

SECOND EDITION

MODERN LINGUISTICS

# MORPHOLOGY

FRANCIS KATAMBA  
AND JOHN STONHAM



## MORPHOLOGY

## MODERN LINGUISTICS SERIES

### Series Editors

Professor Noël Burton-Roberts  
*University of Newcastle upon Tyne*

Each textbook in the **Modern Linguistics** series is designed to provide a carefully graded introduction to a topic in contemporary linguistics and allied disciplines, presented in a manner that is accessible and attractive to readers with no previous experience of the topic, but leading them to some understanding of current issues. The texts are designed to engage the active participation of the reader, favouring a problem-solving approach and including liberal and varied exercise material.

### *Titles published in the series*

*English Syntax and Argumentation* (second edition) Bas Aarts  
*Phonology* Philip Carr  
*Linguistics and Second Language Acquisition* Vivian Cook  
*Sociolinguistics: A Reader and Coursebook*  
Nikolas Coupland and Adam Jaworski  
*Morphology* (second edition) Francis Katamba and John Stonham  
*Semantics* Kate Kearns  
*Syntactic Theory* Geoffrey Poole  
*Contact Languages: Pidgins and Creoles* Mark Sebba

### *Further titles in preparation*

---

#### **Modern Linguistics Series**

#### **Series Standing Order**

ISBN 978-0-333-71701-1 hardcover

ISBN 978-0-333-69344-5 paperback

*(outside North America only)*

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in the case of difficulty, write to us at the address below with your name and address, the title of the series and the ISBN quoted above.

Customer Services Department, Palgrave Distribution Ltd  
Houndmills, Basingstoke, Hampshire RG21 6XS, England

---

# Morphology

Second Edition

Francis Katamba and John Stonham



© Francis Katamba 1993, Francis Katamba and John Stonham 2006,  
under exclusive licence to Springer Nature Limited 2019

All rights reserved. No reproduction, copy or transmission of this  
publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted  
save with written permission or in accordance with the provisions of the  
Copyright, Designs and Patents Act 1988, or under the terms of any licence  
permitting limited copying issued by the Copyright Licensing Agency,  
Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication  
may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this  
work in accordance with the Copyright, Designs and Patents Act 1988.

First edition published 1993

Second edition published 2006 by  
RED GLOBE PRESS

Red Globe Press in the UK is an imprint of Springer Nature Limited,  
registered in England, company number 785998, of 4 Crinan Street,  
London, N1 9XW.

Red Globe Press® is a registered trademark in the United States,  
the United Kingdom, Europe and other countries.

ISBN 978–1–4039–1644–0

ISBN 978–1–137–11131–9 (eBook)

This book is printed on paper suitable for recycling and made from fully  
managed and sustained forest sources. Logging, pulping and manufacturing  
processes are expected to conform to the environmental regulations  
of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

To Janet and Eun-Sook

# Contents

<i>Preface</i>	xii
<i>Preface to the Second Edition</i>	xiv
<i>Acknowledgements</i>	xv
<i>Abbreviations and Symbols</i>	xvi
<i>The International Phonetic Alphabet</i>	xvii
<b>PART I BACKGROUND</b>	<b>1</b>
<b>1 INTRODUCTION</b>	<b>3</b>
<b>1.1 The Emergence of Morphology</b>	<b>3</b>
<b>1.2 Morphology in American Structural Linguistics</b>	<b>3</b>
<b>1.3 The Concept of Chomskyan Generative Grammar</b>	<b>5</b>
1.3.1 The place of morphology in early generative grammar	10
1.3.2 The morphology–phonology interaction	12
1.3.3 The morphology–syntax interaction	13
1.3.4 The morphology–semantics interface	13
1.3.5 The lexicon	14
<b>1.4 Organisation of the Book</b>	<b>14</b>
<b>2 INTRODUCTION TO WORD-STRUCTURE</b>	<b>17</b>
<b>2.1 What is a Word?</b>	<b>17</b>
2.1.1 The lexeme	17
2.1.2 Word-form	18
2.1.3 The grammatical word	19
<b>2.2 Morphemes: The Smallest Units of Meaning</b>	<b>19</b>
2.2.1 Analysing words	22
2.2.2 Morphemes, morphs and allomorphs	23
2.2.3 Grammatical conditioning, lexical conditioning and suppletion	30
2.2.4 Underlying representations	32
<b>2.3 The Nature of Morphemes</b>	<b>35</b>
<b>2.4 Summary</b>	<b>39</b>
<b>Further Reading</b>	<b>40</b>
<b>Exercises</b>	<b>40</b>
<b>3 TYPES OF MORPHEMES</b>	<b>42</b>
<b>3.1 Roots, Affixes, Stems and Bases</b>	<b>42</b>
3.1.1 Roots	42
3.1.2 Affixes	44

---

3.1.3	Roots, stems and bases	46
3.1.4	Stem extenders	47
<b>3.2</b>	<b>Inflectional and Derivational Morphemes</b>	<b>48</b>
<b>3.3</b>	<b>Multiple Affixation</b>	<b>54</b>
<b>3.4</b>	<b>Compounding</b>	<b>55</b>
<b>3.5</b>	<b>Conversion</b>	<b>56</b>
<b>3.6</b>	<b>Morphological Haplogy</b>	<b>57</b>
<b>3.7</b>	<b>Morphological Typology</b>	<b>58</b>
<b>3.8</b>	<b>WP and the Centrality of the Word</b>	<b>63</b>
	<b>Exercises</b>	<b>65</b>
<b>4</b>	<b>PRODUCTIVITY IN WORD-FORMATION</b>	<b>67</b>
<b>4.1</b>	<b>The Open-Endedness of the Lexicon</b>	<b>67</b>
4.1.1	What is productivity?	68
4.1.2	Semi-productivity	73
4.1.3	Productivity and creativity	74
<b>4.2</b>	<b>Constraints on Productivity</b>	<b>75</b>
4.2.1	Blocking	75
<b>4.3</b>	<b>Does Productivity Separate Inflection from Derivation?</b>	<b>81</b>
<b>4.4</b>	<b>The Nature of the Lexicon</b>	<b>83</b>
4.4.1	Potential words	83
4.4.2	Knowledge of language and the role of the lexicon	84
	<b>Further Reading</b>	<b>85</b>
	<b>Exercises</b>	<b>85</b>
<b>PART II</b>	<b>MORPHOLOGY AND ITS RELATION</b>	
	<b>TO PHONOLOGY</b>	<b>87</b>
<b>5</b>	<b>INTRODUCING LEXICAL MORPHOLOGY</b>	<b>89</b>
<b>5.1</b>	<b>The Lexical Phonology and Morphology Model</b>	<b>89</b>
<b>5.2</b>	<b>Lexical Strata</b>	<b>89</b>
5.2.1	Derivation in lexical morphology	92
5.2.2	Inflection in lexical morphology	100
<b>5.3</b>	<b>Lexical Rules</b>	<b>104</b>
<b>5.4</b>	<b>Differences between Lexical and Post-Lexical Rules</b>	<b>106</b>
	<b>Further Reading</b>	<b>109</b>
	<b>Exercises</b>	<b>109</b>
<b>6</b>	<b>INSIGHTS FROM LEXICAL MORPHOLOGY</b>	<b>111</b>
<b>6.1</b>	<b>Introduction</b>	<b>111</b>
<b>6.2</b>	<b>Insights</b>	<b>111</b>
6.2.1	Stratum ordering reflecting morpheme sequencing	112
6.2.2	Stratum ordering and productivity	117

---

6.2.3	Stratum ordering and conversion	118
6.2.4	The Strict Cycle Condition	121
	<b>Further Reading</b>	<b>130</b>
	<b>Exercises</b>	<b>131</b>
<b>7</b>	<b>LEXICAL MORPHOLOGY: AN APPRAISAL</b>	<b>133</b>
<b>7.1</b>	<b>Introduction: The Claims Made by Lexical Phonology</b>	<b>133</b>
<b>7.2</b>	<b>Criticisms of Lexical Phonology</b>	<b>133</b>
7.2.1	Are lexical strata determined by affixes rather than roots?	134
7.2.2	Do affixes uniquely belong to one stratum?	135
7.2.3	How many strata are needed?	139
7.2.4	Are phonological rules restricted to one stratum?	140
7.2.5	Are morphological rules restricted to one stratum?	142
<b>7.3</b>	<b>Conclusion</b>	<b>151</b>
	<b>Exercises</b>	<b>151</b>
<b>8</b>	<b>TEMPLATIC MORPHOLOGY</b>	<b>154</b>
<b>8.1</b>	<b>Introduction</b>	<b>154</b>
<b>8.2</b>	<b>Phonological Prelude: Autosegmental Phonology</b>	<b>154</b>
8.2.1	Autosegmental phonology: mapping principles	155
8.2.2	The skeletal tier	160
<b>8.3</b>	<b>Root and Pattern Morphology</b>	<b>162</b>
8.3.1	Arabic Binyanim	163
8.3.2	Prosodic morphology and non-concatenative morphology	165
8.3.3	The morpheme tier hypothesis	172
<b>8.4</b>	<b>Conclusion</b>	<b>177</b>
	<b>Exercises</b>	<b>178</b>
<b>9</b>	<b>TEMPLATIC AND PROSODIC MORPHOLOGY</b>	<b>180</b>
<b>9.1</b>	<b>What is Reduplication?</b>	<b>180</b>
<b>9.2</b>	<b>Is Reduplication Constituent Copying?</b>	<b>182</b>
<b>9.3</b>	<b>CV Templates and Reduplication</b>	<b>183</b>
9.3.1	Underspecification	184
9.3.2	Reduplication as prefixation	185
9.3.3	Reduplication as suffixation	189
9.3.4	Internal reduplication	190
9.3.5	Reduplication and fixed segmentism	192
<b>9.4</b>	<b>Prosodic Morphology</b>	<b>194</b>
<b>9.5</b>	<b>Other Prosodic Phenomena</b>	<b>199</b>
9.5.1	Subtractive morphology	199
9.5.2	English expletive infixation	201

---

9.6	<b>Conclusion</b>	<b>201</b>
	Exercises	<b>202</b>
<b>10</b>	<b>OPTIMALITY THEORY AND MORPHOLOGY</b>	<b>204</b>
10.1	<b>Introduction</b>	<b>204</b>
10.2	<b>The Basics</b>	<b>204</b>
	10.2.1 Notation	205
	10.2.2 The interaction of constraints	206
10.3	<b>Morphology in Optimality Theory</b>	<b>213</b>
10.4	<b>Stratal Optimality Theory</b>	<b>215</b>
10.5	<b>Conclusion</b>	<b>219</b>
	Further Reading	<b>219</b>
	Exercises	<b>220</b>
<b>PART III</b>	<b>MORPHOLOGY AND ITS RELATION TO THE LEXICON AND SYNTAX</b>	<b>221</b>
<b>11</b>	<b>INFLECTIONAL MORPHOLOGY</b>	<b>223</b>
11.1	<b>Introduction</b>	<b>223</b>
11.2	<b>Inflection and Derivation</b>	<b>223</b>
	11.2.1 Differentiating between inflection and derivation	224
	11.2.2 Relevance and generality	229
	11.2.3 Is morphology necessary?	234
11.3	<b>Verbal Inflectional Categories</b>	<b>237</b>
	11.3.1 Inherent verbal properties	237
	11.3.2 Agreement properties of verbs	242
	11.3.3 Configurational properties of verbs	244
11.4	<b>Inflectional Categories of Nouns</b>	<b>250</b>
	11.4.1 Inherent categories of nouns	250
	11.4.2 Agreement categories of nouns	253
	11.4.3 Configurational categories of nouns	254
11.5	<b>Conclusion</b>	<b>261</b>
	Exercises	<b>262</b>
<b>12</b>	<b>MORPHOLOGICAL MAPPING OF GRAMMATICAL FUNCTIONS</b>	<b>267</b>
12.1	<b>Introduction</b>	
12.2	<b>Predicates, Arguments and Lexical Entries</b>	<b>267</b>
12.3	<b>Theta-Roles and Lexical Entries</b>	<b>268</b>
12.4	<b>Grammatical Relations</b>	<b>274</b>
12.5	<b>Grammatical Function-Changing Rules</b>	<b>276</b>
	12.5.1 Passive	279
	12.5.2 Antipassive	281

---

12.5.3	Applicative	282
12.5.4	Causative	286
<b>12.6</b>	<b>The Mirror Principle</b>	<b>287</b>
<b>12.7</b>	<b>Incorporation</b>	<b>294</b>
12.7.1	Noun incorporation	295
12.7.2	Verb incorporation	296
12.7.3	Preposition incorporation	297
<b>12.8</b>	<b>Conclusion</b>	<b>298</b>
	<b>Exercises</b>	<b>300</b>
<b>13</b>	<b>THE INTERPRETATION OF THE LEXICON, MORPHOLOGY AND SYNTAX</b>	<b>304</b>
<b>13.1</b>	<b>Introduction: The Interface Between Modules</b>	<b>304</b>
<b>13.2</b>	<b>Phonological Factors In Compounding</b>	<b>305</b>
<b>13.3</b>	<b>Are Compounds Different from Syntactic Phrases?</b>	<b>306</b>
13.3.1	The notion ‘word’ revisited	307
13.3.2	Listemes	308
13.3.3	Unlisted morphological objects	309
13.3.4	Syntactic objects and syntactic atoms	310
<b>13.4</b>	<b>The Character of Word-Formation Rules</b>	<b>315</b>
13.4.1	Headedness of compounds	316
13.4.2	The right-hand head rule (RHR)	324
13.4.3	Left-headed compounds	328
13.4.4	Headless compounds	331
<b>13.5</b>	<b>Compounding and Derivation</b>	<b>334</b>
13.5.1	Cranberry words	334
13.5.2	Neoclassical compounds	336
<b>13.6</b>	<b>Clitics</b>	<b>337</b>
<b>13.7</b>	<b>Conclusion</b>	<b>343</b>
	<b>Exercises</b>	<b>345</b>
	<i>Glossary</i>	348
	<i>References</i>	357
	<i>Language Index</i>	371
	<i>Subject Index</i>	373
	<i>Author Index</i>	381

# Preface

This book is an introduction to morphology that presupposes little previous exposure to linguistics. It is meant to be useful both to students of English and to those of linguistics. Most of the first half of the book, as well as the final chapter, is devoted mainly to problems of English word-formation. The remaining chapters cover a range of morphological phenomena in other languages, but even the parts dealing with English raise issues of a general theoretical interest. The detail in which different parts are studied will vary, depending on the kind of student that uses the book.

I present morphology from the standpoint of current, mainstream generative grammar. My main concerns are the nature of word-formation processes and the ways in which word-formation interacts with phonology, syntax and the lexicon. I hope that the reader will come away not only with an understanding of the descriptive problems in morphology but also with a firm grasp of the theoretical issues and the analytical tools that are available within the model of generative grammar. On completing a course in morphology based on this book, students should be equipped to tackle the growing morphological literature that has appeared in recent years.

There are many people whom I must thank for the help they have given me in writing this book. The book grew out of my morphology course at Lancaster University. I must thank the students who have taken this course over the last four years. Special thanks go to Elena Semino and Saleh al-Khateb, whose Italian and Syrian Arabic data I have used here.

I have benefited from discussions with a number of Berkeley linguists, especially Sharon Inkelas, Sam Mchombo and Karl Zimmer. Above all, I must thank in a special way Larry Hyman, with whom I have collaborated on Luganda morphology and phonology for the last ten years. I have learned much of what I know about phonology/morphology through our collaboration.

There are also many other linguists whose theoretical and descriptive studies I have drawn on. They have all contributed in an obvious way to my writing this book.

I also owe a special debt of gratitude to Noel Burton-Roberts, the editor of this series. His rigorous critical comments and positive suggestions have enabled me to avoid some of the pitfalls I would otherwise have encountered. There are two other people at Macmillan that I wish to thank for their technical support: they are Doreen Alig and Cathryn Tanner. I should also

like to thank Valery Rose and David Watson, who both helped with the production of this book.

Finally, I thank my wife Janet for her support during the long months and years of writing this book.

*Lancaster*

FRANCIS KATAMBA

# Preface to the Second Edition

The first edition of this book, published in 1993, has been an extremely popular textbook for over a decade and has been the introduction to morphology for many students. There have been a number of developments in morphological theory over that period and we felt that it was time to provide an updated version of the book for the next generation of students.

This second edition is closely based on the original, but with changes in several areas: first, we have updated the discussion in a number of areas, providing more recent references where appropriate; second, we have added a chapter on Optimality Theory, which has become the dominant paradigm in phonology, although perhaps not yet in morphology; third, we have added new exercises to complement the existing ones in places; and finally, we have taken the opportunity to correct a number of typos/oversights that appeared in the first edition.

Hopefully, this new edition will enhance the understanding of another generation of students' introduction to morphology.

In addition to all those who contributed to the first edition, in particular Gyung-ran Kim and Jin-hying Kim, we would also like to thank our editorial consultant, Penny Simmons, for her careful reading of the entire manuscript, and Noel Burton-Roberts, for his encouragement in pursuing this second edition. We would also like to thank Joel Nevis and Eun-Sook Kim for extensive discussion of the manuscript for this second edition.

FRANCIS KATAMBA and JOHN STONHAM

# Acknowledgements

The authors and publishers wish to thank the following for permission to use copyright material:

Cambridge University Press for Fig. 3.34 from P. Matthews, *Inflectional Morphology* (1972) 7.4.3, p. 132;

The International Phonetic Association [[www.arts.gla.ac.uk/IPA/ipa.html](http://www.arts.gla.ac.uk/IPA/ipa.html)] for the International Phonetic Alphabet, revised 1989, from *Journal of the International Phonetic Association*, vol. 19, no. 2.

Every effort has been made to trace the copyright holders but if any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangement at the first opportunity.

Finally, we would like to thank the many colleagues and their students who have given us much useful feedback on the first edition. We would especially like to single out Gyung-ran Kim and Jin-hyung Kim, who translated the first edition of the book into Korean, for their numerous constructive comments.

# Abbreviations and Symbols

ACC	accusative
ADJ/Adj	adjective
AdjP	adjectival phrase
ADV/Adv	adverb
AdvP	adverbial phrase
APPL	applicative
ASP	aspect
BVS	Basic verbal suffix (in Bantu)
CAUS	causative
DET/De	determiner
ERG	ergative
GF	grammatical function
GVS	Great Vowel Shift
FUT	future
Inf.	infinitive
N/n	noun
NP/Np	noun phrase
OBJ	object
OCP	Obligatory Contour Principle
OED	<i>Oxford English Dictionary</i>
P/Prep	preposition
PP	prepositional phrase
Pron	pronoun
RHR	Right-Hand Head Rule
S	sentence
<i>SPE</i>	<i>The Sound Pattern of English</i>
SUBJ/Sub	subject
V	verb
VP	verb phrase
V <sub>(intr)</sub>	verb <sub>(intransitive)</sub>
V <sub>(tr)</sub>	verb <sub>(transitive)</sub>
VP	verb phrase
WFC	Well-Formedness Condition
WP	Word and Paradigm (morphology)



# **Part I**

## **Background**

# 1 Introduction

## 1.1 THE EMERGENCE OF MORPHOLOGY

Although students of language have always been aware of the importance of words, **morphology**, the study of the internal structure of words did not emerge as a distinct sub-branch of linguistics until the nineteenth century. Nevertheless, its importance has always been assumed, as attested by its central role in Pāṇini's fourth-century BC grammar of Sanskrit, the *Aṣṭadhyayi*, for instance.

Early in the nineteenth century, morphology played a pivotal role in the reconstruction of Indo-European. In 1816, Franz Bopp published the results of a study supporting the claim, originally made by Sir William Jones in 1786, that Sanskrit, Latin, Persian and the Germanic languages were descended from a common ancestor. Bopp's evidence was based on a comparison of the grammatical endings of words in these languages.

Between 1819 and 1837, Bopp's contemporary, Jacob Grimm, published his classic work, *Deutsche Grammatik*. By making a thorough analytical comparison of sound systems and word-formation patterns, Grimm showed the evolution of the grammar of Germanic languages and the relationships of Germanic to other Indo-European languages.

Later, under the influence of the Darwinian theory of evolution, the philologist Max Müller contended, in his Oxford lectures of 1899, that the study of the evolution of words would illuminate the evolution of language just as in biology, morphology, the study of the forms of organisms, had thrown light on the evolution of species. His specific claim was that the study of the 400–500 basic roots of the Indo-European ancestor of many of the languages of Europe and Asia was the key to understanding the origin of human language (cf. Müller, 1899, cited in Matthews, 1974).

Such evolutionary pretensions were abandoned very early on in the history of morphology. Since then morphology has been regarded as an essentially **synchronic** discipline, that is to say, a discipline focusing on the study of word-structure at one stage in the life of a language rather than on the evolution of words. But, in spite of the unanimous agreement among linguists on this point, morphology has had a chequered career in twentieth-century linguistics, as we shall see.

## 1.2 MORPHOLOGY IN AMERICAN STRUCTURAL LINGUISTICS

Adherents to American structural linguistics, one of the dominant schools of linguistics in the first part of the twentieth century, typically viewed

linguistics not so much as a ‘theory’ of the nature of language but rather as a body of descriptive and analytical procedures. Ideally, linguistic analysis was expected to proceed by focusing selectively on one dimension of language structure at a time before tackling the next one. Each dimension was formally referred to as a **linguistic level**. The various levels are shown in [1.1].

<b>[1.1]</b>	<u>Semantic level</u> :	↑	deals with meaning
	<u>Syntactic level</u> :	↑	deals with sentence-structure
	<u>Morphological level</u> :	↑	deals with word-structure
	<u>Phonology (or phonemics)</u> :	↑	deals with sound systems

The levels were assumed to be ordered in a hierarchy, with phonology at the bottom and semantics at the top. The task of the analyst producing a description of a language was seen as one of working out, in separate stages, first the pronunciation, then the word-structure, then the sentence-structure and finally the meaning of utterances. It was considered theoretically reprehensible to make use of information from a higher level, for example, syntax, when analysing a lower level such as phonology. This was the doctrine of **separation of levels**.

In the early days, especially between 1920 and 1945, American structuralists grappled with the problem of how sounds are used to distinguish meaning in language. They built upon nineteenth-century work, such as that of Dufriche-Desgenettes (Joseph, 1999) and Baudouin de Courtenay (1895), and further developed and refined the theory of the **phoneme** (cf., Sapir, 1925; Swadesh, 1934; Twaddell, 1935; Harris, 1944).

As time went on, the focus gradually shifted to morphology. When structuralism was in its prime, especially between 1940 and 1960, the study of morphology occupied centre stage. Many major structuralists investigated issues in the theory of word-structure (cf. Bloomfield, 1933; Harris, 1942, 1946, 1951; Hockett, 1952, 1954, 1958). Nida’s coursebook entitled *Morphology*, published in 1949, codified structuralist theory and practice. It introduced generations of linguists to the descriptive analysis of words.

The structuralists’ methodological insistence on the separation of levels that we noted above was a mistake, as we shall see below in Sections 1.3.2 and 1.3.3. Despite this flaw, there was much that was commendable in the structuralist approach to morphology. One of the structuralists’ main contributions was the recognition of the fact that words may have intricate internal structures. Whereas traditionally linguistic analysis had treated the word as the basic unit of grammatical theory and lexicography, the American structuralists showed that words are analysable in terms of

**morphemes.** These are the smallest units of meaning and/or grammatical function. Previously, word-structure had been treated together with sentence-structure under grammar. The structuralists viewed morphology as a separate sub-branch of linguistics. Its purpose was ‘the study of morphemes and their arrangements in forming words’ (Nida, 1949: 1). The contribution of the structuralists informs much of the discussion in the first part of this book.

### 1.3 THE CONCEPT OF CHOMSKYAN GENERATIVE GRAMMAR

The bulk of this book presents morphological theory within the linguistic model of **generative grammar**, initiated by Chomsky. Before considering how this theory works, we will sketch the background assumptions made by generative grammarians so that we can place the theory of morphology in the wider theoretical context of generative linguistics.

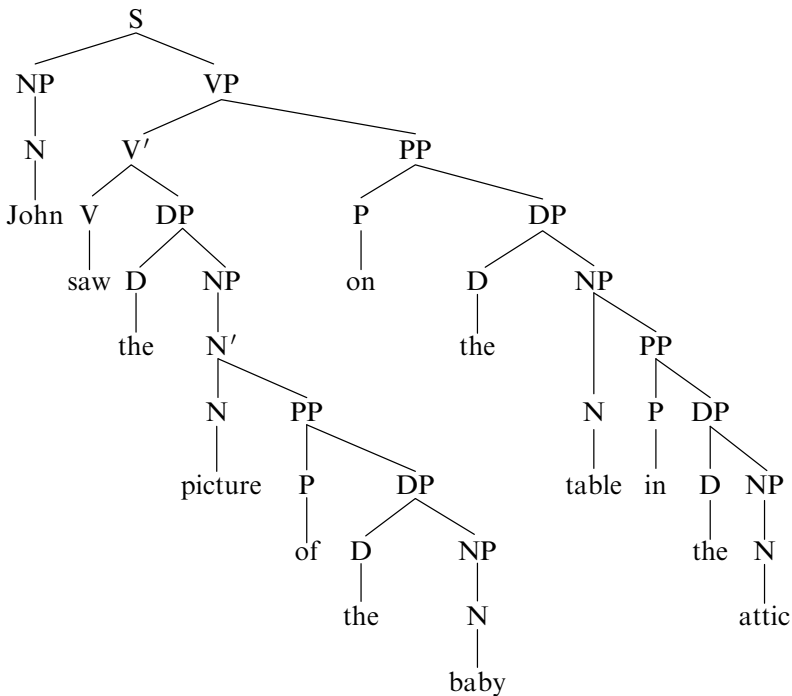
The central objective of generative linguistics is to understand the nature of linguistic knowledge and how it is acquired by infants. In the light of this objective, a fundamental question that a theory of word-structure must address is, ‘what kinds of information must speakers have about the words of their language in order to use them in utterances?’ Attempts to answer this question have led to the development of sub-theories of the **lexicon** (i.e., mental dictionary) and of morphology.

According to Chomsky (1980, 1981, 1986), the central goal of linguistic theory is to determine what it is people know if they *know* a particular language. Chomsky observes that knowing a language is not simply a matter of being able to manipulate a long list of sentences that have been memorised. Rather, knowing a language involves having the ability to produce and understand a vast (and indeed unlimited) number of utterances of that language that one may never have heard or produced before. In other words, creativity (also called productivity or open-endedness) is an aspect of linguistic knowledge that is of paramount importance.

Linguistic creativity is, for the most part, rule-governed. For instance, speakers of English know that it is possible to indicate that there is more than one entity referred to by a *noun* and that the standard way of doing this is to add *-s* at the end of a noun. Given the noun *book*, which we have all encountered before, we know that if there is more than one of these objects we refer to them as *books*. Likewise, given the nonsense word *smilts* as in the sentence *The smilts stink*, you know *smilts* would refer to more than one of these smelly things. Speakers of English have tacit knowledge of the rule that says ‘add *-s* for plural’ and they can use it to produce the plural form of virtually any noun. We emphasise the notion of rule, taking the existence of rules for granted.

We will now explain why a generative grammar is a system of explicit rules that may apply recursively to generate an indefinite number of sentences that can be as long as one wants them to be. **Recursiveness** has the consequence that, in principle, there is no upper limit to the length of sentences. A grammatical constituent like a noun phrase (NP) or prepositional phrase (PP) can directly or indirectly contain an indefinite number of further constituents of that category as in the sentence *John saw the picture of the baby on the table in the attic*. The recursion can be clearly seen in the tree diagram in [1.2], which represents that sentence. On the one hand, DPs can have an optional determiner followed by NP and PP, and on the other hand, PPs are expanded as P followed by NP. So, NPs can indirectly contain other NPs since NPs can contain PPs which contain NPs:

[1.2]



Note: S – sentence; N – noun, NP – noun phrase; V – verb, VP – verb phrase; P – preposition, PP – prepositional phrase; DP – determiner phrase; DET – determiner.

One of our concerns will be to determine whether morphology should be recognised as a separate linguistic level (or **module**) that is independent of syntax and phonology (see [1.1] above and [1.3] below). Do morphological rules have certain properties that they do not share with rules in other parts of the grammar? Are recursive rules of the kind found in syntax needed in

morphology? This book will address these issues in depth. Here we only attempt to give you a flavour of one of the issues to be explored.

There are morphological processes that are similar to syntactic processes. For instance, certain adjectives that describe periods in history, such as *industrial*, can have the prefix *post-* before them as in *post-industrial*. Furthermore, given the adjective *post-industrial*, we can place another *post-* before it to yield *post-post-industrial*. Clearly, the word-formation process we witness here is recursive. We have the rule attaching *post-* to a word reapplying to its own output. This raises an interesting question: if morphological rules that build words are similar to syntactic rules that build sentences, what reason is there for assuming that morphology is essentially different from syntax?

Before we go any further we need to clarify the terms **grammar** and **rule of grammar**. These terms are used by linguists in at least four distinct senses. First, in generative linguistics ‘grammar’ can refer to the implicit, totally unarticulated knowledge of rules and principles of their language that people have in their heads. This tacit knowledge enables them to distinguish between well-formed and ill-formed words and utterances in their language. For example, many English speakers may not be able to explain in an articulate manner why it is ‘correct’ to say *a grain* but ‘incorrect’ to say *\*a oat*. (An asterisk indicates a disallowed or ungrammatical form.) Nevertheless, their knowledge of English grammatical structure enables them to determine that the former is correct and the latter is not.

Second, whereas in traditional approaches ‘grammar’ only includes morphology and syntax, in generative linguistics the term ‘grammar’ is employed in a much wider sense. It covers not only morphology and syntax but also semantics, the lexicon and phonology. Hence, there are rules of grammar in every linguistic module. Phonological rules, morphological rules, syntactic rules and semantic rules are all regarded as rules of grammar.

Third, grammar and rules of grammar may refer to a book containing a statement of the rules and principles inferred by linguists to lie behind the linguistic behaviour of speakers of a particular language. These rules simply describe regular patterns observed in the linguistic data.

Lastly, some grammars are books containing prescriptive statements. Such grammars contain rules that *prescribe* certain kinds of usage. Outside linguistics, this view of grammar is still prevalent. The reason for this is clear. In everyday life, rules are normally mechanisms for regulating behaviour – the behaviour of pupils in a school, members of a club, inmates of a prison, etc. In many traditional pedagogical grammars, rules serve the same purpose. They are statements like ‘A sentence must not end with a preposition.’ They prescribe what the ‘officially or socially approved’ usage is – in the opinion of the grammarian.

In much of modern linguistics, however, rules have a different function. They are not prescriptions of behaviour that the grammarian imposes on

speakers, but rather they are statements of principles responsible for the observed regularities in the speech or writing of users of a particular language. The characterisation of regularities in observed patterns of usage is what the American structuralists regarded as the primary objective of linguistic investigations. Their grammatical rules were descriptive statements like ‘The article precedes the noun in the English noun phrase.’ This statement reflects the fact that *the book*, as in *I read the book*, is allowed whereas *\*book the*, as in *\*I read book the* is disallowed.

Chomsky has shifted the focus of linguistic theory from the study of observed behaviour to the investigation of the knowledge that underlies that behaviour. In generative linguistics, rules are intended to go beyond accounting for patterns in the data to a characterisation of speakers’ linguistic knowledge. The primary objective of generative grammar is to model a speaker’s linguistic knowledge.

Chomsky characterises linguistic knowledge using the concepts of competence and performance. **Competence** is a person’s implicit knowledge of the rules of a language that makes the production and understanding of an indefinitely large number of new utterances possible while **performance** is the actual use of language in real situations. Chomsky proposes that competence, rather than performance, is the primary object of linguistic inquiry. Put simply, knowledge of a language entails mastery of an elaborate system of rules that enables a person to encode and decode a limitless number of utterances in that language. One subset of this rule system is the rules of word-formation that this book introduces you to. In Section 4.1.3 and Section 13.3.3 it will be shown that speakers of a language do not just commit to memory all the words they know. Their competence includes the ability to manipulate rules in order to create new words and to unscramble the meanings of novel or unfamiliar words that they encounter.

If knowing a language essentially involves mastering a system of rules, how do humans accomplish this task? Chomsky contends that the linguistic capacity of humans is **innate**. The general character of linguistic knowledge is determined by the nature of the mind, which is endowed with a specialised **language faculty**. This faculty is determined in turn by the biology of the brain. The human child is born with a blueprint of language that is called **Universal Grammar**.

According to Chomsky, Universal Grammar is the faculty of the mind that determines the nature of language acquisition in the infant and of linguistic competence. The properties that lie behind the competence of speakers of various languages are governed by restricted and unified elementary principles rooted in Universal Grammar. This explains the striking underlying similarity between languages in their essential structural properties. Admittedly, languages differ from each other, but the structural differences between them occur within the fairly narrow range sanctioned by Universal Grammar. As we shall see (especially in Chapters 3, 8, 9 and 13)

with regard to word-formation, very similar word-building principles recur in language after language. The language faculty of the mind is essentially the same in all humans. Hence, languages can only differ from each other within the limits predetermined by the neurology and physiology of the human brain, which determine the nature of Universal Grammar. Universal Grammar in turn determines the kinds of grammars of particular languages that can be acquired by infants.

The differences between the grammars acquired by individual speakers of, say, English and Arabic can be attributed to experience. An individual's experience serves to specify a particular grammar for the particular language which that individual is exposed to – within the range permitted by Universal Grammar.

How is Universal Grammar structured? It is modular in structure: it consists of various sub-systems of principles. Many of its principles consist of **parameters** that are fixed by experience on the basis of simple evidence of the kind available to the child. Chomsky compares Universal Grammar to an intricate electrical system that is all wired up, but not switched on. The system contains a finite set of switches, each one of which has a restricted number of positions. Exposure to a specific language experience is required to turn on these switches and give them the appropriate setting.

The basic idea of parameters is meant to capture the fact that many rules are interdependent. If one choice is made, it may either preclude some other choices or set in motion other related choices. This makes the task of language acquisition simpler than it would be if each rule had to be worked out independently of all other rules. The parametric approach assumes that the infant acquiring a language makes very clever guesses or hypotheses about the rules of the grammar being acquired on the basis of rules already acquired after experience of a particular language.

For a concrete example of a parameter, we will consider the **Right-Hand Head Rule** to be discussed in Chapter 13. This parameter is concerned with the position of the **head** of a grammatical constituent. Some languages, like English, normally place the head on the right, that is, it is the last element of a constituent. For example, in the noun phrase *these big books* the right-handmost word, the noun *books*, is the head. It must come last. (Alternatives like *\*books big these* and *\*these books big* are forbidden.)

As a rule, the head is the only obligatory element of a constituent like an NP. *Books* is a well-formed NP but neither *these* nor *big* is a permissible NP on its own. Furthermore, in terms of meaning, the head *books* is the key word in this NP. The function of *these* and *big* is merely to specify further the particular books referred to. Note also that *these* agrees in number with *books*, that is, they are both plural.

Likewise, at the word level, in a compound like *farmhouse*, the head, *house*, is the last element and it is the pivotal element from a semantic point of view. (*A farmhouse* is a kind of *house*.) Once an infant has worked out the

position of the head for one construction, this can be generalised with a considerable degree of success to other constructions.

Universal Grammar consists of a number of modules that are inter-related. This is shown in [1.3] (which you should compare with [1.1] above):

- [1.3]** (i) Lexicon and Morphology  
(ii) Syntax  
(iii) Phonetic Form (PF) (which deals with the representation in speech)  
(iv) Logical Form (LF) (which deals with meaning)

As can be seen, Universal Grammar includes the lexicon and morphology. Knowledge of word-structure is a central aspect of linguistic competence. A case can be made for recognising morphology as a separate module of Universal Grammar. Yet, at the same time, morphology (and the lexicon) is like a bridge that links the other modules of the grammar. It is therefore necessary to examine morphology not in isolation, but in relation to the other modules. Morphology interacts with both phonology and syntax, as well as semantics. So, it can only be studied by considering the phonological, syntactic and semantic dimensions of words.

### **1.3.1 The Place of Morphology in Early Generative Grammar**

Today the place of morphology in generative grammar is secure, but this is a recent development. After being in the limelight when structuralism peaked in the 1950s, morphology was at first eclipsed when generative grammar came on the scene. Generative grammarians initially rejected the validity of a separate morphological module.

From the point of view of advancing our understanding of word-structure, this stance was unfortunate. Since generative grammar has been the dominant school of linguistics since the second half of the twentieth century, it meant that the study of word-structure was in the shadows for more than a decade. Morphology did not re-emerge until the mid-1970s. Fortunately, a few isolated (for the most part non-generative) scholars such as Robins (1959) and Matthews (1972, 1974) made important contributions to morphology during this time, as we shall see.

Part of the reason for the widespread neglect of morphology during the early years of generative grammar was the belief that word-formation could be adequately covered if it was partitioned between phonology and syntax. It was argued that no separate morphological level or component was needed in the grammar. Ways were found of describing the structure of words in a model of language that had a phonological component, a syntactic component and a semantic component but no morphological component. Those aspects of word-structure that relate to phonology

(e.g., the alternation between *sane* [seɪn] and *sanity* [sænɪti], would be dealt with using devices found in the phonological component. And those aspects of word-structure that are affected by syntax would be dealt with in the syntactic component.

The job of the syntactic component of the grammar was thought of as being to generate (i.e., to specify or enumerate explicitly) all the well-formed sentences of a language, without generating any ill-formed ones. Significantly, generating all the sentences of a language was seen as meaning generating all the permissible sequences of morphemes (not words), and showing which morpheme groupings formed syntactic constituents like noun phrases and verb phrases. A specialised morphological component and a properly, articulated lexicon were not part of the picture. Thus, Lees (1960), in the first major descriptive study produced by a generative linguist, used syntactic rules to create derived words like the noun *appointment* from the verb *appoint*. As seen in [1.4a], Lees derived the sentence containing the noun *appointment* from a source sentence with the verb *appoint*. Likewise, he derived the abstract noun *priesthood* from a source sentence with the noun *priest*, as indicated in [1.4b]:

- [1.4] a. The committee appoints John.  
The committee's appointment of John. (Lees, 1960: 67)
- b. John is a priest.  
John's priesthood. (Lees, 1960: 110)

We will not examine the particulars of the syntactic rules that Lees uses. Our concern is that Lees saw this type of word-formation as taking place in the syntax and believed that he could dispense with morphology. We will revisit this issue in Chapter 13.

Let us now turn our attention to questions of phonological realisation. **Readjustment rules** (which were morphological rules in disguise) played a key role in this area. They operated on the final output of the syntactic component, making whatever modifications were necessary in order to enable phonological rules to apply to the representation obtained after all syntactic rules had applied.

Unfortunately, there seems to have been no constraint on the power of readjustment rules. For instance, in *The Sound Pattern of English (SPE)*, which appeared in 1968 and was the pivotal work in the development of generative phonological theory, Chomsky and Halle proposed that the syntax should generate both the regular past tense form *mended* [<sub>V</sub>[<sub>V</sub> mend]<sub>V</sub>past]<sub>V</sub> and the irregular past tense form *sang* [<sub>V</sub>[<sub>V</sub> sing]<sub>V</sub> past]<sub>V</sub>. These bracketed strings, which were the output of the syntactic component, would form the input to the readjustment rules. Next, the readjustment rules would remove

all the brackets associated with the past tense. In the case of *mend*, a general readjustment rule would replace *past* by *d*, while in the case of *sing* a special readjustment rule would delete the item *past*, along with the associated bracket labels, giving  $[_V \text{ sing}]_V$ . The same readjustment rule would also attach the diacritic mark \* to the vowel /i/ indicating that eventually a phonological rule would change it into /æ/. The readjustment rules would give the forms  $[_V[_V \text{ mend}]_V d]_V$  and  $[_V s^*ng]_V$ . These representations – and all other such representations yielded by readjustment rules – were referred to as **phonological representations**. Finally, phonological representations would be converted into the phonetic representations [mɛndɪd] and [sæŋ] by rules in the phonology module.

With the benefit of hindsight, we can see that readjustment rules were a mistake. They were rules with unbridled power. They could make whatever modifications were deemed necessary to enable phonological rules to apply to strings of morphemes produced by the syntax. It is very undesirable to have a batch of rules that empower us linguists to do whatever we like, whenever we like, so long as we come up with the answer we like. A theory becomes vacuous if it has rules that can insert all manner of elements, remove all manner of elements and make all manner of elements exchange places whenever we choose to, with no principles restricting our freedom. Effectively, this means that we are given *carte blanche* to start off with any arbitrary input, apply the rules, and come up with the ‘correct’ answer.

Furthermore, readjustment rules were a bad idea because they are evidence of a lack of interest in words *qua* words and in morphology as a linguistic level. Using rules of the syntax to specify permissible sequences of morphemes, regardless of whether they occurred in words or sentences, and using readjustment rules to turn strings generated by the syntax into strings that the phonology could process and assign a pronunciation to was merely skirting round the problem. Words are a central dimension of language. They have certain unique properties that they do not share with other elements of linguistic structure like sentences and speech sounds. A theory of language must include a properly developed model of word-formation that enables the linguist to describe words on their own terms – without overlooking the ways in which word-formation rules interact with rules in other modules. As time went by, this became clear to generative linguists who, in increasing numbers, began to explore more satisfactory ways of dealing with word-structure.

### 1.3.2 The Morphology–Phonology Interaction

As regards the interaction of morphology with phonology, the selection of the form that manifests a given morpheme may be influenced by the sounds that realise neighbouring morphemes. Take the indefinite article in English. It has two manifestations. It is *a* before a word that begins with a consonant

(e.g., *a pear*) and *an* before a word that begins with a vowel (e.g., *an orange*). The same alternation occurs with the prefix *a/an* that occurs in forms such as *a-sexual* and *an-aerobic*. We cannot describe the phonological shape of the indefinite article without referring to the sound at the beginning of the word that follows it.

### 1.3.3 The Morphology–Syntax Interaction

As regards the interaction with syntax, the form of a word may be affected by the syntactic construction in which the word is used. For instance, the verb *walk* has a number of forms including *walk*, *walks* and *walked*. The selection of a particular form of this verb on a given occasion is dependent on the syntactic construction in which it appears. Thus, in the present tense, the choice between the forms *walks* and *walk* depends on whether the subject of the verb is third-person singular (in which case *walks* is selected as in *he/she/it walks*) or not (in which case *walk* is selected as in *I/you/we/they walk*). In the past tense, *walk* is realised as *walked* in all cases.

Chomsky (1957: 39) deals with all these facts as uncontroversial syntactic phenomena, using the phrase structure rule below:

$$[1.5] \quad C \rightarrow \left\{ \begin{array}{l} s \quad \text{in the context NP}_{\text{sing-}} \\ \emptyset \quad \text{in the context NP}_{\text{pl-}} \\ \text{PAST} \end{array} \right\}$$

Note: (i) ‘ $\rightarrow$ ’ stands for ‘expand’ or ‘rewrite as’; (ii) C stands for the various verbal suffixes that may be realised as *-s* (as in *walks*),  $\emptyset$  (i.e., zero) as in *walk* and *-ed* as in *walked*.

Chomsky’s analysis does not separate phrase structure rules (e.g., Sentence  $\rightarrow$  NP + VP; VP  $\rightarrow$  Verb + NP) which enumerate permissible combinations of words in phrases and sentences from rules of word-structure like the one in [1.5] that gives *walks* from *walk*. All these rules are banded together because they are concerned with enumerating permissible combinations of morphemes (see above). Note, however, that this treatment of syntactically motivated alternation in the form of words is controversial. We have merely aired the problem for the present. We postpone detailed discussion until Chapter 11.

### 1.3.4 The Morphology–Semantics Interface

Turning to semantics, the connection between morphology and the lexicon on the one hand with meaning on the other is obvious since a major role of the lexicon or dictionary is to list the meanings of words. This is because normally the relationship between a word and its meaning is arbitrary.

There is no reason why a word has the particular meaning that it has. For instance, you just have to memorise the fact that the word *faillie* refers to a kind of headdress worn in the seventeenth century. There is no way that you could discover this fact from the sounds or the structure of the word. We will come back to this topic in Section 13.3.2.

### 1.3.5 The Lexicon

It is less immediately obvious that, in addition to indicating the meaning of words and morphemes, the lexicon must also store other kinds of information relevant to the application of syntactic and phonological rules. Syntax needs to have access to **morphosyntactic properties** (i.e., properties that are partly morphological and partly syntactic) such as whether a noun is countable like *spades* or uncountable like *equipment*. This affects its behaviour in phrases and sentences. We may say *this spade* or *these spades* but we can only say *this equipment* (not *\*these equipments*).

Furthermore, some phonological rules apply to words differently depending on their morphosyntactic properties. For example, some phonological rules are sensitive to the difference between nouns and verbs. Thus, in the word *permit*, the main stress falls on the first syllable if the word functions as a noun (*permit<sub>noun</sub>*), but if it functions as a verb (*per'mit<sub>verb</sub>*), main stress falls on the last syllable. Obviously, for phonological rules that assign stress to apply correctly, access to such morphosyntactic information is essential. This information must form part of the entry of the word in the lexicon.

The study of morphology, therefore, cannot be self-contained. The structuralist doctrine of the rigid separation of linguistic levels sketched in [1.1] is untenable. True, there are some issues that are the internal concerns of morphology, but many morphological problems involve the interaction between morphology and other modules of the grammar. For this reason, much of the space in the chapters that follow is devoted to the interaction between the lexicon and morphology with the other modules.

## 1.4 ORGANISATION OF THE BOOK

The book is organised as follows:

**Part I** (Chapters 1–4) introduces basic concepts and traditional notions which are fundamental to all morphological discussions.

**Part II** (Chapters 5–10) explores the relationship between morphology, phonology and the lexicon in current generative theory, examining several models of morphology.

**Part III** (Chapters 11–13) deals with the relationship between morphology and syntax in current generative theory.

Over the years, there have been several morphological theories that have been proposed by linguists. One way of introducing you to morphology would be to present a historical and comparative survey. We could have examined various theories in turn, and perhaps compared them. Or, alternatively, we could have been polemical and proselytising, trying to persuade you that a preferred theory is the best theory. That is not what we shall do in this book.

Instead, we present you, sympathetically, but at the same time critically, with one theoretically coherent approach to morphology, namely the theory of morphology in current mainstream generative grammar. This decision is sensible not only because this is the dominant model in the field today, but also because we think it offers the most promising solutions to the perennial problems in morphological analysis.

Even so, the book is inevitably selective. We have not attempted to represent every shade of opinion within the generative school. Rather we have focused on ideas and practices that seem to form part of the emerging 'canon' in mainstream generative morphology. Obviously, to some extent this is a matter of subjective judgement. In some cases, the judgements here may not be the same as those of some other linguists.

In particular, we have only briefly discussed what is now the dominant model of phonology, Optimality Theory, in Chapter 10. The main reason for this is that the treatment of morphology in Optimality Theory remains in a rather nascent state with many questions remaining unaddressed, or even unexplored. Where current Optimality Theoretical analyses are available, principally with respect to reduplication, they will be discussed in the text.

Of course, morphological theory in current mainstream generative grammar does not enjoy a monopoly of insight. The debt owed to other approaches will be evident in this book, especially in the early chapters and in the recommended readings and bibliography.

A major feature of the book is that you will be asked to be an active investigator, not a passive reader. We have endeavoured to engage you actively and practically in *doing* morphology rather than in merely learning about its history and watching from the stalls how it is done. As you read each chapter, you are asked to pause at places and answer in-text questions and exercises before proceeding (the questions and exercises are signalled by lines across the page). Each chapter (after this one) ends with further exercises dealing with points raised in the body of the text. This insistence on getting you to analyse data is due to our firm conviction that the best initiation for anyone who wishes to become a linguist is to *do* linguistic analysis right from the start rather than to read about it.

In the text new morphological terms appear in bold type and they are explained when they are first introduced. They will also appear in the Glossary at the end of this book where key terms from other branches of linguistics are also explained. Transcriptions of Standard R.P. English are

drawn from Jones and colleagues (2003). For any other linguistic terms that are unfamiliar, a good dictionary of linguistics, such as David Crystal's *A Dictionary of Linguistics and Phonetics* (2003) or Laurie Bauer's *A Glossary of Morphology* (2004), should be consulted.

# 2 Introduction to Word-Structure

## 2.1 WHAT IS A WORD?

The assumption that languages contain words is taken for granted by most people. Even illiterate speakers know that there are words in their language. True, sometimes there are differences of opinion as to what units are to be treated as words. For instance, English speakers might not agree whether *all right* is one word or two and as a result disputes may arise as to whether *alright* is the correct way of writing *all right*. But, by and large, people can easily recognise a word of their language when they see or hear one. And normally their judgements as to what is or is not a word do coincide. English speakers agree, for example, that the form *stlody* in the sentence *The stlody cat sat on the mat* is not an English word – but all the other forms are.

### 2.1.1 The Lexeme

However, closer examination of the nature of the ‘word’ reveals a somewhat more complex picture than painted above. What we mean by ‘word’ is not always clear. As we shall see in the next few paragraphs, difficulties in clarifying the nature of the word are largely due to the fact that the term ‘word’ is used in a variety of senses that usually are not clearly distinguished. In taking the existence of words for granted, we tend to overlook the complexity of what it is we are taking for granted.

---

#### Exercise

What would you do if you were reading a book and you encountered the ‘word’ *pockled* for the first time in this context?

[2.1] He went to the pub for a pint and then pockled off.

---

You would probably look up that unfamiliar word in a dictionary, not under *pockled*, but under *pockle*. This is because you know that *pockled* is not going to be listed in the dictionary. You also know, though nobody has told you, that the words *pockling* and *pockles* will also exist. Furthermore,

you know that *pockling*, *pockle*, *pockles* and *pockled* are all in a sense different manifestations of the ‘same’ abstract vocabulary item.

We shall refer to the ‘word’ in this sense of abstract vocabulary item using the term **lexeme**. The forms *pockling*, *pockle*, *pockles* and *pockled* are different **realisations** (or representations or manifestations) of the lexeme POCKLE (lexemes will be written in capital letters). They all share a core meaning although they are spelled and pronounced differently. Lexemes are the vocabulary items that are listed in the dictionary (cf. Di Sciullo and Williams, 1987).

### Exercise

Which ones of the words in [2.2] below belong to the same lexeme?

<b>[2.2]</b>	see	catches	taller	boy	catching	sees
	sleeps	woman	catch	saw	tallest	sleeping
	boys	sleep	seen	tall	jumped	caught
	seeing	jump	women	slept	jumps	jumping

We should all agree that:

<u>The physical word-forms</u>	<u>are realisations of</u>	<u>the lexeme</u>
<i>see, sees, seeing, saw, seen</i>		SEE
<i>sleeps, sleeping, slept</i>		SLEEP
<i>catch, catches, catching, caught</i>		CATCH

<u>The physical word-forms</u>	<u>are realisations of</u>	<u>the lexeme</u>
<i>jump, jumps, jumped, jumping</i>		JUMP
<i>tall, taller, tallest</i>		TALL
<i>boy, boys</i>		BOY
<i>woman, women</i>		WOMAN

### 2.1.2 Word-Form

As we have just seen above, sometimes, when we use the term ‘word’, it is not the abstract vocabulary item with a common core of meaning, the lexeme, that we want to refer to. Rather, we may use the term ‘word’ to refer to a particular physical realisation of that lexeme in speech or writing, that is, a particular **word-form**. Thus, we can refer to *see*, *sees*, *seeing*, *saw* and *seen* as five different words. In this sense, three different occurrences of any one of these word-forms would count as three words. We can also say that the word-form *see* has three letters and the word-form *seeing* has six. And, if

we were counting the number of words in a passage, we would gladly count *see*, *sees*, *seeing*, *saw* and *seen* as five different word-forms (belonging to the same lexeme).

### 2.1.3 The Grammatical Word

The ‘word’ can also be seen as a representation of a lexeme that is associated with certain morphosyntactic properties (i.e., partly morphological and partly syntactic properties) such as noun, adjective, verb, tense, gender, number, etc. We shall use the term **grammatical word** to refer to the ‘word’ in this sense.

#### Exercise

Show why *cut* should be regarded as representing two distinct grammatical words in the following.

- [2.3] a. Usually I cut the bread on the table.  
b. Yesterday I cut the bread in the sink.

The same word-form *cut*, belonging to the verbal lexeme CUT, can represent two different grammatical words. In [2.3a], *cut* represents the grammatical word *cut*<sub>[verb, present, non-3rd person singular]</sub>, that is, the present tense, non-third person singular form of the verb CUT. But in [2.3b] it represents the grammatical word *cut*<sub>[verb, past]</sub> which realises the past tense of CUT.

Besides the two grammatical words realised by the word-form *cut* which we have mentioned above, there is a third one which you can observe in *Jane has a cut on her finger*. This grammatical word is *cut*<sub>[noun, singular]</sub>. It belongs to a separate lexeme CUT, the noun. Obviously, CUT, the noun, is related in meaning to CUT, the verb. However, CUT, the noun, is a separate lexeme from CUT, the verb, because it belongs to a different word-class (see Section 3.5 below).

The nature of the grammatical word is important in the discussion of the relationship between words and sentences and the boundary between morphology and syntax.

## 2.2 MORPHEMES: THE SMALLEST UNITS OF MEANING

Morphology is the study of word-structure. The claim that words have structure might come as a surprise because normally speakers think of

words as indivisible units of meaning. This is probably due to the fact that many words are morphologically simple. For example, *the*, *fierce*, *elephant*, *eat*, *boot*, *at*, *fee*, *mosquito*, etc., cannot be segmented (i.e., divided up) into smaller units that are themselves meaningful. It is impossible to say what the *-quito* part of *mosquito* or the *-erce* part of *fierce* means.

But very many English words are morphologically complex. They can be broken down into smaller units that are meaningful. This is true of words like *desk-s* and *boot-s*, for instance, where *desk* refers to one piece of furniture and *boot* refers to one item of footwear, while in both cases the *-s* serves the grammatical function of indicating plurality.

The term morpheme is used to refer to the smallest, indivisible units of semantic content or grammatical function from which words are made up. By definition, a morpheme cannot be decomposed into smaller units which are either meaningful by themselves or mark a grammatical function like singular or plural number in the noun. If we divided up the word *fee* [fi:] (which contains just one morpheme) into, say, [f] and [i:], it would be impossible to say what each of the sounds [f] and [i:] means by itself, since sounds in themselves do not have meaning.

How do we know when to recognise a single sound or a group of sounds as representing a morpheme? Whether a particular sound or string of sounds is to be regarded as a manifestation of a morpheme depends on the word in which it appears. So, while *un-* represents a negative morpheme and has a meaning that can roughly be glossed as ‘not’ in words such as *un-just* and *un-tidy*, it has no claim to morpheme status when it occurs in *uncle* or in *under*, since in these latter words it does not have any identifiable grammatical or semantic value, because *-cle* and *-der* on their own do not mean anything. (Morphemes will be separated with a hyphen in the examples.)

Lego provides a useful analogy: morphemes can be compared to pieces of Lego that can be used again and again as building blocks to form different words. Recurrent parts of words that have the same meaning are isolated and recognised as manifestations of the same morpheme. Thus, the negative morpheme *un-* occurs in an indefinitely large number of words, besides those listed above. We find it in *unsafe*, *unclean*, *unhappy*, *unfit*, *uneven*, etc.

However, recurrence in a large number of words is not an essential property of morphemes. Sometimes a morpheme may be restricted to relatively few words. This is true of the morpheme *-dom*, meaning ‘condition, state, dignity’, which is found in words like *martyrdom*, *kingdom*, *chiefdom*, etc. (Glosses, here and elsewhere in the book, are based on definitions in the *Oxford English Dictionary OED*.)

It has been argued that, in an extreme case, a morpheme may occur in a single word. Lightner (1975) has claimed that the morpheme *-ric* meaning ‘diocese’ is only found in the word *bishopric*. But this claim is disputed by Bauer (1983) who suggests instead that perhaps *-ric* is not a distinct

morpheme and that *bishopric* should be listed in the dictionary as an unanalysable word. We will leave this controversy at that and instead see how morphemes are identified in less problematic cases.

### Exercise

List two other words that contain each morpheme represented below.

[2.4] a. -er as in play-er, call-er  
 -ness as in kind-ness, good-ness  
 -ette as in kitchen-ette, cigar-ette

b. ex- as in ex-wife, ex-minister  
 pre- as in pre-war, pre-school  
 mis- as in mis-kick, mis-judge

- (i) Write down the meaning of each morpheme you identify. (If you are in doubt, consult a good etymological dictionary.)
- (ii) What is the syntactic category (noun, adjective, verb, etc.) of the form which this morpheme attaches to and what is the category of the resulting word?

Your answer should confirm that, in each example in [2.4], the elements recognised as belonging to a given morpheme contribute an identifiable meaning to the word of which they are a part. The form *-er* is attached to verbs to derive nouns with the general meaning ‘someone who does X’ (where X indicates whatever action the verb involves). When *-ness* is added to an adjective, it produces a noun meaning ‘having the state or condition (e.g., of being *kind*)’. The addition of the diminutive morpheme *-ette* to a noun derives a new noun which has the meaning ‘smaller in size’ (e.g., a *kitchenette* is a small *kitchen* and a *cigarette* is smaller than a *cigar*). Finally, the morphemes *ex-* and *pre-* derive nouns from nouns while *mis-* derives verbs from verbs. We can gloss the morpheme *ex-* as ‘former’, *pre-* as ‘before’ and *mis-* as ‘badly’.

So far we have described words with just one or two morphemes. In fact, it is possible to combine several morphemes together to form more complex words. This can be seen in long words like *unfaithfulness* and *reincarnation*, which contain the morphemes *un-faith-ful-ness* and *re-in-carn-at-ion* respectively. But on what grounds do we divide up these words in this fashion? In the following sections, we will examine the basis on which morphemes are identified.

### 2.2.1 Analysing Words

Up to now, we have used the criterion of meaning to identify morphemes. In many cases, forms that share the same meaning may be safely assigned to the same morpheme. Where the meaning of a morpheme has been somewhat obscure, you have been encouraged to consult a good etymological dictionary. Unfortunately, in practice, appealing to meanings listed in etymological dictionaries has its problems.

#### Exercise

Consider the following words.

[2.5] helicopter    pteropus        diptera  
          bible            bibliography    bibliophile

Historically *pter* was borrowed from Greek, where it meant ‘feather or wing’. The form *bibl-* also came from Greek where it meant ‘papyrus, scroll, book’. Do you think *pter-* and *bibl-* should be recognised as morphemes in modern English?

It seems questionable whether *pter-* is a morpheme of modern English. A *helicopter* is a kind of non-fixed-wing aircraft that most speakers of English know about; *pteropus* are tropical bats with membranous wings popularly known as ‘flying foxes’ and *diptera* are two-winged flies (which few of us who are not entomologists know about). Obviously, *pter-* does occur in modern English words that have the meaning ‘pertaining to wings’. What is doubtful is whether this fact is part of the tacit knowledge of speakers of English who are not versed in etymology. Most people probably go through life without seeing a semantic connection between ‘wings’ and ‘helicopters’.

Similarly, as we have already noted, the words *bible*, *bibliography* and *bibliophile* have to do with books. Probably many English speakers can see the book connection in *bibliography* and *bibliophile*. But it is unlikely that anyone lacking a profound knowledge of English etymology (and a classical education) is aware that the word *bible* is not just the name of a scripture book and that it contains a morpheme which is found in a number of other words pertaining to books.

Clearly, we need to distinguish between etymological information, whose relevance is essentially historical, and synchronic information that is part of speakers’ competence. Our primary task as morphologists is to investigate speakers’ tacit knowledge of the rules of their language rather than to

perform historical reconstruction. We shall discuss this further in Chapter 4.

The cases in [2.5] raise another important point: speakers may have different mental lexicons, based on their personal experience. An entomologist, for instance, may have a morpheme, *pter*, in their mental lexicon, based on their knowledge of the names of numerous insects containing this morpheme, whereas the average speaker may not.

Over-reliance on meaning in the isolation of morphemes puts us in a quandary in cases where etymological meanings are shrouded in the mists of history and lose their synchronic relevance. The common definition of the morpheme as the 'minimal meaningful unit' implies the claim that every morpheme has a readily identifiable meaning. But this is problematic. There are cases where we can justify recognising a recurrent word-building unit as a morpheme although we cannot assign it a consistent meaning.

This is true of *-fer* in words like *pre-fer*, *in-fer*, *de-fer*, *con-fer*, *trans-fer* and *re-fer*. An etymological dictionary will tell us that *fer* comes from the Latin word meaning 'bear, bring, send'. However, we would be hard-pressed to identify a consistent meaning like 'bring' attributable to *-fer* in every instance above. For this reason, some linguists, such as Aronoff (1976), have argued that it is the word in its entirety rather than the morpheme *per se* that must be meaningful. Whereas all words must be meaningful when they occur on their own, morphemes need not be. Some morphemes, like *ex-* 'former' as in *ex-wife* and *pre-* 'before' as in *pre-war*, have a transparent, unambiguous meaning while others like *-fer* do not. Their interpretation varies depending on the other morphemes that occur together with them in a word.

In view of the above, while semantic considerations must play a role in the identification of morphemes, given the pitfalls of a *purely* semantic approach, linguists tend to give a higher priority to more formal factors.

### 2.2.2 Morphemes, Morphs and Allomorphs

At one time, establishing mechanical procedures for the identification of morphemes was considered a realistic goal by structural linguists (cf. Harris, 1951; Longacre, 1964). But it did not take long before most linguists realised that it was impossible to develop a set of **discovery procedures** that would lead automatically to a correct morphological analysis. No scientific discipline purports to equip its practitioners with infallible procedures for arriving at correct theories. Creative genius is needed to enable the scientist to make that leap into uncharted waters that results in a scientific discovery. What is true of science in general is also true of linguistics (cf. Chomsky, 1957). Writing a grammar of a language entails constructing a theory of how that language works by making generalisations about its structure that go beyond the data that are observed.

Nevertheless, although there are no effective mechanical procedures for discovering the grammatical structure of a language in general or, in our case, the structure of its words, there exist reasonably reliable and widely accepted techniques that have been evolved by linguists working on morphology. These techniques are outlined in this section.

The main principle used in the analysis of words is the principle of **contrast**. We contrast forms that differ: (i) in phonological shape due to the sounds used; and (ii) in meaning, broadly defined to cover both lexical meaning and grammatical function. Thus, the phonological difference between /bɔɪ/ and /gɜ:l/ correlates with a semantic difference. The difference in meaning between the two sentences *The boy plays* and *The girl plays* is attributable to the difference in lexical meaning between /bɔɪ/ and /gɜ:l/. Likewise, the difference in grammatical function between *play-s* (present tense) and *play-ed* (past tense) is responsible for the difference in meaning between *The girl plays* and *The girl played*.

**DEFINITION:** The **morpheme** is the smallest difference in the shape of a word that correlates with the smallest difference in word or sentence meaning or in grammatical structure.

The analysis of words into morphemes begins with the isolation of **morphs**. A morph is a physical form representing some morpheme in a language. It is a recurrent distinctive sound (phoneme) or sequence of sounds (phonemes).

### Exercise

Study the data in [2.6] and identify the morphs.

- [2.6] a. I parked the car.      e. She parked the car.  
 b. We parked the car.      f. She parks the car.  
 c. I park the car.          g. We park the car.  
 d. He parks the car.        h. He parked the car.

The morphs are:

<u>Morph</u>		<u>Rekurs in</u>
/aɪ/	'I'	[2.6a] and [2.6c]
/ʃi:/	'she'	[2.6e] and [2.6f]
/hi:/	'he'	[2.6d] and [2.6h]
/ðə/	'the'	in all the examples
/kɑ/	'car'	in all the examples