure 5: *Some caudal Vv's*

3.1.1.3 The medial Vv's.

The medials are the most numerous and heterogeneous of the subclasses. They have highly specific meanings (e.g. *late*, *dare*, *busy*, *easy*) and, like the juxtacapitals, are mutually exclusive. This class has a number of adjectival verbs, and is open, since it includes several loanwords from Tai. (See Figure 6.)

VERB	MEANING AS Vh	MEANING AS Vv	EXAMPLES	
yàʔ	'commit an offense; be at fault'	'Vh by mistake; offend by Vh'ing'	*dɔ̂ yàʔ*	'misunderstand'
			šɛ yàʔ	'lead astray'
bûʔ	'be immoderate; be harsh'	'Vh to satiety; Vh to excess'	*yɨ̀ʔ bûʔ*	'sleep one's fill'
			dɔ̀ bûʔ	'drink plenty'
bɔ̀	'be lazy'	'be tired of Vh'ing'	*chɛ̂ bɔ̀*	'tired of living'
			g̈ɨ̂ bɔ̀	'tired of running'

Figure 6: *Some medial Vv's*

3.1.1.4 The variable Vv's.

The variables are among the most abstract of the Vv's. They have the interesting property of relatively free permutability in "multiversatile" strings (i.e. concatenations with at least 2 versatile verbs), and may occur before or after medials or caudals — with alternative orderings entailing differences in semantic interpretation (see (44) and below 3.13):

(44a) *câ phɛ̀ʔ chɛ̂*
eat POT PROG
Vh Vv Vv
'is still able to eat'

(44b) *câ chɛ̂ phɛ̀ʔ*
eat PROG POT
Vh Vv Vv
'is able to continue eating'

The meanings of the variable Vv's are best characterized as aspectual. (See Figure 7.)

chɛ̂	dwell/*continuative*	*ni*	look at/*tentative*
ci	send/*causative*	*pə̀*	finish/*completive*; *exhaustive*
qay	go/*continuative*; *inchoative*	*pî*	give/*benefactive*; *permisso-causative*
mɔ	be a long time/*durative*	*lɔ̀ʔ*	enough/*sufficitive*

We may illustrate the grammaticized meanings of these verbs by combining them all with the same Vh, e.g. *và?* 'put on; wear':

và? chɛ̂	'is wearing'	*và? cɨ*	'make/let smn wear'
và? qay	'goes on wearing'	*và? mɔ*	'has worn for a long time'
và? ni	'wear and see; try on'	*và? pə̀*	(1) 'has already put on' (2) 'everybody wears'
và? pî	'dress smn'	*và? lɔ̀ʔ*	'enough to wear; wear enough'

Figure 7: *The aspectual nature of the variable Vv's*

3.1.2 Pre- vs. post-headedness and the continuum of abstractness.

The reality of the left- vs. right-headedness distinction may be dramatically demonstrated by considering three types of binary concatenations (i.e. concatenations consisting of only 2 verbs, the Vh and a single versatile verb).

3.1.2.1 Where both verbs can function as pre-head versatiles.

qhɔ̂ (Vh) 'steal' (vV) 'Vh stealthily'
phôʔ (Vh) 'pile up; assemble' (vV) 'Vh in a group'
qhɔ̂ phôʔ (vV + Vh) 'assemble secretly'
phôʔ qhɔ̂ (vV + Vh) 'steal in a group'

lɔ̀	(Vh) 'beg, ask for'	(vV) 'ask to Vh'
tà	(Vh) begin'	(vV) 'begin to Vh'
	lɔ̀ tà (vV + Vh) 'ask to begin'	
	tà lɔ̀ (vV + Vh) 'begin to ask'	

In cases like these, permutation shows that it is the *second* of the two verbs that is interpreted as the head.

3.1.2.2 Where both verbs can function as post-head versatiles:

chɛ̂	(Vh) 'be in a place; dwell'	(Vv) 'progressive'
bɔ̀	(Vh) 'be lazy, unwilling, bored'	(Vv) 'tired of Vh'ing'
	chɛ̂ bɔ̀ (Vh + Vv) 'be tired of living; be weary of life'	
	bɔ̀ chɛ̂ (Vh + Vv) 'is bored; is being lazy'	
ni	(Vh) 'look at'	(Vv) 'Vh and see; try Vh'ing'
cɨ	(Vh) 'send on an errand'	(Vv) 'causative'
	ni cɨ (Vh + Vv) 'cause to look at; show smn sthg'	
	cɨ ni (Vh + Vv) 'try sending smn on an errand'	

In these cases it is the *first* of the two verbs that is interpreted as the head.

3.1.2.3 Where the first verb is a vV and the second is a Vv:

Concatenations like these are inherently ambiguous, EVEN WITH NO PERMUTATION, according to which verb is interpreted as the head. (See Figure 8.)

	vV + Vh	Vh + Vv
lɔ̀ chɛ̂	'beg to be there'	'is begging'
gä kì	'must be busy'	'is busy getting'
tà ša	'begin to be easy'	'easy to begin'
ga cɨ	'help to send'	'cause to help'
ca ni	'go and look'	'try and search for'
gä mɔ̀	'get to see; find'	'see (smn) get'
lɔ̀ lɛ̀?	'beg to eat'	'earn a living by begging'
šú gɔ̀ɛ́ bû?	'be terribly nauseated by tobacco'	'have enough of tobacco smoking'

Figure 8: *Inherent ambiguity of vV + Vv concentrations*

3.1.2.4 The constituent structure and semantic hierarchization of multiversatile concatenations.

Even though there is no correlation between semantic content and pre-headedness vs. postheadedess per se, within the pre-head or the post-head group there is definitely a correlation between position in the verb-string and relative abstractness — this can only come out clearly in concatenations of at least 3 verbs (i.e. containing 2 versatile verbs, both of which are on the same side of the head):

Vh + Vv + Vv or vV + vV + Vh .

Let us here repeat the previous example:

(44a) *câ phɛ̀ʔchɛ̂*
eat POT PROG
Vh Vv Vv
'is still able to eat'

(44b) *câ chɛ̂ phɛ̀ʔ*
eat PROG POT
Vh Vv Vv
'is able to continue eating'

The general principle is, the closer to the Vh a versatile verb is, the less abstract it is — *the more abstract modifies, the more concrete is modified.* If all the versatile verbs in a concatenation are on the same side of the head, there is thus no problem in determining the hierarchy of modification. Every outer verb modifies the ensemble of the more inner constituents taken as a unit. Thus, (45) is a pure pre-head concatenation (vC), while (46) is a pure post-head one (Cv):

(45) *ŋà-hi ɡa qɔ̀ʔ phôʔ lɔ̀ câ ve*[15]
1PL vV vV vV vV Vh PTCL
must again together ask eat
'We've got to ask (them) again together (for food) to eat.'

The most deeply embedded of the vV's, *lɔ̀*, is subordinate to its head *câ* (the act of begging is secondary to its envisaged goal, eating): 'beg to eat'. *lɔ̀ câ* as a whole is the head of *phôʔ*: the 'begging to eat' is performed 'together'. *phôʔ lɔ̀ câ* as a whole is the head of *qɔ̀ʔ*: the action of 'together begging to eat' is performed 'again'. Finally, the string *qɔ̀ʔ phôʔ lɔ̀ câ* as a whole is the head of *ɡa*: it is the 'repetition of the communal act of

supplication to relieve hunger' which is deemed to be 'necessary'.[15] (See Matisoff, 1973: 214–5)

It can readily be seen that the structure of (46) is the right-branching mirror image of (45):

(46) *cì ǧɔ̀ tɔ̂ʔ mā pî cɔ̂*
Vh Vv Vv Vv Vv
tooth pull emerge teach BENEF ought
'(They) ought to show them how to pull out teeth.'

The futility of trying to write formal rules to capture semantic interpretations is particularly apparent in the case of fore-and-aft concatenations. where there is at least one grammaticized verb on each side of the Vh. Here there is no mechanical way to predict from the syntactic structure what the hierarchy of modification is. Consider the pair of 3-verb concatenations in Figure 9, each of which has the superficially identical constituent structure vV + Vh + Vv :
Those interested in the analysis of maximally complex 5-verb fore-and-aft concatenations like (47) and (48) are urged to consult Matisoff, 1973: 260–4.

(47) *ŋà-hi ga yù qay ci cɔ̂ ve lâ*
vV Vh Vv Vv V Q
help take go send ought
'Should (we) make (them) help to take (it) away?'
1. *cɔ̂* outranks *ga* 2. *ci* outranks *ga* 3. *ga* outranks *qay* (See Matisoff, 1973: 251, Fig. 27)

(48) *ŋà-hi ǧa qɔ̀ʔ chî tɔ̂ʔ pî ve*
vV vV Vh Vv Vv
must return lift emerge give
'(We) had to lift (it) out again for (them).'
1. *ǧa* outranks *pî* 2. *qɔ̀ʔ* outranks *pî* 3. *pî* (automatically) outranks *tɔ̂ʔ* (See Matisoff, 1973: 262, Fig. 28)

3.1.3 Other types of multiverbal constructions in Lahu

3.1.3.1 Fortuitous concatenations

Not all Lahu multiverbal sequences represent true versatile concatenations. Sometimes verbs come to be juxtaposed "accidentally", even though

pre-primacy

ga	*yù*	*tɔ́ʔ*
vV	Vh	Vv
help	take	emerge

'help ‖ to take out'

post-primacy

ga	*yù*	*bɔ̀*
vV	Vh	Vv
help	take	tired of

'tired of ‖ helping to take'

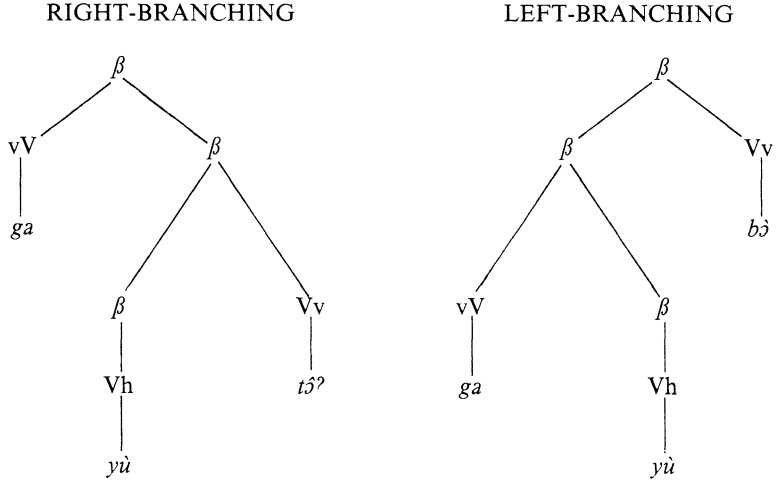


Figure 9: *Contrasting transcapital relationships: Pre-and post-primacy*

they belong to separate underlying VP's. In most of these cases the verbs refer to a series of temporally consecutive actions, as in (49):

(49) *lâ pɔ́ʔ chèʔ câ pə̀ šē ve cê*
tiger jump bite eat finish PTCL PTCL QUOT
Vh Vh Vh Vv Pv Puniv Puf
'The tiger jumped (out) and bit (into them) and ate (them) all up!'

A good rule of thumb to test whether a verb-string is really multi-clausal is to try inserting the "suspensive" non-final unrestricted particle (Punf) *lɛ*, as in (49a), and see if the meaning is thereby changed significantly:

(49a) *lâ pɔ́ʔ lɛ chèʔ lɛ câ pə̀ šē ve cê*
PTCL PTCL

This particle is an explicit demarcator of the end of a non-final clause, functioning similarly to the "gerundive" or "-TE form" of a Japanese verb:

(49b) *tora ga tobidashite kamitsuite kutte shimatta to-yuu*
tiger PTCL jump out bite onto eat complete QUOT

In a true versatile concatenation it either makes no sense, or else drastically changes the meaning to insert *lɛ* between the verbs (see, e.g. (43a,b) above).

Yet here too the interclausal bonding is a matter of degree. Although the jumping, biting, and eating are 3 successive actions, clearly the latter two verbs, *chè? câ* are more closely related to each other (almost forming a lexical compound) than either is to *pɔ̂?* 'jump'. On the other hand, the 4th verb in the string, *pə̀* 'finish', is definitely forming a true concatenation with *câ*, with the grammaticized aspectual meaning 'completive' (see above, Figure 7) — so *lɛ*-insertion is impossible between them.[16]

3.1.3.2 Resultative complements

Some resultative verbs are truly versatile, and can occur with grammaticized meanings after an indefinite number of Vh's. Thus *cɨ́* 'stick to; stick fast' appears both in more "literal" concatenations (e.g. *jû? cɨ́* (Vh + Vv) 'stab home (so the knife sticks in the wound)') and more "figurative" ones (e.g. *hà? cɨ́* (Vh + Vv) 'love unshakably').

Other resultative verbs are restricted to collocations with only one or two specific Vh's. Thus *tò?* 'catch fire; burn' occurs only after *tú* 'kindle'; *dɔ* 'fit inside' only after *kə* 'insert' or *dô?* 'pack into'; *mi* 'catch' only after *g̈à?* 'chase'; *kɨ̀* 'be melted' only after *lɔ̄* 'cause to melt'. Verbs of this type, which fail the criterion for "juxtapository productivity", we call *resultative complements* (Cr's). Combinations of Vh + Cr resemble lexical compounds, but differ from the latter in that they must be negated "intrusively": *tú mâ tò?* 'does not catch fire' ("light-not-burn"), *g̈à? mâ mi* 'cannot catch' ("chase-not-catch"), etc.

3.1.3.3 Lexical compounds

Lexical compounds composed of 2 free verbs are usually readily distinguishable from versatile concatenations: *nù-qhâ* 'smell acrid' ("stink" + "bitter"), *phɛ-chɨ̀?* 'tie up' ("restrain" + "bind"), *chî-mu* 'praise' ("lift" + "high"). Neither of the elements in a true compound is juxta-productive —

each occurs at most in a few combinations with verbs of compatible semantic content. It is as difficult to invent a comprehensible and acceptable Lahu compound as it is to create any neologism. Versatile concatenations, on the other hand, are freely inventable as the communicative need arises, and are not likely to be recognized as novel.

3.1.3.4 Borderline cases.

As always, however, there are borderline cases, where verbal collocations appear to be intermediate between lexical compounds and versatile concatenations. These *'idiomatic concatenations'* and *'productive compound-formations'* are discussed briefly in the context of *'continua of grammatico-semantic categories'* (below 4.3).

3.1.4 Polysemy and grammatization.

Not every meaning shift or *glissement sémantique* is in the direction of greater abstraction/bleaching/grammatization. Sometimes we just find "lateral slides" in semantic space. Consider the following two Lahu verbs, *phê* 'release' and *pò(n)* 'pass by' (the latter a loan from Tai).

3.1.4. phê *'send forth' > violent action*

The verb *phê* 'release' has developed many shades of meaning, only one of which can be said to represent a "grammaticization" — its use as an emphatic Vv, where it lends a vague nuance of outward-thrusting vigor to the verbal action. See Figure 10.[17]

3.1.4.2 pò(n) *'pass by; be spared' > ago; except*

The Tai-derived verb *pò(n)* 'pass by; go past' seems to have originally had spatial/temporal reference. From there it easily passed over into the social ("be spared") and moral ("transgress") realms, and finally into the field of abstract logic ("omit in thought"), achieving grammaticized status in a phrase that translates the English conjunction/preposition *except*. (See Figure 11.)

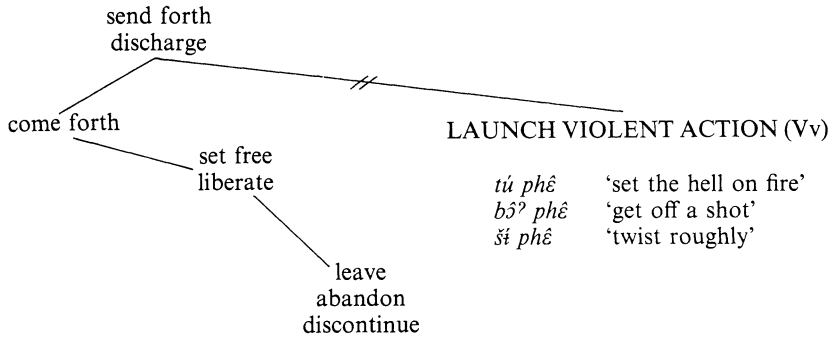


Figure 10: *Polysemy and grammatization of* phê

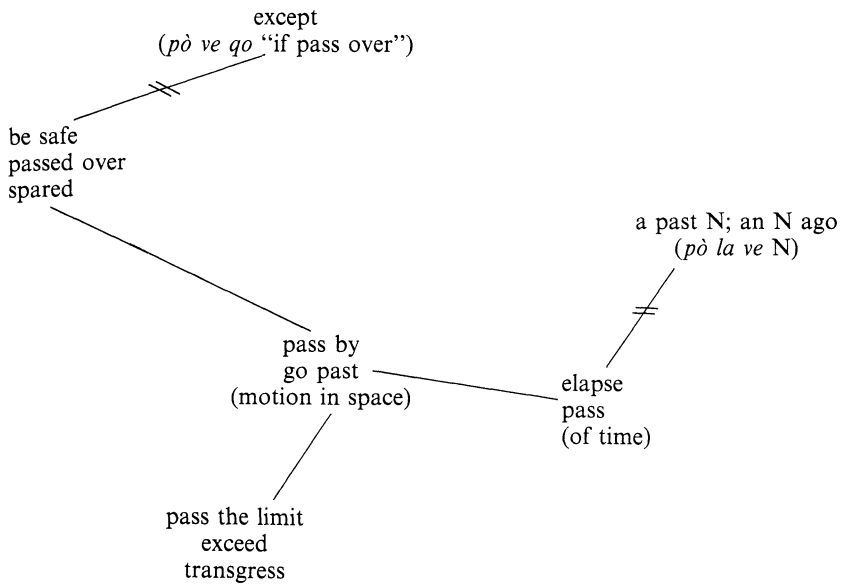


Figure 11: *Polysemy and grammatization of* pò(n)

3.2 Areal and universal patterns of verbleaching

We cannot attempt here to trace all the interesting semantic developments along the continuum of verb-grammaticization in SEA'n languages. Let us pick a few which Lahu shares broadly with other East and Southeast Asian

languages, and which seem to have universal resonances as well: *dwell/be at*; *get/obtain*; and *give*.

3.2.1 'dwell/be at' > continuative/progressive → in/at

Verbs meaning 'dwell; be in/at a place' typically develop into progressive auxiliaries. In some of our SVO languages such morphemes can also function as locative prepositions (see the discussion of "verpositions", below 4.1).

3.2.1.1 Lahu chê

a. As main verb:

(50) *ŋà nâʔ ɔ̀-ha mâ* ***chê***
1SG gun soul NEG dwell
'There dwells no soul in my gun.'

b. As post-head versatile verb:

(51) *qa-mi* ***chê*** *šē*
sing *** PTCL
Vh Vv
'Go on singing!'

(52) *yâ-pā chi ɡ̈î tô* ***chê*** *ve tí yò*
boy this run roam *** PTCL only PTCL
Vh Vv Vv
'This boy just keeps running around.'

As we have noted (above 3.1, [43a,b]), sometimes there is a real ambiguity according to whether the *chê* is functioning as an auxiliary or merely as a full verb in fortuitous concatenation:

(43a) *yɔ̂ yɛ̀ te* ***chê*** *ve*
3SG house make PROG
Vh Vv
'He is building a house.'

(43b) *yɔ̂ yɛ̀ te (lɛ)* ***chê*** *ve*
3SG house make live
Vh Vh
He built a house and is living (in it).'

3.2.1.2 Burmese nei

a. As main verb:
(53) *θu ʔeĩ* ***nei*** *te*
3SG house be at PTCL
'He is at home.'

b. As post-head auxiliary:
(54) *θu zəgâ pyô* ***nei*** te
3SG words speak *** PTCL
'He is speaking.'

3.2.1.3 Thai jùu

a. As main verb:
(55) *khun phôɔ mâj* ***jùu*** *bâan*
HON father NEG be at home
'Father is not at home.'

b. As post-verbal auxiliary:
(56) *khăw rian phasăa ʔaŋkrìt* ***jùu***
3SG study language English ***
'He's still studying English.' (cf. Eng. "He's still at it.")

3.2.1.4 Chinese zài

a. As main verb:
(57) *tā xiànzài* ***zài*** *bù* ***zài*** *jiā*
3SG now be at NEG be at home
'Is he at home now?'

b. As pre-head auxiliary verb:
(58) *tā* ***zài*** *fúshuǐ*
3SG *** swim
'She is swimming.'

(59) *gémìng* ***zài*** *fāzhǎn, rénmín* ***zài*** *qiánjìn*
revolution *** develop people *** advance
'Revolution is on the ascendant and the people are on the march.'

c. As preposition:

(60) *tā* ***zài*** *jiā* *lǐ* *kàn* *zhe* *zázhì*
3SG *** house interior read PROG magazine
'He's reading magazines in the house.'

(*zhe* is another progressive marker, grammaticized from a full verb *zháo* 'touch'.)

3.2.1.5 Vietnamese ở,

a. As main verb:

(61) *ông chủ* ***ở*** *nhà*
HON boss be at home
'The boss is at home.'

b. As pre-head auxiliary verb:

(62) *nhà* *tôi* ***ở*** *gần* *đây*
house 1SG *** be near here
'My house is near here.'

c. As preposition (post- or pre-verbal):

(1) Post-verbal

(63) *nhiều khi* *tôi* *thích ăn* ***ở*** *ngoài*
many time 1SG like eat *** outside
'Often I prefer to eat out.'

(2) Pre-verbal

A phrase with ở may also be preposed to the verb, especially to indicate the source of an action:

(64) ***ở*** Mỹ *qua* *Việtnam*
*** America go over to
'come to Vietnam from America'

3.2.1.6 Yao Samsao yiəm

a. As main verb:

(65) *ń-ñɛ maa* ***yiəm*** *wúə nɔm pyáw*
my mother *** DEM CLF house
'My mother lives in that house.'

b. As pre-verbal preposition:

(66) *yiə* ***yiəm*** *Samsàaw tshuət-sə̌y* *tâay*
1SG *** emerge world come
'I was born in Samsao.'

(67) *yiə* ***yiəm*** *wúə* ʔii *hñə̃ŋ* *tòoʔ* *səw*
1SG *** DEM two year read book
'I studied during those two years.'

(68) *yiə* ***yiəm*** *kàdíə míə* *bɔ́ʔ* *naaŋ*
1SG *** grass inside hit snake
'I, being in the grass, hit the snake.'

c. As post-verbal preposition:

(69) *yiə* *bɔ́ʔ* *naaŋ* ***yiəm*** *kàdíə míə*
1SG hit snake *** grass inside
'I hit the snake that was in the grass.'

d. As postposition:

(70) *yiə* *pyə̌w* *jìə* *Nâmtūəy* ***yiəm***
1SG flee descend ***
'I fled down to Namtuay (village) to live.'

3.2.1.7 Hmong nyob

a. As main verb:

(71) *kuv txiv tsis* ***nyob*** *hauv tsev*
1SG male NEG be at inside house
'My father is not at home.' (Clark 1979: 7)

b. As verposition:

(72) *nws pw* ***nyob*** *hauv txaj*
3SG lie *** inside room
'He's sleeping in the room.' (ibid.: 6)

3.2.2 'get/obtain' → MANAGE/GET TO; HAVE TO/MUST; ABLE TO

Verbs meaning 'get; obtain' show striking parallelism in the way they grammatize in Southeast Asian languages. Of particular interest is a widespread pattern whereby this verb functions as both a pre-head and a post-head auxiliary, with concomitant meaning difference: in pre-verbal position

meaning 'have a successful past experience; have managed to Vh; have gotten to Vh' and/or 'have to Vh; must Vh' and in post-verbal position meaning 'able to Vh'.

3.2.2.1 Lahu ğa

This is the only Lahu versatile verb that can function either as a vV or a Vv.

a. As main verb:

(73) *mâ* ***ğa*** *e kàʔ mâ bəʔ — **ğa***
NEG get PTCL even NEG angry get
e kàʔ mâ ha-lɛ̀
PTCL even NEG happy
'If I don't get it I won't be mad — if I do get it I won't be happy.'

b. As pre-head versatile verb:

(1) 'manage to Vh; get to Vh'

(74) *ɔ̄-chî mɛ-mɛ ve kàʔ mâ **ğa** câ ve yò*
curry tasty PTCL even NEG *** eat PTCL PTCL
'We didn't get to eat very tasty food either.'

(2) 'must Vh'

(75) *chi-bəʔ ŋà **ğa** qay ve yò*
now 1SG *** go PTCL PTCL
'I must leave now.'

c. As post-head versatile verb: 'get by Vh'ing; able to Vh; completed Vh'ing'

(76a) *jûʔ* ***ğa ve***
stab ***
'get by impaling'

(76b) *Kâlâ-phu ve ɔ̄-chî ŋà câ mâ **ğa***
whiteman GEN food 1SG eat NEG ***
'I just can't eat white men's food.'

3.2.2.2 Thai dâj

a. As main verb: get/obtain

(77) *khăw* ***dâj*** *sətaaŋ maa càak khun phɔ̂ɔ léɛw*
3SG obtain money come from HON father already
'He's already gotten some money from his father.'

b. As pre-head auxiliary verb: 'get to Vh'

(78) *nŭu* ***dâj*** *paj Kruŋthêep măj*
2SG *** go Bangkok Q
vV Vh
'Did you get to go to Bangkok?'

c. As post-head auxiliary verb: 'able to Vh'

(79) *nŭu paj Kruŋthêep* ***dâj*** *măj*
Vh Vv
'Can you go to Bangkok?'

3.2.2.3 Yao (Mien) Samsao tú$^{?}$

a. As main verb:

(80) *pyăw jìə Hùəy-*$^{?}$*ɔ̀ɔ yiəm, nɔ̂ɔ ḿ-ñɛ*
flee descend live then my
$^{?}$*áaw* ***tú***$^{?}$ *kûu-ŋwăa*
wife ***baby
'My wife had ("got") a baby after we had fled down to Hùəy-$^{?}$ɔ̀ɔ to live.'

b. As pre-head auxiliary:

(81) *yiə* ***tú***$^{?}$ *tòo*$^{?}$ $^{?}$*ii hñăŋ səw*
1SG *** read two year books
vV Vh
'I had the opportunity to study for two years.'

c. As post-head auxiliary:

(82) *yiə tòo*$^{?}$ ***tú***$^{?}$ $^{?}$*ii hñăŋ səw*
Vh Vv
'I've completed two years of study.'

(83) *yiə mwăŋ* ***tú***$^{?}$ *híw mîən wàa*
1SG listen *** know Mien language
Vh Vv Cr
'I can understand Mien.'

It seems likely that this Mien verb stands in a contact relationship with a number of phonologically and functionally similar verbs in the coterritorial languages, especially Vietnamese *đ*ư*ợ-c*, Thai *thùuk*, and Chinese **tək*. See below 3.2.2.5.

3.2.2.4 Hmong tau

The syntactic and semantic behavior of the Hmong grammatized verb *tau* is notably similar to what we have seen for Lahu *ɡ̈a* and Thai *dâj*. Here too, *tau* means 'get; obtain' as a main verb, and may occur as an auxiliary either before or after another verb, with correspondingly different shades of meaning.

This morpheme seems ultimately to be borrowed from Chinese (Mand. *dào*) 'arrive; reach', a verb which has developed a number of grammaticized functions within Chinese itself, e.g. as a potential complement. (The examples in this section are translated from Mottin, 1978, pp. 84, 102–3.)

a. As main verb:
(84a) **tau** nyiaj 'get money'
(84b) **tau** kev 'find a way'
(84c) dev **tau** npluas. 'The dog caught leeches.'

b. As pre-head auxiliary (vV):
In this usage, *tau* functions as a *signe du passé*, but Mottin stresses that this is optional "*si l'on sait déjà qu'on est au passé*". It is used especially in negative sentences, after the adverb *tsis* 'not':

(85) *nws* ***tau*** *mus tsev*
3SG *** go home
vV Vh
'He went home.'

(86) *nyob lawv kuv* ***tau*** *noj ib pluag*
be at 3PL 1SG *** eat one meal
vV Vh
'I've eaten at their place once.'

(87) *koj noj mov-tas lawm los tsis* ***tau***
2SG eat finish already CONJ NEG ***
'Have you finished eating yet (or not)?'

The sequence V + *lawm los tsis tau*, as in (87), corresponds exactly to the Thai construction V + *lέɛw rɤ̌y jaŋ* used in asking "yes-no" questions about the completion of an event.

c. As post-head auxiliary (Vv)

In post-head position *tau* (like Thai *dâj* and Lahu *ɡ̈a*) indicates ability, possibility, permissibility:

(88) *kuv mus* ***tau***
Vh Vv
'I can go there.'

(89) *los tsev puas* ***tau***
come house Q ***
Vh Vv
'May I come in?'

(90) *hom txiv no noj* ***tau***
fruit CLF DEM eat -able
Vh Vv
'This fruit may be eaten.'

In close constituency with a preceding verb, *tau* may indicate the real accomplishment of the purpose of an action (*faire et obtenir réellement quelque chose*): *nrhiav* 'look for' > *nrhiav tau* 'find sthg.'; *yuav* 'buy' > *yuav tau* 'possess'; *caum* 'chase' > *caum tau* 'catch'.

There is a related verb under a different tone, *taus*, with meanings quite similar to those of Lahu *pɨ́* (below 4.41) '*avoir la possibilité physique ou matérielle de; avoir l'habitude de*'. Note that *ɡ̈a* (mid-tone), the Lahu functional equivalent of Hmong *tau*, also has a morphotonemically related "allofam" (word-family alternant; see Matisoff 1978) *ɡ̈â* (high-falling tone) 'win; overcome'.

3.2.2.5 Vietnamese đươc *and its analogues (especially Thai* thùuk *and Chinese* *tək*)*[18]

a. As main verb:

As a main verb , Vietnamese *đươc* means (1) 'be acceptable; correct; fine; O.K; all right' (2) 'obtain; get' (3) 'win'.

đươc phép 'obtain authority'
đươc mùa 'get a harvest',
đươc trận 'win a battle'

b. As pre-head auxiliary:

In this position, *được* is grammaticized into a passivizer (but in the oppositely valued sense from its Thai phonological lookalike, *thùuk!*), conveying the notion of 'being benefited by, affected favorably by some action, state or factor':[19]

(91) *tôi* ***được*** *đi*
1SG *** go
vV Vh
'I was allowed to go.'

(92) *tôi* ***được*** *ăn*
vV Vh
'I am permitted to eat.'

(Contrast (92a): *tôi ăn* ***được***
Vh Vv
'I am edible' (see 3.225c, below).)

(93) *em Toàn* ***được*** *thầy giáo khen*
y.bro. vV teacher Vh
'Brother Toan was praised by the teacher.'

(94) *mười lăm giáo sư* ***được*** *chọn đi Mỹ*
10 5 teachers vV Vh go America
'Fifteen teachers were chosen to go to America.'

c. As post-head auxiliary:

In post-head position, *được* may indicate ability (in a manner similar to Lahu *ɡa*, Thai *dâj*, Hmong *tau*):

(95) *tôi đi* ***được***
Vh Vv
'I can go.'

Either verb of (95) may be negated, with a different nuance of meaning:

(96) *tôi đi không* ***được***
Vh NEG Vv
'I can't go' (because of physical or compelling material reasons).

(97) *tôi không đi* ***được***
NEG Vh Vv
'I can't go' (because I'm too busy, it's not convenient, etc.).

With some preceding verbs, *được* is a neat translation of the English suffix *-able/-ible*:

ăn được ("eat + -able") > 'eatable, edible'
làm được ("do + -able") > 'doable, feasible'

As a post-head auxiliary *được* may also have another shade of meaning: That the action of the verb has resulted in something real or good, or has been carried out to a successful conclusion:

(98) *mua* ***được*** *nhà*
buy *** house
Vh Vv
'bought a house'

(99) sinh ***được*** *một đứa con-giai*
born *** one/a CLF son
Vh Vv
'had a son born to us'

The position of the object can also affect the semantic interpretation of a post-verbal *được*:

(100a) *tôi bắt hai con cá* ***được***
1SG catch two CLF fish ***
Vh Vv
'I'm able to catch two fish.'

(100b) *tôi bắt* ***được*** *hai con cá*
Vh Vv
'I caught two fish.'

Contrast a similar sentence where *được* appears in the pre-head position:

(100c) *tôi* ***được*** *bắt hai con cá*
vV Vh
I'm permitted to catch two fish.'

Phonologically, *được* resembles the Thai word *thùuk*, and there is a considerable semantic and grammatical overlap between these two important words. The lexical (fully-verbal) meaning of Thai *thùuk* ranges from the moral/abstract sense 'be right; correct' to the physical sense of 'hit the mark; touch; come in contact with.' There is a homophonous and homographic

verb meaning 'be cheap; inexpensive' which must also ultimately be related, via the notion *the price is right*. (See Li Fang Kuei, 1977, p. 102, #16.) As a grammatized morpheme, *thùuk* forms adversative passive constructions with a following verb or clause: *thùuk kàt* 'be bitten', *thùuk thɔɔramaan* 'be tortured', *thùuk khăa hàk* 'get one's leg broken', *thùuk rót chon* 'be hit by a car'. This is in sharp semantic contrast to the Vietnamese '*benefactive passive*' construction, *đ*ư*ợc* + Vh (above (b).

There is another close phonological resemblance between *đ*ư*ợc* and Chinese (Old Chinese **tək*) 'get; obtain', a word with a variety of grammaticized uses in Chinese dialects. This character has several Mandarin pronunciations. As *dé* it is a full verb 'get; obtain'. With a different tone and vowel, *děi*, it is a modal auxiliary meaning 'must; have to', or expressing an unpleasant eventuality. In unstressed form, *-de-* it marks a positive resultative complement (*kàn-de-jiàn* 'look-can-see', here being opposed to the negative resultative marker *-bu-* (*kàn-bu-jiàn*'' look-not-see').

The phonological and semantic interrelationships among these Vietnamese, Thai, and Chinese words — and undoubtedly the Yao (Mien) word *tú*^{9} as well — bespeak a long period of profound language contact.

We may note in passing that there is also a Khmer verb, *trəw*, which fills a range of syntactic/semantic roles quite similar to Thai *thùuk* and Viet. *đ*ư*ợc*, even though it seems quite unrelated phonologically to the latter two. Thus *trəw* also can mean 'be correct' or 'to hit (the mark)' as a main verb[20], and is also used as a passive complementizer:

(101) *pii-msəl-məñ kñom* ***trəw*** *ckae kham*
yesterday 1SG *** dog bite
'I was bitten by a dog yesterday.' (Huffman 1970: 302)

3.2.2.6 Khmer baan

Finally, the astounding parallelism in the syntactic/semantic details of the grammatization of words for *get* in SEA'n languages is also manifested in the Khmer verb *baan*:

a. As main verb:
(102) *look cʊng* ***baan*** *chəə-kuh tee*
2SG want *** matches Q
'Do you want to get some matches?'

b. As pre-head auxiliary:

In this position, *baan* expresses successful past experience ('to have managed to, been able to; to have had a chance to'):

(103) *kñom* ***baan*** *tiw pii dɒɒng*
1SG *** go two time
vV Vh
'I've gotten to go twice.'[21]

c. As post-head auxiliary:

In post-head position, *baan*, like its equivalents in our other languages, refers to ability:

(104) *kñom sdap* ***baan***
Vh Vv
'I can understand.'

(105) *kñom sdap min* ***baan*** *tee*
Vh NEG Vv PTCL
'I can't understand.'

3.2.2.7 English 'get'

The modal polysemy of *get* comes as no surprise to English speakers. Cf. the ambiguity of (106):

(106) Tomorrow you can get to work.
a. begin working
b. have the privilege of working
c. be transported to work

Sometimes stress helps to disambiguate:

107a) I gotta [*ajgʌrə*] swim today. (=I must swim today.)
(107b) I got to [*ajgátta*] swim today. (=I managed to swim today.)

At any event, the constellation of ideas emanating from the basic notion of 'get; obtain' seems to have much more than areal significance. See the diagram in Figure 12.

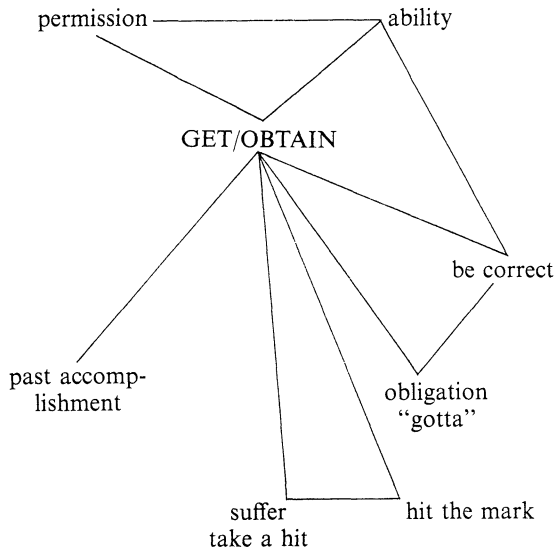


Figure 12: *Semantic extensions of GET/OBTAIN*

3.2.3. 'give' > causative/benefactive

Though many verbs are typically grammatized for causative purposes (e.g. TAKE, DO, MAKE, LET, HAVE, SEND (on an errand)), the benefactive verb *par excellence* is, of course, GIVE.

The Thai causative/benefactive morpheme *hâj* is discussed in connection with postpositions (below 4.2.2). In this section we shall briefly consider some examples from Lahu, Yao, Vietnamese, Khmer, and Mandarin.

3.2.3.1. Lahu pî *and its competitors.*

Several verbs have vied for causativizational honors in Lahu, including *yù* 'take', *te* 'do', *ci* 'send on an errand' (see Matisoff, 1976). The most highly grammatized of them all is *pî* 'give':

a. As main verb:
(108) *phu* ***pî*** ve 'give money'
OBJ V
b. As post-head versatile verb:

(1) benefactive

We have already discussed the paradigmatic opposition between *pî* (Vv) '3rd person beneficiary' and *lâ* (P_v) 'non-3rd person beneficiary' in the Lahu VP. (See above 2.32 and Figure 1.)

(109) *šu à? tâ dê* ***pî-?***
3SG PTCL ADV Vh Vv
'Don't scold other people!'

(110) *yɔ̂ à? chi qhe qô? pî-?*
3SG PTCL this like Vh Vv
'Tell that to him.'

(2) causative

In many contexts, the benefactive meaning of *pî* shades into that of a causativizer (3rd person causee):

ce pî	(Vh + Vv)	'make it fall; drop sthg'
chu pî	(Vh + Vv)	'make him fat'
ši pî	(Vh + Vv)	'make/let him die'

3.2.3.2. Yao Samsao pun

a. As a main verb:

(111) *nîn* ***pun*** *pɔ́w yiə*
3SG Vh axe 1SG
'He gave me an axe.'

b. As a benefactive preposition:

(112) *yiə tsiá? nɔm daan* ***pun*** *nîn*
1SG weave CLF basket *** 3SG
'I wove a basket for him.'

(113) *maa cáp bùdò?-gwǎy* ***pun*** *fù?-cúəy*
mother cut fingernails *** child
'The mother cut the child's nails for him' ("cut the nails for the child").

c. As a causative complementizer (with sentential object):

(114) *maa* ***pun*** *fù?-cúəy cáp bùdò?-gwǎy*
mother *** child cut nails
'The mother let the child cut his nails.'

This sentence is understood as a permissive causative. To express the coercive

causative equivalent, a different higher verb, *bŭa* 'tell' must be used, along with a pre-head auxiliary *a-tsù?* expressing obligation:

(115) *maa bŭa fù?-cúəy a-tsù? cáp bùdò?-gwăy*
mother tell child must cut nails
'The mother made the child cut his nails.'

3.2.3.3. Vietnamese cho

a. As a main verb:
(116) *ông ấy vừa* ***cho*** *con gái chiếc xe*
HON 3SG just *** daughter CLF car
'He just gave his daughter a car.'

b. As benefactive/causative pre-head auxiliary verb:
cho ăn (vV + Vh) 'feed; have someone eat'
cho hay (vV + Vh) 'let someone know; inform smn'

c. As benefactive preposition:
(117) *đưa cái chổi đây* ***cho*** *tôi*
hand over CLF broom DET *** 1SG
'Please hand me the broom.'

d. As permissive-causative complementizer:
(118) *ông ấy không* ***cho*** *tôi thôi*
HON 3SG NEG *** 1SG resign
'He wouldn't let me resign.'

e. As benefactive postposition:
(119) *để tôi viết* ***cho***
let 1SG write ***
'Let me write it for (you).'

(Cf. the postpositional use of Thai *hâj*, below 4.22.)

3.2.3.4. Khmer qaoy

a. As main verb:
(120) *mənuh prɔh baan* ***qaoy*** *siəwphəw təw mənuh srɒy*
person male PAST *** book to person female
'The man gave the book to the woman.'

b. As a pre-head auxiliary before adjectival verbs:

This is just a special case of the pre-verbal causative effect of *qaoy*, but this construction translates most naturally by English adverbs (Huffman, 1970: 153): [22]

(121) *riəp-cɒm kluən* ***qaoy*** *chap*
prepare self *** quick
'Get ready quickly.' ("making it quick")

(122) *sliəq-peəq* ***qaoy*** *sqaat-baat*
dress *** careful
'Dress carefully.'

c. As a causative complementizer (with sentential object):

(123) *qəyləw maq cɒng* ***qaoy*** *kñom twəə qəy*
now mother want *** 1SG do PTCL
'What would Mother like to have me do now?'

(124) *kñom* ***qaoy*** *koət ruət*
1SG *** 3SG run
'I had him run.' (intentionally)

In this function, *qaoy* is often used together with the preceding verb *twəə* 'do', to express a more coercive causative.[23]

(125) *kñom* ***twəə*** ***qaoy*** *koət ruət*
*** ***
'I made him run.' (maybe by scaring him inadvertently)

(126) *kñom baan* ***twəə*** ***qaoy*** *koət phək sraa klah*
PAST *** *** 3SG drink liquor some
'I made her drink some wine.'

3.2.3.5. Mandarin gěi

a. As main verb:

(127) *wǒmen yīnggāi* ***gěi*** *ta yí ge xīngqī de jià*
1PL ought *** 3SG one CLF week GEN vacation
'We ought to give him a week's vacation.'

b. As benefactive preposition:

A benefactive prepositional phrase with *gěi* may either follow or precede the verb of the clause:

(1) Post-verbal prepositional phrase:

(128) *xìn yǐjing jiāo* ***gěi*** *ta le*
letter already hand over *** 3SG PTCL
'I've already handed the letter to him.'

(129) *wǒ bǎ yàoshi liú* ***gěi*** *nǐ*
1SG OBJ key leave *** 2SG
'I'll leave the key with you.'

(2) Pre-verbal prepositional phrase:

(130) *tā* ***gěi*** *lǚkè sòng shuǐ dào chá*
3SG *** passenger bring water serve tea
'She brought water and served tea for the passengers.'

Sometimes the meaning of *gěi* in this construction is more like 'let; allow':

(131) ***gěi*** *wǒ kàn-kan*
*** 1SG look-look
'Let me have a look!'

If there is a sentential complement, however, *gěi* cannot be used to convey permission, and the verb *ràng* (lit. "give way; yield") is required instead:

(132) *tā hái bù kěn* ***ràng*** *xiǎoháir gen*
3SG still NEG willing *** child with
Yóutàirén jiéhūn
Jew marry
'She's still not willing to let her child marry a Jew.'

c. As passive complementizer:

In this construction, *gěi* functions much like the adversative passive marker *bèi* (see above 3.2.2.5, note 19), following the topicalized patient and introducing the clause which expresses the agency that brought about the unpleasant event:

(133) *zhè gǔ dírén quán* ***gěi*** *yóujī-duì xiāomiè le*
this gang enemy all *** guerrilla-band wipe out PTLC
'The whole horde of enemy soldiers was wiped out by the guerrillas.'

(134) *wǒmen de yīfu* ***gěi*** *hàn-shuǐ shī-tòu le*
1PL GEN clothes *** sweat soak through PTLC
'Our clothes got soaked with sweat.'

3.2.4. Excursus on Lahu te *'do'*

Not surprisingly, *te* 'do; make' is the Lahu verb with the most abstract meaning and the widest range of grammatical functions. Even in the simplest cases, where it occurs as the only verb in its VP, it collocates with preceding NP's or adverbials in a wide variety of semantic relationships which can only be deduced from the context. To these one may give discrete names (though of course they shade imperceptibly into each other), e.g.: "abstract" (*kán te ve* 'do work'), "factitive" (*ú-gɛ̂ te ve* 'make a pillow'), "functional" (*qhâ?-šɛ te ve* 'act as headman'), "exploitative" (*ú-gɛ̂ te ve* 'use for a pillow'), "appurtenative" (*ɔ̀-qhɛ̀? te ve* 'have ridges, be corrugated'), "locomotive" (*khi-nê? te ve* "limp"), "adverbial" (*qɔ̀? ɛ̀ te ve* 'be crooked'), etc.[24]

More abstractly, *te* may serve as the higher verb governing embedded clauses of various types, e.g. as a "resumptive pro-verb" after two correlative or contrasting verbs:

(135) *g̈ɨ̂ yà? g̈ɨ̂ tâ? **te** ve*
run descend run ascend *** PTCL
'keep running up and down'

This is quite reminiscent of the Japanese construction V_1 *-tari* V_2 *-tari* + *suru*, with the verb *suru* 'do' used "resumptively" after a pair of correlative verbs, e.g. *ittari kitari suru* 'be coming and going (all the time)'.

Closer to our present concerns is the use of *te* as a pre-head versatile verb. Sometimes it retains a more literal meaning in this construction: 'Make something and Vh with it' (*te kə ve* 'make and put in', *là-g̈ì te tɛ ve* 'make tea and set it down'), but often it serves to transitivize or causativize the following verb, whether it be an adjectival (= stative) verb (*te qɛ̀ ve* 'widen; make wide'), an intransitive action verb (*te pɔ̄ ve* 'cause to go through (as a road)'; cf. French *faire pénétrer*), or a verb which is already transitive (*te tē ve* 'cause to crush'; cf. French *faire écraser*). Most elusively, it sometimes is purely redundant, just adding a bit of phonological bulk but bringing no perceptible increment of meaning. Thus in combination with *tâ?* 'carry', *te* has neither factitive nor causative force: *te tâ? ve* can just mean 'carry something along' — it does not necessarily mean either 'make sthg. and carry it', or 'cause to carry'. Similarly, *te dê ve* just means 'scold smn.' ("do-scold"), not 'cause to scold'.

4.0 CONTINUA OF GRAMMATICO-SEMANTIC CATEGORIES

There seems to be a fundamental temperamental difference among linguists, according to whether they like neat or messy classifications.[25] In her otherwise perceptive paper on Hmong verb-derived prepositions (1979), Clark comes down squarely on the side of the rigid categorizers:

> The question here is, at what point can a word be said to be a preposition and not a verb? The notion of category 'squish' (Ross 1972) or gradual shift from one category to another is as absurd for lexical change as the notion of gradual shift from one sound to another is absurd for sound change. (p. 8)

Somewhat similar is the reductionist position taken in Warotamasikkhadit (1988), who, having made the valid observation that almost all Thai prepositions can be derived historically from full nouns or verbs, goes on to claim that even synchronically there is no such thing as a class of "prepositions" in Thai.

Some grammarians who have dealt with Chinese despair of imposing ready-made "Standard Average European" category-labels on all form-classes and construction types, and have been reduced (or inspired?) to create their own terminology, in a way that some find distasteful.[26] It is as if a proliferation of terminology were some kind of violation of Occam's razor.

Certainly neologistic terminologizing can be carried too far — since linguistic phenomena are continua, it is impossible to find discrete names for every point on each (or any) continuum. One must use heuristic principles — create new terms if necessary, but only if none of the old ones will do. The motive should not be to dazzle or obfuscate by new coinages, but rather to clarify.

Yet I personally would like to come down wetly on the side of the squishers. There are duck-billed platypuses in the animal world, and there are borderline entities in grammar as well. I have never shrunk from coining conflationary neologisms to characterize entities which straddle categories: e.g. my Lahu nadverbs, a class of expressions with both nominal and adverbial characteristics (Matisoff, 1973: 306–15). It is far more interesting to set these anomalous phenomena into high relief than to try and sweep them under the rug, or make it look as if they behaved like the non-problematic cases.[27]

4.1 "Verpositions" in SVO languages

In this non-confrontational spirit I would like to propose the euphonious term *verpositions* for verb-derived morphemes that have come to function like

prepositions.[28] A few random examples:

(a) Viet. *vào* 'enter' (V) ↔ 'into NP' (*inessive Prep.*)
(136) *ði sau* ***vào*** *xã-hội Việt-Nam*
go deep *** society
'go deeper into Vietnamese society'

(b) Viet. *vớ́i* 'join' (V) ↔ 'together with NP' (*comitative Prep.*)
(137) *tôi không thích làm* ***vớ́i*** *ông ta*
1SG NEG like work *** 3SG
'I don't like to work for him.'

(c) Thai *khâam* 'cross over' (V) ↔ 'across NP' (*transitive Prep.*)
(138) *phûuujĭŋ dəən* ***khâam*** *thənŏn paj lέεw*
woman walk **** street go already
'The woman went off across the street already.'

(d) Mand. *yòng* 'use' (V) ↔ 'with NP' (*instrumental Prep.*)
(139) *tā* ***yòng*** *kuàizi chī bái-cài*
3SG *** chopstick eat cabbage
'He eats cabbage with chopsticks.'

It is interesting to compare and contrast similar serializations in SOV languages. The Lahu translation of (139) is:

(140) *yɔ̂ á-cu-ka yù lɛ ɡ̈ɔ̂-cá câ ve*
3SG chopstick take PTCL cabbage eat PTCL
(lit. 'He, taking chopsticks, eats cabbage.')

Just as the Mandarin serialization with *yòng* 'use' is the only way to express the instrumental notion of this sentence, Lahu has no other means than the *yù lɛ* construction to do so.

But there are several differences here, over and above the relative ordering of verb and object.[29] Chinese can use the *yòng*-construction even with an instrument that is a part of the body:

(141) *tā yòng shŏu chī bái-cài*
*** hand
'He eats cabbage with his hands.'

Here Lahu cannot use the *yù lɛ* construction, but must use the postposed noun-particle *gɛ* instead:[30]

(142) *yɔ̂ là?-šɛ gɛ ɡ̈ɔ̂-cá câ ve*
hand Pn

In Lahu the use of the serialized construction implies volitionality on the part of the agent:

(143) *yɔ̂ á-thɔ yù lɛ là?-nɔ tɔ̂? ve*
3SG knife take PTCL finger cut PTCL

This can only mean 'He cut his finger with a knife (on purpose).' To express the corresponding accidental event the sentence must be recast somehow, e.g.:

(144) *ŋà ve là?-nɔ á-thɔ tɔ̂? ve*
1SG GEN finger knife cut PTCL
(lit. 'As for my finger, knife cut (it).')

The most interesting overt difference between (139) and (140) is the presence of the particle *lɛ* in the Lahu sentence. The function of this "non-final unrestricted particle" is to indicate that the preceding clause is NOT THE LAST in the sentence.[31] The presence of *lɛ* makes explicit the biclausal nature of the sentence. (As we have noted (above 3.141), the insertibility of *lɛ* is the chief criterion for distinguishing "fortuitous" from real concatenations.) Yet one cannot help wondering whether the Lahu speaker really considers the verbal event of cabbage-eating to be more of a two-part sequence than the Chinese speaker does. Surely sentence (140) is less biclausal than a sentence like (146):

(146) *yɔ̂ á-cu-ka ɔ́-thɛ* ***lɛ*** *ɡ̈ɔ̂-cá câ ve*
bend-break
'He broke the chopsticks (by bending them) and ate the cabbage.'[32]

There must in fact be a *continuum of clausality* — a continuum of closeness of relationship or unitarity more subtle than the minimal morphological/syntactic apparatus of these languages can overtly indicate. (Cf. the discussion of the relative clausal autonomy of pairs of verbs in "fortuitous concatenation" above 3.141, ex. (49).)

4.2. Verbs as postpositions in SVO languages

We'll just give a few quick examples:

4.2.1 'finish' > perfective

4.2.1.1 Lahu pə̀

As a main verb this word means 'spend; use up' (*phu pə̀ ve* 'spend money') or 'finish':

(147) *pə̀ thâ qha-gà ga̱ te qay ve yò*
finish time until must do go on PTCL PTCL
Vh Puniv Next vV Vh Vv
'We've got to keep on doing it until it's finished.'

As post-head versatile verb (Vv), *pə̀* means 'finish Vh'ing' (see above 3.1.2.4):

(148) *ɔ̄ câ **pə̀** ò lâ*
rice eat *** PTCL Q
'Have you finished eating?' (common greeting)

In this post-verbal position *pə̀* has developed into a more general perfective marker, in which capacity it can now follow the perfective Pv *tā* (see 2.3.3 above). Here we should probably consider *pə̀* to have become a true verb-particle in its own right:

(149) *ô tê ni thâ ŋà pa pɔ **tā** **pə̀***
that one day PTCL 1SG father born PTCL PTCL
'On that day my father was born.'

4.2.1.2. Thai lɛ́ɛw

This word does not usually occur as a main verb in Thai, except in some common idioms like *pen lɛ́ɛw kan* 'That's settled; That does it', *kɔ̂ lɛ́ɛw paj* 'Let it go; Leave it at that'.

In clause-final position this word has been grammaticized into one of the language's most important aspectual postpositions, signifying 'completed action; change of state':

(150) *klàp paj bâan lɛ́ɛw khráp*
return go home *** POLITE
'He's gone off home already, sir.'

(151) *thâa khrŷaŋdiim mòt* ***lɛ́ɛw*** *raw cə tham ʔəraj*
if beverage used up*** 1PL PTCL do what
'What'll we do if the drinks are all finished?'

This seems certainly to be an early loan from Chinese (Mand. *liǎo*, also a former verb (now literary) meaning 'to finish', but grammaticized in unstressed form, *lə*, into a perfective particle. Also undoubtedly due to contact influence from Chinese is the functionally and phonologically similar Hmong aspectual particle *lawm* (cf. sentence (87) above; the *-m* represents a creaky tone in Hmong orthography, while the *-w-* is barred-i).

4.2.2. GIVE/TAKE↔outer-directed/inner-directed

The Thai verbs *hâj* 'give' and *ʔaw* 'take' play several interrelated roles in the grammar, where they are more-or-less neatly, paradigmatically opposed to each other.[33]

As prepositions, *hâj* indicates intentional benefaction, while *ʔaw* signals that an unintentional effect has been produced on something:

(152) *tôŋ jŷyn ŋəən hâj phənákŋaan*
must handmoney*** clerk
'You've got to hand the money over to the clerk.'

(153) *mɛɛw təkuj ʔaw sŷa*
cat scratch *** mat
'The cat scratched up the mat.'

As complementizers (clause-introducers), *hâj* has a causative function (human or at least animate causee),[34] while *ʔaw* has an instrumental one (inanimate instrument):

(154) *mêɛ-khrua* ***hâj dèk*** *tàt nŷa pen chín lék-lék*
cook *** child cut meat be slice small-small
'The cook had the child cut the meat into tiny slices.'

(155) *mêɛ-khrua* ***ʔaw mîit*** *tàt nŷa pen chín lék-lék*
*** knife
'The cook used a knife to cut the meat into tiny slices.'

This latter construction is substantially identical to the Lahu instrumental serialization with *yù lɛ* (above 4.1). We may translate (155) into Lahu as:

(156) *te-câ-šɛ̄-ma*[35] *á-thɔ* ***yù lɛ*** *šā* *thà?*
cook knife *** meat PTCL
ɔ̀-dì ɔ̀-dì ɛ́ *te* *tɔ̂?* *ve*
hunks little make cut PTCL

Finally, what we are interested in here is the postpositional use of these words, where they contrast on the dimension of inner- vs. outer-directedness: *hâj* being used for actions performed for someone else ("givingly"), while *?aw* is used for actions done for oneself/itself ("takingly"):

(157) *phûujɨ̌ŋ* *kèp* *dɔ̀ɔkmáaj* ***hâj***
woman pick flower ***
The woman picked flowers for (somebody).'[36]

(158) *phûujɨ̌ŋ kèp dɔ̀ɔkmáaj ?aw*
'The woman picked flowers for (herself).'

4.2.3. COME/GO←→toward/away from the deictic center (Postposition)

We have already touched on some of the multifarious grammaticizations of *come* and *go* (above 2.3.1, 2.4.2). Here we will just give some examples of the use of these verbs as postpositions, with either temporal or spatial or "more abstract" reference.

(a) *Temporal deictic function*:

The Yao Samsao expression for 'hitherto' (159) is exactly parallel in its internal tri-verbal structure to similar expressions in Thai (160) and Mandarin (161):

(159) *yiəm* *thaaw* *?îi-tsàn* ***tâay***
be reach now COME
V_1 V_2 V_3

(160) *jùu* *thɨ̌ŋ* *dǐawníi* ***maa***
be reach now COME

(161) *zài dào* *xiànzài* ***lái***
be reach now COME

(b) *Spatial deictic function*:

Expressions for *give* and *take* in East and SE Asian languages tend to

be 2-verb sequences composed of the same direction-neutral main verb meaning 'hold', plus either of the pair of grammaticized postpositional verbs meaning *come/go* to specify the directionality. See Figure 13.

	BRING = HOLD + COME		TAKE = HOLD + GO	
Yao	*tɔ̂ɔ*	*tâay*	*tɔ̀ɔ*	*mîing*
Thai	*ʔaw*	*maa*	*ʔaw*	*paj*
Mandarin	*ná*	*lái*	*ná*	*qù*
Japanese	*motte*	*kuru*	*motte*	*iku*[37]

Figure 13: *Bring/take and come/go*

4.2.3.1. Thai maa *'come' and* paj *'go'*[38]

(162a) *laa wîŋ khâw* ***paj*** *naj pàa*
donkey run enter **go** into forest
V_h V_v V_v PREP
'The donkey ran off into the forest (e.g. to get away).'

(162b) *laa wîŋ khâw* ***maa*** *naj pàa*
donkey run enter **come** into forest
V_h V_v V_v PREP
'The donkey came running into the forest (e.g. to rejoin his waiting wife and children).'[39]

In this connection we should cite the common Thai greeting, so hard for foreigners properly to understand at first:

(163) *paj nǎj maa*
go where come
'Where are you coming from?; Where have you been?'
("Having gone where do you come?")

4.2.3.2. Yao Samsao tâay *'come' and* mîing *'go'*

The difference in deictic viewpoint exemplified by (162a,b) is exactly paralleled in these Yao sentences:

(164a) *buə tɔ̀ɔ ʔâw-cúəy piəʔ kêem* ***tâay***
1PL *** family enter forest ***
'We brought our wives and kids into the forest (to join us there).'

(164b) *buə tɔ̀ɔ ʔâw-cúəy piəʔ kêem* ***mûng***
*** ***

'We brought our wives and kids into the forest (and left them there).'

4.3 Verb-compounds vs. verb concatenations

Certain Lahu verb sequences seem intermediate between lexical compounds and verb concatenations (see Matisoff, 1973, pp. 210–211).

4.3.1. "Idiomatic" concatenations

Occasionally the collocation of a versatile verb with a particular Vh acquires an unpredictable or "idiomatic" meaning. Thus *ša* (Vadj) 'pleasant; comfortable' participates as a Vv in many normal concatenations, where it means 'easy to Vh', e.g. *tâʔ ša* (Vh + Vv) 'easy to climb', *í-kâʔ qho ša* (OBJ + Vh + Vv) 'easy to fetch water', etc. In a few combinations, however, the meaning has become idiomatic:

pɔ-ša 'be wealthy; prosperous' ("born-easy")
chê-ša 'be healthy' ("dwell-easy")

These are more like lexical compounds than like true concatenations, and it is best to write them with hyphens.

Another interesting example is the compound *te-câ* 'cook' (lit. "do + eat"), which is homophonous with the true concatenation *te câ* (Vh + Vv) 'do for a living' (see 4.32).

4.3.2. Productive compound-formations

The opposite kind of problem is exemplified by the data in Figure 14. The verb *câ* 'eat' appears as an indubitable Vv in collocations where it has the bleached meaning 'Vh for a living'. However, in a number of other collocations with verbs relating to the preparation of food or the killing of an edible animal, *câ* maintains its literal meaning, i.e. 'Vh to eat'. These sequences are like fortuitous concatenations in that the two verbs represent temporally successive actions (though it is not natural to insert *lɛ* between the elements). They are like versatile concatenations in that the verbs which may precede *câ* in this construction are actually considerable in number. (Lahu happens to abound in verbs relating to the preparation of food.) But

they are most like lexical compounds in that the possible partners of *câ* are restricted to a narrow semantic range.

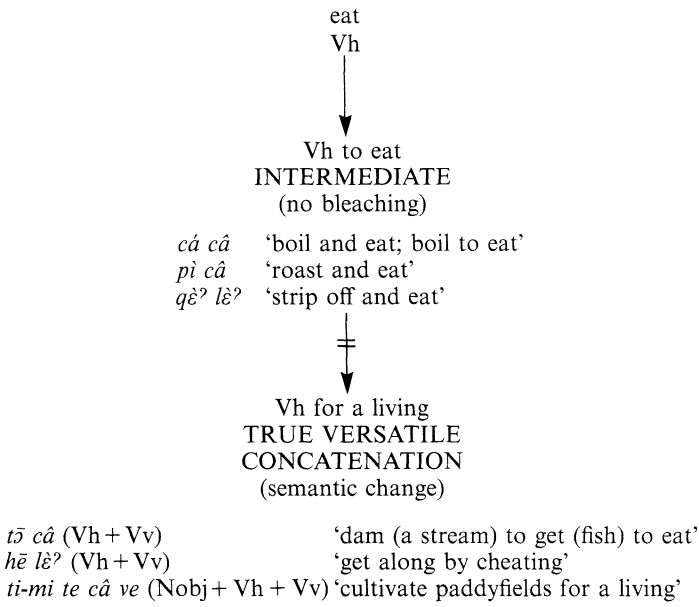


Figure 14: *Stages in the grammatization of* câ *'eat'*

4.4 Versatile verbs vs. verb-particles

4.4.1. Lahu pɨ *'be able'*

The high-frequency morpheme *pɨ* 'be able to V; possess the skill to V; possess the quality of Vh to a notable extent') is definitely a verb, since it can be negated (*yɔ mâ pɨ* 'cannot speak'). Yet it never seems to occur without another verb preceding, e.g.:

kâ pɨ	(Vh + Vv)	'be good at hearing; have sharp ears'
hàʔ-qá pɨ	(Vh + Vv)	'be compassionate' ("able to pity")

This makes *pɨ* unique among Lahu posthead versatile verbs, a strange animal less verby than a verb that can constitute a verbal nucleus all by itself, but more verby than a verb-particle (which cannot be negated). (See Matisoff, 1973: 679.)

4.4.2. Lahu 'perfective' tā *revisited*

We have noted (above 2.3.3) how this durative/perfective morpheme seems intermediate between a Vv and a Pv. In *The Grammar of Lahu* I decided (for complicated reasons involving ordering in multiparticle sequences) to consider *tā* to be a Pv and not a Vv, though I have had misgivings ever since. This problem is aggravated by the fact that as a head verb, *tā* (whose original meaning is 'place, put, set upright') seems to be conflating with another, accidentally homophonous verb of Tai origin, *tā(n)* 'set up, establish' (cf. Siamese *tâŋ*), in expressions like, e.g. *vê-bâ? tā ve* 'accuse' ("set up a sin").[40]

Curiously, a similar intermediate place on the Vv↔Pv continuum is occupied by another perfective morpheme, *pɔ̀* 'finish' (see above 4.2.1.1).

4.5 The syntagmatic continuum: loss of syllable/morpheme boundary

4.5.1. Lahu qay *'go' < *qa + e*

Certain verbs seem to attract certain particles as collocation-partners — conversely, certain particles occur with particular frequency after a handful of very high frequency verbs. Thus the asseverative particle *à* is very common after *dà?* 'be good', or *šī* 'know' (*dà?-à* 'Fine! O.K.!'; *šī à* 'I know it (already)!') The strong declarative particle *ɔ* is frequent after *phè?* 'be; be able' (*phè? ɔ* 'Sure I can do it!') and *chê* 'be in a place' (*chê ɔ* 'Here he is!; It's here all right!').

Just about the only native monomorphemic Lahu word that has the diphthong *-ay* is the verb *qay* 'go'.[41] Internal reconstruction, as well as comparative/historical evidence strongly indicates that this is an old fusion of a now obsolete verb **qa* (< PTB **ka* 'divaricate; spread the legs in striding, go' (STC #469) plus *e.* (Pv) 'transitive motion; motion away from deictic center'. (This is the antonymous counterpart of *la* (Pv) 'motion toward deictic center' (above 2.3.2).) This particle *e* must itself once have been a full verb, related to two other words meaning 'go' that survive in limited contexts, viz. *i* (archaic) as in the ritual phrase *ši-mɔ̀ tê phɔ̂ nɔ̀ tâ i* 'Wend not thy way to the land of the dead!', and *yì* (used in set expressions like *á-qà? yì ve* 'defecate (polite)' (lit. "go outside") and *yì-yì là-là te ve* 'keep going to and fro'). There is also an Akha cognate, *í* 'go'.

From a broad historical perspective we may thus assume a gradual development from two independent verbs in series, to a sequence of head-

verb plus grammaticized directional verb, to verb plus verb-particle, to complete incorporation of the particle into the verb:

V1 + V2 → Vh + Vv → V + Pv → V

4.5.2. Thai khráp *'polite particle'* < khɔ̌ɔ ráp

Equally dramatic is the case of the Thai word *khráp*, which is used by men as a sentence-final politeness particle (and as the polite way of saying "Yes" as an independent utterance):

(165) ***khráp, khráp*** — *cə* *paj thanthii* ***khráp***
PTCL go right away ***
'Yes, yes — I'll go right away, sir!'

This is historically a fusion of a two-verb sequence, part of the humilific expression *khɔ̌ɔ ráp thôot* 'beg to receive punishment' (*khɔ̌ɔ* (V) 'beg', *ráp* (V) 'receive').

This is quite analogous to the radical reduction undergone by the old English greeting *God be with ye!* (> *good-bye!* > *bye!*).

4.6. Cheshirization

Finally, we should mention a couple of cases where the phonological reduction of a grammatical morpheme has gone so far that all of its segmental phonemes have disappeared, leaving only a perturbation of the neighboring syllable's tone as the reminder of its former presence — the smile of the Cheshire cat, gradually fading away.

4.6.1. PTB *sya *'meat; flesh'* > *sə- *'body-part prefix'* > *HIGH TONE*

The Lahu root for 'eye' is *mɛ̂ʔ*, under the high-stopped tone. If this syllable had descended from an ordinary (fully voiced) nasal it would have acquired the *low*-stopped tone. We therefore infer the presence of an **s*-prefix at the Proto-Lolo-Burmese level. (This is amply confirmed by forms in the other Loloish languages; see Matisoff, 1972, #145.) The source of this prefix, which can be reconstructed for many other body-part etyma, was undoubtedly the independent root **sya* 'flesh; animal':

**sya-myak* → **səmyak* → *mɛʔ* (+HIGH TONE)

4.6.2. Atsi and Mpi: citation-particle on verb → special tone on verb

Burling (1967) noted the interesting fact that in Atsi (=Zaiwa), a language in the Burmish branch of Lolo-Burmese, words descending from PLB Tone *2 developed different tonal reflexes according to whether they were nouns or verbs. Later a similar phenomenon was observed in the S. Loloish language Mpi, where all three of the PLB open tones split according to the form-class of the etymon, yielding a 6-tone system (Srinuan, 1976; Matisoff 1978).

What undoubtedly occurred was the absorption into the verb-root of a citation-particle (actually a nominalizer, like Eng. *to* in *to go*). Such a particle is ubiquitous in TB languages, the Lahu form being *ve* (see Matisoff 1972b, and 2.2 above). This particle in many languages is a form with initial vowel — e.g. Akha *ɔ* — eminently fusable into the syllabic nucleus of the preceding verb. The vowel of the particle must eventually have been swallowed up, leaving only a perturbation of the tone to establish a form-class identity for its head-syllables.

5. Concluding remarks

5.1. Areal and typological differences in grammatizational tendencies.

Although many grammatizational tendencies are doubtless universal, there are certainly areal differences of nuance. Bréal (1897: 16 and passim) points out that Indo-European prepositions usually develop from *adverbs* (an independent and well-defined form-class). In the SVO languages of SE Asia, as we have seen, prepositions derive typically from full verbs.

Brute facts of word-order can be decisive in determining the paths grammatization will take. In the SE Asian linguistic area, verb concatenation has reached its fullest flowering — i.e. strings of up to 5 verbs in a row — in the SOV languages of the Tibeto-Burman family (especially in the phonologically reduced Sinospheric branches, like Lolo-Burmese). These languages are accustomed to bunching up long series of nominal arguments, without verbal intrusions. In their verb-phrases as well they are free to string together long sequences of verbs, without nominal interruptions. On the other hand, the SVO languages (Chinese, Vietnamese, Thai, Hmong-Mien, and the Karenic branch of TB) must break up verb-strings by inserting NP's as close to their governing verbs as possible, so that even their longest concatenations

can only achieve a length of 3 verbs (e.g. motion-verb sequences like Chinese *fēi jìn qù* 'go flying in' (see note 39)).

5.2. Predictability and controllability of grammatization

It would seem to be a truism that the course of semantic change is unpredictable in principle, partly because it is at the mercy of extralinguistic factors. We live in a world where a word meaning 'befigged; stuffed with figs' (Vulgar Latin **fīcātu*) could come to mean 'liver' (It. *fegato*, Sp. *higado*, Fr. *foie*), because of the Mediterranean custom of force-feeding geese with figs in order to cause their livers to hypertrophy into delectable *foie gras*.

The question then arises whether grammatizational meaning shifts are more predictable than semantic shifts in general? Joseph Greenberg has claimed (at the Symposium on Grammaticalization, May, 1988) that this is indeed the case, and that one might safely predict, e.g. that a verb meaning 'sneeze' or 'buy fish' would never be pressed into grammatical service. While we can certainly agree that such a development is much less likely than for a verb meaning 'give' to develop into a benefactive marker, one should perhaps never say *never* when talking about hypothetical semantic developments. Surely in some language or other a verb meaning 'sneeze' might acquire a figurative sense like 'to treat lightly; treat with contempt' (cf. Eng. *That's nothing to sneeze at*), giving rise to such grammaticized collocations as 'to sneeze-talk' (i.e. "talk disparagingly") or 'to sneeze-walk' (i.e. "walk insouciantly; strut")!

Equally unpredictable are the occasional cases of *degrammaticization* that may be observed — instances where an "empty morpheme" has filled itself up and acquired a true lexical root-meaning. Perhaps the best-known example in English is the word for the large conveyance we call *bus*, which is derived ultimately from the Latin dative plural ending of the form *omnibus* 'for all; for everyone'.

Attempts are sometimes made to intervene directly in the grammatizational process by deliberate language planning. Chinese linguist/administrators have recently taken a paternalistic approach toward "improving" the minority languages of China by encouraging the introduction of new functors from Chinese, e.g. in order to make inter-clausal semantic relationships more precise. Actually, however, the borrowing or calquing of functors from more prestigious languages is something that has gone on for millennia even without official encouragement![42] We have noted the areal diffusion of several functors above (see e.g. the discussions of Vietnamese *được*, Thai *thùuk*, Yao *tú²* and

Chinese **tək* (3.2.2.5); the Thai perfective particle *léɛw* and Chinese *liǎo* (4.2.1.2), Lahu *šē* 'attitude of regret' and Thai *sǐa* (2)).

Let us close by trying to predict the fate of two Slavic derivational suffixes in English.

The Russian agentive suffix *-nik* first impinged on the general American consciousness in 1957, with the successful flight of the first manned satellite, the *isskustvennyj sputnik zemli* 'artificial satellite of the earth', quickly dubbed *sputnik* (a tri-morphemic form also comprising *s-* 'with; together' and *-put-* 'journey'; i.e. "fellow-traveler, satellite"). That might have been the end of the saga had not the way already been paved for *-nik* by words introduced into American English from Yiddish, especially the useful word *nudnik* 'pest; pain-in-the-ass' (from a Slavic root meaning 'to bother; nag; get on the nerves of'). There is something ingratiating about the sound of *-nik*, which had long before ensured its success in Yiddish, leading to such made-in-America coinages as *allrightnik* 'somebody who's doing all right for himself; a prosperous person'; *neatnik* 'a compulsively neat person', and *nogoodnik* 'scoundrel; sonofabitch'. Yiddish influence was probably cooperating with *sputnik* to stimulate the brain of the popular columnist for the San Francisco Chronicle, Herb Caen, when he coined the word *beatnik* around 1960, to refer to members of the "beat generation". This coinage was spectacularly successful, opening the floodgates for *-nik*'s further penetration into the lexicon. Soon we had *refusenik* 'a Soviet Jew who was refused a visa to emigrate' — and most recently (February 1989) *waitnik*, 'an émigré who is still waiting in Italy for permission to come to the U.S.' It seems safe to predict that the future holds limitless promise for *-nik*. It is only a matter of time before such locutions as *government-and-bind-nik* will catch on.

To this success story we may contrast the bleak future that doubtless awaits another Slavic suffix, even though two words containing it are already familiar to millions of Americans. This is the noun-forming derivational morpheme represented by Polish *solidarność* 'solidarity; name of a labor union' and Russian *glasnost'* 'openness; frankness'. We may confidently venture to predict that this morpheme will never be attached to English roots. In the first place it takes a bit of sophistication to identify the Polish and Russian forms as cognate, so that they do not reinforce each other. Secondly, they are hard to pronounce (especially the Polish variety). Thirdly, they do not sound particularly cute. And finally, they do not fill a derivational or semantic lack — English has plenty of its own nominalizing suffixes, including the phonologically similar *-ness*.

In general, however, we must content ourselves with sitting back and

watching how languages change syntactically and semantically according to their own inner, inscrutable laws. The early generative conception of language change as involving "the addition of rules to the end of a grammar" naively implied that one could completely characterize all the rules of a grammar in the first place. Students of grammatization realize that worrying about where one grammar ends and the next grammar begins is a totally meaningless and futile pursuit. For "the new grammar is constantly being created on top of the willing and yielding ruins of the old."[43]

SYMBOLS AND ABBREVIATIONS

A ⨛ B	A is an allofam of B; A and B belong to the same word family
CLF	classifier
Cr	resultative complement
HON	honorific
Next	extentive noun
Nh	head noun; noun-head; head of a NP
Pn	noun-particle
Pu	unrestricted particle
Puf	final unrestricted particle
Punf	non-final unrestricted particle
Puniv	universal unrestricted particle
Pv	verb-particle
Vh	head verb; verb-head; head of a VP
Vint	intentional verb
Vmod	modal verb
vV	pre-head versatile verb
Vv	post-head versatile verb
PLB	Proto-Lolo-Burmese
POT	potential
PST	Proto-Sino-Tibetan
PTB	Proto-Tibeto-Burman
STC	Sino-Tibetan: a Conspectus (Benedict/Matisoff, 1972)
WB	Written Burmese
WT	Written Tibetan

NOTES

1. They eat bugs (as well as wounded members of their own species). The only bad habit they have is laying their droppings on you from the ceiling. This is probably why newly arrived foreigners in Thailand are urged to avoid sleeping with their mouths open.

2. One could point to many similar grammatizations in the SEA area — e.g. Thai *tua* 'body', *tua ʔeeŋ* 'by oneself' — but the grammatization of a word for BODY into a pronoun is more a universal than an areal semantic development: cf. Eng. *body* → indefinite pronouns like *somebody*, *anybody* (with reduced stress).

   Lahu has a separate word *ɔ̀-šɛ̄-phâ* for 'the male body', which also means 'male owner' or 'he himself':

   (14c) *yɔ̂ ɔ̀-šɛ̄-phâ dàʔ jâ* 'He has a fine physique.'

3. Incidentally, this expression contains a 4-verb concatenation (see below 3.1), with the main verb *pɔ* 'be born' preceded by two pre-head versatile verbs (*ǧa* 'OBTAIN/must', *ga* 'HELP/auxiliative'), and followed by the post-head versatile verb *pî* 'GIVE/benefactive.'

4. The various semantic developments of this etymon are diagrammed in a "metastatic flowchart" in Matisoff, 1985: 56.

5. *gâ* may have other shades of meaning as a verb-particle, e.g. 'like to V; be fond of V'ing': *ŋà á-pɔ̂-qu câ gâ ve* "I like to eat bananas" (i.e. habitual desire); or "imminent likelihood" (*mû-yè là gâ ve* 'It looks like it will rain'). There is a similar constellation of meanings shared by such particles as Burmese *hcɪ̃*, Yao *ʔɔ̃ɔy*, and Chrau *co·nh* (above 2.3).

6. The WB cognate *la* 'come' shows that Lahu *là* has the etymologically correct tone for the full verb.

7. Verbs meaning 'come' (and 'go') are eminently grammatizable, in a variety of syntactic contexts. For a discussion of the Thai postpositions *paj* 'GO; away from deictic center' and *maa* 'COME; toward deictic center', see below 4.1.2. For the use of Chinese *lái* 'come' as an introducer of purpose-clauses, see below 2.42.

8. In this usage it closely parallels Mandarin *lái* appearing as a post-verbal auxiliary: *fēi lái* 'come flying', *pǎo lái* 'come running', etc.

9. We shall see, however, that in certain contexts it has conflated itself with a distinct Tai-derived morpheme (below 4.4).

10. In this usage it closely parallels the Japanese verb *oku* 'put; place', which has acquired a perfective meaning as an auxiliary (governing the -TE form or gerund of the preceding verb): *kangaete oku* 'think sthg. over', *oite oku* 'put sthg. down to keep', *totte oku* 'take sthg. to keep', etc.

11. What have traditionally been called *verba dicendi* and *verba cognoscendi*. Cf. Frajzyngier, 1988.

12. *qôʔ* could descend from PLB **kap*, while *qo* could be from PLB **kam*. (The alternation of homorganic stops and nasals in syllable-final position is one of the most common morphophonemic patterns in Sino-Tibetan word-families.) It is possible that these words are to be related to the etyma reconstructed as PTB **gam* (STC #491) 'put into the mouth; seize with the mouth' and/or **hap* (STC #89) 'bite; snap at; mouthful' (both with Chinese cognates). On the other hand, there is a Proto-Tai form **gam* 'mouthful; bite' (Li Fang Kuei, 1977:198), underlying Thai *kham* 'word; speech; classifier for words'. A Tai donor language might well be the immediate source of the Lahu forms. The ultimate relationship of the Proto-Tai and Proto-Sino-Tibetan forms is undoubtedly one of borrowing in one direction or the other, though the resemblance might also be coincidental, involving sound-symbolism (closing the lips at syllable-final to suggest a bite).

13. The benefactive word *gěi*, here translated 'for', is a grammaticalized usage of the verb GIVE (below 3.2.3, 4.2.2).

14. I expanded somewhat on those reasons in a footnote: "It would require a great clanking of generative machinery: obligatory deletions of dummy NP 'subjects' (and other ghostly nominal arguments), rules specifying that no adverbs or particles may occur in any of the underlying sentences but the one underlying the Vh, etc. Most crucially, how would it help us to capture the fact that the meanings of the verbs as Vh's are systematically different from their meanings as versatile verbs?" (Matisoff, 1973: 552, note 43.)

15. To achieve other hierarchical orders of modification, the sentence must be recast, typically by breaking up the concatenation into two pieces, nominalizing one of them and embedding it as the object of the other. (See Matisoff, 1973: 218, Fig. 20.)

16. Despite the ending on *kutte*, the Japanese verb sequence *kutte shimatta* 'ate up; ate completely' (lit. "having eaten, finished"), also expresses a unitary monoclausal idea. The *-te* ending (unlike Lahu *lɛ*) appears on non-final members of Japanese verb-sequences whether or not the following verb has been "grammatized into" the same clause with it.

17. The theory of these crude semantic diagrams ("metastatic flowcharts") is developed in Matisoff, 1978. Representations like Figures 10–12 are merely intended to suggest pathways of association in "semantic space", and are not to be taken in too literal-minded a way. Whatever semantic space is, it is certainly at least non-Euclidean. We can hardly do justice to its unimaginable convolutions by pathetic symbols like arrows and straight lines. A double bar across a line is meant to show that the "boundary of grammatization" has been crossed, so that on the other side of the bar one is approaching the void of emptied-out, bleached semantic space.

18. The examples in this section are from Thompson, 1965 (pp. 228–9, 309–10, 314, 344–5) and Nguyen Dinh-Hoa, 1966 (p. 147), sometimes with slight changes.

19. Vietnamese does indeed have an adversative passive construction, but this is marked by the auxiliary verb *bị* 'suffer; undergo'. This word looks very much like the Mandarin adversative passive marker *bèi* (which originally meant 'to cover, as with a quilt'), as in:

(90a) *Tā bā-ba shì* ***bèi*** *dìzhǔ hài-sǐ de*
3SG father COP *** landlord murder PRT
'His father was murdered by a landlord.'

20. The Khmer expression *trəw haəy 'That's right!' ("correct already")* corresponds exactly to the Thai phrase *thùuk lɛ́ɛw*.

21. Huffman (1970) observes (p. 75) that in sentences like (103) that include a quantified expression *baan* may also occur in post-verbal position with little or no meaning difference:
(103a) *kñom tiw baan pii daaŋ* 'id.'

22. Huffman (1970) transcribes glottal stop by "q". Some of the examples in this section are from my own field-notes, but they have been more or less regularized into Huffman's transcription.

23. This is exactly analogous to the causative use of the Thai verb sequence *tham* + *hâj* + Clause (lit. "make + give + Clause").

24. *ve* is a nominalizing particle used in the citation form of verbs or noun-plus-verb collocations; *è* is an adverbializing particle.

25. We might dub these camps the *neatniks* vs. the *messniks*. For a discussion of our new affix *-nik*, see 5.2 below.

26. Cf. Alleton's (1978) review of Hagège, 1975, where she criticizes such terminological

innovations of the author as *adverbal* (pl. *adverbaux*) vs. *adverbial* (pl. *adverbiaux*), *nominoïdes*, *prédicatoïdes*, etc. Similar reservations (from an early generativistic point of view) were expressed by Lehman in his review (1978) of *The Grammar of Lahu*.

27. E.g., the distinction between Lahu verb-particles (which can only follow verbs) and unrestricted particles (which can follow nouns as well as verbs) is basic to my analytical framework (see above, 2.0). But there are two morphemes I wish to consider verb-particles (*šɛ̄* and *šɔ̄*) which can sometimes occur after nouns as well. The pros and cons of the various analytical possibilities are discussed in GL, 339–41.

28. These items have sometimes been called co-verbs in the literature (e.g. Li and Thompson (1974) for Mandarin, Clark (1978) for Vietnamese. Noss (1964: 154–6) recognizes a class of prepositions that are "weak-stressed forms of verbs". Parallel examples may be found at will in Hmong-Mien and the Karenic branch of Tibeto-Burman. The most perceptive analysis of Karenic multiverbal constructions yet achieved is to be found in an unpublished dissertation by D.B. Solnit (1986).

29. The first difference lies in the choice of verb each language made to grammaticalize into the instrumental sense. The Lahu word for 'use' is *yɛ̂* (< PLB **zum*[2]), which is actually ultimately cognate to Mand. *yòng* (< Old Chin. *d̰iung*), but this verb plays no grammatical role in Lahu. On the other hand Chinese does have a co-verb that derives from an old verb meaning 'take', none other than *bǎ*, but this has been grammaticized in a different direction, and is now the highly abstract marker of a pre-verbal "disposed-of" object. It has in fact become the most famous of the co-verbs of the language, with no fewer than 16 papers having been devoted to it during the first 21 years of the annual Sino-Tibetan Conferences!

30. The Pn *gɛ* is most often used comitatively, expressing 'with' in the sense of accompaniment:

(141a) *chi tê há ŋà-hɨ gɛ chê mɛ̄*
DEM one night 1PL *** stay PTCL
'Please stay with us tonight.'

31. This is very similar to the role of the "-TE form" or "gerund" in Japanese verb morphology (see (49b), above). Instrumental NP's are typically marked with the postposition ***de***, as in:

(145a) *naifu* ***de*** *yubi wo kitta*
knife *** finger OBJ cut-PAST
'(He) cut (his) finger with a knife.'

This sentence is neutral with respect to intentionality, as is the alternative construction with the gerundive form of ***tsukau*** 'use':

(145b) *naifu wo* ***tsukatte*** *yubi wo kitta.*
OBJ use -TE

To indicate intentionality, an adverb like *waza to* 'on purpose' must be added.

32. *ɔ́-thɛ* is a verb-compound of the loose, *resultative* type, formed from the verbs *ɔ́* 'bend' and *thɛ* 'break (of an elongated object)'. It may be negated as *ɔ́ mâ thɛ* 'not break when bent' ("bend not break").

33. These are cogently discussed in Noss, 1964: 159–62.

34. For a more general discussion of causative/benefactive grammatizations of GIVE-verbs, see above 3.2.3.

35. Note the use of the feminine agentive nominalizer *śɛ̄-ma* (above 2.1.4), nominalizing the verb-compound *te-câ* 'cook' (below 4.3.1).

36. There is an identical construction in Vietnamese, with *cho* used as a postposition (see above 3.2.3.3e).

37. The first verb in the Japanese examples are in the gerundive (-TE form). See note 31.

38. Cf. the treatment in Noss (p. 185), and Bickner's extensive study (1978) of the discoursal dynamics of these two words.

39. This sort of Thai 3-verb sequence is exactly paralleled in Mandarin:

verb of specific motion	+ IN/OUT UP/DOWN	+ COME/GO
V1	V2	V3

where the first verb expresses the specific manner of locomotion ('fly', 'run', 'swim'), the second signals whether the motion is IN or OUT (Mand. *jìn* vs. *chū*) or UP or DOWN (Mand. *shàng* vs. *xià*), and the third tells whether the whole verbal action is performed toward (*lái*) or away from (*qù*) the deictic center, e.g. *fēi xià lái* 'come flying down', *pǎo chū qù* 'go running out'.

The ordering of similar verbs in Yao is somewhat different, e.g.

Vh + *mîiŋ* + *tshuet* (*mîiŋ* 'go', *tshuet* 'emerge').

The latter verb is a loanword < Chinese (cf. Mand. *chū*).

40. This involves still another continuum, that of *etymological identity*. At what point do two originally distinct lexemes conflate into one? Cf. the conflation of English 'let' (as in 'let ball' (tennis)) < Mid.Eng. *letten* < OE *lettan* 'delay; hinder' and let 'permit' (< ME *leten* < OE *lætan*).

41. This verb has itself been grammatized, with several shades of meaning as a Vv: "go and Vh" (*lɛ̄-gî qay* 'go and play'); "Vh along/away' (*mɛ́ qay* 'disappear'); 'keep on Vh'ing' (*te qay* 'keep doing'); 'become more and more Vh' (*chu qay* 'get fatter and fatter').

42. As one random example, the high-frequency Hmong conjunction *tabsis* 'but' seems to have been borrowed long ago from Chinese *dàn-shì*. See C. Li (1988).

43. Matisoff, 1979: 2.

REFERENCES

Alleton, Viviane. 1978. "Compte-rendu de Claude Hagège (1975)" q.v. *Cahiers de Linguistique Asie Orientale* 3:85-92.

Becker, Alton L. and Karen Adams. 1969. "Prepositions as predicates." *Chicago Linguistic Studies* V.

Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus*. Contributing editor, James A. Matisoff. Cambridge, UK: Cambridge University Press.

Bickner, Robert J. 1978. *Directional Modification in Thai Fiction: The use of **come** and **go** in Text Building*. Monograph circulated at 11th Sino-Tibetan Conference, University of Arizona. 106 pp. MS.

Bréal, Michel. 1897. *Essai de sémantique*. 1st English ed., Henry Holt and Co.,1900, translated by Nina Cust; reprinted (1964) by Dover Publications.

Burling, Robbins. 1967. "Proto-Lolo-Burmese." *International Journal of American Linguistics 33(2): Part II. (= Publication 43 of Indiana Publications in Anthropology and Linguistics*, Mouton and Co., The Hague.)

Clark, Marybeth. 1978. "Coverbs and case in Vietnamese." Canberra: Australian National University. [*Pacific Linguistics* Series B, 48.]

Clark, Marybeth. 1979. "Synchronically derived prepositions in diachronic perspective: some evidence from Hmong." Paper presented at Twelfth International Conference on Sino-Tibetan Languages and Linguistics, Paris. 12 pp. MS.

Crisfield, Arthur G. 1974. "Lao final particles." In *Southeast Asian Linguistic Studies.* Nguyen Dang Liem (ed). Canberra Australian National University.

Frajzyngier, Zygmunt. Volume I. "The *de dicto* domain in language."

Givón, Talmy . 1979. *On Understanding Grammar.* New York Academic Press.

Haas, Mary R. 1964. *Thai-English Students' Dictionary.* Stanford: Stanford University Press.

Hagège, Claude. 1975. *Le problème linguistique des prépositions et la solution chinoise. Société de Linguistique de Paris Collection Linguistique* 71.

Huffman, Franklin E. 1970. *Modern Spoken Cambodian.* New Haven: Yale University Press.

Lehman, F.K. 1978. "Review of J. Matisoff, *The Grammar of Lahu." Journal of the American Oriental Society* 98(3): 296–7.

Li, Charles N. 1988. "The origin and function of switch reference in Green Hmong." Paper presented at the Symposium on Grammaticalization, Eugene, Oregon.

Li, Charles N. and Sandra A. Thompson. 1974. "Co-verbs in Mandarin Chinese: verbs or prepositions?" *Journal of Chinese Linguistics* 2(3): 257–78.

Li, Fang Kuei. 1977. *A Handbook of Comparative Tai.* Honolulu: University Press of Hawaii.

Matisoff, James A. 1969. "Verb concatenation in Lahu." *Acta Linguistica Hafniensia (Copenhagen)* 12(1): 69–120.

Matisoff, James A. 1972a. *The Loloish Tonal Split Revisited.* Center for South and Southeast Asia Studies, University of California, Berkeley. Research Monograph 7.

Matisoff, James A. 1972b "Oy, ve! Lahu nominalization, relativization, and genitivization." In *Syntax and Semantics, Volume I,* John Kimball (ed.), 237–57. New York: Seminar Press.

Matisoff, James A. 1973. *The Grammar of Lahu.* University of California Publications in Linguistics 75. Berkeley: University of California Press. (2nd Printing, 1982.)

Matisoff, James A. 1976. "Lahu causative constructions: Case hierarchies and the morphology/syntax cycle in a Tibeto-Burman perspective." In *The Grammar of Causative Constructions*, Masayoshi Shibatani (ed.), 413–42. New York: Academic Press.

Matisoff, James A. 1978. *Variational Semantics in Tibeto-Burman.* Philadelphia: Institute for the Study of Human Issues (ISHI).

Matisoff, James A. 1978. "Mpi and Lolo-Burmese microlinguistics." Tokyo: Institute for the Study of the Languages and Cultures of Asia and Africa. Monumenta Serindica 4.

Matisoff, James A. 1979. *Blessings, Curses, Hopes, and Fears: Psycho-ostensive Expressions in Yiddish.* Philadelphia: Institute for the Study of Human Issues (ISHI).

Matisoff, James A. 1981 "The bulging monosyllable, or 'the mora the merrier': echo-vowel adverbialization in Lahu." In *Contributions to Southeast Asian Linguistics:*

Essays in Honour of Eugénie J.A.Henderson, J. Davidson (ed.). London: School of Oriental and African Studies.
Matisoff, James A. 1983. "Linguistic diversity and language contact in Thailand." In *Highlanders of Thailand*, John McKinnon and Wanat Bhruksasri (eds), 56-86. New York: Oxford University Press.
Matisoff, James A. 1985. "God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes." *Journal of Asian and African Studies* 29: 1-81.
Matisoff, James A. 1988. *The Dictionary of Lahu*. Berkeley: University of California Press. University of California Publications in Linguistics 111.
Mazaudon, Martine. 1973. *Phonologie Tamang (Népal)*. Paris: Société d'Etudes Linguistiques et Anthropologiques de France.
Michailovsky, Boyd. 1989. *La langue Hayu*. Paris: Editions du Centre National de la Recherche Scientifique.
Mottin, Jean. 1978. *Eléments de grammaire Hmong Blanc*. Bangkok: Don Bosco Press.
Nguyen Dinh-Hoa. 1966. *Vietnamese-English Dictionary*. Rutland, VT: Charles E. Tuttle Co.
Noss, Richard B. 1964 *Thai Reference Grammar*. Washington, D.C.: Foreign Service Institute.
Ross, John Robert. 1972. "The category squish: Endstation Hauptwort." *Proceedings of the Eighth Meeting of the Chicago Linguistics Society*, 316–28.
Smeall, Christopher. 1976. "Grammaticalized verbs in Lolo-Burmese." *Linguistics of the Tibeto-Burman Area* 2(2): 273-87.
Solnit, David B. 1986. *A Grammatical Sketch of Eastern Kayah (Red Karen)*. University of California at Berkeley dissertation.
Srinuan Duanghom. 1976. *An Mpi Dictionary*. Bangkok: *Indigenous Languages of Thailand Research Project*, Vol. I.
Stern, Gustaf. 1931. *Meaning and Change of Meaning*. Bloomington: Indiana University Press.
Thomas, David D. 1971. *Chrau Grammar*. Honolulu: University of Hawaii Press. [Special publication 7. Oceanic Linguistics.]
Thompson, Laurence C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.
Warotamasikkhadit, Udom. 1988. "There are no prepositions in Thai." Paper presented at 21st International Conference on Sino-Tibetan Languages and Linguistics, Lund, Sweden.
Watters, David E. "Emergent word tone in Kham: The evolution of a Tibeto-Burman halfway house." *Linguistics of the Tibeto-Burman Area* 8(2): 36–54.
Wu Jingrong, et al (eds). 1979. *The Pinyin Chinese-English Dictionary*. Beijing: The Commercial Press.

Ways to go in Rama: A Case Study in Polygrammaticalization

Colette G. Craig
University of Oregon

1. INTRODUCTION

The term 'polygrammaticalization' is used here to describe the phenomenon by which a single morpheme is the source of multiple 'grammaticalization chains'[1] where the expression 'grammaticalization chain' itself needs explaining. 'Grammaticalization' is the evolutionary process by which grammatical morphemes arise. Although it is assumed that the original source of all grammatical morphemes is ultimately a lexical morpheme, the term grammaticalization often covers instances of the evolution of free to bound grammatical morphemes:

1. lexical ...> free grammatical ...> bound grammatical

FIRST TYPE OF GRAMMATICALIZATION	SECOND TYPE OF GRAMMATICALIZATION

Instances of the first type of grammaticalization, from lexical to free grammatical morpheme, are the widely documented cases of lexical verbs as sources of adpositions and verbal auxiliaries. Instances of the second type of grammaticalization, from free to bound grammatical morpheme, are found in the formation of verbal affixal morphology, such as bound subject (and object) person agreement markers arising from free personal pronouns, and bound tense/aspect markers from free verbal auxiliaries. The term 'chain' is used to evoke the step-by-step nature of the grammaticalization process, which is most directly observable in the pairing of two morphemes through a scenario of change, creating links with the internal structure outlined below:

2. [SOURCE ... pathway ... OUTCOME]

Chaining happens when the outcome of a link becomes the source component of another link, the notion of source being a relative one:

SOURCE pathway OUTCOME/SOURCE pathwayOUTCOME/SOURCEpathwayOUTCOME

This is of course a rather simplistic account of what is meant by grammaticalization chain, but it should suffice for the purpose of this paper, which is to focus primarily on determining the immediate and distant sources of particular grammatical morphemes, and on reconstructing, at least in bare outline form, several grammaticalization chains in a specific language. The principle aim of this paper is to account for an additional characteristic of chaining: the fact that a source morpheme may be at the origin of multiple grammaticalization chains which may develop in separate functional domains within the same language. Additional issues associated with discussions of grammaticalization and chaining — unidirectionality of the process, search for parameters of correlation, continuous or discontinuous nature of the evolutionary scheme — will be acknowledged only as they bear on the task at hand.

Incorporating the notions of grammaticalization and chaining into the descriptive grammar of a language has the effect of adding both a dynamic and a diachronic dimension to the task. In the synchronic study of a particular language, the phenomenon of 'chaining' surfaces in morphemes related by polysemy which can be interpreted as traces of links of various chains. The degree of semantic connectedness between polysemous morphemes can then be seen to correspond to the relative stages of the morpheme on a grammaticalization chain. Although it is rare to have all the links of a grammaticalization chain still present at any one time in a language, comparative as well as typological data can usually help reconstruct the chaining process.

This paper is a case study of grammaticalization chains in Rama, a Chibchan language of Nicaragua. It responds to the needs for an adequate description of the persisting recurrence of a particular morpheme *BA(NG)* which appears in various shapes and in numerous functions. Some of its grammatical functions are illustrated in the following set of examples:

(1) a. postposition
naas sii ba aa taak-iikar
I water PSP neg go-want
'I don't want to go get/for water.'

b. conjunction of subordination
tiiskama ni-sung-bang taak-i
baby 1-see-SUB go-TNS
'I am going (in order) to see/look at the baby.'

c. prospective aspect
tiiskama ni-tanang-bang
baby 1-look at-ASP
'I am going to look at the baby.'

In the course of accounting for the various uses of the grammatical morpheme *BA(NG)*, most of the major grammatical structures of the language will need to be considered, including verb words, adpositional phrases, main and subordinate clauses, since *BA(NG)* functions either as a postposition or as one of a number of verbal affixes identifiable as argument, subordination, aspect or mood marker.

The claim of the paper is that the multiplicity of grammatical functions of *BA(NG)* observed in the synchronic grammar of Rama can be plotted on several grammaticalization chains, the components of which have already been recognized in the verbal morphology of other languages of the world. This case study from Rama will hopefully contribute to a better understanding of how such grammaticalization chains evolve, by adding evidence of likely sets of source and target morphemes and possible paths of evolution. However, the most important contribution of this study will lie in the description of the rich network of grammatical constructions which tightly holds many parts of the Rama grammar together. This network consists of parallel and interconnected chains of grammaticalization which have all ultimately evolved from a single lexical source — in this instance the lexical verb *BANG* 'GO' through different grammatical domains of the language.

The paper is organized as follows. The major typological characteristics of Rama are given in Section 2 as a background on the major grammatical structures of the language. Next, all identified instances of the grammatical morpheme *BANG* are examined. They are presented in pairs, according to the structure of links they form, as mentioned above [SOURCE ...pathway... OUTPUT]. The first two sets of pairs belong to the domain of argument marking: one relates the goal postposition *BANG* to a certain type of verbal prefix *BA-* called 'relational preverb' (section 3), the other relates the same goal postposition to a subordinating marker *BANG* (Section 4). The other sets of paired *BANG* morphemes belong to the domain of tense-aspect-

modality marking and relate a form of the lexical verb 'GO' to aspectual auxiliaries and suffixes as well as mood markers (Section 5).

A number of hypothesized grammaticalization chains are then reconstructed from the pairings of morphemes presented above. Section 6 reconsiders the grammaticalization chains of the domains of argument marking and tense-aspect-modality marking. All these data are brought together to offer an overview of the interactions postulated between the various grammatical constructions.

2. THE RAMA LANGUAGE: A TYPLOGICAL OVERVIEW

Rama is an obsolescing language spoken today by a few dozen native speakers on the Atlantic Coast of Nicaragua.[2] The earliest linguistic material on the language is from the early part of this century and is limited to word lists and short grammatical sketches (Lehmann 1914, 1920; Conzemius, 1927). These early Rama references mention only some of the material covered in this paper. The majority of the postpositions are identified, the relational preverbs are not recognized as such, although some appear in examples, some of the imperative forms are given but the majority of the aspect markers are not. The discrepancy between the earlier data and that which are presented here begs the question of whether it is due to the way the data were gathered, or whether it is an indication of linguistic change, or both.

The early data were collected entirely through direct elicitation and contain no text material, while the data on which the present analysis is built consists primarily of a collection of spoken narrative texts. Complementary direct elicitation was used mostly to clarify the morphemic analysis of the texts. Although texts were collected from two speakers, an older near-native speaker and a younger native speaker, the language described in this paper is that of the native speaker alone.

Rama has a basic SOV word order and exhibits a number of the typological characteristics of verb final languages, such as postpositional phrases, preverbal argument marking and postverbal tense-aspect marking, all illustrated below:

(1) SOV
kumaa bauli ku-u
woman bowl take-TNS
'the women took bowls.'

(2) POSTPOSITIONS

ipang ika	*nguu ki*	*kumaa u*
island from	house to	woman with
'From the island'	'in the house'	'with the woman'

(3) FULL LEXICAL SUBJECT VS SUBJECT PROCLITIC:[3]

a. *namangku kruubu tabi-i amaing; kiikna isii*
now tiger come out-TNS again; man like
i-kalmk-i
3-stands-TNS
'Now the tiger comes out again; he stands up like a man.'

b. *nah ngriing-i tatah ni-uung-i*
I beat-TNS mashed 1-do-TNS
'I beat it and I mash it up.'

(4) POSTVERBAL TENSE-ASPECT-MODALITY MARKING

a. *an-upluul-u* b. *an-alpaik-i*
3PL-dream-TNS 3PL-think-TNS
'they dreamt' 'they think'

c. *kruubu kiikna kwis-atkul-u*
tiger man eat-ASP-TNS
'The tiger ate the man all up.'

d. *ma kalma apaak-ut naing isii*
you dress sew-MOD mine like
'Can you sew a dress like mine?'

(5) POSTVERBAL MARKING OF SUBORDINATION[4]

kruubu an-sung-ka an-alpaik-i kiikna
tiger 3PL-see-SUB 3PL-think-TNS man
'When they see the tiger, they think it is a man.'

All the above examples point to Rama having a morphosyntax rather typical of verb final languages. However, at the clause and sentence levels, Rama cannot be said to be a strictly verb final language in that it takes both preverbal and postverbal oblique arguments, and has both subordinate/main and main/subordinate clausal order.

(6) OBLIQUE ARGUMENTS

a. V obl: *nah ngalbi-u [naing taata kang]*
I run-TNS my father from
'I ran away from my father.'

b. Obl V: *[salpka taara sauk u] i-ku-u*
fish big hook with 3-catch-TNS
'He caught a big fish with a hook.'

(7) SUBORDINATE CLAUSES

a. subordinate + main clauses
[nah maa alkuk-kata], nah uwaik siik-ut
I you hear-SUB I long time come-IRR
'If I had heard you, I would have come a long time ago.'

b. main + subordinate clauses
kruubu tamaaski ui tabi-i [kumaa i-sung-bang]
tiger morning every come out-TNS woman 3-see-SUB
'The tiger comes out every morning in order to see the woman.'

This section has established the basic traits of Rama morphosyntax needed to proceed with the study of a number of interlocking grammaticalization chains. Postpositional phrases and bound verbal morphology define two functional domains: a preverbal subject and object marking and a postverbal tense-aspect-modality marking, a distribution consistent with the morphological typology of verbs in verb-final languages.

3. FROM POSTPOSITION TO RELATIONAL PREVERB

The first pairing of morphemes to be examined is that of the goal postposition BANG and the relational preverb *BA-*, as an instance of grammaticalization from a free grammatical morpheme to a bound verbal affix. Rama relational preverbs, of which *BA-* is the most commonly used, were initially described in Craig (1987) and then brought into typological and diachronic perspective with other cases of the phenomenon from various languages of the Americas in Craig and Hale (1988).

The expression 'relational preverb' is meant to be self explanatory: 'preverb' because it is prefixed to the verb and 'relational' because it marks a relation between an argument of the proposition and the verb. The sentences

below illustrate how these two types of argument marking co-exist, with examples of the two types of instrumental marking, the postposition (PSP) *u*, as shown in 8b, and the related relational preverb (RP) *yu-* found in (8a):

(8) Text excerpt: the 'kiskis' (tongs)

a. *ung-ki yaadar tkua yu-nsu-uung-kama*
pot-in thing hot RP/with-1PL-make-SUB
'For us to make hot things in the pot with (it).'

b. *nsu-suluk u angka nsu-uung-i*
1PL-finger PSP/with cannot 1PL-make-TNS
'With our fingers we can't do it.'

The hypothesis developed in Craig (1987) and Craig and Hale (1988) is that Rama relational preverbs come from postpositions through a process of progressive incorporation of the postpositional element into the verb word — from discourse motivated cliticization to morphological lexicalization.

The interesting aspect of relational preverbs in Rama is that they co-occur in the language with the postpositions from which they are derived, as shown in the inventory of relational preverbs given below:

(9)

POSTPOSITIONS		RELATIONAL PREVERBS
ba(ng)#	'goal, target'	*ba-*
u#	'comitative, instrumental'	*yu-*
ka(ng)#	'ablative, source	*k(a)-*
su#	'locative'	*su-*
aa(k)#	'object'	*yaa-*
ki#	'locative'	
kama#	'beneficiary'	
ki(ng)#	'beneficiary'	
aing#	'genitive'	

Postpositions and relational preverbs differ in their phonological shapes in that most postpositions appear in one of two forms, either a long form with a final velar consonant, which is used in postverbal or left dislocated postpositional phrases, or an open syllable form without the velar. Relational preverbs are always in the short form, conforming to the phonological reduction which is said to correlate with the process of grammaticalization of a free morpheme into a bound morpheme. In particular, while the postposition *BANG* exists in both long and short form *BANG* and *BA*, the relational preverb is always the short form *BA-*.

There are various patterns of use of postpositions and relational preverbs attested in Rama: pre- or postverbal postposition with an overt NP (10a,

10a'), relational preverbs with or without the NP with which they are construed (10b,c), as outlined below:

(10) a. [NP PSP] V a'. X V [NP PSP]##
b. 0 [RP-V]
c. NP [RP-V]

The sets of examples in (11, 12) below illustrate how not all the patterns shown in (10) are found for any specific pairing of postposition, relational preverb and verb. For example, the verb *ngalbi-* 'to run' cannot take the relational preverb *ka-*, but can take the relational preverb *ba-* in the pattern given in 11c. Conversely, it will take the postpositional constructions with *ka(ng)* but not the ones with *ba(ng)* of patterns a. and a'.

(11) a. *naing taata ka na-ngalbi-u*
my father PSP/from 1-run-TNS
'I ran away from my father.'

a'. *na-ngalbi-u naing taata kang*
1-run-TNS my father PSP/from
'I ran away from my father.'

b. *ka-na-ngalbi-u*
RP/from-1-run-TNS
'I ran away from (him).'

c.* *naing taata ka-na-ngalbi-u*
my father RP/from-1-run-TNS
(I ran away from my father.)

(12) a.* *paalpa ba aa an-alpi-u*
manatee PSP/for NEG 3PL-look-TNS
(They did not look for a manatee.)

a'.**aa an-alpi-u paalpa bang*
PSP/for

b. *ba-an-alpi-u*
RP/for-3PL-look-TNS
'They looked for (it).'

c. *paalpa aa ba-an-alpi-u*
manatee NEG RP/for-3PL-look-TNS
'They did not look for a manatee.'

The discrepancy of behavior illustrated by the ungrammatical examples (11c* and 12a*) is indicative of two characteristics of the phonemenon of grammaticalization of postposition into relational preverbs in Rama. One is that the grammaticalization process is postposition and verb specific. The other is that it proceeds in stages, from simple cliticization to the verb to complete lexicalization with the verb, yielding two types of relational preverbs: proclitized and lexicalized.

The differences between patterns of procliticized and lexicalized relational preverbs emerged from a study of the uses of relational preverbs in narrative texts. The results of the text counts are contained in Table 1 below which is based on Craig and Hale (1988: 324). The postpositions are ordered vertically in their increasing order of frequency of use as postposition, and, conversely, their decreasing frequency of use as relational preverb:

Table 1. *Text counts of Rama postpositions and relational preverbs*

PSP/PV	*a. [NP PSP];*	*b. [0 PSP];*	*c. 0 [PV-verb];*	*d. NP [PV-verb]*
ba(ng)- PURPOSE	— (0)	— (0)	39% (13)	61% (20)
u/yu— ASSOC/ INST	31% (35)	— (0)	51% (57)	18% (20)
a(ak)ya- DATIVE	49% (21)	— (0)	39% (17)	12% (5)
su/su- LOC	89% (34)	— (0)	11% (4)	— (0)
ki/ki- LOC	98% (90)	— (0)	2% (2)	— (0)

The table emphasizes the need to consider each morpheme separately, as each one displays a different distribution pattern. The most glaring fact to be gleaned from Table I is the series of zeros in column b. which creates a gap in the pattern of possible constructions.

The significance of the gap is that there are no stranded postpositions in Rama: in case of zero anaphora affecting the object NP of a postposition, the postposition occurs attached to the verb as a relational preverb. This analysis was later confirmed through direct elicitation.[5]

The relational preverbs which result from the systematic avoidance of stranded postpositions are analyzed as being cliticized, and are to be found in column c.[6] Meanwhile, the relational preverbs found in column d. represent cases of lexicalization of the relational preverb with the verb, as instance of advanced incorporation.

As discussed in greater detail in Craig and Hale (1988), one can distinguish between cliticized and lexicalized relational preverbs in Rama by:

a) the nature of the conditions under which they occur;
b) their degree of productivity;
c) the semantic relation that holds between preverbs and postposition.

All the criteria will be reviewed in turn.

3.1. The nature of the conditions under which they occur

The cliticized relational preverbs are those which arise in case of zero anaphora of the NP and the lexicalized ones are those which are used in conjunction with an overt NP. The fact that, on the rare occasion that a pronoun is used — such as in first and second person or in the third person for emphasis — it is accompanied by a postposition and not a relational preverb supports further the claim that the trigger of the procliticization of a postposition is the non-overtness of a third person object NP itself.[7]

3.2. Their degree of productivity

The cliticization of the stranded postposition is productive in the sense that the procliticized relational preverb may occur with all sorts of verbs, and that it affects all postpositions whose NP is non overt, the limit to the freedom of association of a cliticized relational preverb and a verb being of a pragmatic nature. In contrast, the incorporated relational preverbs are lexically restricted to combinations of specific postpositions with specific verbs. A text example (13) will be given to illustrate the productivity of the cliticized relational preverb. The text, which is a fuller version of (8) above shows the procliticization of the instrumental relational preverb *yu-* in every instance of zero anaphora of its object NP, once the NP *kiskis* 'tongs' has been established in

(13a) as the referent NP for the oblique instrument arguments of the different verbs of (13b-e). In all four clauses, the anaphoric instrumental NP is non-overt and the verb takes an instrumental relational preverb *yu-*. However note how, in (13f), an instrumental postposition appears, specifically when a new instrumental NP is introduced, as a full lexical NP which establishes a contrast between two instruments, the tongs '*kiskis*' and *nsu suluk* 'our fingers'. This switch to the postpositional phrase pattern supports the analysis that the relational preverbs in the preceding clauses are cliticized to the verb because of the non-overtness of their object NP.

(13) Cliticization of stranded PSP with anaphora marked ( ).

a. *Nainguku kiskis nsu-kuaakar-i,*
so tongs 1PL-have-TNS
That's why we have the tongs,

b. *suuli-kaas yu-nsu-auk-kama,*
animal-meat ( ) RP/with-1PL-roast-SUB
for us to roast meat with (it),

c. *sumuu yu-nsu-apii-kama*
banana ( ) RP/with-1PL-extract-SUB
for us to take out banana with (it),

d. *an ungi karka salpka-kaas yu-nsu-kaniis-kama*
and pot out:from fish-meat ( ) RP/with-1PL-fry-SUB
and for us to fry fish from the pot with (it),

e. *ungi yaadar tkua yu-nsu-uung-kama,*
pot thing hot ( ) RP/with-1PL-make-SUB
for us to do hot things in the pot with (it).

f. *Nsu-suluk u angka nsu-uung-i,*
our-finger PSP/with can't 1PL-make-TNS
With our fingers we can't do it,

g. *nsu-suluk y-auk-baing-uting*
our-finger 3-burn-too:much-TNS
we would burn our fingers too much.

The text sample shows a procliticized instrumental relational preverb associating with a variety of verbs: 'to roast' in (13b), 'to take out' in (13c), 'to fry' in (13d) and 'to make' in (13e), i.e. the process is productive. It also demon-

strates how the instrumental oblique argument of the same verb 'to make' can also be expressed through a postpositional phrase, as in (13f), the choice between postposition and relational preverb being a matter of the overtness status of the instrumental NP itself, and not of the combination of specific postpositions with specific verbs.

The major lexicalized patterns are combinations of the comitative *yu-* with motion verbs, to be considered in the next section, and combinations of a limited number of verbs with the relational preverb *BA-*. There were 17 examples of the *BA-* relational preverb with an overt NP in the text study. Only two verbs accounted for 16 of the 17 cases: They were [NP *BA-alpi-*] 'to look for something' (11 occurences) and [NP *BA-ting-*] 'to want something' (5 occurrences). These two verbs are characterized by the fact that they are only found in the relational preverb construction with a direct object NP, or conversely, the objects of these two verbs can never appear as the objects of free standing postpositions. Both verbal roots exist by themselves, with associated but different meanings: *aapi-* means 'to find' and *ting-* means 'to happen'.

Examples of two incorporated relational preverbs are found in the following text sample (14): The *BA-* of the construction [NP *BA-alpi-*] in (14a) (which contrasts with the bare verb *aapi* of the next clause (14b)) and the *yu-* of the construction [NP *yu-siik-*] found in clauses (14c and d):

(14) a. *ngulkang ba-an-alpi-u*
wild pig RP/for-3PL-look-TNS
'They looked for the wild pig.'

b. *an-aapi-u an-aung-i*
3PL-find-TNS 3PL-say-TNS
'They found (it), they say.'

c. *naiksangku ning ngulkang yu-an-siik-u*
that way the wild pig RP/with-3PL-come-TNS
'So they brought the wild pig.'

d. *ngulkang bangku yu-an-siik-u kaing*
wild pig now RP/with-3PL-come-TNS DISC.
'They brought the wild pig now.'

e. *an-angsk-u*
3PL-clean-TNS
'They cleaned (it).'

As will be discussed in the section on tense-aspect-modality below, the relational preverb *BA-* is found elsewhere in other instances of lexicalization such as the formation of the progressive aspect marker *baakar*, a compound of *ba-aakar* [RP/for-V/be].

With the instrumental/comitative relational preverb *yu-*, the most common instance of lexicalization involves verbs of motion, *siik* 'come' or *taak-* 'go'. A comparison of the cases of lexicalization of *BA-* and *yu-* brings out the need to introduce the notion of degrees of lexicalization. While the *BA-* lexicalizations discussed above represent a fairly advanced stage of the process, the instances of *yu-* incorporation represent a more incipient degree of lexicalization. The pair of examples below show a clearly instrumental postposition *u* and an incorporated relational preverb which, although it may still be interpreted as an instrumental, exhibits semantic drift as will be discussed shortly:

(15) a. PSP: *Maukala i-park-i seem ngabang u*
net 3-make-TNS same silkgrass PSP/with
'He makes nets with silkgrass.'

b. RP: *Naing taata ngabang yu-i-siik-i nguu ki*
my father silkgrass RP/with-3-come-TNS house in
'My father brings the silkgrass in the house.'

3.3. The semantic relation that holds between each type of relational preverb and its postposition of origin

The third factor that distinguishes procliticized relational preverbs from incorporated relational preverbs is their stable semantics. While cliticized relational preverbs maintain the interpretation that they would have as postpositions, fully incorporated relational preverbs exhibit semantic drift from their associated postpositions, as was just mentioned earlier.

The best illustration of this difference can be found in examples of the intrumental/comitative relational preverb *yu-*. The above text sample (13) showed how the postposition in (13f) and the cliticized relational preverbs of (13b-e), were all given the same instrumental interpretation, i.e. there was no semantic drift of the relational preverb.

Meanwhile, examples (15) above and (16,17) below illustrate how there is a semantic difference between constructions with a free postposition *u* and constructions with an incorporated relational preverb *yu-*. One way to charac-

terize the semantic difference is in terms of the relative control of the agent over the movement of the object, with the incorporated form corresponding to a reading of increased control of the agent:

(16) a. PSP: *... barka aa i-taak-u baaning anul u*
but NEG 3-go-TNS then them PSP/with
'... but she did not go with them.'

b. PSP: *Taa u m-taak-u*
who PSP/with 2-go-TNS
'With whom did you go?'

c. RP: [+agent] *Tiiskama taa yu-taak-u*
child who RP/with-go-TNS
'Who took (carried) the child?'

(17) a. RP: *ngurii psutki yaing tiiskama yu-i-taak-u*
hole inside her child RP/with-3-go-TNS
'Inside the hole she took her child.'

b. PSP: [+agent] *namaa y-aakar ngurii ki yaing tiiskama u*
still 3-stay hole in her child PSP/with
'She stays still in the hole with her child.'

The complex verbs *yu-taak* [with-go] 'to take, carry' and *yu-siik* [with-come] 'to bring' are the most common instances of incorporated relational preverbs: the comitative postposition with either of the two basic verbs of movement. In fact, of the 20 tokens of the preverb *yu-* used in combination with an overt NP found in the text study, 18 are in combination with either *siik-* 'come' or *taak-* 'go'.[8]

The goal of this section was to establish the existence in Rama of postpositions and bound relational preverbs paired through a process of progressive grammaticalization. The phenomenon has been documented for a number of languages of the world, including many languages of the Native Americas. The contribution made by the Rama data is two-fold. It constitutes a good case for showing the origin of such relational preverbs, since the postpositions from which the relational preverbs are derived co-exist in the language. It is also an interesting case for exploring the process by which such relational preverb constructions may arise through a series of steps from cliticization of stranded postpositions to complete lexicalization into the verb word.[9] If a distinction is established between the two types of relational

preverbs, cliticized and lexicalized, one can begin to talk about a grammaticalization chain:

Postposition ---→ cliticized RP ---→ lexicalized RP

A discussion of the nature of the chain and the heterogeneity of its component links will be postponed until the final section.

4. FROM POSTPOSITIONS TO SUBORDINATING MARKERS

The second instance of grammaticalization to be considered in this paper relates subordinating morphemes to postpositions. The phenomenon has been observed in other languages and does not exhibit very different characteristics in Rama, as will be shown below, although it provides further evidence for establishing certain paths of evolution linking certain postpositions to certain subordinate markers. Of specific interest for this paper is the pairing of the goal postposition *BANG* and the purpose subordinator *BANG*.

The structural characteristics of the subordinate clause are the absence of tense marking and the presence of a subordinating morpheme which expresses the meaning relation of the dependent clause to the main clause.

The set of subordinating morphemes includes:

(18)

purposive	*bang*	'in order to, for'
purposive	*kama*	'in order to, for'
temporal/conditional	*ka*	'when/if'
counterfactual	*kata*	'if'
succession	*su*	'upon, and then, since'

These subordinating morphemes are suffixed to tenseless verbs:[10]

(19) a. *kruubu tamaaski ui tabi-i kumaa i-sung-bang*
tiger morning every come out-TNS woman 3-see-SUB
'The tiger comes out every morning to see the woman.'

b. *kruubu an-sung-ka an-alpaik-i kiikna*
tiger 3PL-see-SUB 3PL-think-TNS man
'When they see the tiger, they think it is a man.'

The presence of a postverbal postpositional phrase in the following example underlines the fact that the subordinating morphemes are verbal suffixes rather than clausal conjunctions of subordination:

(20) *kiskis yu-i-siik-ka nguu ki*
tongs RP/with-3P-come-SUB house in
sut kumaa kaa abung su kiibi uung-kama
we woman put fire in straight make-SUB
'when he brings the tongs into the house we women put it in the fire to straighten it.'

The most interesting aspect of subordination in Rama is the nature of the subordinating suffixes themselves. All the Rama subordinators are related to postpositions, as described below.

(21)

Subordinators		Postpositions	
bang	'purpose'	*ba(ng)*	'goal'
kama	'purpose'	*kama*	'goal'
ka	'time, condition'	*ka(ng)*	'ablative'
kata	'counterfactual condition'	*ka(ng)*	,,
su	'time, after/upon'	*su*	'locative'

As a verb final morpheme, the purpose subordinator *BANG* is always in the long form.[11]

Examples illustrating the use of the pairs of postpositions and subordinators follow:

(22) *ba(ng)* a. PSP *naas sii ba aa taak-iikar*
I water PSP/for NEG go-want
'I don't want to go for water.'

b. SUB *tiiskama ni-sung-bang taak-i*
baby 1-see-SUB/to go-TNS
'I am going to see/look at the baby.'

(23) *kama* a. PSP *tamaik suuli-kaas ni-auk-ut maing kama*
tomorrow animal-meat 1-cook-IRR you PSP/for
'Tomorrow I will cook some meat for you.'

b. SUB *nah suuli-kaas ba-alp-i ni-paya-kama*
I animal-meat RP-look-TNS 1-buy-SUB/for
'I am looking for meat to buy.'

(24) *su* a. PSP *ipang su an-siik-u*
island PSP/on 3PL-come-TNS
'They came to the island.'

b. SUB *nais tum-ting-atkut-su y-aakir-i*
right so dark-happen-ASP-SUB/upon 3-stay-TNS
'Upon getting dark, he stays.'

(25) *ka(ng)* a. PSP *nah altawa-i naing taata kang*
I afraid-TNS my fatherPSP/from
'I am afraid of my father.'

b. SUB *nah kaafi ngu-atkut-ka kalma ni-sku-ut*
I coffee drink-ASP-SUB clothes 1-wash-TNS
'When I have drunk up my coffee, I will do the wash.'

c. SUB *nah maa alkuk-kata, nah uwaik siik-ut*
I you hear-SUB I long time come-IRR
'If I had heard you, I would have come a long time ago.'

The hypothesized grammaticalization of postpositions into subordinators raises issues of why and how postpositions become subordinators — assuming the evolution always goes in that direction — and which postpositions become which subordinators — assuming that some patterns can be established across languages of the world where similar evolutions have arisen.

The issue of why postpositions would be the source of subordinators is not a very complex one. Structurally, it has to do with subordinate clauses being treated like nominal constituents and the analogical use of the marking of oblique nominal arguments of a simplex sentence (monoclausal) for oblique clausal arguments of a complex sentence. In Rama there are few overt features of clausal nominalization — besides the stripping of tense on the embedded verb and the structural parallelism between postpositional phrases and adverbial subordinate clauses. Both types of clauses have the same word order, both within their respective phrasal structure and within the sentence structure; the same way postpositional phrases are either post- or preverbal, subordinate clauses either follow or precede the main clause.[12] More interesting is the semantic relation that holds between postpositions and subordinators, and the way in which postpositions match subordinators. The matchings observed in Rama support the patterns of prototypical matching described in Genetti (1986) for data from completely unrelated languages (Tibeto-Burman) and reproduced below, with the Rama cases in bold letters:

(26) postpositions: subordinators

locative	if/although, when/while/after
ablative	when/while/after, because, non-final
allative	purpose
dative	purpose
ergative/instrumental	because/when/while

Rama provides additional evidence, therefore, for the well-documented extension of the dative and allative case markers into subordinators of purpose clauses, the language specific version being the extension of the goal postposition *BANG* into the subordinator of purpose *BANG*.

The Rama pattern of matching for the other cases of grammaticalization of postpositions into subordinators falls within the patterns found in Tibeto-Burman languages. For instance, one of the extensions of the locative case is a time subordinate clause: the Rama locative *su* 'on' is the source of the time subordinator *su* 'upon, after'. The extension of the ablative postposition *ka(ng)* into the time/conditional subordinator *ka* 'when/if' and the counterfactual conditional subordinator *kata* 'even if' is not semantically as simple as the preceding ones, in that the semantic analysis of the path of extension from the concrete nominal case to the abstract clausal subordinator is not as straightforward as for the other matchings. It falls, however, within the range of expectations.

This section has presented a second process of grammaticalization involving postpositions, this time relating them to subordinators. While this phenomenon does not represent typologically new data, it provides additional empirical data about the source and evolution of subordinators in some languages of the world, and does so in a rather clear fashion since postpositions and subordinators in Rama are still very close, both structurally and semantically.

One additional argument may be proposed to support the analysis of a close link between subordinators and postpositions in this language. The argument rests on a variety of patterns encountered with the morpheme *BANG* in a subordination context, as shown in example 27 below:

(27) a. SUB *tiiskama ni-sung-bang taak-i*
baby 1-see-SUB go-TNS
'I am going to see/look at the baby.'

b. RP *ni-tanang ba-n-taak-i*
1-look RP-1-go-TNS
'I am going to look at it.'

c. SUB + RP *ni-ku-bang ba-n-taak-i*
1-grab-SUB RP-P-go-TNS
'I am going to go get him.'

As shown in (27a and b) the subordinator *BANG* of the embedded verb may alternate with a relational preverb *BA-* on the main verb which gives the same semantic reading of purpose subordination. Both subordinators — *BANG* and subordinating preverb *BA-* may in fact co-occur (27c). This double marking might be a type of speech "error" found in casual speech and may not qualify as an instance of "grammar" proper, but it is documented in recorded narratives. In any case, such examples may indicate the possibility of some relational preverbs coming indirectly from postpositions via shift to subordinator. The postpositions therefore are linked to both relational preverbs and subordinators in Rama, as indicated by the arrows below. The dotted line indicates the possibility of an additional indirect link between postposition and relational preverb mentioned above.

POSTPOSITIONS ------→ RELATIONAL PREVERBS
.
.
.
SUBORDINATORS ------→ RELATIONAL PREVERBS

5. ON THE VERBAL ORIGIN OF SOME TENSE/ASPECT/ MODALITY MARKERS

While the preceding sections dealt with instance of grammaticalization in the domain of argument marking, this section deals with parallel instances of grammaticalization which take place in the domain of aspect and modality marking. The parallelism consists of the fact that some of the morphemes involved in the inventory of aspect and modality marking are homophonous with some of the postpositions, relational preverbs and subordinators considered in the above sections. This is particularly the case with the morpheme *BANG* which is the focus of this paper.

Once again the main purpose of the discussion is to relate grammatical

morphemes to their sources, both immediate and original, in order to sketch out the grammaticalization chains along which these morphemes may be plotted. The discussion of the origin of most of the aspect and modality markers in Rama will lead to the original lexical source of the multiple chains of grammaticalization to which the morpheme *BANG* belongs.

And once again, other interesting issues such as the nature of the pathways or the relative timing of the various evolutions will be mentioned only in passing here.

5.1. The inventory of Rama aspect and mood markers

Rama has a set of tense markers which are independent of the aspect and modality markers considered here, and which appear in verb final position:

(28)

simple tense markers	Complex tense markers (with aspect or modality overtones)
-i 'present'	*-aing* 'now, for sure'
-u 'past'	*-uing* 'used to'
-ut 'future'	*-uting* 'will for sure/every time'

Aspect/mood markers are closer to the verb stem, in the typologically expected relative order [V-ASP-TNS]. The inventory of postverbal aspect and mood markers is given below:

(29) Aspect markers:

1. *-atkul-*	'complete'
2. *-atkar-/itkr-*	'repetitive'
3. *aakar*	'stative/resultative'
4. *-bang*	'1pl Imperative'
5. *-bang*	'prospective: to be going to'
6. *aakar/baakar*	'progressive'
7. *baakar*	'prospective: to be about to'
8. *-kama aakar*	'prospective: to be ready to'
9. *bating-*	'prospective: to be going to'
10. *bating baakar*	'prospective: to get ready to'

A word of caution is needed here — the above inventory may look more definitive and categorial than it should. It represents the list of postverbal forms identified so far as functioning as aspect or mood markers. It is a working tool for the purpose at hand, shaped by the awareness of the gradual nature of the process of grammaticalization itself and of the textual nature of the data base, where variation in the use of forms abounds. The inventory

is functionally oriented and regroups various types of constituents, such as suffixes (1,2,4,5,8) and free auxiliaries (3,6,7,8,9,10). This issue of morpheme boundaries bears on the later discussion of the process of grammaticalization. The main point to notice in this inventory is the recurrence of two morphemes, *BA(NG)* and *AAKAR* in both simple and polymorphemic markers. In particular note that *BA(NG)* appears as a simple morpheme marking imperative and prospective aspect markers, as well as in the complex forms *bating*, *baakar* and *bating baakar*.

5.2. Deverbal origin of aspect and modality markers

It is not difficult to argue in Rama that aspect and modality markers have lexical verbs as their sources. When compared to the inventory found in the literature on the basic concepts which are sources of grammaticalization (see in particular Heine, Claudi and Hünnemeyer, Volume I), the Rama inventory offers no surprises:

(30)	*AAKAR*		BE/LOC
	ATKUL		FINISH
	BANG	suppletive form of *TAAK*	GO
	TING	HAPPEN, in complex form *BATING*	GO-HAPPEN = WANT

All four verbs are well within the semantic range of verbs from which tense, aspect and modality markers are derived across languages.

The case of *ATKUL* 'finish' being the source of a completive aspectual marker constitutes a typical example of the phenomenon of grammaticalization of a lexical verb into an aspectual marker. It will be used to illustrate some of the correlates of the process of grammaticalization, such as the switch of morphological category (from main verb to suffix) which is accompanied by a shift of semantics, from the act of finishing set in time to the more abstract concept of completeness. The semantic shift is noticeable in Rama because the use of the suffix *-atkul* emphasizes the completeness of the impact on the object rather than a notion of end point in time:

(31) VERB

a. *atkut-king*
finish-make
'get it done!'

b. *tabulaak tkeeruk nsu-atkul-u*
evening grave 1PL-finish-TNS
'We finished (digging) the grave in the evening.'

(32) ASPECT

a. *paalpa an-sung-atkul-u*
manatee 3PL-see-ASP-TNS
'They saw the *whole* manatee.'

b. *dor y-aakang-atkul-u*
door 3-shut-ASP-TNS
'She shut the door *tight*.'

c. *kauling an-apulk-atkut-su ba-an-taak-u*
people 3PL-gather-ASP-SUB RP/for-3PL-go-TNS
'Upon having gathered *all* the people, they went (to look) for him.'

(33) VERB + ASPECT

y-ausa atkut-atkul-u y-aung-i
3-strength finish-ASP-TNS 3-say-TNS
'Her strength is totally gone, she says.'

The example of the lexical verb 'finish' giving raise to a completive aspect marker in Rama is one of the most common examples of grammaticalization in the domain of tense-aspect-modality.

The grammaticalization of the lexical verb 'go', which is also fairly common across languages, exhibits various language particular characteristics in Rama. One has to do with the actual nature of the morpheme *BANG* and its relation to the Rama verb *taak* 'go'. Another is the fact that the original morpheme *BANG* has given rise to a number of grammaticalization processes, some of which may have developed in parrallel, but independent fashion, yielding a number of aspect and modality markers.

It happens to be that, in the grammar of present-day Rama, the main lexical form of the verb GO is *taak*, which is phonologically unrelated to the morpheme *BANG*. The link between the two morphemes is that *BANG* is a suppletive form of *taak* in the second person imperative form. As a main verb, at least, *BANG* has a free variant *mang*:

(34) a. *i-taak-u, tawan ki yu-i-taak-u*
3-go-TNS town to with-3-go-TNS
'He went, he took it to town.'

b. *mang, tawan ki yu-mang*!
go town to with-go
'go, take it to town!'

It is therefore what surfaces today as the suppletive imperative form of the verb GO which is the source of the grammatical morphemes functioning as aspect markers.[13] As a grammatical morpheme of the domain of tense-aspect-modality, the simple suffixal form *BANG* carries two irrealis functions. One is that of prospective aspect marker 'to be going to':

(35) a. *bal-n-aating-bang*
PREF-1-talk-ASP
'I am going to talk.'

b. *tiiskama ni-tanang-bang*
baby 1-look at-ASP
'I am going to look at the baby.'

That the function of the suffix *BANG* is aspectual is clearly shown by examples such as the one below, in which the lexical verb stem is a verb of movement and the reading of *BANG* is not that of a verb of movement anymore:

(36) *i-traat-bang*
3-walk-ASP
'he is going to walk.'

The other aspectual function of the suffix *BANG* is that of first person plural imperative, which takes a first person plural subject marker.[14]

(37) a. *mwaing yairi s-tuk-bang*
1PL soup 1PL-drink-IMP
'Let's drink our soup.'

b. *m-kuiik si-suk-bang*
PL-hand 1PL-wash-IMP
'Let's wash our hands.'

The modality function of this instance of *BANG* is also made clear with an example where the lexical verb is itself the verb *taak* 'go':

(38) *ka-s-taak-bang*
RP/from-1PL-go-IMP
'Let's go away from it.'

BANG is also found as a component of two bimorphemic aspectual forms, *BAAKAR* and *BATING*, formed with *BA-* and either the stative auxiliary *AAKAR* 'be' or the helping verb *TING* 'happen'.

There appear to be two *BAAKAR* forms which are distinguished by their semantics and the optionality of the combination. *BA-* may optionally combine with *AAKAR* in its function of progressive marker:

(39) a. *nah paun baakir-i*
1 cry AUX-TNS
'I am crying.'

b. *tiiskiba-lut kaun aakur-u*
child-PL holler AUX-TNS
'The children were hollering.'

BA- is not optional in the other *BAAKAR* form, a prospective marker signalling the imminence of an action 'to be about to' :

(40) *yaadar yu-n-siik baakur-u, barka n-suleer-uung-u*
smthg RPL-1-bring AUX-TNS but 1-forget-do-TNS
'I was about to bring in something, but I forgot.'

The combination of *BA-* with the helping verb *TING* 'happen' which yields the form *BATING* actually exists as a main verb meaning 'want'. It is therefore already a derived lexical verb which is the source of the other prospective aspect marker 'to be going to':

(41) a. *sumuu i-bating-i*
banana 3-want-TNS
'He wants a banana.'

b. *i-traat bating-i*
3-walk ASP-TNS
'He is going to walk.'

c. *sumuu y-angtik bating-i*
banana 3-cut ASP-TNS
'He is going to cut bananas.'

d. *nah yuwa-ting bating-i*
1 old-get ASP-TNS
'I am going to get old.'

The meaning of *BATING* as an aspectual marker is very close to the simple prospective marker *BANG*.

This is not the end of the aspectual forms involving the morpheme *BANG*. The two complex auxiliary forms just presented may also combine, yielding another prospective aspect, this one with two instances of *BA-*, as in *bating baakar* 'to be getting ready to':

(42) *na ba i-tkit bating baakur-u*
1 PSP/on 3-jump ASP ASP-TNS
'He got ready to jump on me.'

(43) *n-alngu bating baakir-i*
1-drink ASP ASP-TN
'I get ready to drink.'

(44) *abingka tabi bating baakir-i*
sideways come out ASP ASP-TNS
'I am getting ready to come out on the side now.'

The last aspectual forms just described, *BAAKAR* and *BATING* are the result of rather complex stepwise processes of grammaticalization. Although the morpheme *BA-* found in those bimorphemic forms derives ultimately from the lexical verb *BANG* 'go', it is likely that its most immediate source is the relational preverb *BA-*, itself derived from a goal postposition, which is the one which is originally derived from the verb 'go'. If this analysis is correct, this particular chain of grammaticalization would involve the steps sketched below:

VERB ---→ postposition ---→ relational preverb ---→ aspect

Several arguments can be advanced to support the analysis of the *BA-* of *BAAKAR* and *BATING* having as its source an already grammaticalized relational preverb rather than the lexical verb 'go'. One is the position of the affix in the word, in a preverbal position characteristic of relational preverbs. Another is a parallelism between the prospective *BAAKAR* aspect 'to be about to' and the prospective *KAMA AAKAR* aspect 'to be ready to', as in:

(45) ASP *n-taak-kama aakar-i*
1-go-SUB ASP-TNS
'I am ready to go.'

Both prospective aspects have the same components: a goal/purpose morpheme combining with the auxiliary *AAKAR* 'be'. The difference in the morphological category of the purpose morpheme, a suffixed subordination marker in the case of *kama* and a prefixed relational preverb in the case of *BANG* is related ultimately to different behavioral properties of the two purpose morphemes. There happens to be no relational preverb equivalent for the postposition-subordinator *KAMA*, as established in the early section on relational preverbs.

(46) *VERB BA-AAKAR* vs *VERB-KAMA AAKAR*
VERB RP-AUX VERB SUB AUX
'to be about to V' 'to be ready to V'

This section has introduced several more instances of the morpheme *BANG* as it pertains to the verbal morphology of the language from the domain of tense-aspect-modality marking. It established that the lexical origin of *BANG* is in the paradigm of the verb 'go' surfacing today as the suppletive imperative form. It also outlined various evolutions of *BANG* into grammatical markers of aspect and modality.

6. POLYGRAMMATICALIZATION: OVERVIEW OF INTERCONNECTED RAMA GRAMMATICALIZATION CHAINS

The previous sections have described a number of grammaticalization chains which all involved verbal morphology; in some instances preverbal morphology, as with relational preverbs; and in others postverbal morphology, as with subordinators and aspect/modality markers. The interest of the Rama data lies in the fact that related morphemes are found in both the pre- and the post-verbal morphology; as described in the various sections, some of the same morphemes appear to fulfill multiple functions such as relational preverbs, subordinators, and aspect/modality markers.

This final section will discuss the possible scenarios by which the various grammaticalization chains considered in this paper may have evolved. The earlier emphasis on a consideration of the relative notion of source in a chain will then be complemented by an attempt at identifying some of the pathways linking sources to target verbal affixes. It will be assumed, as stated at the beginning of this paper, that the overall direction of a grammaticalization

chain is of course from lexical to grammatical, and from free to bound morphology. However, one of the notions that emerges from the exercise of sketching out full grammaticalization chains is the fact that some of them reveal a somewhat heterogenous nature, to the extent that not all their hypothesized component links will fall within the range of prototypical instances of grammaticalization.

The types of verbal affixation discussed in this paper belong to two grammatical domains: the domain of argument marking and the domain of tense/aspect/modality marking, which, conforming to typological tendencies of verb final languages, align themselves in the following fashion for main verbs: argument marking in preverbal position and tense/aspect/modality marking in postverbal position. Each domain will be considered in turn, before attempting a final sketch of the multiple interconnections the morpheme *BANG* creates in the grammar of the language.

6.1. Argument marking domain

In the argument marking domain, the two grammaticalization chains are a) the one accounting for the evolution of relational preverbs, and b) the one accounting for subordinators. These two grammaticalization processes have a common local source — the set of postpositions — and follow two different paths of evolution — one to preverbal and the other to postverbal position.

The path of evolution from postposition to relational preverb is well-documented in Rama. As argued in Craig and Hale (1988), the different types of relational preverbs found in Rama, together with the types of relational preverbs of other languages of the Americas, can be placed on a continuum that provides a sketch of the possible evolutionary scenario of such verbal morphology.

The contribution of the Rama data to the study of the evolution of relational preverbs is in the documentation of an initial stage of simple cliticization. At this early stage, all that seems to be involved is a surface reanalysis which does not affect semantics and is discourse dependent, in the sense that zero anaphora is discourse dependent. Whether the cliticization is more than a surface phenomenon, that is to say whether it affects grammatical relations in such a way that the oblique NP could be said to have been promoted to the status of direct object, is not easily settled for the Rama data. In other languages of the Americas, the presence of relational preverbs can be positively related to changes in grammatical relations, through overt morphology characteristic of certain grammatical relations, but the lack of

relevant morphology in Rama does not allow for language internal argumentation and the analysis must remain theory-dependent for that language. The points to be made here are that Rama has a much less syntactically involved type of cliticization of relational preverb than is commonly found in languages with such verbal prefixes, and that this cliticized type of relational preverb represents a possible first step in the path of evolution of relational preverbs.

Within a typological/functional framework the difference between discourse cliticized relational preverbs and syntactically incorporated (grammatical relation changing) relational preverbs need not be an absolute categorial one. Such a framework allows for the gradual evolution of one type into the other. Within such a framework, the type of cliticized relational preverb found in Rama could be considered as a relational preverb which has advanced on the chain of grammaticalization less than its syntacticized counterpart in other languages.

Further down the chain of grammaticalization are other types of incorporated relational preverbs, the ones which in Rama allow for a full, overt NP to be construed with them. As described earlier, these incorporated relational preverbs are characterized by the specific pairing of certain relational preverbs with certain verbs, and a change in semantics accompanying the change from free postposition to bound relational preverb. These Rama incorporated relational preverbs align themselves on a continuum from incipient to advanced lexicalization. The selection of relational preverbs and verbs that are lexicalized together is assumed to be a reflection of the cognitive and cultural salience of the combination which translates into high frequency of use in natural texts. Of interest here is the decomposition of the link from postposition to relational preverb into several steps, two of which are attested in Rama: cliticization and lexicalization. The different manifestations of the morpheme *BANG* are the best examples of lexicalization in Rama.

While the above remarks are a consideration of how postpositions are the source of grammaticalization chains, they do not address the issue of where the postpositions themselves ultimately came from in Rama. Evidence for postulating a verbal lexical source rather than another source for Rama postpositions can only be pointed out in the synchronic grammar of Rama for the case of the morpheme *BANG*. As discusssed earlier, there exists a verbal morpheme *BANG* which is a suppletive form of the verb 'to go'.[15]

When all these aspects of the morpheme *BANG* are assembled, a rather complete grammaticalization chain from free lexical to bound grammatical morphemes can be constituted, as follows:

(47)

V---→ PSP---→ CLITICIZED RP-(SYNTACTICIZED RP?)---→ LEXICALIZED RP

GO GOAL GOAL Direct Object

It may be added that the initial step of the grammaticalization of the postposition *BANG* into a bound relational morpheme, through cliticization, may not really constitute a bonafide instance of grammaticalization. It may fall rather within the scope of reanalysis. However, this link appears to belong to the chain by virtue of being the most likely source for the lexicalized relational prefixes.

The chain of grammaticalization which links postpositions to subordinators constitutes another path of evolution taken by the postpositions in the language. By structural analogy between the two types of oblique arguments of a verb, the nominal and the clausal type, the concrete markers of oblique NP are extended to become the markers of the more abstract set of relations that hold between main verbs and their adverbial clausal complements. The extension of postpositions to subordinators does not entail much change. The embedded verb of the adverbial complement clause is already a nominalized construction in that it is stripped of its tense marker, and both nominal and clausal markers follow the element with which they are construed: NP for the postposition and final verb of the embedded clause for the subordinator.

The parallel between the markings of these two oblique arguments — nominal and clausal — extends to the fact that subordinate clauses may also, under certain circumstances, be marked by relational preverbs on the main verbs, instead of by a subordinator on the embedded verb, the same way oblique NPs may be marked either by postpositions or relational preverbs. Such a parallel exists only with the morpheme *BANG* of goal and purpose, which is found in pairs of patterns outlined below, with the same meaning:

(48) embedded verb-*bang*/SUB MAIN VERB
embedded verb *ba*/RP-MAIN VERB

There is little data available on the use of the relational preverb *BA-* as a subordinating morpheme. An example of it is given again below (cf. 27b):

(49) *ni-tanang ba-n-taak-i*
1-look RP/for-1-go-TNS
'I am going to look at it.'

Pairs of the type shown in (48) would argue therefore for two different paths of evolution of relational preverbs, directly from postpositions and indirectly from postpositions through subordinators:

(50) postpositions ---→ relational preverbs

:

:

:

subordinators (---→ relational preverb)

The evolution of subordinating relational preverbs is presented in parenthesis to emphasize its ephemeral nature in the language. However, if its existence is acknowledged, it constitutes a challenge to the established view in studies of grammaticalization that grammaticalization involving the emergence of subordinating morphology is unidirectional: From main clause to subordinated clause. In the case of the *BA-* subordinating relational preverb of Rama, the direction would appear to be from subordinate morphology to main clause morphology.

6.2. Aspect/modality domain

The grammaticalization chains in the domain of aspect/modality marking exhibit many more cases of the morpheme *BANG*. Within the aspect/modality marking system, *BANG* functions either as a verbal suffix (prospective aspect or imperative mood: V-*bang*), or as a relational preverb which is part of auxiliary type markers (progressive and prospective aspect auxiliaries: V *baakar*. It is tempting to consider the path of evolution from the verb 'go' to the prospective aspect suffix *-BANG* as passing through a stage of verb serialization. One could probably construct a case for verb serialization in Rama on the basis of several arguments. For one, such constructions do exist in the language; they are limited to combinations of action verb and a few final verbs, most of them motion verbs:

(51) a. *naas ngulkang alais-traal-i*
I wild pig hunt-walk-TNS
'I go hunt the wild pig.'

b. *an-kuu-alngul-u*
3PL-put-reach-TNS
'They managed to put him (in the hospital).'

c. *nsul u y-alpaak-alkung-u*
we PSP/with 3-meet-go-TNS
'They went to meet together with us.'

While more work needs to be done on these constructions, one remark can be made now: these three action verbs themselves do not have exactly the same status. *Traal* is the freest of the three, being used either in serial verb constructions or as a free lexical verb, while the other two, *alngul* and *alkung*, are already constrained in their use to this type of construction and do not seem to have the status of free lexical morphemes. Their behavior is reminiscent of directional morphemes. What the three have in common with *BANG* is that they are movement verbs appended to lexical verbs, with the additional characteristic that a certain process of grammaticalization may have already been set in motion for at least two of them.

Another argument for taking a serial verb construction as the conduit for the grammaticalization of *BANG* into an aspect and modality marker would be its particular use in the first person plural imperative, where the meaning of motion is still retrievable:

(52) *nsu-kami-bang*
1PL-sleep-IMP/GO
'Let's (go) sleep.'

The use of *BANG* in first person imperative, if it is to be analyzed as a grammaticalized imperative marker, is the closest link between the free lexical motion verb *BANG* and a bound aspect/mood marker *-BANG*.[16]

It is assumed that verb serialization is the most likely structural source of the first person plural imperative suffix *BANG* and the prospective *BANG*. However, it is not the only possible path of evolution for the instances of the morpheme *BANG* in the bimorphemic prospective and progressive aspects. As pointed at by the form of the complex auxiliary forms *kama aakar / baakar*, the source of these prospective aspect markers may be a complex construction with the copula *aakar* as the higher verb and *kama/BA-* as markers of embedding.

6.3. Polygrammaticalization

This paper has presented data from a language in which various chains of grammaticalization interlocked through a common set of morphemes. All the chains presented have already been identified in other languages. Verbal prefixes derived from postpositions, subordinators derived from postpositions, aspect markers derived from verbs, auxiliaries derived from subordinators and relational preverbs. The density of the interconnections between the chains is what makes the Rama data interesting, as well as some aspects of the internal make-up of the chains.

This case study ends up raising more questions than it answers about the nature of grammaticalization chains and how to conceive of them in the detail, including the issue of the relation of the process of grammaticalization to that of reanalysis. It has hopefully provided a good example of "polygrammaticalization", defined as a multiplicity of grammaticalization chains that may originate in one particular lexical morpheme. I will leave questions of the relative timing of the various evolutions and of the nature of the pathways followed for further research.

The following table illustrates clearly, for the morpheme *BANG*, the notion of 'polygrammaticalization', which I have argued for in this paper.[17]

*1 The lexical item exists only as an imperative suppletive form in today's language.

*2 The argument for claiming that the source of the PSP is the verb GO is primarily typological. The long form of the postposition is for left or right dislocated PP.

*3 The cliticized RP associates with any verb and maintains its semantics of GOAL.

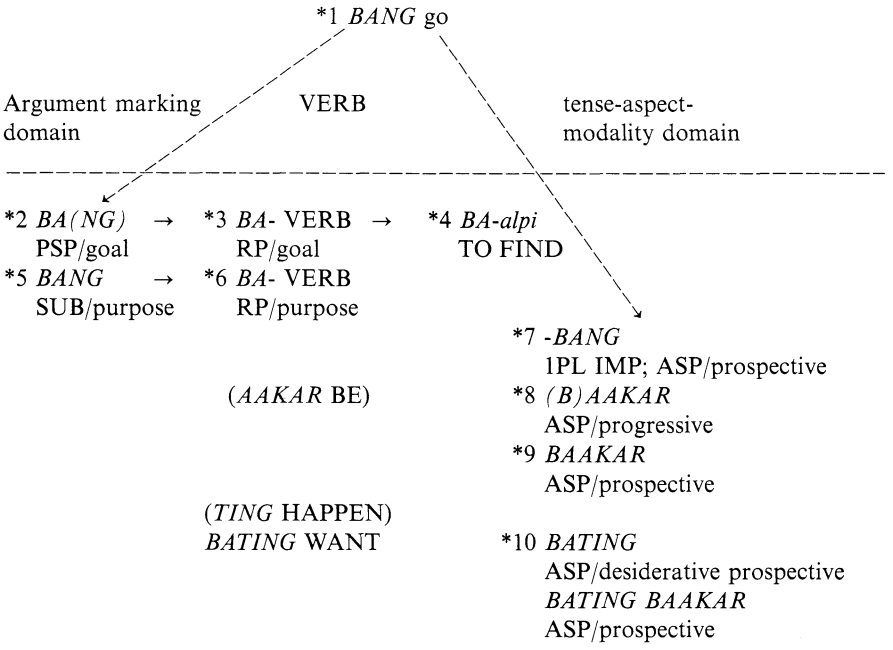
*4 There are few examples of lexical verbs formed with a lexicalized RP *BA-*. One of them is (NP) *BA-alpi* 'to look for (NP)'. The verb stem **alpi* does not exist by itself, nor can it be used with a postpositional phrase with *BA*: *(NP *BA*) **alpi*. The other is *BA-TING* 'to want', see *10. For other instances of lexicalization of the RP *BA-* see *8 and *9.

*5 The subordinating *BANG* is itself derived from the PSP. It is suffixed to the bare stem of an embedded verb.

*6 The subordinating RP *BA-* clitized on the main verb alternates with the subordinating suffix of the embedded verb, with equivalent semantics. The subordinating RP is much less common than the subordinating suffix of *5.

*7 *-BANG* combines with first person plural subject marker to form the

Table 2. Polygrammaticalization of Rama *bang* go



1PL Imperative form. In non-imperative context, *-BANG* corresponds to a future tense or prospective aspect.

*8 *(B)AAKAR* stands for the optional presence of the RP *BA-* on the verb *aakar* 'to be (at)'. As auxiliary verbs both *aakar* and *BAAKAR* carry a meaning of progressive aspect.

*9 *BAAKAR* is another auxiliary verb derived from the lexicalization of a RP *BA-* with the verb *aakar*, to give a prospective aspect 'to be about to'.

*10 Auxiliary verb *BATING*, from the lexical verb *BATING* 'to want', itself derived from the lexicalization of a RP *BA-*.

One of the goals of this study was to make the case for a functional typological approach to grammatical description, an approach which incorporates the phenomenon of grammaticalization as an integral part of the analysis. In the case of Rama, it offered the means to handle an essential component of the grammar of the language — its extensive

morphological interconnection through the morpheme *BA(NG)* across grammatical structures.

NOTES

1. The term was suggested to me by Alan Timberlake. I am grateful for the careful reviews of the draft of this manuscript by the co-editors and Gillian Ramchand.

2. Rama is presently the object of a language revitalization project carried out within the context of the new Autonomous Statute of the Atlantic Coast of Nicaragua. Research on the Rama language has been supported by grants from the National Science Foundation (BNS 851156), the Wenner Gren Foundation for Anthropological Research #4906, the Failing Fund and Research funds from the Graduate School of the University of Oregon. The data were collected in Bluefields, Nicaragua, in the course of various field trips in 1985, 86, 87, 88. I would like to thank here the various people who have contributed to the development of the Rama Language Project: the Rama speakers, Nora Rigby and Cristina Benjamins, the research assistants Bonny Tibbitts and Barbara Assadi, and my colleagues from Linguists for Nicaragua, in particular Ken Hale, Danilo Salamanca, Susan Norwood and Wayne O'Neil. The research in Nicaragua is sponsored by the Center of Documentation and Research on the Atlantic Coast (CIDCA, Galio Gurdian, director).

3. The present paper will only deal with a particular set of interconnecting grammaticalization chains, although more examples of grammaticalization chains could be studied. Another case of grammaticalization chain, for instance, is that which links independent subject personal pronouns and preverbal subject clitics, also a common one in the languages of the world.

   All dependent subject (and object) pronominal forms are morphologically related to the independent personal pronouns, as shown here with an example of the third person plural:

   a. Full NP: *kiikna-dut apaik-i* 'the men paddle'
      man- pl paddle-tense
   b. Independent PRO: *anulut apaik-i* 'THEY paddle'
      they paddle-tense
   c. Pro-clitic: *an-apaik-i* 'they paddle'
      they-paddle-tense

4. Rama has an additional type of postverbal marker, *kaing*, which is found in relative clauses:

   suulikas ni-paaya-u kaing neli aa ni-tang-u
   meat 1-buy-tns REL Nelly obj 1-give-TNS
   'The meat I bought, I gave it to Nelly.'

   However, this marker is not strictly a 'relativizer' (REL) as glossed in the example. *kaing*, which is optional in relative clauses, is a discourse marker of presupposed information and may in fact appear on the end of main clauses too. This morpheme is not a verbal suffix as much as a clause final particle, which is evidenced by the presence of a tense

marker on the verb and the possiblity of elements intervening between it and the verb. The reason it is brought up here is simply to emphasize the general post verbal placement of all clause embedding marking.

5. Direct elicitation also showed that the ten occurrences of the postposition *BA(NG)* of column a. is an artifact of the texts that have been collected — the postposition *BA(NG)* was found in direct elicitation, see example 1a.

6. Going back to column c. now, note that not all instances of relational preverbs in that column are of the cliticized type. That column contains also the instances of lexicalized relational preverbs whose objects happen to be zero anaphora too.

7. The supporting evidence for the claim can be found in examples such as:

   a. *ma kang n-altawa-i*
   you PSP/from 1-afraid-TNS
   'I am afraid of you.'

   b. **ma ka-n-altawa-i*
   you RP/from-1-afraid-TNS
   (I am afraid of you.)

8. The other two cases of overt NP construed with the preverb *yu-* which do not involve either of these two verbs turn out to be, upon closer look, cases of left dislocation of the NP; accordingly, they are to be analyzed as instances of procliticization triggered by the stranding effect of the dislocation.

9. One could argue that the cases of overt NPs construed with the relational preverb *yu-* are instances of syntactic incorporation which results in the NP being the object of the verb. The absence of extensive morphosyntax in Rama — the language has no object or number agreement, no object case marking, no voice system — deprives us of very convincing arguments for such an analysis, although a few standard ones may be advanced. One is the structural argument of word order. Rama has a strict SOV word order and both pre- or post-verbal postpositional phrases. In the constructions involving a relational preverb and an overt NP to be construed with it, that NP is always in the preverbal object position. One could also advance a semantic argument to the effect that the patient semantic relation of the overt NP to the verb is more typical of object NPs. Text counts of topic chains, as the ones done in Tibbitts (1988) would also provide some argument in support of a promotion analysis, although this whole issue deserves a paper of its own and will not be pursued any further here.

10. There is one additional conjunction of subordination, *kaing* 'because'. This morpheme was already introduced as a marker of relative clauses, as well as a discourse marker of independent presupposed clauses (see fn. 4 above). Unlike the other subordinators, it follows a fully inflected verb.

11. The postposition is in the long forms *bang* when part of a dislocated postpositional phrase, either post- or pre-verbally. If anything needs explaining in the forms of the subordinatory it is the presence of the short form *-ka* of the time subordinator, not the morphologically expected post verbal long form of *-bang*.

12. Additional evidence of the closeness of postpositions and subordinators is the existence of complex subordinating morphemes of the type [borrowed N + PSP], as in *taim ki* 'at the time, when' or [borrowed N + SUB], as in *taim ka* 'time when, when'

13. The argument for ascribing to *BANG* a lexical verb as its source is primarily a typological one. Note how it behaves like a verb in taking a relational preverb.

14. Lehmann and Conzemius had noted the imperative first person plural marker *bang*. The apparent lack of first person plural marker in Conzemius may be due to defective phonological transcription since the only example given is of a verb with an initial *s-* which would make hearing the first person plural *s(u)* difficult:

    a. *siik-bang* (Conzemius, 1927:336)
       come-IMP
       'Let us come.'

    b. *sa-siik-bang* (Lehmann, 1914:15)
       1PL-come-IMP
       'Let us come.'

15. There is actually a verb *kaa* meaning 'to put', whether it can in any way be shown to be related to the postpositions *ka(ng)* 'from' and *kama* 'for' remains to be seen. The fact that *kama* does not behave like the other goal postposition *bang* with respect to the possiblity of becoming a relational preverb may argue for a different origin for the two goal postpositions.

16. The serial verb analysis of the imperative first person plural corresponds to the encoding of a particular component of a prototypical imperative situation: the fact that imperatives are requests to move or change from one state to another. Rama is an interesting language in that its grammar encodes the fact that the intended effect of the imperative speech act is the creation of a sequence of two actions or states: a first state or action to be abandoned and moved away from and a second state or action to be embraced and moved toward. While the serialized motion verb *bang* marks the motion toward the desired new action or state, there is, in Rama, the possibility of encoding in an imperative form the move away from the first action or state, through the use of the relational preverb *kaa-*, from the ablative *kang* 'from', as is illustrated below:

    a. *kaa-yuk-ting*
       RP/from-sit-happen
       'sit down'

    b. *kaa-s-alamki*
       RP/from-1PL-turn back
       'let's turn back'

17. Examples to illustrate the table are re-assembled below:

                                  BANG/BA
    PSP a) *naas sii ba aa taak-iikar*
           I water PSP NEG go-want
           'I don't want to go get/for water.'

    RP b) *yaing aa ba-y-alamk-i; aa y-alamk-i*
          he NEG RP-3-turn back-TNS NEG 3-turn back-TNS
          'He never turned back for (it/him); he never turned back.'

    RP c) *ni-tanang ba-n-taak-i*
          1-look RP-1-go-TNS
          'I am going to look at it.'

RP d) *tamaik siita sut-aapluk ba-taak-i ma u*
tomorrow oysters 1-pick RL-go-TNS you PSP/with
'Tomorrow we'll go pick oysters with you.'

SUB e) *tiiskama ni-sung-bang taak-i*
baby 1-see-SUB go-TNS
'I am going to look at the baby.'

ASP f) *tiiskama ni-tanang-bang*
baby 1-look at-ASP
'I'm going to look at the baby.'

VB g) *i-taak-u; tawan ki yu-i-taak*
3-go-TNS town to with-3-go-TNS
'He went' 'he took it to town.'

mang!; tawan ki yu-mang
go! town to with-go
'Go!' 'take it to town.'

h) *ni-ku-bang ba-n-taak-i*
I-grab-ASP RP-1-go-TNS
'I'm going to go get him.'

KAMA

PSP a) *tamaik suuli-kaas ni-auk-ut maing kama*
tomorrow animal-meat 1-cook-IRR you PSP/for
'Tomorrow I will cook some meat for you.'

SUB b) *nah suuli-kaas ba-alp-i ni-paya-kama*
I animal-meat RP-look-TNS 1-buy-SUB
'I am looking for meat to buy.'

ASP/MOD c) see (20a,b) *kama aakar* 'be supposed to'

MOD d) *tiiskiba umling taak-kama skuul ki*
child all go-MOD school to
'All the children must go to school.'

SU

PSP a) *ipang su an-siik-u*
island PSP/on 2PL-come-TNS
'they came to the island.'

b) *walsa anut su tabi-u*
tiger them PSP/on come out-TNS
'The tiger came out at them.'

RP c) *paalpa suuli taara nainguku mliika su-an-aakar*
manatee animal big so happy RP-3pl-be
'the manatee is a big animal, and so they are happy about that.'

SUB d) *yaing ipang aap-su y-aakur-u nuunik puksak*
she island find-SUB 3p-be-tns say two
'upon finding the island, she stayed two days.'

e) *tawan ki y-aakit y-aakit-su y-almaling-u*
town in 3p-be 3p-be-SUB 3p-die-tns
'she stayed and stayed in town until she died.'
(upon staying and staying in town, she died.)

U

a) *namaa y-aakar ngurii ki yaing tiiskama u*
still 3p-be hole in her child PSP/with
'she stays still in the hole with the child.'

b) *ngurii ki yu-i-taak-u*
hole in RP/with-3p-go-tns
'she takes (him) in the hole.' (lit: goes with him)

REFERENCES

Breugelman, R. 1987. "Tense markers in Rama narratives." Ms. University of Oregon.
Conzemius, Edouard. 1927. "Die Rama Indianer von Nicaragua." *Zeitschrift für Ethnologie* 59:291–362.
Craig, Colette G. 1987. "Rama relational prefixes: From postpositions to verbal prefixes." In *Proceedings of the Second Annual Meeting of the Pacific Linguistics Conference*, S. DeLancey and R. Tomlin, (eds), 51–71. Eugene, University of Oregon.
Craig, Colette G. and Ken Hale. 1988. "Relational preverbs in some languages of the Americas: Typological and historical perspectives." *Language* 64(2):57–78.
Genetti, Carol. 1986. "The development of subordinators from postpositions in Bodic languages." In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society* 12:387–400.
Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. Volume I. "From cognition to grammar — evidence from African languages."
Lehmann, Walter. 1914. "Vokabular der Rama-sprache." In *Abhandlungen der königlich Bayerischen Akademie der Wissenschaften*. Munich.
Lehmann, Walter. 1920. *Die Sprachen Zentral-Amerikas in ihren Beziehungen zueinander sowie zu Sud-Amerika und Mexico*. 2 vols. Berlin.
Tibbitts, Bonny. 1988. "Cliticization of adpositions in Rama." Ms. University of Oregon.

Grammaticalization and Related Changes in Contemporary German

Christian Lehmann
University of Bielefeld

1. INTRODUCTION[1]

1.1. In a sense, every grammatical change leads to a new grammar. The term *grammaticalization* could be interpreted as 'creation of new grammar' (where *grammar* is taken as a mass noun). Then grammatical change would equal grammaticalization. Extension is a fate that traditional terms commonly undergo when they are taken up in modern linguistics. Here, 'grammaticalization' will be used in a narrower sense. A variation affecting a linguistic unit will be considered a case of grammaticalization if it is directed towards subjecting that unit more to rules of grammar. This proposition is not meant to carry any implications as to whether rules pre-exist units or vice versa. Among the implications that it does carry are the following:[2]

— The element in question turns from a less grammatical (initially, lexical) element into a more grammatical one.
— The element loses substance both on the phonological and the semantic sides; its selection restrictions are loosened.
— The freedom to manipulate the element decreases. It is integrated into a paradigm, it becomes increasingly obligatory in certain constructions, and it occupies a fixed (ultimately, morphological) position.

A paradigm case which illustrates most of these features is the grammaticalization of the Latin preposition *ad* 'to' to the Spanish case marker *a* 'accusative'. Phonological attrition goes hand in hand with the loss of concrete semantic features. The original sense of direction towards an entity or a place is extended to mark a being that is reached by the verbal action. This yields an indirect object marker. As such, its locus is with human complements. Then

it is extended towards a direct object marker, first only of human, then increasingly of other animate objects. At the same time, it regularly fuses with the definite article *el*; and it enters a minimal paradigm with *de* (which itself is grammaticalized from a Latin local preposition).

Analogical change is the other principal kind of grammatical change. The weak conjugation of erstwhile strong verbs in several Germanic languages is a well-known example. Grammaticalization and analogy have some points in common, among them the expansion of the distribution of an element and the introduction of a subregularity into the grammar. Also, grammaticalization may involve reanalysis of an underlying construction. Reanalysis, in turn, presupposes a pattern which can serve as its model. For instance, the grammaticalization of German *haben* 'have' to a perfect auxiliary involves, as an early stage, the construction [*hat* NP_{acc} $V_{part.pass.}$], where the NP_{acc} is the direct object of *hat* and the $V_{part.pass.}$ is a predicative complement of the direct object. This is reanalyzed as a periphrastic form of V, which takes NP_{acc} as its direct object. This reanalysis, however, is rendered possible only because German already possessed the periphrastic verb form [*ist* $V_{part.pass.}$], which thus could serve as an analogical model. For such reasons, it is often difficult to distinguish grammaticalization from analogical change.

Finally, grammatical change is to be distinguished from lexical change. The latter often also involves reduction and coalescence, as when Anglo-Saxon *hlāf-weard* 'bread-warden' becomes Modern English *lord*. Changes like *terrific* 'horrifying' > 'remarkable' have both the generalization of meaning and use and the codification in certain grammatical positions (here, the position of an evaluator/intensifier) in common with grammaticalization. Therefore, there will also be borderline cases between grammatical and lexical change.

1.2. The purpose of this contribution is to review a variety of grammatical constructions in contemporary non-literary German which are currently in vogue but are either not considered or at least not (yet) accepted by standard grammars. All of them share at least some of the features that are characteristic of grammaticalization. Some of them can be classed as cases of grammaticalization beyond doubt. Others are not so clear and may show more aspects of analogical grammatical change or of lexical change. They will be presented nonetheless, because they occasion some discussion on the limits of grammaticalization. There will also be the question as to whether the phenomena in their totality indicate a drift in the contemporary language.

Not being a Germanist, I am not in the best position to write on

grammaticalization in German. The following survey is not intended to be exhaustive or balanced. Some changes will be highlighted, others mentioned only in passing. For the past three years or so, I have been gathering data on the present topic. They are instances of ongoing changes or current fashions, occasionally encountered and unsystematically noted down. Some of them are tied up with my own dialect environment, which is Northwest Germany. For several of the phenomena to be mentioned below, I have no spontaneous data at all. Thus, my data cannot claim to be representative or to show the extent to which a phenomenon is or has been gaining ground. Furthermore, no arguments *ex silencio* are allowed: It is quite possible that the most interesting ongoing changes have escaped my notice. To the extent, however, that there are data, they prove that the phenomenon exists at a supra-idiolectal level.

There is also the methodological difficulty of chronology. Historical linguists and philologists are used to working through the relevant corpus in order to ascertain the earliest occurrence of a given phenomenon and to trace its tradition. I have not been able to do this. All I can say is that, except where stated otherwise, the phenomena treated have aroused my attention during the past ten years or so as things that were not formerly in the language. Many of them are not in the grammars. Nevertheless, should anyone be inclined to conclude that just because something is currently in the colloquial language but condemned by normative grammar, it must be a recent development or indicative of an ongoing change, it is beneficial to read Sandig, 1973. There it is shown that some of the nonstandard features of contemporary German have been in the language since the Old High German stage. Even a phenomenon that has been spreading only since very recently does not need to have been invented very recently. It is quite possible that it can be found in earlier literature, and such earlier occurrences may or may not be the source of the recent development.

2. NOMINAL CATEGORIES

2.1. Gender and number

2.1.1. There are currently changes in all of the traditional nominal categories, i.e. gender, number and case. The change in gender concerns the nature of the opposition between the masculine and feminine genders. As in many other languages, it used to be a privative opposition, with masculine

as the unmarked member. For about fifteen years, this relationship has been gradually shifted in favor of an equipollent opposition between the two. This manifests itself clearly in the use of the derivational suffix *-in* '-ess'. The most obvious examples occur in job announcements. For a position admitting both women and men, formerly a *Lehrer* 'teacher' would have been sought; now invariably a post for a *Lehrer/in* (i.e. *Lehrer/Lehrerin*) is announced. This is a clear case of a language change resulting from a change in society.

The phenomenon has some of the characteristics of grammaticalization, above all the quantitative expansion of the feminine suffix to the detriment of the zero form.[3] So far, however, the suffix remains a derivational one; it has not become a gender marker. Hence, we should probably not class this change as a case of grammaticalization, but instead learn from it that expansion should not be used too naively as a criterion of grammaticalization.

2.1.2. One of the German plural allomorphs is *-s*. While this has a stronger basis in the Low German dialect, it used to be one of the rarest in standard German. It has occurred for a long time in certain forms ending in posttonic *-el/n* (*Jungens* 'boys', *Mädels* 'girls', *Kumpels* 'miners, buddies'), which may be said to belong to the colloquial register of standard German. Apart from such marginal exceptions, the suffix originated in loans with one of two non-German phonological traits: Nouns ending in a vowel other than *ə* (*Mamas*, *Tees* 'teas', *Nazis*, *Autos*, *Zulus*), and certain non-feminine nouns stressed on the last syllable (*Kartons*, *Parks*).

It is apparently during the course of this century that *-s* has been continually gaining ground (see Wurzel, 1984, ch. 4 for details and explanations). It is penetrating into nominal morphology by a relaxation of the above compound condition: It now suffices that the noun in question be either a loan or have a stress or final full vowel in the last syllable. First, while most loan nouns still form their plural in one of the genuinely German allomorphs (*Visionen*, *Passagiere*, *Computer*), *-s* is increasingly used on non-Germanized loans with various phonological structures (*Sputniks*, nonstandard *Lexikons*). Second, it is gradually becoming the usual plural allmorph for words ending in a long vowel or in a stressed syllable. Native German nouns with such a phonological structure are rare. *Uhus* 'eagle-owls' is the only plural form for this noun. If *wo* 'where' were substantivized, it would form an *-s*-plural: *die wo's*. However, the relevant phonological classes are fed by abbreviations and acronyms (*PCs*, *Profs*), which are very frequent and where *-s* is currently most productive both in spoken and in written language. Other words are changing their plural marking (*Tablette* > *Tabletts* 'trays', *Konten* > *Kontos*

'accounts'). Yet others take the *-s* on top of an irregular plural which is probably not felt as a morphological plural: nonstandard *Spaghettis, Brokkolis.*

Paradigm instances of the development of plural morphemes by grammaticalization include Mandarin *men* 'class' and Bengali *jana* 'host', both of which have become plural suffixes. The case of German *-s*, instead, is another example of analogical expansion. It shows how problematic the distinction between the latter and grammaticalization can be. Clear cases of grammaticalization, such as the two mentioned, show a loosening of selection restrictions. This is parallel to the loosening of the restriction on plural *-s* that the noun so marked should be foreign and end either in a long vowel or on a stressed syllable. The reason why we nevertheless distinguish between grammaticalization and analogical extension is that the latter fails on the central criterion of grammaticalization, which is that the unit in question should become more subject to rules of grammar.

2.2. Case relations

The development of case markers is one of the standard examples of grammaticalization. Its reverse side is the loss of old inflectional case markers. Reduction to zero is the logical end phase of grammaticalization. Let us first look at the relevant phenomena and then discuss the problems behind them.

2.2.1. The possessive attribute relation

Case endings have been reduced and lost in German from Proto-Indo-European times. Feminine nouns, except proper names, no longer inflect for case. Neuter nouns form the genitive in *-s*. Masculine nouns fall into two declension classes and accordingly form the genitive either in *-s* or in *-en*. These markers have proved astonishingly resistant. Nevertheless, those nouns that take the *-s* allomorph lack it in a variety of cases, all of which have in common an article form — definite *des* or indefinite *eines* — which shows the genitive. Two of these cases are provided for by standard grammar:

a) Foreign nouns, especially if they end in *s*, often lack genitive marking: *des Atlas/Numerus*, but also *des Kazike*.
b) Those proper names which are generally articleless (chiefly, person and place names) do not take their genitive *-s* after an article. Such names normally take the definite article only if they are expanded by an attribute, as in *des*

jungen Goethe 'of young Goethe', *des alten Berlin* 'of old Berlin'. Names which are by themselves complex allow the definite article more easily: *des Ludwig van Beethoven*. However, this construction is not colloquial and connotes distance between speaker and referent.

The genitive attribute with article, but no mark on the noun(s) is currently extended in two directions: First, to other composite attributes, especially those including appositions. Thus, while *des Konjunktiv* is nonstandard, *des Konjunktiv Imperfekt* is all right. And examples like (1) and (2), although not (yet) standard, are increasingly found.

(1) *... ist eine Wanderausstellung des Koblenzer Bundesarchiv geplant.* [radio, 1988]
'... a flying exhibition of the Koblenz Federal Archive is planned.'

(2) *Anstelle des traditionell gewählten Mann des Jahres prangt ein Computer auf dem Titel.* [computer magazine, 1988]
'Instead of the traditionally elected man of the year, the title page boasts a computer.'

Second, the postnominal possessive attribute without genitive marking extends to such simple proper names as are generally accompanied by the definite article: *des Saturn/Mai/Hunsrück* (mountain range).[4]

Inside possessive attribution, there are currently two ways of renewing the genitive. We first turn to the preposition *von* 'from, of', which has been slowly making its way into the possessive attribute relation for a long time. Most genitive attributes are paraphrasable by a prepositional attribute involving *von* (*die Oper von Ludwig van Beethoven* 'Ludwig van Beethoven's opera'), although this is often stylistically clumsy and not high style. Thus, *die CPU von dem Computer* will be heard, but not read.

However, there are at least two cases where *von* must be used to signal the possessive relation. The first involves a personal possessive relation in an indefinite NP. The indefinite counterpart to the definite NP *mein Schwager* (=nonstandard *der Schwager von mir*) 'my brother-in-law' is *ein Schwager von mir*, which, keeping *Schwager* as the head constant, has no alternative.

Second, no noun has a distinctive genitive plural morpheme. If the genitive noun is modified by determiners or adjectives, these will have different patterns of case syncretism. For example, the masculine noun form *Effekte* can be nominative, genitive or accusative plural. The adjective form *stilistischer* can be nominative singular masculine, genitive and dative singu-

lar feminine and genitive plural of any gender. If the two form a phrase, as in *Erzielung stilistischer Effekte* 'attainment of stylistic effects', the case of this phrase is calculated as the intersection set of the case meanings of the two constituent forms, reducing to the genitive plural in this case (see Werner 1979). When the noun stands alone, case ambiguity is not resolved by morphological means. Now, while the other cases do not require morphological disambiguation, so that *Effekte* may constitute an NP in the nominative or accusative, it may not constitute an NP in the genitive. The modifier in *Erzielung stilistischer Effekte* cannot be dropped. Therefore, if *Effekte* alone is to constitute an attribute, *von* has to be used: *Erzielung von Effekten.*[5]

The other possibility of formally renewing the possessive attribute is by means of the possessive pronoun. There is a cross-linguistically widespread construction which combines the possessed noun with a possessive pronoun (referring to the possessor) and an (optional) noun phrase in some appropriate syntactic relation (representing the possessor). In German, this construction is nonstandard and dialectal. There are variants of it which differ as to case inflection of the possessive NP and possessor gender marking in the possessive pronoun. A mapping of this construction into Standard German would look as in (3)a.

(3) a. *der Lehrerin ihr Hut*
'the (female) teacher's hat'
b. *die Lehrerin ihr Hut*
c. *die Lehrerin sein Hut*

In (3)a, the case relation of the possessive NP is the dative (just as in the corresponding Hungarian construction), and the possessive pronoun duly reflects the possessor's gender. However, the dialects and sociolects which actually use this construction (for instance, Ruhrdeutsch, a variety of Lower Frankish) often also reduce inflection in various ways. Thus, the possessive NP may appear in the nominative, as in (3)b. Finally, the possessive pronoun is limited to the masculine-possessor variant, as in c. This is, incidentally, also what happened in the development of the English genitive *'s*.

Beside its function in attribution, the genitive has another one in verbal valency. There, however, it has retreated to a very few verbs, most of which have alternative valencies or enough synonyms to allow their avoidance. Typical examples include *sich erinnern* 'remember', which can take a genitive object, but more commonly takes a prepositional object with *an*, or *bedürfen* 'need', which has an almost perfect synonym in the transitive *brauchen*.

2.2.2. The dative relation

In the plural, the dative is the only case with a morphological mark of its own (*-(e)n*). In the singular, neuter nouns have *-e*/∅, while masculine nouns have either *-e*/∅ or *-(e)n*, according to declension class. The two allomorphs differ markedly in their stability. The *-en* allomorph is fairly stable. *Dem Herr* 'to the sir' is nonstandard and characteristic of child language. Yet, as in the instance of the loss of the genitive (*-s*), omission of the dative mark is encouraged if the noun has postnominal attributes; cf. (4).

(4) *Flucht vom Planet der Affen* [German movie title, 1971]
'Escape from the Planet of the Apes'

The *-e* allomorph has all but disappeared from the colloquial and even non-literary written language. Most nouns do not admit it at all (see the list in Heidolph et al., 1981:598), and those that do admit it normally do not carry it: *dem Manne/Hause/Gewichte* 'to the man/house/weight' are stilted. The suffix survives only in some phraseologisms such as *am Tage* 'in the daytime', *zu Hause* 'at home'. Fossilization is a common symptom of grammaticalization in its terminal phase.

Unlike the English and Latin-Romance dative cases, and unlike the German, English and Latin-Romance genitive cases, there is no single preposition taking the place of the dative inflection. The following verbs can all take simple indirect objects, but alternatively allow prepositional objects as indicated: *senden/schicken an* 'send', *weiterleiten an* 'forward', *genügen für* 'suffice'. Other verbs taking indirect object have derived counterparts taking direct object: *(be-)drohen* 'menace', *(be-)dienen* 'serve', *(be-)raten* 'advise', *(an-)schreiben* 'write'. Earlier yet, experiencer verbs which used to take the perceived entity as the subject and the experiencer in the dative changed their valence, taking the experiencer as the subject and the perceived entity as the direct object: *mir ahnt* > *ich ahne* 'I forebode', *mir träumt* > *ich träume* 'I dream'. On the other hand, many verbs such as *geben* 'give', *fehlen* 'be absent', *gehorchen* 'obey', have a stable indirect object construction without an alternative. Outside of verbal valency, the dative is mostly substitutable by prepositions, especially *für* 'for'. Thus, *ich kaufe ihm ein Buch* 'I buy him a book' = *ich kaufe ein Buch für ihn*, *das klingt mir merkwürdig* 'that sounds weird to me' = *das klingt merkwürdig für mich*. It would be a rewarding study to disentangle and then recombine change in inflection paradigms, in syntactic relations and in semantic relations in this area.

2.2.3. Loss by grammaticalization

As I have argued in earlier publications (e.g., Lehmann 1985), it is futile to ask whether the grammaticalized prepositions oust the older case suffixes or whether, instead, the attrition of the case inflection attracts the prepositions to fill the gap. It seems more appropriate to view adpositions and case affixes as being at two stages of the same scale and both being grammaticalized simultaneously and in parallel.

It has long been clear that one of the most palpable symptoms of grammaticalization is reduction on both the semantic and the phonological sides of the language sign. The logical endpoint of this is complete loss. However, not every case of loss of a grammatical element can be regarded as purely a matter of grammaticalization. This is also true of the reduction of case inflection. Both the examples of the genitive and of the dative show that factors at very different levels contribute to this reduction. At the level of the grammaticalization scale which includes adpositions and case affixes, it is, loosely speaking, a question of fulfilling a given function with an alternative device. If, however, case forms disappear because there are no longer any verbs which require or even admit them, it is rather a question of no longer fulfilling a given function. To the extent that this is true of German genitive and dative inflections, their fate is not exclusively one of grammaticalization.

2.3. Complex prepositions

The formation of new prepositions through the combination of a (mostly relational) noun with either a preposition or a case suffix is one of the most common grammaticalization processes in the world. It has been done in German, as in other Indo-European languages, for millennia. Since about 1975, there has been a new wave of complex prepositions. The most common of these are the following:

im Zuge 'by, during, in',
im Wege 'by (way/means of)',
im Vorfeld 'on the eve (of), before',
im Anschluß 'following, after',
im Gefolge 'in the sequel (of), after',
im Verfolg 'in the course (of)'.

They are mostly promoted in bureaucratic and journalistic jargon; none of them are really colloquial. Here are some illustrative examples:

(5) *Im Zuge der Regierungsumbildung wurden zwei Ministerien zusammengelegt.* [example constructed on a common pattern]
'In the course of the reshaping of the government, two ministeries were conflated.'

(6) *wenn die Bußgeldstelle im Wege der automatischen Datenverarbeitung arbeitet* [radio interview, 1987]
'if the fines department works by means of automatic data processing'

(7) *Im Vorfeld seiner Reise hatte der pakistanische Außenminister vor einigen Tagen erklärt, ...* [radio news, 1987]
'On the eve of his journey, the Pakistan Secretary of State had declared some days ago ...'

(8) *Im Anschluß an die Nachrichten der Wetterbericht.* [daily news]
'After the news, the weather forecast.'

(9) *Die vorliegende Untersuchung ist die revidierte Fassung einer Arbeit, die im Gefolge eines Seminars über 'Sprachliche Variation und Diversität' entstand ...* [linguistic paper, 1981]
'The present study is the revised version of a paper that originated in consequence of a seminar on "Linguistic variation and diversity"'

(10) *Im Verfolg der Untersuchungen ergab sich, ...* [scientific article, 1989]
'In pursuance of the investigations, it turned out ...'

The following features of this set are noteworthy:

1. All the cases involve a metaphor; the respective literal meanings are as follows:
im Zuge 'in the train/procession (of)',
im Wege 'in the way (of)',
im Vorfeld 'in the fore-field (of)',
im Anschluß 'in connection (with)',
im Gefolge 'in the suite (of)',
im Verfolg 'in pursuance (of)'.

2. There is a common formal pattern here: The basic nouns are bi- or trisyllabic, with final or penultimate stress. They combine with the form *im* (elementary locative preposition *in* plus definite article) to yield a complex preposition which (except for *im Anschluß*) requires the genitive of its complement. The pattern has been productive for a certain period (since about 1975, I guess). Earlier, such complex prepositions as *aufgrund* 'on the basis', *anhand* 'with the aid', and even the postposition *zufolge* 'according to' came into usage. I am not aware of any formation parallel to the latter in the last years.[6] That the above is a currently productive pattern is confirmed by an additional consideration: What is meant by *im Wege* would normally be expressed by *auf dem Wege*. *Weg* after the preposition *in* occurs exclusively in this complex preposition. There appears to be analogical pressure at work.

3. An examination of my data reveals a feature that so far seems to have gone unnoticed in research on grammaticalization: The complex prepositions arise in a narrowly circumscribed set of contexts. The complement of *im Zuge* and *im Verfolg* is a verbal noun, normally an action noun. The complement of *im Wege* is a new technology in all three recorded examples; in two of them it is automatic data processing.[7] The complement of *im Vorfeld* is a social activity, often some sort of meeting. The complement of *im Gefolge* is usually a trend-setting, preferably scientific activity. This is a preliminary hint that phraseology[8] is at work in grammaticalization: When a newly coined periphrastic expression is received by other members of the speech community, it will not be in isolation, but in the context in which it was originally coined. It will not then spread at once to all kinds of contexts which, given the rules of grammar, would admit it, but will initially be restricted to certain collocations which come close to being phraseologisms.

The interplay of grammaticalization and lexicalization can also be seen inside each of the complex prepositions. First of all, *im Zuge* and *im Wege* show the dative ending *-e* which is falling out of use in free syntax (cf. section 2.2.2), but survives in phraseological contexts. Second, the typical symptom of lexicalization, namely a disturbance of compositionality by either semantic or formal changes, is again observable in *im Wege*, since its internal structure does not obey the respective subrule of syntax. This is why complex prepositions are termed 'prepositional phraseologisms' in Burger et al., 1982:23. If they remain in the language, one day each of them will become an entry in the German lexicon, just as the earlier *auf Grund* has now been entered as *aufgrund*.

4. According to their external syntactic behavior, the six complex prepositions fall into two groups: Those that are used exclusively as prepositions (i.e. they require a complement), and those which allow adverbial use. *Im Zuge* and *im Wege* belong in the former group, *im Vorfeld* and *im Anschluß* belong in the latter. (11) and (12) illustrate the latter.

(11) ... *ist die Verfahrensweise der LABK auch darauf gerichtet, im Außenverhältnis fachaufsichtliche Interventionen grundsätzlich im Vorfeld abzuwehren.* [administrative circular, 1987]
'the procedure of the LABK is also directed, in its external relations, principally towards precluding in advance interventions of disciplinary supervision.'[9]

(12) *Mehr darüber gleich im Anschluß.* [radio lead-in, 1989]
'More on this directly hereafter.'

(12) should be compared with (8). The rhetorical purpose of (11) is apparently to avoid being precise on what exactly might provoke an intervention. It may be remarked at this point that, in general, periphrases do not seem to be used for the sake of greater precision. The law of iconicity according to which more expression generally corresponds to more content here appears to be suspended: The periphrases add vagueness by being intercalated into, and thus dilating, relations that would otherwise be direct.[10]

(13) shows *im Gefolge* without a complement.

(13) *Sie haben also auf Edelmann und in dém Gefolge auf Dieckmann hingewiesen.*
'you have referred to Edelmann and, in consequence of that, to Dieckmann.' [radio interview, 1987; a research tradition is being referred to.]

In dém Gefolge must mean either 'thereafter' or 'in the succession of Edelmann'. To the extent that grammaticality judgements are possible in the case of "trendy" phrases, (13) is probably ungrammatical in regard to the variety to which it belongs. I have no example of *im Gefolge* or *im Verfolg* in simple adverbial use.

Taken by themselves, the lexical nouns *Weg* and *Zug* would probably be classed as non-relational, while *Vorfeld*, *Anschluß*, *Gefolge* and *Verfolg* appear to be relational. The first guess would therefore be that the former set could be used adverbially, the latter not. The data point in the opposite direction. A possible explanation might be that only a relational noun gives the informa-

tion that reference to something else is implied, whereas adverbial *im Zuge, im Wege* would not link up with the semantic context in this way (cf. Seiler, 1983).[11]

2.4. *von ... her* and other limitative constructions

2.4.1. By far the most fashionable of those grammatical constructions which are unknown to normative grammar is the *von-her* construction. It has been in everybody's mouth since about 1980, and its new function seems so natural to speakers that most of the people whose attention I drew to its "trendy" character were completely unaware of it.

Before the change, the relevant situation was roughly as follows: The discontinuous preposition *von ... her* (preposition *von* 'from' plus deictic particle *her* 'hither') had the strictly concrete-local meaning '(coming) from', as in (14).

(14) *Von draußen vom Walde komme ich her.* [folk poem]
'I am coming from outside, from the woods.'

(15) shows a slight metaphorical extension in an adverbial phrase at the sentence level.

(15) *Er war vermögend, schon von den Eltern her, besaß zudem ein Herrenmodegeschäft ...* [*Schöne Welt*, 1987]
'He was already affluent from his parents, and in addition owned a gentlemen's fashion shop.'

The discontinuous preposition was also commonly used in a set of metaphoric phraseologisms centering around the notion *eine Sache von einem Standpunkt her betrachten* 'to view a matter from a standpoint', such as *von X her gesehen/betrachtet/beurteilt*, etc. 'seen/viewed/judged from X'. Such phrases were and are used as sentence-level adverbials to express such notions as 'with reference to, in/with regard to/of, with respect to, as regards, considering X'. Other devices to express this have been such clumsy circumlocutions as *was X betrifft, im Hinblick auf X, hinsichtlich X*, but also simpler constructions with *in* 'in' and *nach* 'after'. We will return to these in section 2.4.2.

The phrase *von X her gesehen/betrachtet* is already fairly fixed and depleted. Otherwise it would not be possible to express the same semantic relation to X either in an ablative (*von X her gesehen*) or in an allative (*im Hinblick auf X, hinsichtlich X*) metaphor. The first and decisive step in the recent development of *von-her* is the omission of the past participle of a verb

of viewing which formerly closed the adverbial. (16) is an example which might as well contain such a participle (e.g., *von heute her gesehen*) and would then illustrate the typical use of the latter.

(16) *Solche Sätze wirken von heute her nichts weniger denn kabarettistisch überzogen.* [*Die Zeit*, 1986]
'Such sentences, viewed from today's standpoint, seem anything but cabaretishly exaggerated.'

The omission of the participle is a remarkable feat, since it is the head of the phrase. It seems natural to assume that constructions such as the one in (15) served as a formal model. At this stage, we have a discontinuous preposition which expresses a limitative relation between a proposition and a factor which restricts the scope for which the proposition is valid. (17) is a very early example.

(17) *Auch ist die Zahl der bisher in Linear A aufgefundenen Inschriften nicht groß (etwa 300) und ihr Textumfang zumeist recht gering, desgleichen lassen sie von ihrem Inhalt her (Abrechnungsquittungen, kurze Weihaufschriften) keine breitere Erfassung der Sprache erwarten.* [linguistic book, 1970]
'Also, the number of the inscriptions discovered so far in Linear A is not large (around 300), and their text length is, for the most part, rather small; likewise, considering their content (clearing receipts, short votive inscriptions), they do not make one expect any broader conception of the language.'

(18)–(23) show how much the initial meaning of *von ... her* has broadened.

(18) *Der ist nicht der allerschönste, den ich vorführen kann, aber von der Tendenz her wird sicher klar, was gemeint ist.* [free speech, 1987]
'That is certainly not the most beautiful one that I can demonstrate; but the tendency of what is meant will surely get clear.'

(19) *Wenn der Minister in dieser Sache Wort hält, kann ich nur sagen, daß wir von der Zielrichtung her einer Auffassung sind.* [radio interview, 1987]
'If the minister keeps his word in this matter, I can only say that we are of one opinion as regards the goal.'

(20) *Also von den Bildern her, mir persönlich sagt der nich so zú, der Macke.* [conversation at a Macke exhibition, 1987]

'These pictures, me personally Macke does not appeal to that much.'

(21) *[Unter dem Fuß des Rechners] duckt sich der vorderste Teil der Mutterplatine. Von der Breite her schließt er schön ordentlich mit der Querleiste der Tastatur.* [computer magazine, 1986]
'Under the foot of the computer is located the foremost part of the motherboard. In its width, it is exactly flush with the cross-bar of the keyboard.'

(22) [Ad 'move alpha':] *Nun, dieser Ausdruck ist ein zeitlicher und suggeriert sozusagen vom Ausdruck her, daß da etwas passiert.* [free speech, 1987]
'Well, this expression is a temporal one and implies, you know, by this expression that something is happening here.'

(23) A: *Treffen die hauptsächlich auf Substantive zu?* — B: *Ja,* [C: *Auf Prädikatoren.*] — *also von den Wortarten her auf Prädikatoren.* [discussion, 1987]
'A: Do those apply principally to nouns? — B: Yes [C, interrupting: to predicators] — well, in regard to word classes, to predicators.'

Some of these examples show that the ablative meaning of the metaphorical base is completely lost. This is particularly clear in (19), which features *von der Zielrichtung her* 'from the goal-direction'. Others show that *von ... her* is used as a convenient device for introducing concepts which are somehow felt as relevant but whose exact relation to the point is not (to be made) clear. This can be seen in (23). Most important, however, are the roles of *von-her* phrases as contrastive topic (in the sense of Chafe 1976), as simple topic and as shadow actant. These will now be taken up in turn.

A contrast between the NP in the *von-her* phrase and a concept in the context can already be seen in (17). (24) and (25) are unmistakable examples of *von-her* phrases functioning as contrastive topics.

(24) *Es ist diese und nur diese Dimension, die seine Kunst als subversiv vorkommen läßt und indizierende Saubermänner auf den Plan ruft. Vom bloßen Wortlaut her sind die Formulierungen nämlich eher von der verbindlichen Art.* [*Die Zeit*, 1986]
'It is this and only this dimension [of interactive speech] that makes his art appear subversive and that evokes indexing expurgators.

For, as regards the wording alone, the formulations are rather of the diplomatic kind.'

(25) *Herbert Hupka ist auf Ceylon geboren. Die ersten drei Lebensjahre verbrachte er in Australien, aber von Herkunft, Elternhaus und Schule her sind Ratibor und Oberschlesien seine Heimat.* [book, 1980]
'Herbert Hupka was born in Ceylon. The first three years of his life he spent in Australia, but as for origin, home and school, Ratibor and Upper Silesia are his homeland.'

The use of *von ... her* to mark simple topic can be regarded as a further step in its depletion. Relevant examples are (21) and (23). By far the most common phrase of this type, and the most common use of *von ... her* in general, is the phraseologism *von daher*, as in (26).

(26) *... Also von dáher weiß ich nicht, was das sóll.* [from a discussion, 1986]
'In view of this, I wonder what that is supposed to be (good for).'

Von daher is a sentence introducer that links the sentence with the preceding context. The nature of the suggested relation is as vague as in the other cases.

Loss of communicative force in the sense of functional sentence perspective has often been regarded as a factor in grammaticalization, above all in the various hypotheses about a development from topic to subject. In the present connection, it suffices to realize that when the communicative function of an element fades, its syntactic cohesion with the context tends to tighten. The limitative force of the *von-her* phrase then relates directly to a particular sentence actant. In a sense, it represents that actant at a different syntactic level. A shadow actant is, then, a syntagm which bears a limitative semantic relation to a certain actant. Semantically, it is part of that actant or on equal footing with it as regards the predication. It does not bear itself the syntactic function of that actant, but a small transformation, with no change in voice and little change in meaning, could put it into that actant's position. (18) and (20) show what is meant more effectively than further explanation.

Given both the general prominence of the subject over the other actant functions and the natural affinity between subjects and topics, it becomes intelligible that the shadow subject is by far the most common shadow actant function of *von ... her*. Typically, the syntactic subject is either empty, as in (18), or it bears a relation of factual connection to the NP in the *von-her* phrase, e.g. a possessive relation in (17), (21) and (24), a source relation in (20). An extreme example of this kind is (22), in which the syntactic subject

is repeated in a *von-her* phrase. An example of a shadow object will be found below.

Concomitant with the development from contrastive topic via simple topic to shadow actant, we find an increasing integration of the *von-her* phrase into the clause. In some of the examples, especially (17), (19) and (22), it is no longer a sentence-level adverbial, but well integrated into the clause. It may be recalled that decrease of structural scope is a symptom of grammaticalization (cf. Lehmann 1982, ch. IV.3.1).

Although German allows a variety of discontinuous constructions, it does not exactly favor them. One may therefore wonder whether discontinuity would not be an obstacle in the grammaticalization of an element. Actually, the second part of *von ... her* is dropped more and more often,[12] so that the result is the simple preposition *von*, which is in the center of the system of grammatical relations, anyway. Here are some examples:

(27) *Lisp ist nach FORTRAN die zweitälteste Programmiersprache, von der Sprachkonzeption aber völlig anders.* [computer magazine, 1987]
'Lisp is, after FORTRAN, the second-oldest programming language, but completely different in its language-conception.'

(28) *weil sie eben in Relation zu den anderen Kandidaten von den Vorleistungen nicht das aufzuweisen hatte* [from a discussion, 1986]
'because, in relation to the other candidates, she did not have that much to offer from her previous performance'

(29) *Die Gesellschaft ist inzwischen auf über 800 Mitglieder angewachsen, und das ist natürlich von der Zielsetzung der DGfS äußerst erfreulich.* [free speech, 1987]
'The society has meanwhile grown to over 800 members, and this is, of course, highly gratifying, in view of the goal of the DGfS.'

(30) *besonders raffiniert von der Farbe* [sticker at a dress in a show-window, 1987]
'particularly sophisticated in its color'

(31) *Man müßte nun für jeden dieser zusammengesetzten Datentypen ... Prozeduren schreiben, was allein vom Schreibaufwand sehr aufwendig wäre.* [programming handbook, 1984]
'One would now have to write procedures for each of these complex data types, which would be very expensive considering the writing expenditure alone.'

(27) exemplifies the contrastive topic, (30) the shadow subject (the color of the dress is sophisticated), (31) a limitative relation which is close to the shadow subject. (28) shows a particularly high degree of the integration of the *von (- her)* phrase, which in this case yields a shadow object; an adequate paraphrase would be *weil sie eben in Relation zu den anderen Kandidaten nicht die Vorleistungen aufzuweisen hatte* 'because ... she did not have that previous performance to offer'. Incidentally, a comparison of (29) with (19) confirms the point made in section 2.3 about the phraseological locus of many newly grammaticalized elements.

2.4.2. The examples are from all kinds of non-literary texts, from informal talk to journal articles to scientific books. There can be no doubt that (quite independently from the use of a "trendy" expression) several of them evidence unpremeditated, negligent talk or bad style, e.g. (18), (20), (21), (22), (23), (30),[13] (31). To complete the picture, we should ascertain that the alternatives to *von (- her)* are alive, although they are not used as spontaneously as their competitor. Those that I have data for are *hinsichtlich* 'in regard to' ((32)), *gemessen an* 'measured against' ((33)), *in* 'in' ((34)) and *nach* 'after' ((35)).

(32) *Der mit 12 MHz getaktete ARC Turbo 12 kann sich hinsichtlich seiner Leistung mit dem PS/2 Modell 60 von IBM messen.* [computer magazine, 1987]
'The ARC Turbo 12, tacted with 12 mhz, can compete with the PS/2 model 60 of IBM in performance.'

(33) *Der neue Tischcomputer aus Texas stellt nämlich, gemessen an Preis und Leistung, alles bislang Dagewesene in den Schatten.* [*Capital*, 1986]
'The new desk computer from Texas indeed overshadows, in regard to price and performance, everything that has appeared up to now.'

(34) *Wir sind ja doch eben gerade in der Téchnik, von der Téchnik her, wesentlich weiter.* [radio interview, 1987]
'For we are substantially more advanced particularly in technology, viewed technically.'

(35) *Noch ein Wort dazu, was Beispiele ihrer Funktión nach sein können.* [free speech, 1987]
'Another word on what examples can be as to their function.'

In each of these examples, the phrase in question could easily be substituted by a *von-her* phrase, with no change in meaning. (34) is particularly intriguing as the speaker offers the paraphrase himself, apparently thinking that the received construction does not sufficiently bring out his point. Compare, moreover, (35) with (17) and (32)/(33) with (28). If *von (-her)* should be further grammaticalized, then the above alternatives will probably survive, as they allow the expression of more specific limitative relations.

It is very hard to prognosticate where this development can lead. On the one hand, it is probably no accident that it is just the preposition *von* which acquires the function of shadow subject. It is well established in the function of marking the passive agent, which probably has semantic affinities with the shadow subject as illustrated above. In view of this, there is little to prevent German from getting double subject constructions just as Japanese. On the other hand, it might be an obstacle for such a development that *von* already has so many functions, in particular the ablative and possessive functions, so that it might prove undesirable to freight it with an additional one.

2.4.3. Finally, yet another limitative construction should be mentioned, which I have singled out because it is a word formation process. It combines the topic (or shadow actant) noun with the complex derivational suffix *-mäßig* ($[[x]_N$-*mäßig*$]_{Adj}$ 'by the measure of x') and thus converts it into an adjective which, for the present purpose, is used adverbially. This construction alternates to a large extent with the *von-her* construction and is insofar synonymous with it. Examples can be derived from (28) by replacing *von den Vorleistungen* by *vorleistungsmäßig*, or from (32) by replacing *hinsichtlich seiner Leistung* by *leistungsmäßig*.

This process is fashionable, too. However, as is typical in word-formation, it is restricted to single-noun topics which, in one of the alternative constructions, would be accompanied at most by a definite determiner. The complex suffix *-mäßig* is in a paradigm with a few other adjective-forming suffixes such as (N)-*förmig* 'of the form (of N)', (N)-*haltig* 'containing (N)' and (N)-*artig* 'of the kind (of N)'. These are currently productive and illustrate the development of new derivational suffixes from formerly free lexical morphemes. The phenomenon has been much investigated in German linguistics (cf., e.g., Inghult 1975). It illustrates the transition from compounding to derivation. It might be subsumed under grammaticalization. This would entail that derivation is more grammatical than composition. The latter problems are discussed in Lehmann 1989. Both these and the question of the function of all the devices mentioned in this section need much more investigation.

3. VERBAL CATEGORIES

3.1. Modal verbs

There is a long tradition in German of creating verbal categories by embedding a non-finite form of the full verb in a periphrasis with another verb which then degenerates to an auxiliary. The most recent of these which is well established as an auxiliary, but still expanding, is *würde* 'would', which forms an analytic subjunctive which can substitute for the synthetic past subjunctive. The relationship between the two is very similar to the one between the synthetic genitive and the analytic *von* construction (cf. section 2.2.1). Normative grammar recommends the use of the synthetic form wherever possible, but the lower the level of formality, the more the analytic form prevails. Also, there are cases where the synthetic form is insufficient and thus impossible. For example, originally *möchte* was the past subjunctive of *mag* 'like'. Now, however, this category must be paraphrased by *würde mögen*, because *möchte* has lexicalized into a polite form of *will* 'want'.[14] I hypothesize that the grammatical necessity of using an analytic construction instead of a competing synthetic one is a safe hint as to the direction in which the development goes. However, on the whole there have been no revolutionary developments in our century as regards the expansion of *würde*.

More recent is the development of *brauchen* 'need, require' into a modal verb. As this is well studied by now (cf., e.g., Wurzel, 1984:149), a few remarks will suffice here. The established modal verbs, such as *wollen* 'will', *können* 'can', *müssen* 'must' and a couple of others, form a very homogeneous group. They differ from lexical verbs in such respects as being defective as to verbal categories, having different personal endings and taking their infinitive complement without *zu* 'to'. All of these are common symptoms of grammaticalization.

In standard German, the verb *brauchen* belongs in the class of verbs which take an infinitive complement with *zu*, as in (36).

(36) *Das beabsichtigt/braucht er nicht zweimal zu sagen.*
'He does not intend/need to say that twice.'

In colloquial German and various dialects, it is making its way into the group of modal verbs by adapting itself semantically and formally to its constraints (see Lehmann 1982, ch. IV.4.2.2 on paradigmaticization). Semantically, it functions as the counterpart of *müssen* in negative contexts, as illustrated in (37).

(37) a. *Das muß er zweimal sagen.* 'He must say that twice.'
b. *Das brauch er nicht zweimal sagen.*
'He need not say that twice.'

Incidentally, it is not quite clear why this should happen, since *das muß er nicht zweimal sagen* means the same as (37)b. Anyway, if the former alternative did not exist, *brauchen* would be, as part of the paradigm of modal verbs, in a suppletive relation with *müssen*. The formal adjustment of *brauchen* involves loss of the final *-t* in the third person singular present and omission of the *zu* before the infinitive; cf. (37)b with (36). It is a safe bet that *brauchen* will become a fully integrated modal verb if it survives.

3.2. Progressive aspect

Other European and Non-European languages have a progressive aspect of the type *I am working*, *estoy trabajando* or even *je suis en train de travailler*. Standard German, which is, perhaps, one of the poorest languages as regards the category of aspect, has nothing of the sort. However, in a couple of West German dialects, in particular Lower Frankish (spoken, e.g., in Cologne), a progressive aspect has come into use.[15] This is expressed by a periphrasis which involves *sein* 'be' plus the infinitive, which is governed via the preposition *an* 'at' fused with the definite article of the infinitive to yield *am*. (38) is a typical and, at the same time, the earliest example I know of.[16]

(38) *Ich bin am Arbeiten.* [letter by August Macke, 1908]
'I am working.'

There are a couple of possible models for this construction. On the one hand, there is the gerund of the type *ich habe mich beim Arbeiten verletzt* 'I have hurt myself at work/while working', which also features the infinitive depending on a preposition fused with the definite article. This can also occur as the predicate of a copula sentence: *ich bin beim Arbeiten* 'I am at work/working'. On the other hand, there is the model expression *ich habe zu arbeiten* 'I have to work,' which features the infinitive depending, via the preposition *zu*, on an auxilliary verb and thus forming a periphrastic verb form. Except for the preposition *an* itself, which is not otherwise used in standard German for the embedding of non-finite verbs, all the structural features of the progressive construction are well established in standard German. It is currently spreading into the colloquial language of Northwest Germany. The interesting feature about the relative chronology of the spread is that intransitive verbs allow

the progressive aspect much more easily than transitive or other plurivalent ones. Thus, while many speakers would accept (38) at the colloquial level, examples like (39) are still rare outside of the original dialect area.

(39) *Ich bin jetzt dieses Orientierungs-Papier am überarbeiten.* [Colonese speaker, 1988]
'I am currently reworking this orientation paper.'

As (39) shows, the internal syntax of the verbal group in the progressive follows the pattern of other periphrastic verbal categories: Auxiliary verb in second position, then all the complements and adjuncts (in partly variable order), finally the non-finite form of the full verb. The preposition-plus-article portmanteau morph *am* forms a constituent with the non-finite verb form.

There is an alternative construction, exemplified in (40)b.[17]

(40) a. *Martha ist Flaschen am sortieren.*
'Martha is sorting out bottles.'
b. *Martha ist am Flaschen-Sortieren.*
'Martha is bottle-sorting.'

Here the complement comes between *am* and the non-finite verb form. This construction is obligatory if the direct object is indefinite and non-specific; cf. B41.

(41) *Jupp ist am Radfahren / *Rad am fahren.*
'Joe is bicycle-riding.'

Again, examples like (40)b and (41) are more widely accepted than those like (40)a. They have to be analyzed as instances of incorporation. This is in consonance with the above observation that the progressive is much more common with intransitive than with transitive verb constructions: The incorporation construction is treated like an intransitive one.

If one wants an explanation for the order in the spread of the progressive, there is a structural and a semantic factor that can be adduced. First, the relevant shallow[18] syntactic structure of constructions such as those in (39) and (40)a. is as in figure 1. This means that the material depending on the non-finite verb form is outside the scope of the definite article which determines it. This is non-iconic and unprecedented in German syntax.

Second, the semantic locus of the progressive aspect is in intransitive verbs. Progressivity and intransitivity have a natural affinity (see, e.g. Hopper & Thompson, 1980), as the progressive aspect centers the attention on the situation of the subject. Given such conditions, one would predict that if a

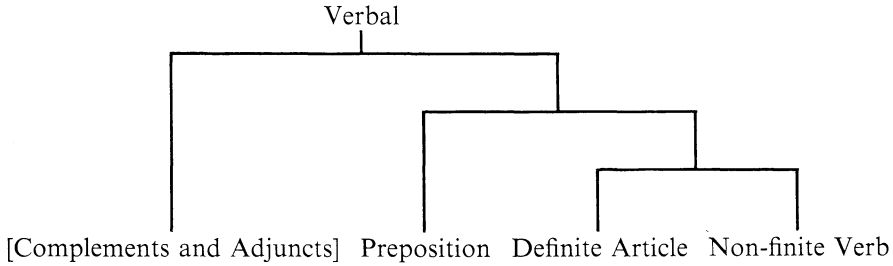


Figure 1: *Progressive of syntactically complex verbal*

grammatical category is introduced into a language, it will start from its natural locus and only later penetrate into less prototypical domains. This is borne out by the diffusion of the German progressive aspect. In the meantime, it is intriguing to observe the conspiracy of the structural and the semantic factors, which are completely independent.

3.3. Andative and venitive

Motion deixis is well established in German in the form of the particles *hin/her* 'away from/towards the deictic center'. These combine freely with motion verbs; cf. *hinschwimmen* 'swim towards a point outside the deictic center' vs. *herschwimmen* 'swim towards the deictic center'. Their combination with other verbs, however, essentially leads to lexicalization, with the deictic feature being lost. Cf., from *sagen* 'say', *hinsagen* 'say carelessly' and *hersagen* 'rehearse', or, from *machen* 'make', *hinmachen* 'hurry up' and *hermachen* 'produce an impression'. The idea of the grammatical categories andative vs. venitive, however, is that any verbal action can be characterized as having its origin in the deictic center (and, thus, being directed away from it) vs. having its origin outside the deictic center (and, thus, preferably being directed towards it). Cf. the Italian periphrastic passive with *andare* vs. *venire* as a paradigm example.

Again in the colloquial language of Northwest Germany, there has been an andative for quite some time. It is formed by coordinating the full verb with *hingehen* 'go away from the deictic center', as in (42).

(42) *Nun könnte man ja hingehen und das Ganze nochmal schreiben.*
'Now one [including the deictic center] could go right ahead and write the whole thing again.'

The range of functions of this construction is not yet understood. I realized its andative function only when I first heard its venitive counterpart, as in (43).

(43) *Was wir nun nicht verstehen, ist, warum man nicht herkommt und das schutzgeimpfte Tier dem antibakteriell positiven Tier gleichstellt.* [radio interview, 1987]
'What we now do not understand is why one [i.e. persons outside the deictic center] does not put the protectively inoculated animal on a par with the antibacterially positive animal.'

Herkommen und, as in (43), is much rarer than *hingehen und*. Its venitive function, however, is unmistakable. Lacking sufficient data, I cannot say more about this.

3.4. The addressee passive

German passives demote the active subject and signal that whatever is the subject is not the agent. The standard passive involves the use of *werden* 'become' and *sein* 'be' as auxiliaries. If applied to transitive verbs, it involves the promotion of the active direct object to subject.[19] The colloquial language has, in addition, an addressee passive (see Heidolph et al. 1981: 556f). This involves *bekommen/kriegen* 'get' as auxiliaries. Its locus of application is ditransitive verbs with an indirect object, as in (44).[20]

(44) a. *Die Arbeiter zeigten ihr den ganzen Betrieb.*
'The workers showed her the whole operation.'
b. *Sie kriegte (von den Arbeitern) den ganzen Betrieb gezeigt.*
'She was shown the whole operation (by the workers).'

The active indirect object is promoted to subject. Other complements remain unaffected. Just as in the standard passive, the active subject can be adjoined in a *von*-phrase.

The construction is currently expanding in the following directions:

1. The variant employing *bekommen* penetrates into the standard language.[21] Thus, a version of (44)b with *bekam* instead of *kriegte* appearing in the radio news is not unthinkable.

2. The addressee need not be an indirect object. NPs in benefactive or possessive functions, as long as they are datives in the active version, can also be promoted, as in (45).

(45) *Sie kriegte den Wagen repariert.*
'She got the car repaired.'

3. The verb does not need to have a non-dative complement. The addressee passive is becoming possible with verbs that take only an indirect object.

(46) *Sonja bekommt immer von euch geholfen.* [conversation, 1987]
'Sonja always gets helped by you.'

4. The complement to be promoted does not even need to be in the dative. Other syntactic functions are allowable, as long as they express something like an addressee semantic function, as in (47).

(47) *Dann kriege ich immer geschimpft.* [conversation, 1987]
'Then I always get scolded.'

(46) and (47) are observed data. Speaker judgements vary on such examples; they are outside my own idiolect. Steps 2–4 of the expansion nicely illustrate the dropping of selection restrictions which stem from the etymology of the construction.

Restricted to the colloquial language, there is yet another periphrastic construction involving *bekommen/kriegen*, with a meaning that Seiler (1973) calls 'possessor of an act', as exemplified in (48).

(48) *Kriegst du das bis heute abend geschrieben?*
'Can you (manage to) get that written by this evening?'

The construction is formally identical to the addressee passive. Here, however, there is no implication that the subject cannot be agent. Instead, there is an attribution of the act to the subject as the one who is responsible for it, so that he will either have to do it himself or else see to it that it is done. Probably, this auxiliary function of *bekommen/ kriegen* is based on (interference with) a different meaning of *kriegen*, namely 'manage, succeed'.

3.5. Function verbs

A function verb construction (*Funktionsverbgefüge*, also *[verbale] Streckform* 'stretched [verbal] form') is a syntagm consisting of a moderately grammaticalized verb and an abstract verbal noun as its complement such that the syntagm bears a paradigmatic relation to the verb underlying the verbal noun. The paradigmatic relation may consist of a valency change or a (some-

times very slight) change in the *aktionsart* of the full verb. The grammaticalized verb takes on those grammatical properties of the corresponding full verb which can only be expressed by a finite (as opposed to nominalized) verb, including subject and direct object government, while the verbal noun retains the lexical properties of the full verb, including that part of its valency which is not affected by nominalization. The grammaticalized verb is called a function verb (*Funktionsverb*). Function verb constructions have been a favorite topic in German linguistics (cf. Helbig, 1979; Heidolph et al., 1981:431–442); thus I can be brief here.

Function verb constructions have their stylistic locus in bureaucratic, scientific and journalistic jargon. Some function verbs are more grammaticalized and form highly regular constructions; also, several function verbs are synonyms and thus mutually exchangeable with a given verbal noun. Other function verb constructions are more lexicalized and have the character of phraseologisms (cf. Burger et al., 1982:37); they may not even have a simple verb corresponding to them. There are transitions between the two groups. Thus, some of the highly regular function verb constructions such as *Anwendung finden* 'find application', as in (52), are fairly common and to that extent approach the status of phraseologisms. The whole strategy, thus, is productive more or less like a derivative pattern. I will here concentrate on the grammaticalized cases. Some of the most common function verbs are *kommen* 'come', *gelangen* 'arrive, reach', *bringen* 'bring', *finden* 'find', *erfahren* 'experience, undergo', *unterziehen* 'subject', *zuführen* 'bring to', *vornehmen* 'execute'. Here are some recorded examples:

(49) *eine Notationsvereinfachung, bei der nur mehr die Endeinheiten D und 1/2 zur Darstellung kommen.* [linguistic book, 1984]
'a notational simplification in which only the terminal units D and 1/2 are represented'

(50) *Wir haben dieses Konzert aufgenommen und bringen es heute zur Sendung.* [radio, 1988]
'We have recorded this concert and are broadcasting it today.'

(51) *bis ein Stapelprogramm geladen und zur Ausführung gebracht wird* [computer magazine, 1987]
'until a stack program is loaded and executed'

(52) *Für Sendungen nach den österreichischen Zollausschlußgebieten ... finden die für den Verkehr innerhalb des Bereichs der Deutschen*

Bundespost geltenden Gebühren Anwendung. [mail information booklet, 1982]
'For shipments to the Austrian Customs-Exclusion Territories, the fees which are valid for traffic within the domain of the Deutsche Bundespost apply.'

(53) *Einige Bilder erfahren auch eine neue Interpretation* ... [book commercial, 1987]
'Also, some pictures get a new interpretation'

(54) *Diese Regelungen sollen noch einer vollständigen Überarbeitung unterzogen werden.* [radio interview, 1987]
'These regulations are yet to be subjected to a complete revision.'

(55) *daß sie einer geregelten Bejagung zugeführt werden können.* [hunting journal, 1987]
'that they [the kinds of game] can be subjected to a controlled hunting'

(56) *Herr N.N., der seinerzeit die Aufstellung des Glasbruchschadens vorgenommen hat* [business letter, 1988]
'Mr. N.N., who at that time made the statement of the glass damage'

The most productive function verbs form morpho-syntactic paradigms. Figure 2 presents the commonest pattern (VN = verbal noun).

version / construction	active	passive
simple full verb	x_{sbj} V y_{obj}	y_{sbj} wird (von x) V_{pass}
function verb constructions	x_{sbj} bringt y_{obj} zu VN_{dat} x_{sbj} unterzieht y_{obj} VN_{dat}	y_{sbj} kommt/gelangt (durch x) zu VN_{dat} y_{sbj} findet/erfährt (durch x) VN_{acc}

Figure 2: *Terminative function verb constructions*

These constructions all involve terminative *aktionsart*. As is apparent from Figure 2, the periphrasis of an active simple verb generally involves transitive function verbs, while the periphrasis of a passive simple verb generally involves intransitive function verbs. Here as elsewhere in the grammar of German and other languages, *bringen* is the causative-transitive counterpart

of *kommen*. The same lexical relationship obtains between *unterziehen* and *erfahren*. The choice between the two function verb patterns in Figure 2 depends on the degree of affectedness of the object of the simple verb; *finden/erfahren/unterziehen* are appropriate with more strongly affected objects.

4. INTENSIFICATION

Intensification is probably an area which continuously undergoes change in most languages; people are never content with available means to express the importance of what they have to say. Very much goes on here at the nonstandard level, where what was a strong expression yesterday is antiquated today. However, these are chiefly lexical changes. We will look at two subareas. Other languages have grammatical means of expression in these subareas so that there is a chance of the German phenomena being cases of grammaticalization. Moreover, the two phenomena described in this section are typical of the sometimes stilted register of journalistic prose.

4.1. *mehr als*

The history of modern *mehr als* 'more than' remains to be written. Probably the first and decisive step in its development into an intensifier was a syntactic reanalysis. Cf. (57)a. and b.

(57) a. *Das ist mehr als das Doppelte.*
'That is more than double.'
b. *Das ist mehr als ausreichend.* [computer magazine, 1987]
'That is more than sufficient.'

In (57)a., *mehr* is the head of the phrase, *als das Doppelte* is some kind of complement. (57)b. can be analyzed analogously. But current usage presupposes its reanalysis with *ausreichend* as the head of the phrase, *mehr als* as a modifier. Cf. Matthews, 1981:156f for similar reanalyses.

Phraseology again plays a major role. The collocation in (57)b was found in two computer journals of 1987. Other common collocations are *mehr als zweifelhaft/fraglich* 'more than doubtful/questionable'. These cases all admit a literal interpretation. That is, *mehr als ausreichend* could signify, on a scale reaching from wanting over insufficient, sufficient, adequate to abundant, a point above 'sufficient'; *mehr als zweifelhaft* could imply 'improbable', assuming that this is the next step on the scale of plausibility; and similarly for all

the other constructions. However, there are reasons for assuming that this is not the intended interpretation. First, how do we know that *mehr als ausreichend* means '*above* "sufficient" on the appropriate scale', whereas *mehr als zweifelhaft* means '*below* "doubtful" on the appropriate scale'? Second, there are collocations with adjectives such that the literal interpretation does not make sense, as in (58) and (59).

(58) *Ungeachtet solcher Erfolge sind aber die Aussichten mehr als gering, daß die Resolution den angestrebten Erfolg ... erreicht.* [written radio commentary, 1987]
'Despite such successes, the prospects are more than dim that the resolution will attain the desired success.'

(59) *Mehr als deutlich führt dieses etwas ausführlich beschriebene Beispielprogramm vor Augen, wie systemnah ein Software-Designer mit Modula-2 programmieren kann ...* [computer magazine, 1987]
'This example program, described in some detail, demonstrates more than clearly how closely to the system a software designer can in fact program with Modula-2'

In (58), a literal interpretation would require a level above *gering* 'low', whereas the intended interpretation is, if anything, a level below 'low'. In (59), there is simply no appropriate scale, so that *mehr als deutlich* cannot refer to a level above 'clear'; it can only mean 'with particular clarity'.

This extension to adjectives for which only an intensifier interpretation is available marks a step in the expansion of *mehr als*. Another step is its transfer to nouns, as in (60) and (61).

(60) *Sein Versprechen ist mehr als eine Versicherung.* [commercial, 1987][22]
'His promise is more than an insurance.'

(61) *Wenn es aber gelänge, die "Verweildauer auf vier Tage auszudehnen", dann wäre das mehr als ein Erfolg.* [*Die Zeit*, 1988]
'But if it were possible "to extend the period of stay [of tourists] to four days," that would be more than a success.'

Before 1980 or so, *mehr als ein Erfolg* could only have meant, with the appropriate intonation, 'two or more successes'. In (61), it apparently means something like 'a good success'. Similarly, *mehr als eine Versicherung* in (60) is not supposed to mean 'two or more insurances', but probably 'a good insurance'.

The last step in the expansion of *mehr als* so far is its transfer to verbs, as in (62)–(64).

(62) *Der Nachteil, daß das Fehlen eines Linkers die Entwicklung größerer Programmprojekte erschwert, wird durch die hohe Übersetzungsgeschwindigkeit von Turbo-Pascal mehr als aufgehoben.* [computer magazine, 1987]
'The disadvantage that the development of program packets is impeded by the lack of a linker is more than compensated for by the high compilation speed of Turbo-Pascal.'

(63) *[der Nachteil] läßt sich durch die hohe Druckgeschwindigkeit, also ein Mehrfach-Ausdrucken eines Dokuments, mehr als ausgleichen.* [computer magazine, 1987]
'[the disadvantage] can more than be compensated for by the high printing speed — that is, a repeated printing of a document.'

(64) ... *Vertrauensbeweis, den er mit seinen Forschungen mehr als rechtfertigte.* [written speech, 1987]
'... proof of trust which he more than justified by his research'

Mehr als V means something like 'V and do something in addition', or even simply 'fully V'. Note again the phraseological character of the collocations in (62) and (63). The development is so recent that one cannot yet say which lexical groups of verbs allow *mehr als* as a modifier. What is certain is that only verb forms which are final in the verbal syntagm allow it. This includes non-finite verb forms, as in (62) and (63), and verbs in subordinate clauses, as in (64). This restriction is a consequence of the almighty verb-second principle in independent declarative sentences, which allows no modifiers before the finite verb. Thus, *he more than justified that proof of trust* has no direct translation into German.

The historiography of *mehr als* as an intensifier would probably reveal that it is a loan translation from English, in which *more than* is further advanced in its development. It is not subject to ordering constraints; and it has a negative counterpart, *less than*, as in *the other co-editors have been less than responsive* [letter, 1987], which does not have a German equivalent. The latter point may be interpreted as showing that the English phrase is paradigmaticized and, thus, grammaticalized at least to some extent. In the German case, this is not so clear. *Mehr als* forms a phraseologism, but it is doubtful whether this phraseologism is making its way into the grammar.

4.2. *zunehmend*

Continuous intensification of a property which is expressed by an adjective, as in 'faster and faster', used to be expressed by *immer* 'ever', as in *immer schneller*. For concepts expressed by verbs undergoing a change of continuous increase, the usual expression was *immer mehr* 'ever more'. Recently (estimate: since 1980), *immer (mehr)* has tended to be replaced by *zunehmend* 'increasingly'. (65)–(67) show this for the intensification of adjectives, (68) shows it for verbs.

(65) *Die Nachfrage wird zunehmend schwächer.* [radio commentary, 1987]
'The demand becomes increasingly weaker.'

(66) *Diese Schattenwirtschaft wird zunehmend weniger von der Partei kontrolliert.* [radio, written text, 1988]
'This shadow economy is controlled increasingly less by the party.'

(67) *Das echte Elfenbein wurde Ende des letzten Jahrhunderts zunehmend selten und teuer.* [radio, written text, 1989]
'The real ivory became increasingly rare and expensive by the end of the last century.'

(68) *Die Nachfrage verlagert sich zunehmend ins Ausland.* [constructed on a common pattern]
'The demand increasingly shifts to foreign countries.'

(65)–(67) clearly show the degree of desemanticization which this device has reached in a few years (cf. the semantically parallel (58)). I suspect that this is, again, a loan translation from English. Since *immer (mehr)* is probably a grammatical device of continuous intensification, *zunehmend*, too, may be conceived of as a marginal case of grammaticalization.

5. SENTENCE-LEVEL SYNTAX

The following two subsections deal with changes in main constituent order. Such are not *prima facie* cases of grammaticalization. However, one of the factors in grammaticalization is the reduction of syntagmatic variability (see Lehmann, 1982, ch. IV.3.3). This leads to a fixation of the sequential order of grammaticalized elements in relation to their host elements. A

language that used to have postpositions can acquire prepositions through the grammaticalization of relational nouns or coverbs which precede their complements. The result would be an apparent change in the word order of adpositions.

Again, rigid ordering constraints at low grammatical levels cannot be directly loosened (such loosening would be degrammaticalization). However, they can be replaced by freer ordering if the construction in question is renewed. For example, if there was a rule that possessive attributes must be prenominal, there is no possible rule change that allows the same attributes to alternatively stand in postnominal position. There is, however, the possibility of acquiring, maybe through the grammaticalization of an anaphoric pronoun to an attributor, a new kind of possessive attribute, which would then be postnominal.

Word order change can, therefore, be the consequence of grammaticalization. It is from this perspective that the following two cases should be viewed.

5.1. Verb-second position in subordinate clauses

German is famous for requiring final position of the finite verb in subordinate clauses, an order which never appears in main clauses. There have been enough analyses considering one of the orders as basic and the other as derived. This issue becomes superfluous as soon as one realizes that the orderings in main and subordinate clauses are in a grammaticalization relationship. According to the generalized Penthouse Principle (Ross, 1973), syntactic diversity is reduced at lower grammatical levels. In particular, constituent order is more rigid in subordinate clauses than in main clauses, and it is more rigid in non-finite subordinate constructions than in finite subordinate clauses. If we apply this to the facts of German, the following (simplified) picture of German main constituent structure emerges:[23]

	independent	dependent	
		finite	non-finite
declarative	#X V Y#	*#Z V#*	*#Z V#*
non-declarative	#V Z#		

Figure 3: *Constituent order in clause types*

Whereas there are two alternative orderings in main clauses according to sentence type, there is only one ordering in dependent clauses. In addition

to this greater paradigmatic variability, there is also greater syntagmatic variability in main clauses. There is, to begin with, a ternary subdivision of positional constituents in declarative main clauses, but only a binary one in subordinate clauses (cf. Heidolph et al., 1981:705f).

Ex hypothesi, the verb-final rule for subordinate clauses represents an instance of relative rigidity of sequential ordering. It would therefore be unexpected for non-finite dependent constructions to again show the ordering possibilities of main clauses. Quite to the contrary, the locus of the verb-final constraint is in non-finite subordinate constructions. Finite subordinate constructions display verb-final ordering insofar as they are close to, or modeled on, non-finite ones.

From this reasoning, the expectation arises that if there is variation of main constituent order in dependent clauses, it will be in finite rather than in non-finite ones, and it will show one of the possibilities that obtain for main clauses, rather than yet another one. This is borne out by the facts. There is considerable variation in both colloquial and written German with respect to verb order in finite subordinate clauses; but there is almost no possibility of ending a non-finite subordinate construction by anything but the verb.

First of all, there are subordinate clauses which do not need to be introduced by a conjunction and then regularly show verb-first order. These are conditional clauses and concessive and comparative subtypes of these (cf. Heidolph et al., 1981:718f). Second, if a subordinate clause S_2 is a constituent of a subordinate clause S_1, S_2 will normally follow S_1, so that S_1 does not have verb-final order. The same tends to be true for other heavy constituents of a subordinate clause, which more often than not follow the finite verb even in written German. Thus, (69)b. is a perfectly acceptable alternative to a.

(69) a. *Als er nun endlich nach all der schweren und langweiligen Arbeit Feierabend machte, da kam es ihm vor, als ob ...*
'When he then finally called it a day after all that hard and boring work, it appeared to him as if ...'

b. *Als er nun endlich Feierabend machte nach all der schweren und langweiligen Arbeit, da kam es ihm vor, als ob ...*

While these are cases of free variation which have to be conceived as subcases of the verb-final constraint, there is a third case of non-final position of the finite verb in subordinate clauses which is different from the first two mentioned. This is verb-second ordering, as in declarative main clauses. I am not here referring to asyndetic constructions of the type *ich glaube, er kommt* 'I

think he comes', whose subordinate nature is doubtful. The constructions which are of interest in the present connection are illustrated in (70) and (71).

(70) *Wir konnten den Wagen noch nicht reparieren, weil wir mußten auf die Ersatzteile warten.*
'We could not yet repair the car because we had to wait for the spare parts.'

(71) *Wir können auch noch wárten, obwohl morgen ist er bestimmt áuch nicht weiter.*
'We can wait a little more, although tomorrow he will surely not be any further along.'

Such constructions are not standard, but very common at the colloquial level. They have been in the language since Old High German; cf. Sandig, 1973. So far, they appear to be restricted to causal and concessive clauses. They are not to be interpreted as variations of subordinate clause constituent order. Instead, the subordinative conjunctions *weil* and *obwohl* (and some synonymous ones) are reinterpreted as coordinative conjunctions, on the analogy of *denn* and the adverb *allerdings*, respectively (which could be substituted in (70) and (71), with little or no change in meaning, to produce standard German). After them, main clause syntax is required. This means that subordinate clause order is renewed, at least in certain areas, by main clause order.

An independent clause has its own illocutionary force, whereas a subordinate clause generally has none of its own. Just as main clauses linked by *denn* and *allerdings*, verb-second clauses introduced by *weil* and *obwohl* are subduced from the scope of the would-be main predication and independently express the speaker's point of view. This is illustrated by the following minimal pair:[24]

(72) a. *Maria verehrt ihren Mann, obwohl er es nicht verdient hat.*
'Mary reveres her husband although he does not deserve it.'
b. *Maria verehrt ihren Mann, obwohl verdient hat er es nicht.*
'Mary reveres her husband, but he does not deserve it.'

The concessive clause in (72)a is within the scope both of the main clause illocutionary force and of the main clause predication (i.e. it is a part of the circumstances taken into account by the main clause subject). In (72)b., the concessive clause is outside these two scopes and gives a separate assessment of the situation by the speaker. It is this main clause status of the concessive

clause in b. which also renders possible the expression of emphasis by non-finite predicate fronting.

The main clause hypothesis also explains why the constituent ordering observed in the "subordinate" clauses of (70) and (71) never occurs when the subordinate clause precedes the main clause. Inversion of main and subordinate clause is possible in (72)a. and would be possible in (70) and (71) if the subordinate clauses had verb-final order, but not as they stand. The explanation is that a coordinative conjunction must go between what it coordinates.

5.2. Verb-first position in main clauses

The German constraint on verb-second position in declarative main clauses differs substantially from the situation in English. The English rule may essentially be conceived as a V-3 constraint which says that adverbial, subject and main verb must come in this order at the beginning of the clause. The German constraint presupposes that first and second position are fixed constituent structure slots and that the first can be filled by anything, and determines that the second must be filled by the finite verb.

There are at least two potentially dysfunctional aspects about such a constraint. First of all, it leads to a substantial amount of discontinuity in main clauses. The non-finite verb has to be separated from the finite verb, especially from the auxiliary; an adverb that is part of the verbal lexeme and appears as a verb prefix in subordinate clauses has to be separated from the verb in main clauses; and so forth. The phenomenon has been described by Germanists as *Satzklammer* ('[sentential] braces'; cf. Heidolph et al., 1981:707–709). We will see in a moment that this is not the interpretation given to it by speakers. Second, the constraint forbids clause-initial position of the finite verb, which is necessary if the clause consists only of a verb, and desirable, *inter alia*, in presentative sentences.

The second of these shortcomings is remedied by expletive *es* 'it' and *da* 'there', as illustrated in (73) and (74).

(73) a. *Es regnet.*
'It is raining.'
b. *Es wird getanzt.*
'There is dancing.'

(74) *Da kreuzt plötzlich dieser Holtermann auf.*
'Suddenly old Holtermann shows up.'

To understand the working of main constituent ordering, it is essential to be aware that, in such constructions, *es* and *da* are nothing but dummy slot holders which allow formal observance of the verb-second constraint. In the colloquial language, constructions with expletive *es* are not affected, although they are relatively rare ((73)b. would more commonly be *da wird getanzt*). However, the expletive *da* is increasingly dropped. Here are some recorded examples:

(75) A: *Das is kaputt das Auto.* — B: *Kann man nicht mehr mit fahren.* [1987]
'A: It is broken, the car.' — B [4 years]: 'You can no longer ride in it.'

(76) *Wenn es so weiter geht, können wir pünktlich ankommen. Kommen natürlich nachher noch Geschwindigkeitsbeschränkungen.* [conversation, 1987]
'If it goes on like this, we can arrive in time. There will be, of course, speed limits later on.'

From the point of view of the standard language, an initial *da* is missing in B's utterance in (75),[25] which would anaphorically resume 'the car'; and the second sentence in (76) likewise should begin with either of the expletive particles.

This construction is allowable in a variety of registers and text sorts. It is typical of narrative expositions, both poetic (*Sah ein Knab' ein Röslein stehn* [Goethe] 'a boy saw a little rose') and colloquial, especially in joke-telling (*Kommt eine Frau zum Arzt* ... 'A woman comes to the doctor's'). It is also no accident that one of the examples is from child language, as omission of sentence-initial *da* is very common there. Ten years ago, there used to be a genre of jokes called *Häschenwitze* 'bunny jokes', whose main character was a bunny-rabbit. The punch line was always reserved for him, and it consisted of a declarative main clause introduced by the finite verb. (77) is an example.

(77) *Hattu [hast du] also doch Möhrchen gehabt.*
'So you did have carrots, after all.'[26]

The omission of sentence-initial *da* has been generalized to some of the forms of the anaphorical pronoun *der*. In particular, *das* is as freely omitted as *da*.

(78) *kenn ich, weiß ich, war ich schon* [title of a German movie, 1980]
'I've seen it, I know it, I've been there'

(78) combines some common sayings with omission of both *das* and *da*. In examples such as (75) and (78), the missing form is not merely expletive, but is the sole legitimate representative of a sentence constituent. In such cases, initial position of the element in question is the necessary condition of its omission. Consequently, in the alternative versions (75)' and (78)', where another constituent takes the first position, the anaphoric form has to appear.

(75)'. Man kann nich mehr damit fahren.
(78)'. ich kenn das, ich weiß das, ich wár da schon.

The insensitivity of this process to grammatical relations shows that we are not faced with an instance of 'pro-drop'. Instead, it seems best to interpret the examples as the result of a tendency towards verb-initial order in main declarative clauses. This is also suggested by the sayings in (79), which are extremely common in the colloquial language.

(79) a. *Is[t] ja noch mal gút gegangen.*
'It has turned out well once again.'

b. *Wird schon nicht so schlimm wérden.*
'It won't turn out that badly.'

c. *Hast du eben Péch gehabt.*
'You just have had tough luck.'

d. *Ham [haben] wir ja noch níe gesehen.*
'That we have never seen before.'

e. *Könnte ja jéder kommen.*
'Then anybody could come.'

f. *Mußtu [mußt du] halt noch mál hingehen.*
'You just have to go there again.'

The missing initial pronominal form has the syntactic functions of subject in (79)a. and b., direct object in d. and place adverbial in the others.

Another point is noteworthy about the syntax of verb-initial sentences. While probably any verb can start a sentence of this type, cf. (78), the examples in (75), (77) and (79) are typical in beginning with auxiliaries or modals. It remains to be seen why this should be so.

In this way, both of the limitations of the verb-second constraint are overcome in the colloquial language. Not only can the sentence start with the verb; but the particularly disturbing discontinuities are also removed. The

full verb is relegated, in its non-finite form, to that part of the sentence where the syntagms most closely related to it are assembled. The only remaining discontinuity is the one between the auxiliary or modal and the full verb. However, this may well be merely an apparent one. There are so many languages which separate the expression of such grammatical categories as person, number, tense, mood and aspect from the expression of the lexical verb that it might be an Indo-European preconception to think that they should go together.

Incidentally, if verb-first order in declarative main clauses is colloquial and the replacement of verb-final order in subordinate clauses by main clause order is colloquial, too, one wonders whether it is not an anachronism to introduce verb-second order into subordinate clauses. In fact, it is exactly that hybrid between verb-first and verb-second order which is being used in so-called subordinate clauses in colloquial German. (80) would be an example.

(80) *Das sollten wir nicht tun, weil (da) hát ja keiner was von.*
'That we should not do, since nobody gets anything from it.'

If this development succeeds, German will become a verb-initial language. The constructions which are currently at the center of the movement remind one of the situation in the Celtic languages, where verb-initiality also means, in the first place, initial position of the auxiliary. In the present context, we have yet to consider in which respect the development is one of grammaticalization. Under closer analysis, there are two parallel ways of paying lip service to the verb-second constraint. One is to put an expletive element in the first position which then reduces to zero. What is grammaticalized here is the constraint of filling the preverbal position. The other way is to fill the second position with a grammatical verb, thereby allowing the lexical verb to come later in the clause. Here the constraint of filling the second position by a verb is depleted, grammaticalized. Grammatical rules are changed by obeying them.

6. GRAMMATICALIZATION IN CONTEMPORARY GERMAN

6.1. It is difficult to derive a typologically relevant generalization from the diverse phenomena reviewed above. There is grammaticalization both in the nominal and in the verbal sphere, both at the low and at the high grammatical levels. The advance of prepositions to the detriment of cases, and of modals, auxiliaries and function verbs to the detriment of synthetic

verb forms, bespeaks a general tendency towards analytic morphological structure. Prepositions are separated from the governed noun by the determiners and attributes of the latter. Grammatical verbs are separated from their non-finite or nominalized lexical verb by the rules of main constituent order. There is, therefore, at present no chance that the analytic constructions will coalesce to yield new synthetic ones.

6.2. At a universal level, the data conform to the rule that the greatest freedom reigns at the highest grammatical levels. Therefore, the changes that are potentially most far-reaching are those initiated at the highest levels. At its origin, a change in clause order or main constituent order is merely an instance of free variation and, thus, inconspicuous as regards the typological structure of the language. However, if and when the structures so created grammaticalize, the erstwhile variant will become the rule, and lower level grammar will be shaped in ways that are completely incompatible with its former state.

6.3. While many grammatical changes go on and probably originate at the colloquial level, colloquial speech is far from being the only, or even the principal, source of grammatical change. The mass media play an extremely important role in the coinage and diffusion of new collocations, which then go the way of lexicalization or grammaticalization. Grammatical change is, to some extent, register-bound in the sense that some of the changes are limited to the colloquial level, others to formal style. For some of the grammaticalizational phenomena observed, especially complex prepositions and function verbs, it is not the colloquial, but bureaucratic and journalistic style which promotes them. Some of these innovations may never reach the colloquial level.

6.4. Grammaticaliation manifests itself on the synchronic and diachronic axes. Insofar, phenomena may be, and have been, classed as cases of grammaticalization even if there is no clear evidence of historical change. In fact, the set of phenomena reviewed in the forgoing is quite heterogeneous as regards socio-historical aspects of language change. Some phenomena have appeared in the language only recently, such as the *mehr als* and *zunehmend*-constructions. Others have been well-established in non-standard or dialectal varieties of German for a long time, such as the progressive or the pseudo-subordinate clause with the verb in second position. Yet others have always been possible in the standard language in limited contexts, but have considerably expanded

their domain in recent times. This is true for the main clause with verb in first position.

Again, some of the changes, especially those relating to word order, lead to entirely new grammatical structures. Others, e.g. those producing complex prepositions, are but contemporary instances of a model that has been productive for a long time. In this paper, the common denominator of all the phenomena has been their grammatical character, centering around the concept of grammaticalization, and their recency in colloquial standard German. In all cases, there has been a shift in the pattern of variation which leads to greater prominence of a phenomenon in colloquial standard German.

6.5. The study of grammatical change on the basis of data on synchronic (or, at best, brachychronic) variation in the contemporary stage of a language is subject to a serious problem of verification. Given presently available methodological means, it is next to impossible to know which of the changes that speech habits currently exhibit are synchronic manifestations of ongoing genuine language change, and which of them are but ephemeral fashions. In this situation, a study such as this can hope to elucidate the synchronic dynamism of the language; it cannot hope to tell which innovations will result in changes and which ones will disappear without a trace in future synchronies.

NOTES

1. This paper has greatly benefitted from comments by Ulrike Claudi, Hartmut Haberland, Bernd Heine, Ekkehard König, John McWhorter, Joachim Mugdan, Thomas Stolz, Elizabeth Traugott and Joseph Voyles. Their contributions are acknowledged here with sincere thanks, but not singled out in the following text.

2. Cf. Lehmann 1982 for a systematic treatment.

3. As usual, the ongoing change creates uneasiness, manifesting itself, e.g., in hypercorrection. Thus, a club president addresses the 'dear members', in a circular of 1987, by *Liebe Mitglieder/-innen*, as if *Mitglied* were masculine or *Mitglied(er)innen* could be a grammatical form of German; and a report on women's fight sports features *Der Kung-Fu-Lehrer ... schätzt die Zahl seiner weiblichen Kursteilnehmerinnen auf ungefähr zwanzig Prozent.* [*Zeit-Magazin*, 1988] 'The Kung-Fu instructor estimates the number of his female course participantesses at about 20%.'

4. I even have one isolated example without article: *Die Botschaft Jesus* 'The Message of Jesus' (book title of 1972).

5. Cf. Plank 1980, esp. 296–298, on this matter. Hartmut Haberland draws my attention to the fact that the non-distinctness explanation is not sufficient to generalize to the genitive

singular, since while *das Wahrzeichen Oslos* 'Oslo's landmark' is fine, **die Förderung Erdöls* 'the hauling of petrol' is not, despite its distinctness.

6. One that comes to mind is *auf der Basis* 'on the base', which is parallel to *aufgrund*. However, this is not as recent, and, moreover, it does not yet seem to be a preposition.

7. All of my examples stem from independent discourses. They do not reflect any idiolectal idiosyncrasies or patterns evolved during a particular discourse.

8. In this paper, I am using 'phraseology' and related terms in the broad sense defended in, e.g. Pilz (1983). This includes relatively fixed collocations without regard to their semantic regularity, and also collocations of non-lexical elements.

9. The style of the translation mirrors the original one.

10. Seiler's (1973) interpretation of the periphrasis involving 'have' and 'be' auxiliaries as expressing the configuration 'possessor of an act' would seem to be a case in point.

11. Apart from the fact that adverbial *im Wege* would interfere with the idiom *im Wege sein* 'to be an obstacle'.

12. Given that *her* is not a postposition, the logical alternative of dropping the first part is unavailable.

13. Moreover, (30) shows the extent to which the new function of *von* tends to oust older ones. Ten years ago, this expression would have been interpreted exclusively, with stress on *raffiniert* instead of on *Farbe*, as implying that the color has achieved something sophisticated; cf. *Wie raffiniert von dir!* 'How clever of you!'.

14. Thus, it is not possible to say *wenn nicht so viel Salz drin wäre, möchte ich das wohl* 'if there were not so much salt in it, I would like it'.

15. I am told that it has existed in Pennsylvania Dutch since the 17th century. There it seems possible that a syntactic loan has met with support by a predisposition in certain German dialects.

16. I did not gather data on the progressive aspect because it is too familiar in my dialect environment.

17. This point has been clarified to me by Ulrike Claudi.

18. The surface structure would have to bring out the fact that the preposition and the article form a constituent.

19. Leaving the direct object unpromoted is not standard and uncommon at the nonstandard level.

20. Its narrowest locus is, of course, in constructions which are close to its etymological origin, such as *sie bekam ein Buch geschenkt* 'she was presented a book', etymologically 'she got a book presented'.

21. *Kriegen* and *bekommen*, in- or outside the periphrastic passive, are basically register variants of each other.

22 This is the headline of an extended commercial of an insurance company which also features a trustworthy-looking representative. The headline is not taken up in the commercial text, so that I cannot tell what it means. I take this to be a symptom of the depletion and the extremely phraseological character of *mehr als*.

23. V: verb; X: one constituent; Y, Z: remaining clause constituents.

24. (72) and the relevant observation are due to Ekkehard König. Cf. also König & van der Auwera, 1988.

25. Even then the sentence would be substandard (just as (80) below), because preposition stranding is not standard.

26. Bunny goes to the pharmacist's and asks: 'Do you have carrots?' The pharmacist says: 'No, bunny, this is a drugstore; we don't have any carrots.' Next day, bunny comes to the drugstore and again asks: 'Do you have carrots?', and the pharmacist again explains that he doesn't. Third day, the same story. On the fourth day, the pharmacist has put a sign on his door saying 'Today no carrots'. Bunny appears and reproachfully says to the pharmacist: (77).

REFERENCES

Burger, Harald, A. Buhofer, and A. Sialm, 1982. *Handbuch der Phraseologie*. Berlin & New York: W. de Gruyter.

Chafe, Wallace L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view." In *Subject and Topic,* Charles N. Li (ed.), 25–55. New York: Academic Press.

Heidolph, Karl Erich et al. 1981. *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag.

Helbig, Gerhard 1979. "Probleme der Beschreibung von Funktionsverben im Deutschen." *DaF* 16:273–285.

Hopper, Paul and Sandra A. Thompson, 1980. "Transitivity in grammar and discourse." *Language* 56:251–299.

Inghult, Göran 1975. *Die semantische Struktur desubstantivischer Bildungen auf* -mäßig. Stockholm: Almqvist & Wiksell.

König, Ekkehard and Johan van der Auwera, 1988. "Clause integration in German and Dutch conditionals, concessive conditionals, and concessives." In *Clause combining in Grammar and Discourse*, J. Haiman and S.A. Thompson (eds), 101–33. Amsterdam: John Benjamins (TLS, 18).

Lehmann, Christian 1982. *Thoughts on Grammaticalization: A Programmatic Sketch*, Vol. I. Köln: Institut für Sprachwissenschaft der Universität (AKUP, 48).

Lehmann, Christian 1985. "Grammaticalization: Synchronic variation and diachronic change." *Lingua e Stile* 20:303–318.

Lehmann, Christian 1989. "Grammatikalisierung und Lexikalisierung." *ZPSK* 42:11–19.

Matthews, Peter H. 1981. *Syntax*. Cambridge : Cambridge University Press.

Pilz, Klaus Dieter 1983. "Suche nach einem Oberbegriff der Phraseologie und Terminologie der Klassifikation." In *Phraseologie und ihre Aufgaben*, Josip Matešić (ed.) 194–213. Heidelberg: J. Groos.

Plank, Frans 1980. "Encoding grammatical relations: Acceptable and unacceptable non-distinctness." In *Historical Morphology,* Jacek Fisiak (ed.), 289–325. The Hague: Mouton.

Ross, John R. 1973. "The penthouse principle and the order of constituents." In *You Take the High Node and I'll Take the Low Node, Papers from the Comparative*

Syntax Festival. Corum, Claudia et al. (eds), 397–422. Chicago: Chicago Linguistic Society.

Sandig, Barbara 1973. "Zur historischen Kontinuität normativ diskriminierter syntaktischer Muster in spontaner Sprechsprache." *Deutsche Sprache* (3):37–57.

Seiler, Hansjakob 1973. "On the semantosyntactic configuration 'possessor of an act.'" In *Issues in Linguistics: Papers in Honor of Henry and Renée Kahane,* Braj B. Kachru et al. (eds), 836–53. Urbana: University of Illinois Press.

Seiler, Hansjakob 1983. *POSSESSION as an Operational Dimension of Language*. Tübingen: Narr [LUS 2].

Werner, Otmar 1979. "Kongruenz wird zu Diskontinuität im Deutschen." In *Festschrift for Oswald Szemerényi on the Occasion of his 65th Birthday*, Bela Brogyanyi (ed.), 959-88. Amsterdam: John Benjamins [CILT 11].

Wurzel, Wolfgang Ullrich 1984. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag [Studia grammatica, 21].

Volume II

Index of languages

A
Abipon 29
Abkhaz 21
Afrikaans 318
Akan 222
Akha 444
Akkadian 257, 259, 308
Altaic 61
Alyawarra 48
Am 79
Amharic 299, 307, 308
Anyi 222
Arabic 307
Aramaic 257, 259, 302, 308
Ashante *see* Twi
Atchin 29
Atsi (Zaiwa) 444
Austronesian language 228, 257
Awadhi 62

B
Bahing-Vayu 323
Baining 21
Balti 250
Bambara 213, 215, 219, 222
Bantu language 93
Bari 26
Basque 26
Bemba 307
Bengali 65, 79
Bhojpuri 62
Bodic 228, 238, 242, 248
Bongu 26
Boro 393
Braj 62
Bunan 229, 250
Buriat 26
Burmese 388, 390, 394, 402, 416, 444, 448

C
Cambodian 391
Canaanite 302
Cantonese 26, 29
Catalan 48
Cayuga 12, 161-183
Cebaara 207, 210, 211, 212, 213, 214, 221-222
Cebari 206, 207, 210, 211
Cebuano 105, 108, 111, 14, 115, 116, 117-118, 122
Celtic 530
Chewa *see* Chichewa
Chepang 250
Chichewa (Chewa) 131
Chinese 104, 115, 116-118, 129, 383, 386, 416, 420, 422, 425, 433, 434, 435, 437, 444, 445-446, 448, 450, 451
  Mandarin 82, 89, 386, 400, 425, 427, 430, 434, 438, 448-449, 450, 451
Chinook 238
Chrau 394, 448
Cocama 29

D
Danish 26, 29, 30, 348

Danjongka 250
Dogon 209, 217, 219, 222
Dolakhali 227, 251
Donno So 209
Dravidian 61
Dutch 331, 333, 337, 347-348, 349, 376
Dyirbal 94, 121, 122, 127, 128, 165
Dyula (Jula) 213

E
Engenni 29
Eskimo 121
  Yukip 128
Ethiopic 307

F
Fɔndɔndɔ 207
French 48, 135, 136, 138, 139, 142, 145, 146, 147, 148, 150, 151, 153-154, 257, 259, 262, 302, 348, 349, 445
Frisian 331, 337, 347, 348, 349, 376
Furlan 137, 138, 139, 140, 141, 142, 144, 145, 146, 147, 148, 149, 151, 152, 153

G
Garhwal 250
Gbe *see* Ewe
German 12, 102, 135, 136-137, 149, 152, 153, 257, 259, 302, 303, 331-377, 493-534
Gothic 358, 359, 365, 366, 367, 369
Greek 74, 197, 311, 359, 385
  Old 375-376
Gujarati (Gujerati) 65
Gurung 230, 250

H
Haka 26
Hayu 390
Hebrew 257-308
Hebrew, Biblical 12, 257-308
Hewa 261
Hindi 68, 79
Hindi-Urdu 60, 62, 63, 64, 65, 66, 68, 70, 72, 73, 74, 75, 76, 77, 78, 79, 80
Hittite 359, 375, 376
Hmong 386, 418, 420, 423, 437, 451
Hmong-Mien 444
Hungarian 499

I
Icelandic 135
Igbo 208
Ijo 219, 222
Indic, Old 376
Indo-Aryan 59-87, 319
Indo-European 197, 306, 376, 501
Indonesian 194, 198
Inuit 21
Italian 348, 515
  Northern 135-155

J
Jacaltec 122
Japanese 12, 93, 94, 95, 96, 98, 99, 100, 104, 105, 106, 107, 108-109, 111, 115, 116, 117, 118, 119, 120, 128, 129, 306, 385, 432, 448-449, 450, 451
Jimini 207, 210, 214
Jirel 250
Jula *see* Dyula

K
Kadugli 49
Kagate 250
Kalapuya 193, 198
Kalasha 81
Kalkatungu 127
Kanauri 250
Kannauji 62
Kanuri 21
Kapampangan 115
Karaboro 206, 207, 208, 210, 214, 218
Karenic 450
Karok 159
Kashmiri 65, 75, 76, 78, 81
Kauravi 62
Khaling 250
Khambu 230, 250
Khmer 386, 391, 399, 425, 427, 429, 449
Kinyarwanda 194, 198

Klamath 12, 185-198
Korean 95, 115, 116, 118, 129
Krio 262
Kru 202, 205, 216, 220, 222

L
Ladakhi 250
Ladin 140, 148
Lahu 383-451
Lahul 250
Lakher 393
Latin 49, 150, 153, 161, 197, 359, 367, 375, 376, 445, 494, 500
Lhoke 250
Lhomi 250
Limbu 250

M
Maguindanao 109
Mam 122
Mamara 207, 221
Mandarin *see* Chinese
Maninka 216
Mano 29, 30
Marathi 60, 64, 65, 66, 67, 68, 70, 71-74, 75, 76, 77, 78, 79, 80
Margi 29, 30
Marwari 81
Masai *see* Maasai
Maung 29
Melanau 109
Melanesian Pidgin 7
Mien *see* Yao (Mien) Samsao
Mikir 393
Mon Khmer 121
Mpi 444
Mwera 29

N
Nafaara 206, 207, 212
Nanɛrghɛ 207, 214, 221
Nepali 238
Nez Perce 185-198
Newari 12, 227-252
Niger-Congo 209, 220
Nimboran 26
Norwegian 348, 349
Nung 29
Nyamkat 250

P
Papuan 228
Pali 78
Panjabi 65
Papago 47
Penutian 196
Philipine 12, 94, 105, 106-107, 109, 111, 114, 116, 117, 118, 127, 128
Phoenician 302
Pilara 207, 212
Polish 446
Proto-Niger-Congo 221
Proto-Senufo 206, 217, 218
Purki 250

Q
Quiche 122, 124, 126, 127, 130

R
Rama 6, 12, 455-491
Rhaeto-Romance 148, 149, 151
Romagnol 137, 138, 141, 142, 144, 146
Romantsch 143, 145, 147, 148, 151
Rumanian 347
Russian 63, 64, 78, 79, 161, 197, 446

S
Sahaptin 12, 185-198
Sanskrit 60, 78, 197-198
Selayarese 12, 162, 183
Sɛmɛ 202
Senari 207, 208-209
Seneca 318
Senufo 201-222
Shenara 207, 212, 214, 222
Shɛnpire 207, 212
Sherpa 250, 322
Shina 59, 65
Shina-Gurez 75, 76, 78, 81
Siamese 442

Slave 26
Slavic 446
Sliammon 197
Songay 201
Soninke 216
Spanish 48, 257, 259, 302, 445
Sucite 202, 207, 208-209, 212, 221
Sunwari 250
Supyire 206-209, 211, 212, 213, 214, 217, 218, 221, 222
Surselvan 148, 151, 152, 153
Swedish 349

T
Tagalog 105, 108, 114, 115, 117-118
Tagbana 207, 214, 221
Tahitian 21
Tai 413, 442, 448
Tamang 251, 440
Tamil 78
Tanga 27
Tem 29, 31
Thai 386, 391, 392, 393, 398, 399, 406, 416, 419, 420-421, 422, 423, 424, 425, 427, 429, 433, 436, 437, 438, 439, 443, 444, 445, 448-449, 451
Thakali 231, 251
Thulung 230, 250
Tibetan 238, 257, 390
Tibetan, classical, 229, 250
Tibetan, Lhasa 231, 250, 260
Tibeto-Burman 83, 121, 227, 228, 231, 233-234, 240, 251, 386
Tɔgɔ Kã 209-210
Tok Pisin 26
Tojolabal 29, 30
Toussian 201
Trukese 27
Tucano 29
Turkic, Orkhon 78
Turkish 257
Twi 197
Tyurama 208

U
Umbrian 49
Urdu *see* Hindi-Urdu
Usan 249
Uto-Aztecan 257

V
Vayu 229, 250
Vedic 376
Vietnamese 386, 391, 417, 420, 422, 425, 427, 429, 434, 444, 445, 449, 450, 451

W
Wara 201, 209

Y
Yao 387, 394, 427, 445, 448, 451
Yao (Mien) Samsao 417, 420, 425, 428, 438, 439
Yiddish 446
Yidiny 122, 127
Yukip *see* Eskimo

Z
Zaiwa *see* Atsi
Zeneyze 137
Zuni 29

Volume II

Index of names

A

Abraham, W. 2, 12, 332, 333, 337, 343, 347, 348, 357, 374, 375
Akiba, K. 306
Allen, N.J. 230, 250
Alleton, V. 449
Alphonse, 49
Anderson, J. 122, 160, 165, 231, 238
Aoki, H. 185, 190, 191, 193, 194, 195, 196, 197, 198
Appollonio, B. 136, 141, 144
Arquint, J. 151
Askedal, J.O. 376
Assadi, B. 488
Aston, W.G. 97
Austin, P. 228

B

Babby, L.H. 79
Bacot, J. 229, 250
Bahl, K.C. 81
Bally, C. 138
Barannikov, 79
Barker, M.A.R. 190, 193-194, 198
Bashir, E. 81
Bavin, E. 326
Becker-Donner, U. 30
Beer, G. 308
Bendor-Samuel, J. 220
Benedict, P.K. 388
Benes, E. 136
Benincà, P. 136, 137, 147
Benjamins, C. 488
Benveniste, E. 322
Bickerton, D. 8, 10
Bickner, R.J. 451
Bieri, D. 250
Bird, C. 205, 215, 216
Blake, G.J. 93, 94, 120
Bloomfield, L. 95, 130
Bole-Richard, R. 221, 222
Bolinger, D. 326, 327
Bopp, F. 2
Bosson, G. 160
Bouchard, M. 145
Boutin, P. 207
Bradshaw, J. 85
Brandi, P. 136, 138, 139, 142, 144, 145
Bréal, M. 383, 384, 444
Breivik, L. 136
Bresnan, J. 122, 130, 131
Bright, W. 160, 185
Brown, S.R. 98
Browne, W. 136, 137, 140, 143
Brugman, C. 367
Burger, H. 503, 518
Burling, R. 444
Bybee, J.H. 3, 9, 11, 17, 18, 21, 22, 24, 26, 28, 30, 32-33, 34, 45, 48, 49, 318, 326, 373

C

Campbell, L. 2, 6, 11
Carlson, R. 202, 221, 248
Cauvin, 207
Chafe, W. 94, 96, 100, 173, 317, 318, 326

Chao, Y.R. 104
Chirmule, S.V. 69, 70, 72, 79
Chomsky, N. 347
Chung, S. 22, 198
Cissé, N. 207
Clamens, G. 207, 221
Clancy, P. 100, 120
Clark, M. 450
Claudi, U. 11, 220, 475, 532, 533
Clayre, I. 109
Coates, J. 19, 26, 31
Comrie, B. 79, 95, 104
Cooreman, A. 118, 128
Conzemius, E. 458, 490
Cordin, L. 136, 138, 139, 142, 144, 145
Craig, C. 2, 5, 12, 122, 217, 228, 248, 460, 461, 463, 464, 481
Croft, W. 6

D
Dahl, Ö. 17, 19, 21, 33, 34, 45, 48
Darmesteter, A. 136
Davis, J. 197
DeLancey, S. 197, 198, 234, 239, 248, 250, 251
Dembele, E. 207
Deshpande, M. 78, 80
Devitt, D. 47
Diderichsen, P. 349
Diehl, L. 231, 239, 251-2
Dixon, R.M.W. 95, 121, 122, 127, 128, 160, 165
Downing, P. 100, 120
Dowty, D. 69
DuBois, J. 3, 128, 314, 318, 325, 326
Durie, M. 314

E
Elsness, J. 314
Elwert, W.T. 136, 137, 143
Emonds, J. 314
England, N. 122
Ettmayer, K. von 150

F
Faltz, L.M. 314
Fanselow, G. 345, 375
Fleischman, S. 17, 30, 326
Forsyth, J. 63
Fox, B. 118
Frajzyngier, Z. 448
Franck, D. 350
Francke, A.H. 250
Frau, G. 149

G
Gabelentz, G. von der 2
Garber, A. 207, 208, 221
Garret, A. 251
Genetti, C. 6, 12, 228, 232, 234-8,248, 251, 471
Gerhardt, J. 326
Givón, T. 2, 3, 5, 8, 11, 12, 30, 59, 93, 118, 120, 138, 148, 175, 195, 197, 202, 216, 248, 250, 257-308, 314, 384
Glover, W. 230, 250
Goldstein, M.C. 231, 250
Gornik-Gerhardt, H. 350
Gramley, S. 327
Greenberg, J.H. 2, 5, 6-7, 8, 11, 445
Gregor, D.B. 136, 137, 141, 142, 144, 146
Grierson, G.A. 79
Grimm, J. 365, 367
Grimm, W. 365, 367
Gruber, J. 231

H
Haberland, H. 532
Hagège, C. 449
Haiman, J. 12, 135, 136, 145, 149, 250, 327
Hale, A. 234, 252
Hale, K. 460, 461, 463, 464-481, 488
Hargreaves, D. 234, 240, 250, 251, 252, 326
Hari, A.M. 231, 251
Harris, M. 136
Haspelmath, M. 47
Hawkins, J. 306
Heidolph, K.E. 500, 518, 525, 527
Heine, B. 1, 2, 4, 6, 7, 11, 30, 47, 48,

94, 130, 194, 202, 205, 216, 219, 248, 249, 250, 323, 326, 337, 475, 532
Helbig, G. 333, 353, 354, 518
Helling, C. 348
Helty, G. 148
Hendriksen, H. 60
Hentschel, E. 350, 353, 355, 357, 358, 359, 360, 361-362, 364, 365, 366, 367, 368, 375
Herring, S.C. 3, 5, 11
Hetzron, R. 305, 308
Hirano, T. 130
Hock, H.H. 373
Höhle, T.N. 375
Hook, P. 3, 9, 11, 27, 62, 79, 198, 319
Hooper, J.B. (*see also* Bybee, J.H.) 314
Hopper, P. 1, 3, 4, 5, 8, 11, 12, 74, 203, 205, 221, 248, 314, 318, 320, 324, 325, 326, 514
Huffman, F.E. 425, 430, 449
Humboldt, W. von 2
Hünnemeyer, F. 11, 475
Hyman, L. 202

I

Ikola, O. 289, 290, 296
Iliescu, M. 144, 148
Inghult, G. 511
Inoue, K. 130
Isacenko, A.V. 79

J

Jacobs, J. 346
Jacobs, M. 190, 193, 197
Jäschke, H.A. 250
Jelinek, E. 162
Jespersen, O. 94, 95
Jørgensen, H. 232, 234, 236-237, 239, 244, 245, 247, 250, 251, 252
Johnson, C.D. 326
Jordan, D. 206, 207
Joshi, P.T.R. 250

K

Kapp, D.B. 79
Kaul, 80
Kayne, R. 147
Kendall, M. 39, 40, 43, 45, 46, 205, 215, 216
Keenan, E.L. 95, 104
Keesing, R. 2, 7, 11, 54, 71, 78, 308, 327
Kemme, H.M. 375
Kemmer, S. 6
Kettunen, L. 289, 290
Kibrik, A.E. 130
Kimenyi, A. 198
Kiparsky, P. 145
Kirsner, R. 318
Konow, S. 227, 229, 230, 250
Koshal, S. 250
König, E. 2, 4, 5, 7, 11, 249, 373, 532
Kötz, W. 333, 353, 354
Krivonossov, A.T. 374
Kuen, H. 136, 140
Kuno, S. 96, 122
Kuryłowicz, J. 2, 305-306, 337, 371
Kuroda, S.-Y. 130

L

Labov, W. 4
Lakoff, G. 80
Lambrecht, K. 139, 140, 142, 145, 146, 147, 150
Langacker, R. 7, 9, 318
Larsen, T.W. 122, 130
Laughren, M. 206, 207
Lazard, G. 130
Lee, E.W. 109
Lehmann, C. 1, 2, 4, 6, 7, 17, 20, 21, 25, 28, 77, 248, 337, 372, 501, 509, 511-512, 523, 532
Lehmann, W. 306, 458, 490
Leino, P. 289, 290, 296
Lejeune, 207
Li, C. 3, 93, 130, 160, 308, 450, 451
Li, Fang Kuei, 425, 448
Lichtenberk, F. 3, 4, 11, 12, 228, 248
Lightfoot, D. 293, 296, 306
Linder, K. 139, 143, 151
Lord, C. 59, 216

Luraghi, S. 359, 375, 376
Lütten, J. 350
Lyons, J. 22

M
Madou Kone 207
Magier, D. 81
Maibaum, A. 231, 250
Malla, K.P. 232, 237
Mallinson, G. 93, 94, 120
Malone, J. 305
Marchese, L. 30, 202, 205, 220, 222, 318
Marchetti, G. 136, 137, 139, 146
Margarit, B. 48
Matisoff, J.A. 2,7, 12, 385, 388, 393, 402, 403, 410, 427, 433, 440, 441, 443, 444-448, 449, 450, 451
Matsumoto, Y. 6
Matthews, P.H. 520
Mayr, E. 306
Mazaudon, M. 260, 401
McDavid, V. 314
McFarland, C. 118, 130
Mchombo, S.A. 122, 130, 131
McWhorter, 532
Means, M. 207
Meillet, A. 2, 138, 147, 317, 337, 371
Meinhof, C. 138
Michailovsky, B. 229, 250, 390
Michitoshi, Toganoi 96
Mills, E. 207, 222
Mirikitani, L. 115
Mithun, M. 3, 9, 12, 130, 162
Miyaoka, O. 128
Morrison, H. 191
Moser, H. 366
Mottin, J. 421
Mugdan, J. 532
Muktibodh, G.M. 80
Mulac, A. 3, 12, 326
Myhill, J. 78

N
Nariakira, Fujitani 96
Nay, S. 151
Naylor, P.B. 118
Neethivanan, J. 250
Nespital, H. 79
Nichols, J. 3, 4, 11
Nigam, R.C. 250
Norman, M. 122, 130
Norwood, S. 488
Noske, M. 205, 217
Noss, R.B. 394, 399, 450, 451
Nyholm, L. 376

O
O'Neil, W. 488
Otanes, F. 116

P
Pagliuca, W. 9, 11, 17, 18, 19, 21, 22, 26, 28, 30, 33, 45, 48, 49
Palmer, F.R. 23, 24, 313
Panet, E. 348
Paul, H. 366
Perkins, R.D. 11, 30, 45, 47, 49
Perlmutter, D. 135, 144
Phinney, A. 190, 191, 193, 196
Phuntsok, T.L. 250
Pilz, K.D. 533
Pizzinini, A. 136, 140, 143
Plangg, G. 136, 140, 143
Plank, R. 532
Porizka, V. 79
Prost, A. 207, 208, 209, 217

R
Ramchand, G. 130, 488
Rangan, K. 250
Redeker, G. 326
Read, A.F.C. 250
Reesink, G. 228, 248, 249
Reh, M. 1, 6, 18, 24, 30, 194, 202, 205, 216, 219, 323, 337
Renzi, L. 136
Rickford, J. 10
Rigby, N. 488
Rigsby, B. 197
Rizzi, L. 144, 145

Roerich, de G. 250
Rohlfs, G. 136, 141, 143, 145
Romaine, S. 12
Rosen, H. 302
Ross, J.R. 524
Rude, N. 2, 12, 185, 186, 195, 196, 198

S
Salamanca, D. 488
Salmon, N. 326
Sandig, B. 495, 526
Sankoff, G. 10, 327
Sansom, G. 96, 98
Sapir, E. 42
Saxena, A. 305, 307
Schachter, P. 112, 114, 116, 120, 122
Scheitlin, W. 151
Scherpenisse, W. 375
Schophaus, R. 370
Schottelndreyer, B. 250
Schröbler, I. 366
Schulze, M. 250
Schwartz, A. 326
Searle, J.R. 154
Segert, S. 305
Seidl, K.O. 370
Seiler, H. 505, 517, 533
Shepherd, S.C. 28
Shibatani, M. 3, 12, 106, 112, 116, 118, 130
Shibles, W. 350, 358
Shopen, T. 326
Silverstein, M. 94, 120
Slobin, D.I. 12, 326
Solnit, D.B. 450
Spiess, F. 136, 138, 148
Srinuan, D. 444
Stern, Th. 210, 211
Stolz, T. 532
Strahm, E. 250
Subba, B. 250
Svorou, S. 47, 48
Sweetser, E.E. 8, 59, 77, 80

T
Talmy, L. 8, 80
Taylor, D. 251
Tekin, 61
Thoeni, G. 151
Thomas, D.D. 394
Thomason, S. 368
Thompson, L.C. 449
Thompson, R.M. 326
Thompson, S. 3, 5, 12, 47, 93, 130, 160, 250, 314, 318, 326, 450, 514
Thurgood 228, 248
Thurneysen, R. 136
Tibbitts, B. 488, 489
Tikkanen, B. 61
Timberlake, A. 3, 4, 22, 79, 488
Toba, S. 250
Tomaselli, A. 360
Traugott, E.C. 1, 4, 5, 7, 11, 30, 47, 70, 77, 94, 130, 248-249, 250, 299-300, 305, 313, 324, 326, 532
Tsunoda, T. 128, 130

U
Ultan, R. 17, 18, 30
Underhill, R. 314, 315
Urmson, J.O. 322

V
Vaid, K.B. 60, 69, 70-71, 79
Vajracarya, 232
Van der Auwera, J. 533
Van Valin, R.D. 128
Vanelli, L. 136, 139, 142
Vattuone, B. 136, 137, 140, 143
Vendler, Z. 69
Vennemann, T. 306, 308
Vesalainen, M. 250
Vesalainen, O. 250
Volrath, P. 261
Voyles, J. 532

W
Warotamasikkhadit, U. 433
Weidert, A. 250
Welch, 207
Wells, C.J. 364, 368, 371, 373
Welmers, W. 208, 209, 222

Werner, O. 499
Weydt, H. 350
Wichser, 207
Wiegan, N. 5, 300
Wierzbicka, A. 94, 120
Willett, T. 48
Williamson, K. 219, 221, 222
Wolski, W. 350
Woodbury, A.C. 322
Woodworth, N. 47
Wurzel, W.U. 496, 512

Y
Yallop, C. 48
Yamada, Y. 98, 99
Young, P. 49

Z
Zbavitel, D. 79
Zvelebil, 61

Volume II

Index of Subjects

A

absolutive 160, 163, 164, 167-168, 169, 171-173, 175, 180, 182
absolut(iv)e case 94, 120, 121, 124, 126, 127, 128, 129, 165
abstractification 402
abstraction 373, 413
abstractness 4, 231, 246, 251, 384, 387, 401, 409
accusative language 104, 119, 174, 175, 181, 182
accusative pattern 118
action 240, 246
active case system 161
actor 260
adjunct 340, 349, 514
adjective 523
adjustment 512
adverb
  manner 36, 269
  negative 388
  purpose 270
  temporal 19, 30, 32
affixation 481
afterthought 288
age, semantic 25, 31, 39, 44, 45
agency 165, 175
agent, noun *see* noun
agreement 31, 135, 142, 143, 144, 147, 149, 155, 162, 185, 186, 187, 192, 195, 356, 445
allative 30, 31, 48, 229, 240, 251, 472, 505
allomorphy 18, 35, 36, 37, 38, 500
allophone 237
alphabet, Devanagari 251
ambiguity 246, 408, 415, 426
  case 499
analogical extension, *see* extension
analogical pressure 503
analogy 7, 155, 246, 385, 483, 526
anaphora 348
  causal 365
  temporal 365
  text-deictic 370
  zero 463, 464, 481
andative 515
animacy 245
anterior 21, 79, 274, 307
antipassive 121, 124, 126, 127
apposition 498
appositive 162, 163
approach, functional-typological 487
argument 162, 180, 182, 465
arrangement, linear 249
article (definite) 497, 503, 513, 515
  indefinite 497
aspectogenesis 65-74
associative 191, 192, 215, 216, 233, 236, 239, 247, 252
automatic pattern 159
autonomy 37
  grammatical 275
auxiliarization 203
auxiliary (verb) 18, 25, 30, 47, 59, 61, 66, 67, 72, 76, 202, 203, 207, 208, 211,

215, 216, 219, 220, 221, 222, 240, 384, 386, 394, 402, 416, 417, 420, 422, 423, 424, 426, 430, 448, 449, 455, 475, 478, 479, 480, 484, 485, 486, 512, 514, 516, 517, 527, 529, 530, 533
completive 215, 216
incompletive 216
modal 76, 425

B

backgrounding 176, 177, 181
backness 237
base
long 234, 236
short 234
basic vocabulary, *see* vocabulary
belief-state 249
benefactive (beneficiary) 186, 187, 196, 197, 205, 211, 213-214, 215, 216, 218, 221, 222, 396, 427, 429
bleaching (semantic) 4, 5, 59, 61, 77, 220, 318, 384, 413, 467
blend, syntactic 287
borrowing 154, 303, 332, 336, 350, 357, 371, 448
bondedness
boundary
morpheme 442
syllable 442
case inflection 6, 487, 499
marker 6, 118, 497
160, 162, 163, 181, 205, 216, 217, 218, 308
relation 245, 387

C

category
grammatical 11, 159, 160, 318, 324, 331, 346, 372, 375
paradigmatic 63
semantic 79, 318
category shift 318
causal relationship 236
causativizer 75
cause, 241, 246
center, deictic 442, 448, 515
chain
clause 240, 250, 251
grammaticalization 455-458, 460, 468, 474, 479-482, 484, 486, 488
chaining 456
change
analogical 291, 293, 494
consonant 35
grammatical 248, 249, 298, 494
historical 136
language 3, 4, 7, 532
lexical 494
meaning 5
metaphorical 4
phonological 196
sound 251
semantic 17, 24, 77, 248, 369, 373
syntactic 242, 286, 304
vowel 35
word order 202, 524
channel, grammaticalization 9
chronology 304
class
closed 48
open 36, 48
classifier 386
numeral, *see* numeral classifier
clause
adverbial 81, 258, 267, 268, 269, 271, 281, 299
causal 526
chain 240, 251
complement 168, 169, 170, 273, 313, 315, 316, 326
complex 171
concessive, 283, 526
conditional 300, 525
coordinate 111, 112, 166, 250
copular 219
dependent 370
embedded 203, 219, 314, 315, 360, 362, 366
equative 219
interrogative 364

main 6, 12, 241, 314, 315, 317, 337, 457, 527
matrix 168, 169, 170, 172, 173, 250
nominalized 107, 108, 109, 167, 202, 242, 244, 245, 246, 257, 260
order 531
principal 135
purpose 238, 239, 248, 268, 281, 282, 283, 286, 292, 293, 295, 448
relative, *see* relative clause
subordinate 6, 19, 28, 107, 109, 111, 112, 122, 127, 241, 245, 251, 457, 469, 471, 472, 483, 522, 524, 525, 527, 530
target 322
type 524
cleft construction 122, 123, 124, 126
clefting 122, 126, 127, 128, 131
cline 5
clitic 135-155, 488
cliticization 241, 246, 461, 463, 464, 465, 468, 481, 482, 483
co-occurrence restriction 25
co-verb 450
codification 173
coding condition 251
coding density 288
coding efficiency, *see* efficiency
cognate 193, 205, 214, 217, 221, 308
cognitive factor 159
cohesion 101
collocation 318
colloquial (language) 495, 498, 500, 502, 512, 513, 526, 528, 529
comitative 434, 461, 466, 467, 468
command 164, 165
comment 173
complement
infinitival 512
preposing 314
complementizer 141, 273-275, 286, 369, 398, 399, 425, 437
causative 428-431
modal, 425
passive 431
quotative 399
completeness 475
completive (aspect) 215, 216, 476
complexity, semantic 337, 372
component, meaning 18, 20, 32
compound 260
compounding 236
concatenation 195, 196, 233, 236, 238, 385, 402, 403, 404, 405-408, 410, 412, 413, 415, 435, 440, 444, 448
verbal 12
concept, abstract 9
conceptualization 62, 77
concessive 5, 236, 249, 271, 293
conditional 237, 249, 271, 272, 300, 306, 364, 469, 470, 472
conditional, negative 240, 241
conjunction
coordinating 526, 527
subordinating 398, 402, 457, 489, 526
constituent, *see* dependent constituent
dislocated 139
construction, coordinate 145, 165
contact 301, 302, 308
context 4, 8, 28, 74, 116, 174, 179, 363, 503
clausal 77
discourse 11
semantic 9, 505
syntactic 170
continuous 21
continuum 10, 29, 47, 251, 386, 407, 413, 433, 435, 481, 482
developmental 62
dialect 271, 276, 277, 281, 283, 284-287, 289, 292, 303
semantic 292
syntagmatic 442
universal 301
contrast 179
semantic 62
paradigmatic 65
contrastive 237, 238
control 165, 179, 180, 181, 182

conventionalization 11
co-occurrence 66, 72, 73
coordination 164-166
copula 98, 100, 201, 209, 210, 219, 221, 222, 346, 392
  equative 215
  locative 208, 209, 216, 264
coreference 128, 172, 292
coreferent 166, 168, 169
counterfactual 238, 470
co-verb 524
creativity 10
cycle (of grammaticalization) 114, 153, 228

D

dative 187, 191, 194, 195, 197, 205, 210, 211, 214, 215, 218, 221, 222, 229, 233-234, 238, 240, 242, 247, 248, 251, 252, 268, 271, 472, 498, 500, 501, 503, 517
definite *see* article
definiteness 6
degrammaticalization 445, 524
deixis, *see* center
  motion 515
  social 81
  text 370
deletion 145, 168, 169, 237
deontic modality 333
dependence 34, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46
depletion 508
derivation 10, 176, 180, 511
desemanticization 4, 80, 359, 523
determiner 531
  definite 511
development, diachronic 25
deviation 511
diachrony 3, 11, 237, 298, 350, 374
dialect coninuum, *see* continuum
diffusion 301
diminutive 192
direct quote 240, 241
directional 30, 48, 187, 240
disambiguation 499
discourse
  category 173
  dependence 481
  function 252
  marker 489
  perspective 2
  role 100, 176, 181
  strategy 3
dislocation, left 142, 489
distribution, complementary 151
divergence 221, 324
domain, abstract 246
double-subject 103
drift
  diachronic 258, 301
  language-internal 10
  semantic 467

E

efficacy, scale of 388
embedding 485, 489
emphatic marker 238, 242
emptification 383
enclitic 162, 167, 171, 243
encliticization 216
epistemic
  modality 333
  morpheme 315
  verb 327
  phrase 12, 313-327
epistemicity 323
  marker 322
ergative language 93, 116, 121, 128, 129
ergativity 122
etymology 204, 205, 206, 212, 216, 222, 237, 238, 400, 517
evacuation 383
event 77, 107, 108, 109, 175, 177, 191
evidential 48
evidentiality 387
evidentiality marker 322
evolution 6, 21, 47, 298, 306, 457, 469, 471, 472, 481, 482, 483, 484, 485, 486
expansion 6, 496, 517, 521, 522
  analogical 497

explanation 63, 64, 66, 94, 216, 218, 219, 252, 279, 298, 301, 306, 496, 504, 514, 527
  teleological 9
expressivity 77
extension 150, 196, 229, 246, 246, 252, 306, 375, 521
  analogical 150, 198, 287, 288, 298
  diachronic-semantic 301
  grammatical 251
  metaphorical 80, 505

F

feminine (gender) 161, 172, 174, 302, 392, 496, 497, 499
focus (marking) 105, 106, 109, 112, 114, 124, 126, 209
force, illocutionary/illocutive 333, 339, 350, 353, 357, 359, 373, 526
foregrounding 176, 177, 179, 278, 281
fossilization 500
frequency
  discourse 314
  of use 451
  text 65, 66, 76, 79, 319
frequentative 222
functional, *see* ideational, interpersonal, textual
  semantic 517
  syntactic 517
functor 366, 384, 445
fusion 34, 36-42, 44-46
future
  aspectual 21, 22, 32, 45
  immediate 32
  marking 11
  modality-derived 20
  movement-derived 20, 31

G

gender 143, 163, 495
  grammatical 155
  marker 496
  marking 499
  natural 155
generalization 7, 385
generic *you* 327
genesis 11, 153
genitivizer 393
given information, *see* information
goal (function) 19, 30, 105, 107, 116, 117, 118, 120, 123, 128, 187, 211, 215, 221, 231, 457, 460, 469, 470, 479, 480, 483, 490
grammar, emergence of 314, 326
grammaticalization
  definition of 1-3, 247, 251, 317, 318, 337, 383, 384, 493
  degrees of, 33, 36, 37, 38, 39
  chain, *see* chain
  predisposition for, 159
  process of, 81, 159, 474, 479, 485
grammaticization, *see* grammaticalization
grammatization, *see* grammaticalization

H

habitual 20, 21, 234, 263, 265, 276, 306
head 172, 173
  indefinite 265
hierarchy
  egodeictic 231
  implication 5
homonymy 338, 344, 359, 456
homophony 216, 238
honorification 81
hortative 23
hybrid
  linguistic 5
  mood, *see* mood

I

iconicity 218, 327, 504
idiom, frozen 8
idiomatization 9, 385, 392
illocutionary force, *see* force
imperative 23, 28, 29, 32, 112, 122, 164, 477, 480, 485, 490
incorporation 467
indefinite 172, 175, 179, 265, 498

infinitive 80, 234, 238, 239, 247, 248, 251, 252, 266, 513
inflection 24, 144, 234, 236, 237, 238, 252
information
  given/old 100, 101
  new 175, 176
information processing 11
innovation 11, 531
instrument 194, 218, 240, 249, 260, 267, 434, 465
intensification 520, 523
intensifier 521, 522
interrogative pronoun, *see* pronoun
intonation 322
intransitive 189
irrealis 234, 268, 292, 300, 306, 477
isolation, process of 97

L
laxing 237
layering 203, 248, 325
lexicalization 7, 8, 324, 461, 463, 464, 466, 467, 468, 482, 503, 515, 531
load, functional 288
loan (word) 405, 406, 437, 451, 496
loan translation 522, 523
localist theory 231
location 231, 239
locative 205, 211, 215, 216, 218, 229, 233, 234, 237, 238, 242, 251, 259, 260, 264, 271, 272, 461, 470, 472, 503
locative head 173
logical sphere 251
loss 501, 513

M
manner 240, 241, 264, 269, 270, 299
masculine (gender) 161, 495, 498, 499
meaning, pragmatic 5
meaning component, *see* component
merger 195, 196
metaphor, *see* extension, vehicle
metaphor 11, 249, 383, 384, 392, 502
metonymy, 4, 11
modal particle, *see* particle
modality
  agent-oriented 19, 21, 22, 23, 24, 25, 26, 27, 31, 33, 47
  deontic 23, 24, 25
  epistemic 21, 23, 24, 25, 27, 31, 33
  speaker-oriented 23, 24, 27, 33
mood 457
morphanization 385
morphologization 4
motivation, cognitive 10
motivation for grammaticalization 9, 159-183

N
narrative 100, 109, 178, 240, 278, 528
narrative text 458, 463
nasalization 221
negation 262, 349
negator 366
neuter (gender) 161, 497, 500
new information, *see* information
nominalization 107, 167, 168, 227, 228, 237, 246, 262, 266, 267, 268, 269, 518
  clausal 471
nominalizer 236, 244, 245, 246, 252, 391, 393, 444, 451
nominative 160, 163, 306
non-final clause 239, 251
non-final marker 241, 247, 251, 252
non-past 31, 234, 247, 252
noun
  body part 6
  class system 10, 201
  classifier 260, 262
  head 201, 205, 245, 262, 265
  relational 501, 504, 524
  verbal 517, 518
number 163, 530

O
object
  dative 217
  direct 166, 201, 202, 204, 217, 219, 220, 222

locative 217, 219, 222
oblique 204, 219
operation, mental 322
optative 23
optionality 478
order, *see* word order
constituent 176, 524, 528, 531
orientation, spatial 159, 175
origin 61, 217, 468, 474, 475, 531
lexical 10, 251, 490
nominal 235
verbal 206, 209, 216, 473

P

packaging, temporal 11
panchrony 3, 11
paradigm
closed 336
grammatical 9
morphological 10
paradigmatization 512
parataxis 288
parenthetical, epistemic 313-327
participant 161, 167, 173, 174, 176, 177, 179, 182
new 177
participation 191
participial 265, 270
participle 234, 239, 241, 251, 252, 349, 505
conjunctive 72
particle
definition of 386
epistemic 326
modal 12, 331-377, 393
nominal 12
verbal 12
particularization 387-402
passive 95, 102, 105, 106, 117, 176, 425, 449, 516, 517, 519
adversative 425, 431
agent 511
passivization 106, 180
passivizer 423
past tense 18, 19, 33, 36, 247, 252
pattern 236
perfective 18, 33, 59-87, 260, 265, 274, 277, 278, 281, 325, 388
periphrasis 23, 24, 503, 504, 512, 513, 519
periphrastic (expression) 33, 34, 48, 318, 503
permissive 23
permutation 144, 343
phonologization 4
phraseologism 505, 508, 518, 522
phraseology 503, 520
pidginization 308
plural 161, 172, 222, 251, 496, 498, 500
pluralizer 393
polygrammaticalization 12, 455-491
polysemy 288, 292, 307, 375, 413, 426, 456
portmanteau (morph) 514
possession 167, 209
inalienable 392
postposition 9, 30, 201-222, 227-252
postposition, spatial 249
pragmatic pattern 165
pragmatic significance 337
prediction 19, 21, 22, 24, 26, 27, 28, 48
prepositional verb, *see* verb
presentative 177, 178, 209
pressure, discourse-pragmatic 9
preverb 457, 460-468, 473, 479-484, 486, 489-490
principle 11
relevance 33
procedural 71, 72
process 6, 76, 249, 277, 455, 474, 479, 485
cognitive 142
derivational 222
diachronic 277
evolutionary 455
morphological 76
morphotonemic 221
semantic 300
proclitic 459
procliticization 489

pro-drop 529
productivity 35, 238, 464
progressive (aspect) 30, 33, 195, 207, 208, 211, 219, 274, 325, 417, 466, 474, 478, 484, 485, 513, 514, 533
promotion 516
pronoun
  anaphoric 524, 528
  dummy 137
  interrogative 126, 147, 259
  referential 162
  relative 122, 141, 147, 151, 366
  resumptive 142
proper name 497
proposition 24, 96, 97, 231, 366, 369
  contrasting 96
prospective (aspect) 474, 475, 477-479, 484, 485
protasis 401
proto-language 203, 206, 221
punctual 166
purpose clause, *see* clause

Q

question, rhetorical 367
quotation, direct 71
quote
  direct 240, 241, 307
  indirect 272, 276, 277, 307

R

raising 168, 170
reanalysis 7, 35, 202, 228, 240, 242, 244, 246, 318, 325, 365, 366, 481, 483, 486, 494, 520
  syntactic 369
reason 249
reduction 7, 17, 47, 148, 149, 150, 151, 152, 153, 371, 443, 487, 501, 523
  coordination 144, 146
  phonological 252
  segmental 138
  suprasegmental 37
  in scope 6
referentiality 107
reflexive 146, 392
reflexivization 111
reformulation 318
reinterpretation 7, 28, 149, 150
relative clause 95, 122, 123, 127, 172, 173, 181, 244, 257-308, 366, 401, 488, 489
relative clause
  headless 124, 245
  restrictive 291
relativization 112, 114, 122, 126, 127, 128, 131, 172-173, 260, 261-262, 264-265, 269-271, 282, 306
relativizer 393, 489
repetitive 474
restriction
  selection 497, 517
  syntactic 165, 166, 167
resultative 450, 474
rheme 348, 350, 355, 356, 369, 375, 376
role
  case 191, 194, 196, 197, 259, 260, 262, 263, 264, 266, 270, 299
  discourse 100
  semantic 102, 103, 112, 187, 188, 191, 193, 195, 246, 261
routinized pattern 159

S

scale 520
  grammaticalization 501
scope 344, 350, 357, 366, 375, 506, 509, 515, 526
selection restriction, *see* restriction
  lexical 74
separation 96-98
serial verb 3, 61, 203, 205, 214, 216, 217, 218, 219, 490
  construction 61, 194, 203, 204, 217, 219, 485
serialization (of verbs) 198, 221, 434, 437
  complement 220
sex, *see* gender
shift, *see* topic

category 248
functional 247
functional-semantic 475
syntagmatic 372
shortness 34, 38, 39, 40, 41, 42, 44, 45, 46
significance, statistical 41
similarity 258
singular 161, 166, 172, 302, 322 498, 500, 513
situation, external described 249
situation, internal 249
social sphere 251
source
(function) 231, 240, 241
(lexical) 23, 205, 216, 475, 478, 479, 480, 481, 485
space, topological 8
spatial sphere 251
specialization 74, 320, 325, 383, 384
speech act 23, 241
speech-processing 258
stability 500
stative 174, 234
status, synchronic 239
stem 165, 238, 248
strategy 235, 236, 239, 248, 518
communicative 8
stratum 203
stress 333, 345, 346, 358
contrastive 137, 343, 345
style, formal 116
subject 93-131, 159-183, 247, 314, 315, 320
subjectivization 106
subjunctive 292, 293, 512
subordination 164, 168, 234, 459, 470, 471, 480
marker of 457, 469
subordinator 6, 227-252, 257-308, 393, 469, 470, 471, 473, 480, 481, 483, 486, 489
causal 237
clausal 12
postpositional 480
purpose 242
substrate 7, 11
substratum 308
suffixation 247
superstrate 11, 148
suppletion 76, 222
suppletive (form) 475, 476, 480, 482, 513
syllable 237, 461, 496
syncretism 228, 252, 498
synchrony 237, 238, 239
synonym 499
syntacticization 3, 303

T

target 316, 317, 461
teleology 9
telicity 221
temporal relationship 235
temporal sequence 239, 251
temporal sphere 251
tense, relative 62, 64
text, *see* narrative text
text 235, 237, 239, 248, 482
that-deletion 313, 314, 316
theme 231, 348, 376
tone
floating 212
language 37, 387
underlying 221
topic 12, 93-131, 137, 138, 142, 160, 173, 175, 176, 177, 178, 181, 182, 238, 288, 315, 508
contrastive 507, 509, 510
grammaticalized 94, 95
non-referential 111
topic-comment structure 103
topic continuity 128, 176, 182
topic shift 176, 178, 182
topicalization 106, 112, 195, 238, 340, 356
topicalizer 238, 400
transcription 327
transfer 257, 276, 291, 298, 302
metaphorical 7, 257, 384

metonymic 7
transition 511, 518
transitivity 76, 185, 196, 217, 221
translation, loan 522, 523
typology
language 7, 41, 42, 44, 47, 308
morphological 42

U
unidirectionality 4-7, 11, 18, 47, 248, 466, 484

V
valency 499, 500, 517, 518
variability
paradigmatic 372, 525
syntagmatic 372, 525
variation 474, 525
dialect 147
free 279, 525, 531
vector verb 59, 60, 61, 66, 67, 71, 74, 75, 76, 77, 78, 80, 81
venitive 30, 515
verb
cognition-utterance 272, 297, 306, 308, 398
complement 274
compound 59, 60, 61, 62, 65, 66, 70, 72, 73, 74, 75, 79
desiderative 292, 293, 295
directional 443
epistemic 315, 327
experiencer 500
facilitative 292, 295
factive 296
function 517, 518, 519, 530
intentional 394
locative 209, 219
main 60, 240, 418, 422, 476, 483
manipulative 292, 293, 294, 295, 308
metereological 136
modal 72
movement 28, 29, 30, 468, 477
nominal 246
nominalized 202, 245, 261, 518, 531
parenthetical 322
perception-cognition 288
phasal 72
serial, *see* serial verb
subordinate 246
utterance 272, 275, 398
vector, *see* vector
versatile 402, 404, 406, 407, 409, 415, 419, 427, 432, 436, 440, 441, 448
verb of action 484
of communication 68, 71
of motion 218, 466, 467, 484, 485, 490, 515
vowel harmony 188

W
wh-movement 340
word order 12, 136, 138, 142, 144, 146, 148, 149, 150, 151, 152, 177, 182, 201-222, 277, 278, 279, 280, 306, 346, 359, 360
basic 349
change 524
free 194, 196, 233, 359, 376