Cognitive Foundations of Grammar

Bernd Heine

Cognitive Foundations of Grammar

BERND HEINE

New York Oxford
OXFORD UNIVERSITY PRESS
1997

Oxford University Press

Oxford New York
Athens Auckland Bangkok Bogota Bombay Buenos Aires
Calcutta Cape Town Dar es Salaam Delhi Florence Hong Kong
Istanbul Karachi Kuala Lumpur Madras Madrid Melbourne
Mexico City Nairobi Paris Singapore Taipei Tokyo Toronto Warsaw

and associated companies in
Berlin Ibadan

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press

Library of Congress Cataloging-in-Publication Data
Heine, Bernd, 1939–
Cognitive foundations of grammar / Bernd Heine.
p. cm.
Includes bibliographical references and index.
ISBN 0-19-510251-7; ISBN 0-19-510252-5 (pbk)
1. Cognitive grammar. I. Title.
P165.H45 1997
415—dc20 96-34670

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

PREFACE

During the Australian Linguistic Institute 1994, which took place in July 1994 at LaTrobe University, Melbourne, I gave a course on the cognitive foundations of grammar. I was then asked by students of the course whether what I was saying was available in print. I decided to work on an introductory account that could be of use to students of linguistics, cognitive science, psychology, anthropology, and related disciplines. One year later I had the opportunity to discuss the same subject matter when giving a course at the Institute of the Linguistic Society of America at the University of New Mexico, Albuquerque. This book owes much to discussion with my students in Melbourne and Albuquerque, and it is dedicated to them.

A number of other people have contributed to this book in some way or other. My thanks are due in particular to Jürgen Broschart, Joan Bybee, Ulrike Claudi, Bernard Comrie, Karen Ebert, Suzanne Fleischman, Orin Gensler, Tom Givón, Ingo Heine, Paul Hopper, Christa Kilian-Hatz, Chirsta König, Tania Kuteva, George Lakoff, Dirk Otten, Mechthild Reh, Heinz Roberg, Franz Rottland, Hans-Jürgen Sasse, Mathias Schladt, Fritz Serzisko, Eve Sweetser, and Elizabeth Traugott for critical comments and advice. I also thank Hassan Adam (Swahili), Kossi Tossou (Ewe), and Mohamed Touré (Bambara) for their patience when providing me with information on their mother tongues, and two anonymous reviewers of Oxford University Press for comments on an earlier version of the manuscript. Finally, I thank the Australian Research Council, the Deutsche Forschungsgemeinschaft (German Research Society), the Alexander von Humboldt-Stiftung (Humboldt Foundation), and the Volkswagen-Stiftung (Volkswagen Foundation) for having sponsored the research on which this book is based in some way or other.

Cologne, Germany B. H.
March 1997

ABBREVIATIONS

A	subject of a transitive or ditransitive verb
ABS	absolute
ABSO	Absolutive
ABSTR	abstract possession
ACC	accusative
AL	alienable
ALL	allative
ART	article
ASP	aspect
ASSOC	associative
AUX	auxiliary
CL	noun class
CLASS, CLFR	classifier
CLIT	clitic
COM	comitative
COMPL	completive
CONJ	conjunction
CONN	connector
COP	copula
DAT	dative
DEF	marker of definiteness
DEM	demonstrative
DUAL	dual
EMPH	emphatic marker
ERG	ergative
EXIST	existential marker
F	feminine gender
FOC	focus
GEN	genitive
IMP	imperative
IMPERF	imperfective
INAL	inalienable possession
INDEF	marker of indefiniteness
IN/A	inanimate alienable possession
IN/I	inanimate inalienable possession
INSTR	instrument
IPFV	imperfective
LOC	locative
M	masculine gender
NEU	neuter gender
NF	nonfeminine gender
NOM	nominative
NSP	nonspecific
O	object of a transitive verb
PART	participle
PARTIT	partitive
PAST	past tense
PERF	perfect
PERM	permanent possession
PFV	perfective
PHYS	physical possession

PL	plural
PM	predicate marker
PRES	present tense
PRET	preterite
PRON	pronoun
PRS	presentative marker
PRT	particle
Q	question marker
QUOT	quotative
REC	recipient
S	subject of an intransitive verb
SG	singular
SPEC	specific
STAT	stative
STRESS	stress
SU	subject of intransitive verb
SUBJ	subjunctive
TAM	tense, aspect, and modality
TEMP	temporary possession
TOP	topic
1	first person, class 1
2	second person, class 2
3	third person, class 3
4, etc.	class 4, etc.

CONTENTS

Cognitive Foundations of Grammar

1

THE FRAMEWORK

1.1 Assumptions

Language structure is the product of our interaction with the world around us. The way we build discourses and develop linguistic categories can immediately be derived from the way we experience our environment and use that experience in species-specific communication.

A common human strategy of communication consists in relating different concepts by describing one in terms of the other. This strategy, it is argued here, can be held responsible for much of why grammar looks the way it does, and perhaps also why grammar exists in the first place. And it also constitutes the foundation of the framework used in this book. This framework is based primarily on the following assumptions:

A. The main function of language is to convey meaning. The question of why language is used and structured the way it is must therefore be answered first and foremost with reference to this function.
B. The forms used for expressing meaning are motivated rather than arbitrary (where "motivated" means that linguistic forms are not invented arbitrarily but are, rather, already meaningful when they are introduced for some specific function).
C. Since the motivations for using and developing language are external to language structure, external explanations of language are more powerful than internal ones.

D. Language is a historical product and must be explained first of all with reference to the forces that have shaped it.
E. The synchrony/diachrony distinction derives from the perspective adopted, rather than from the facts considered.
F. Grammatical change is unidirectional, leading from lexical to grammatical, and from grammatical to even more grammatical, forms and structures.

Some of these assumptions are perhaps trivial, others may seem unusual, and still others are hard to reconcile with widely held views of mainstream linguistics. A few elucidating remarks are therefore in order.

Assumption A is based on the observation that when using language, people are less worried about what kind of syntax or phonology to use than they are about how to encode the meanings they want to communicate in the best way possible. This suggests that, first, language use is goal-oriented: People use language to accomplish purposes and goals. Second, linguistic form will tend to adapt to the meaning expressed by it, and not normally the other way round. Third, linguistic explanations in terms of such exponents of language structure as syntax or phonology are likely to highlight peripheral or epiphenomenal rather than central characteristics of language use and language structure (note that we are using the term "meaning" in a wider sense, to include, for example, discourse-pragmatic functions). Furthermore, as we will see, language is not a simple reflection of meaning; content alone is not sufficient to explain why languages look the way they do (cf. Bates & MacWhinney 1989:7).

Assumption B might seem to contradict one of the basic axioms of post-Saussurean linguistics. Following Ferdinand de Saussure (1922:101c, 102b, 180–4), it has become habitual in linguistic works to assume that linguistic signs are "arbitrary" or "unmotivated," where the two terms tend to be used synonymously: There is no natural, inherent connection between a form (*signifiant*) and its meaning (*signifié*)—any *signifié* could be expressed by any *signifiant*. This is proved a posteriori, Saussure argues, by the existence of different languages and by the fact that languages change (cf. Wells 1947).

As a matter of fact, however, B does not contradict Saussure's arbitrariness axiom. There are a number of contrasting senses in which the notions of arbitrariness and motivation can be used; the following example may help illustrate the most common senses:

(1) a. They keep the money.
b. They keep complaining.

It would seem that there are at least three different ways in which the distinction motivated/arbitrary can be used with reference to examples like (1). One approach concerns language structure. The item *keep* in (1) is associated with two different morphosyntactic structures and two different meanings: It functions as the main verb and has the lexical semantics of an action verb in (1a), while it is commonly described as an auxiliary (or catenative) verb that expresses an aspectual notion in (1b). On the

basis of such structural linguistic criteria, one may decide that the phonological identity of *keep* in (1a) and (1b) is coincidental or arbitrary—hence, unmotivated. But it is equally possible to highlight other structural properties (such as shared semantic features or syntactic rules) that suggest the existence of a linguistically definable, motivated relationship.

The second approach has to do with native speakers' intuitions. For example, instead of using structural, linguistic criteria, one could choose a sample of one thousand speakers of English and ask them whether the items *keep* in (1a) and (1b) are related. If a statistically significant majority says yes, then one may conclude that the relationship is motivated.

The third approach relates to diachrony: *keep* in (1a) and (1b) is etymologically the same—that is, both can be assumed to be historically derived from one and the same root—hence, their relationship is motivated rather than arbitrary.

For want of more appropriate terms, the three kinds of motivation just sketched may be called *structural*, *psychological*, and *genetic* motivation. This distinction is not entirely satisfactory, nor is it exhaustive: There are other kinds of motivation that one could think of, such as sociological or areally induced motivation. Like most other linguists, however, I will confine myself to these three notions. Saussure (1922), for example, appears to have been preoccupied with structural and/or psychological motivation (1922:180ff.; see chapter 2).

My interest here is exclusively with genetic motivation since, of the three kinds of motivation discussed, only genetic motivation lies unequivocally within the scope of the linguist. The linguist can make a useful contribution to the study of psychological motivation, but this domain falls essentially within the scope of the psychologist. And, although structural motivation has been a dominant concern of linguistics since Saussure, it is not entirely clear what reason there should be for structural motivation. Finally, genetic motivation has the advantage of being less theory-dependent than the other two kinds of motivation since its reconstruction allows for a straightforward evaluation procedure: It can easily be falsified by means of diachronic evidence. For example, the question of whether the relationship between (1a) and (1b) is genetically motivated is a matter not of the theory or the model adopted but of whether a certain event has or has not taken place (see section 1.2.2). Note, however, that genetic motivation is not an explanatory notion; rather, it has to be explained with reference to other factors, most of all with reference to Assumption A.

Assumption C is based on Assumption A, that the main function of language is to convey meaning and to communicate successfully. Hence, explaining language structure with reference to the goals of communication is likely to yield more insights than explaining it with reference to language-internal mechanisms. For example, an account of lexical borrowing in terms of lexical, syntactic, or morphological structure is probably less "explanatory" than an account in terms of the motivations speakers have for conveying meaning.

Assumption D rests on the observation that language has not been created by the people who are presently using it but, rather, has evolved over the course of centuries and millennia. Grammar, as we now use it, can be described as the conventionalized (and to some extent fossilized) product of earlier patterns of less constrained

language use. Explanations of language in terms of its synchronic structure are therefore likely to account for only a small part of why language is structured the way it is. Many characteristics of language and its uses can therefore be explained satisfactorily only with reference to its diachronic evolution. The following example may illustrate this point. In English, as in a number of other languages, there is an asymmetry in use between definite and indefinite articles: One can utter (2a), (2b), and (3a) but not (3b)—that is, the indefinite article may determine singular nouns but not plural nouns.

(2) a. I see the child.
    b. I see the children.
(3) a. I see a child.
    b. *I see a children.

Any attempt to explain this asymmetry must take account of the historical development of the articles in question. The English indefinite article *a(n)* can be traced back to the numeral *one*. Obviously, numerals for 'one' are inappropriate as modifiers of plural nouns (e.g., **one children*). Although *a(n)* is no longer a numeral, the structural property of incompatibility with plural head nouns has survived its development into an indefinite article. For obvious reasons, such constraints were absent in the genesis of the English definite article, which is thus compatible with both singular and plural nouns (see chapter 4 for more details). Examples like this suggest that what surfaces in synchronic structure is just the tip of the iceberg of what makes up the dynamics of language use.

Assumption D does not mean that explanations based on a synchronic perspective are not meaningful. It does imply, however, that before proceeding to synchronic explanations it is both easier and more efficient to establish to what extent the facts to be explained are due to historical forces. Thus, explaining the said asymmetry in the behavior of the English definite and indefinite article in terms of a synchronic analysis before proceeding to a diachronic analysis is likely to make the task of explanation unnecessarily complicated.

Assumption E rests on the observation that there is no such thing as synchronic or diachronic language or language use: There is just language use. Students of language usually divide their subject matter into synchronic and diachronic linguistics, and this division has turned out to be extremely useful. But for those involved in an individual act of communication—the speaker and the hearer—the distinction is hardly relevant. Whether we adopt a synchronic or a diachronic perspective depends on the goals we want to pursue, not on the subject matter concerned (see section 1.2.1).

Assumption F is by now commonplace in linguistics: The development of grammatical forms proceeds from less grammatical to more grammatical; from open-class to closed-class categories; and from concrete, or less abstract, to less concrete and more abstract meanings (see, e.g., Heine, Claudi, & Hünnemeyer 1991; Traugott & Heine 1991a; Bybee, Perkins, & Pagliuca 1994). A number of exceptions to the unidirectionality principle have been claimed (e.g., Campbell 1991; Greenberg 1991; Ramat 1992), but they have either been refuted or are said to involve processes other than grammaticalization (Hopper & Traugott 1993).

These assumptions will accompany us in the chapters to follow; they will induce us to adopt a perspective on language that differs in a number of ways from that assumed in most works of contemporary linguistics.

1.2 Methodological issues

As may have become apparent in the preceding section, my approach here requires that we look at language structure from a perspective that is not normally found in canonical treatments of language in contemporary linguistics. This means, for example, that theoretical notions that have been crucial in previous accounts are considered here to be epiphenomenal or marginal, while others that have been outside the scope of previous accounts are now interpreted as being central. This perspective also raises problems, however. Some of these are briefly discussed in this chapter, and subsequent chapters provide further details.

1.2.1 Conceptual transfer

The methodology used in this volume rests on the following observation: The presence of one linguistic form with several different meanings may suggest conceptual transfer patterns in which the form was first used to denote one meaning before it was extended to designate one or more additional meanings. Thus, we observed that the English item *keep* has at least two meanings, as illustrated in (1), reprinted here for convenience.

(1) a. They keep the money.
b. They keep complaining.

This fact can be explained as being due to conceptual transfer of the following kind: *Keep* was first used as a main verb in contexts such as (1a); later its use was extended to contexts like (1b), in which it is no longer a main verb but an auxiliary. This transfer has the following properties in particular:

1. It is unidirectional—that is, we do not normally expect a development in the opposite direction, where an auxiliary like *keep* in (1b) develops into a main verb.
2. Unidirectionality leads from concrete, or less abstract, meanings to more abstract meanings. With reference to our example in (1), this means, for example, that *keep* in (1a) is compatible with complements that are visible and tangible, like money, whereas in (1b), complements like *complaining* are more abstract in that, for example, one cannot touch them.
3. The transfer is a historical process and can be accounted for with reference to the principles of diachronic linguistics.

Our *keep*-example concerns the unidirectional transfer leading from lexical items that have a fairly concrete semantics to grammatical categories that express sche-

matic meanings; these latter typically have to do with the relative time, boundary structure, and modality of events.

Conceptual transfer patterns like the one just looked at, as well as many others discussed in the following chapters, have been described variously as involving either figures of speech such as metaphor (e.g., Heine, Claudi, & Hünnemeyer 1991; Sweetser 1990; Stolz 1991, 1994b) and metonymy (Traugott & König 1991) or context-induced processes such as invited inferences, conversational implicatures, and the like. All these notions are relevant for understanding the process concerned, but I will not attempt an evaluation of them here (however, see section 7.4). Following Heine, Claudi, and Hünnemeyer (1991) I argue that the process has both a discontinuous and a continuous component—that is, it can be described variously in terms of both discrete jumps and gradual context-dependent extension of meaning.

1.2.2 On polysemy

The perspective sketched here also suggests an alternative way of dealing with an old and as yet unresolved linguistic issue—how to decide whether two meanings associated with one linguistic form are suggestive of polysemy, rather than monosemy or homonymy. Finding a satisfactory answer to this question is both a central and a controversial problem of linguistics. The answer proposed here is in line with the general theme of this book: In much the same way as we distinguished between three kinds of motivations, we may also distinguish between structural, psychological, and genetic polysemy. The distinction can be illustrated again by means of example (1).

(1) a. They keep the money.
    b. They keep complaining.

Leaving aside various problems, one may say that polysemy has normally been defined by means of a set of three criteria:

1. There are two or more different but related meanings.
2. These meanings are associated with one linguistic form only.
3. The linguistic form belongs to one and the same morphosyntactic category in all its uses.

Whether the item *keep* in (1) is an instance of structural polysemy is not easy to establish; the problems are essentially the same as those mentioned in section 1.1 with reference to structural motivation. Accordingly, while the item appears to comply with criterion no. 2, one may argue, for instance, that criterion no. 1 does not apply. But how the semantic relationship between an auxiliary and the main verb from which it is derived should be defined is hard to answer independent of the theoretical position one decides to adopt. And the same applies to criterion no. 3: Some linguists would argue that *keep* belongs to more than one syntactic category since it is a main verb in (1a) but an auxiliary (or a catenative) in (1b); others claim that auxiliaries and main verbs belong to the same syntactic category (see Heine 1993 for more details). The former would be forced to say that on the basis of the above criteria, *keep*

in (1) is not polysemous, while the latter might say that it is a case of polysemy (provided they can find sufficient formal criteria that allow them to define *keep* in (1) as having "different but related meanings"). Thus, determining structural polysemy is not an easy task. To put it perhaps more seriously, once one has found a convenient way of defining structural polysemy, the question arises as to what one has actually achieved by doing so.

Different problems arise in connection with psychological polysemy. Take, for instance, the following: How does one determine the native speaker's intuitions or awareness of a relatedness of meaning in (1)? Some might say that this question is not within the scope of the linguist's methodology and hence should be left to the psychologist, for example, to answer. Others believe this question can essentially be answered by means of linguistic evidence, though that evidence may not be available as yet (cf. Lyons 1977:552).

Such problems do not exist in the case of genetic polysemy: (1) is unambiguously an instance of genetic polysemy since *keep* in (1a) and (1b) can be traced back historically to one and the same item.

In a number of more recent works, the term *polysemy* is largely used in the sense of genetic polysemy. In such works, polysemy tends to be described as the synchronic reflection of semantic change (Geeraerts 1992:183). What these works appear to have in common is that they do not require polysemy criterion no. 3 to apply—that is, polysemy is not necessarily confined to instances that involve only one morphosyntactic category (cf. Brugman 1984; Traugott 1986; Norvig & Lakoff 1987; Lakoff 1987; Emanatian 1992). Brugman (1984), for example, observes that the English lexeme *over* is an instance of polysemy even if it has prepositional, adverbial, and derivational uses and hence is associated with different morphosyntactic categories.

To avoid such problems surrounding orthodox definitions of polysemy, Lichtenberk introduces the term *heterosemy*. With this term he refers "to cases (within a single language) where two or more meanings or functions that are historically related, in the sense of deriving from the same ultimate source, are borne by reflexes of the common source element that belong in different morphosyntactic categories" (1991:476). Heterosemy as used by Lichtenberk thus is a special case of genetic polysemy: special, since it is confined to instances of genetic polysemy that violate criterion no. 3. Note further that heterosemy is also said to be present when the items concerned are not phonologically identical—that is, when criterion no. 2 is violated. Thus, the English items *have* and *'ve* in (4) are also instances of heterosemy, even if they are not phonologically the same. This is in accordance with the notion of genetic polysemy, which is not confined to cases of linguistic forms observing criterion no. 2.

(4) a. They have two children.
b. They have to come.
c. They've come.

The notion of genetic polysemy is also material for understanding the significance of what has become known as the Typological Convergence criterion (see Croft 1991:166–7; Hopper & Traugott 1993:71). The latter can be illustrated with the fol-