III

Language typology and syntactic description

Grammatical categories and the lexicon

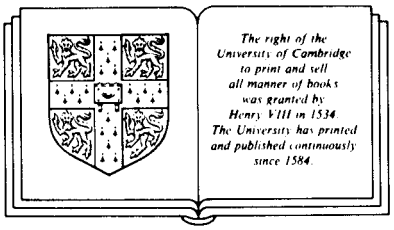
Edited by Timothy Shopen

Language typology and syntactic description

Volume III

Grammatical categories and the lexicon

Edited by
TIMOTHY SHOPEN
Australian National University



CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle
Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1985

Printed in Great Britain at The Bath Press, Avon

Library of Congress catalogue card number: 84-20028

British Library cataloguing in publication data
Language typology and syntactic description.
Vol. 3, Grammatical categories and the lexicon
1. Grammar, Comparative and general – Syntax
I. Shopen, Timothy
415 P291

ISBN 0 521 26859 1 hard covers
ISBN 0 521 31899 8 paperback

Language typology and syntactic description /

Volume III
Grammatical categories and the lexicon

Language typology and syntactic description is published under the auspices of the Center for Applied Linguistics.

Volume I. *Clause structure*
Volume II. *Complex constructions*
Volume III. *Grammatical categories and the lexicon*

Acknowledgements

This work began at a conference on field work questionnaires initiated by Rudolph Troike at the Center for Applied Linguistics (CAL). The participants agreed that the best way to prepare for field work is to develop an idea of what to look for, and this led to the idea of a typological survey that could serve as a reference manual and a textbook for students.

Many people have helped us in the work that we now present. I will name only a few here. Rudolph Troike and John Hammer of CAL, and Alan Bell of the National Science Foundation did much to help in the organization of the project, and the National Science Foundation provided generous financial support without which the work would not have been possible. Diana Riehl of CAL was a reliable and capable intermediary in the complex administration of the project. Carmen Silva-Corvalan and Sandra Thompson deserve special thanks for their work at UCLA, while here in Australia many people provided help. The Australian National University has been very generous in its support of my work. I am grateful to Penny Carter and Julia Harding of Cambridge University Press for the careful work in the production of our books. Three people that have been especially helpful to me in the final stages of the editing are Edith Bavin, Jean Harkins, and above all, Rosemary Butt. My thanks to all.

Timothy Shopen
Canberra, Australia
February 1984

Abbreviations for grammatical terms

The following are abbreviations for grammatical terms used frequently in the glosses for examples. Other abbreviations are explained as they are presented.

ABS	Absolute	IRR	Irrealis
ACC	Accusative	IO	Indirect object
ACT	Actor	LOC	Locative
AG	Agent	NOM	Nominative
ART	Article	NZN	Nominalization
ASP	Aspect	NZR	Nominalizer
ASSOC	Associative	OBJ	Object
AUX	Auxiliary	OBL	Oblique
BEN	Benefactive	PART	Participle
CL	Classifier	PASS	Passive
COMP	Complementizer	PCL	Particle
COMPL	Completive	PERF	Perfective
COND	Conditional	PL	Plural
DAT	Dative	PREP	Preposition
DECLAR	Declarative	PRES	Present
DEF	Definite	PRO	Pro form
DEM	Demonstrative	PROG	Progressive
DET	Determiner	Q	Question marker
DO	Direct object	REFL	Reflexive
DU	Dual	REL	Relativizer
EMPH	Emphasis	RPRO	Relative pronoun
ERG	Ergative	SJNCT	Subjunctive
FUT	Future	SG	Singular
GEN	Genitive	SUBJ	Subject
HABIT	Habitual	TNS	Tense
IMP	Imperative	TOP	Topic
INCOMPL	Incompletive	VN	Verbal noun
INDIC	Indicative	1	First person
INF	Infinitive	2	Second person
INSTR	Instrumental	3	Third person

Contents

Introduction

Grammatical categories and the lexicon is the third of three volumes comprising the work *Language typology and syntactic description*. The first volume is *Clause structure* and the second is *Complex constructions*. Our purpose has been to do a cross-linguistic survey of syntactic and morphological structure that can serve as a manual for field workers, and for anyone interested in relating observations about particular languages to a general theory of language.

The first two chapters of the volume concern the notion of the word. The first is by Stephen Anderson on typological distinctions in word formation. Anderson examines the notion of structure in lexical items with a survey of important grammatical and typological notions that have been applied to word formation. He looks at word formation processes that include stem modification, derivation and compounding.

The second chapter is by Leonard Talmy on what he terms 'lexicalization patterns'. He explores the ways in which languages combine conceptual material into single words, most particularly into verbal roots, their inflections, and a unit that frequently accompanies verbs which he terms the 'satellite'. He proceeds to develop a typology along these lines with special reference to expressions of motion.

The next three chapters concern grammatical categories. The third chapter of the volume is by Stephen Anderson on inflectional morphology. He first clarifies the notion of inflection as a part of word formation and then surveys the grammatical categories encoded in inflections on the major parts of speech.

The fourth chapter is by Sandra Chung and Alan Timberlake on tense, aspect and mood. They take particular care to define the semantic notions involved and then exemplify a typology with sketches of some representative languages.

The fifth chapter is by Stephen Anderson and Edward Keenan on deixis. They concern themselves with person deixis, spatial deixis and temporal deixis. They consider deixis in respect to the speech act, and

'relativized deixis' which keys on other points of space and time established in the discourse.

The last two chapters of the volume are about derivational morphology. The first of these and the sixth of the volume is by Bernard Comrie on causative verb formation and other verb-deriving morphology. Most of the chapter looks at verb formation entailing a change in valency. He considers analytic, morphological and lexical causatives in detail and also considers verbs formed from verbs without a change in valency.

The last chapter of the volume by Bernard Comrie and Sandra Thompson is on lexical nominalization. They explore devices for forming nouns from verbs and adjectives, with a major part of the chapter on the 'action nominal'. They also look at devices for forming nouns from other nouns.

Note: References to chapters in all three volumes of *Language typology and syntactic description* are preceded by the volume number. For example: chapter III.1 (chapter 1, this volume), chapter II.3 (chapter 3, Volume II).

1 Typological distinctions in word formation

STEPHEN R. ANDERSON

0.0 Introduction

The chapters of this volume differ somewhat from much of what has preceded in Volumes I and II. While earlier chapters have been primarily devoted to principles of sentence construction, these chapters are concerned with the principles governing a language's *lexicon* or stock of words. The present chapter deals with the creation of word units out of smaller components, while chapter 2 treats the range of conceptual material that may be grouped together into a single such unit. Chapters 6 and 7 deal with two areas of word formation closely linked to syntactic structure: the formation of causative verbs and related patterns of verbal derivation; and the construction of nominalized forms corresponding to predicative structures.

0.1 *The lexicon*

Our concern here is with the formation of words, and more particularly with the notion of *stems*. Traditionally, the lexicon is thought of as a (more or less structured) list of the form-meaning correspondences (or signs) which speakers have made conventional. With only marginal exceptions, these associations are arbitrary: even onomatopoeic words are arbitrary, at least in part. The fact that English *ear* means what it does and functions as a noun does not follow from any general property of the language: any other combination of English sounds would do as well, or the language might lack such a word altogether. We can oppose this arbitrariness of the sign to other aspects of the word. Thus, given that *ear* is a noun, it follows from the syntax of English that it can appear as a subject, object, etc., but not as an (unsupported) predicate. Similarly, given that it has the phonetic shape /iɹ/, its plural /iɹz/ follows from the general properties of English inflection. This is not to say that it could not – like some lexical items – be exceptional in either morphology or syntax, but nonetheless the contrast is clear: morphology and syntax are areas of overall regularity, as opposed to the brute fact

that in English, *ear* denotes an ear. This fact is completely 'exceptional' in the sense that there is nothing else about the language from which it could have been predicted. Such arbitrariness is typical of the lexicon, which is to this extent the repository of what is idiosyncratic and unpredictable about linguistic forms.

If the lexicon is the locus of unpredictability in form–meaning associations, it follows that it cannot be limited to a list of words alone, but must also include some larger structures. Beyond specifying the forms and meanings of individual words, the lexicon need say nothing about the sentence *He bit my ear*: the structure follows from general principles of English syntax, and the meaning follows from equally general principles for the interpretation of such structures. In *Lend me your ear*, however (in the sense 'listen!'), the meaning of the whole is not *compositional*: that is, it is not a straightforward function of the meanings of its parts. This meaning has much the same arbitrary character as the definition of a simple word, and it is hard to see any principled basis for excluding such phrasal idioms from a lexicon as we have defined it above. Indeed, sometimes the form as well as the meaning of an idiom must be treated as lexical: the structure of *to and fro*, for example, is not syntactically regular in English, and its form as well as its meaning would seem to be a matter for lexical listing.

We will have nothing further to say about the problems of idiom formation here. Our focus is less on the arbitrariness of lexical information than on the possibility of finding some systematic internal structure to the list of forms; for that reason, we will concentrate on the most common kind of lexical form, thinking of the lexicon as a list of words, and ignore the fact that some larger structures ought properly to be included as well.

This brings up the problem of specifying what we mean by a *word*. As will be pointed out in the third chapter of this volume, there really is no satisfactory resolution of this classical problem, since it involves several mutually independent (and sometimes conflicting) criteria. At the risk of some circularity, we will assume that it is the grammatical sense of the notion 'word' (rather than, for example, the phonological notion) which is of interest. 'Words' in this sense are the lexical categories discussed as 'parts of speech' in chapter 1.1. This will include most of the things we usually write between spaces in English, but also some larger structures such as compounds. While there may be cases in which the boundary between 'words' in this sense is unclear, it would appear that the problems we wish to address here are not directly affected by this fact.

0.2 ***The notion of 'structure' in lexical items***

There would be little to study in the domain of word formation principles if all of the items in the lexicon were like *ear* – arbitrary associations between a form and its meaning, where neither has internal structure that is relevant to the existence of the association. Besides such unanalyzable cases, however, other items in the lexicon (perhaps the majority) can be seen as 'partially motivated', in the sense that they involve (individually arbitrary) isolable parts combined in principled ways.

Consider the form *broken-hearted*, for example. It is not enough to list an association between the phonological sequence /browkənhaɪtəd/ and the meaning 'disconsolate'. We must relate the first part of the form to the independent word *broken* (and thus eventually to *break*), and the following part to the independent word *heart*. This is not simply because of the resemblance in sound and meaning (after all, we do not relate *hear* to *ear*, despite the phonetic resemblance and semantic connections between them). Somewhere in the lexicon can be found the information that *break X's heart* means 'render X disconsolate'; but it is probably the existence of other parallel formations (*long-lived*, *three-legged*, *open-mouthed*, etc.) that establishes our sense of a formal pattern into which this relationship fits. What interests us is the sort of principle by which the adjective *broken* (related to the verb *break*) and the noun *heart* are combined to yield the adjective *broken-hearted*.

A number of different processes can be seen at work in this example. *Broken* is related to *break* by a combination of internal vowel change and suffixation found in other deverbal adjectives (*melt/molten*) and participles (*steal/stolen*). It combines with *heart* to form a compound *broken heart*, from which an adjective is formed by adding the suffix *-ed* meaning 'having or endowed with (noun)' (not to be confused with the usual verbal past participle ending: contrast *a learnèd professor* 'one endowed with learning' with *a learned response* 'one which has been learned').

The individual components of the formation can all be seen abundantly elsewhere in the language; but still, the existence of a form *broken-hearted* is a partially arbitrary fact which must be listed in the lexicon. This is because even though we can say a great deal about *how* it is formed, given its components, we cannot predict with certainty *that* it will be formed. Exactly parallel to *break X's heart*, for example, we have *cook X's goose*, *X's goose is cooked*, etc., but no corresponding adjective **cooked-goosed*. The existence of the one but not the other is thus an idiosyncratic fact about the word stock of the language, while the principles by which either is (or could be) formed have a generality