

BASIC LINGUISTIC THEORY

R. M. W. DIXON

VOLUME 2

GRAMMATICAL
TOPICS

Basic Linguistic Theory 2

Basic Linguistic Theory

R. M. W. Dixon

The three volumes of *Basic Linguistic Theory* provide a new and fundamental characterization of the nature of human languages and a comprehensive guide to their description and analysis. The first volume addresses the methodology for recording, analysing, and comparing languages. Volume 3 (which will be published in 2011) examine and explain every underlying principle of grammatical organization and consider how and why grammars vary.

Volume 1

Methodology

Volume 2

Grammatical Topics

Volume 3

Further Grammatical Topics (*in preparation*)

A complete list of R. M. W. Dixon's books may be found on pp. 488–9

Basic Linguistic Theory

Volume 2 Grammatical Topics

R. M. W. DIXON

The Cairns Institute
James Cook University

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford ox2 6DP

Oxford University Press is a department of the University of Oxford.

It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

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

Published in the United States

by Oxford University Press Inc., New York

© R. M. W. Dixon 2010

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2010

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, or under terms agreed with the appropriate reprographics rights organization. Enquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

Typeset by SPI Publisher Services, Pondicherry, India

Printed in Great Britain

on acid-free paper by

CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978-0-19-957107-9 (Hbk)

978-0-19-957108-6 (Pbk)

1 3 5 7 9 10 8 6 4 2

Outline contents for whole work

Volume 1 Methodology

- 1 Basics 1
- 2 Principles to follow 57
- 3 Grammar overview 92
- 4 Analysis, argumentation, and explanation 182
- 5 Terminology 214
- 6 Doing typology 242
- 7 Phonology 264
- 8 Lexicon 289
- 9 Field linguistics 309

Volume 2 Grammatical Topics

- 10 Grammatical word and phonological word 1
- 11 Distinguishing noun and verb 37
- 12 The adjective class 62
- 13 Transitivity 115
- 14 Copula clauses and verbless clauses 159
- 15 Pronouns and demonstratives 189
- 16 Possession 262
- 17 Relative clause constructions 313
- 18 Complement clauses and complementation strategies 370

Volume 3 Further Grammatical Topics

This is planned to include the following chapters

Non-spatial setting
Noun categorization devices
Number systems
Negation
Reflexives and reciprocals
Passives, antipassives, and general intransitivizing derivations
Causatives
Applicatives
Serial verb constructions
Pivots, and switch-reference
Comparative constructions
Noun incorporation
Interrogatives
Imperatives
Culture and language, some correlations

Contents

List of tables and figures	xi
How to read this book	xii
Preface	xiii
Abbreviations and conventions	xv
10. Grammatical Word and Phonological Word	1
10.1. Types of word	1
10.2. What is a word?	3
10.3. Phonological word	7
10.4. Grammatical word	12
10.5. Clitics	20
10.6. Relationship between grammatical and phonological words	22
10.7. Interjections	27
10.8. The social and mental status of words	30
10.9. Summary	32
10.10. What to investigate	33
Sources and notes	34
11. Distinguishing Noun and Verb	37
11.1. Preliminaries	38
11.2. Major functions	41
11.3. Multiple functions or zero derivation?	46
11.4. Structures of noun phrases with different heads	50
11.5. Properties of predicates with different heads	51
11.6. Grammatical categories associated with verbs	52
11.7. Grammatical categories associated with nouns	54
11.8. Further criteria	55
11.9. Summary	57
11.10. What to investigate	59
Sources and notes	60
12. The Adjective Class	62
12.1. Parameters of variation	62
12.2. Why recognize an adjective class?	65
12.3. Criteria for recognition	70
12.4. The semantic content of adjective classes	73
12.5. Distinguishing types of adjective class	76
12.5.1. Distinguishing between adjective and verb classes	77
12.5.2. Distinguishing between adjective and noun classes	84

12.5.3. Adjectives grammatically similar to both verbs and nouns	88
12.5.4. Adjectives grammatically different from both verbs and nouns	90
12.6. Languages with restricted functional possibilities for adjectives	91
12.7. Languages with two adjective classes	93
12.8. Correlations with other grammatical parameters	96
12.9. Semantic overlapping between word classes	99
12.10. Summary	103
12.11. What to investigate	104
Appendix Distinguishing Noun, Verb, and Adjective in Fijian	108
Sources and notes	113
13. Transitivity	115
13.1. Clausal transitivity	116
13.2. Marking of core arguments	119
13.2.1. Recognizing cases	121
13.3. Transitivity classes of verbs	123
13.4. More complex types	125
13.5. The semantic bases	126
13.5.1. Identifying A	126
13.5.2. Identifying O	130
13.5.3. Ways of expressing 'giving'	134
13.5.4. Split systems of marking	137
13.5.5. Transitivity classes	142
13.6. Non-canonical marking of core arguments	147
13.6.1. ATTENTION and LIKING verbs	149
13.7. Summary	152
13.8. What to investigate	153
Appendix 1. Beyond 'accusative' and 'ergative'	153
Appendix 2. Confusing uses of terms 'unaccusative' and 'unergative'	155
Sources and notes	156
14. Copula Clauses and Verbless Clauses	159
14.1. Introduction	159
14.2. Contrasting functions of adjectives and nouns	162
14.3. Syntax	164
14.4. Relational meanings	170
14.4.1. Multiple copulas	175
14.5. Forms	177
14.5.1. Negative copulas	178
14.6. Occurrence and omission	180

- 14.7. Historical development 182
- 14.8. Summary 184
- 14.9. What to investigate 185
 - Sources and notes 186
- 15. Pronouns and Demonstratives 189
 - 15.1. The category of pronoun 189
 - 15.1.1. Person and number 191
 - 15.1.2. 'Me and you' 193
 - 15.1.3. Neutralization 199
 - 15.1.4. Gender 200
 - 15.1.5. Social niceties 201
 - 15.1.6. A further 'person', and impersonal/indefinite 203
 - 15.1.7. Grammatical properties 206
 - 15.1.8. Pronoun elaboration 207
 - 15.1.9. Bound pronouns 210
 - 15.1.10. 'Conjunct/disjunct' marking 222
 - 15.2. Demonstratives 223
 - 15.2.1. Types 225
 - 15.2.2. Forms 231
 - 15.2.3. Functions 234
 - 15.2.4. Reference 239
 - 15.3. Anaphora and cataphora 247
 - 15.3.1. Substitution and textual types of anaphora and cataphora 248
 - 15.3.2. Anaphoric and cataphoric functions of nominal demonstratives 250
 - 15.3.3. Special anaphoric forms 251
 - 15.3.4. Logophoric pronouns 252
 - 15.4. Summary 254
 - 15.5. What to investigate 256
 - Sources and notes 257
- 16. Possession 262
 - 16.1. Introduction 262
 - 16.2. Marking an NP-internal possessive construction 267
 - 16.3. The nature of the possessor (R) 271
 - 16.4. The nature of the possessive relationship 274
 - 16.5. The nature of the possessed (D) 277
 - 16.5.1. Semantic basis 278
 - 16.5.2. Marking 286
 - 16.6. Further functions of possessive markers 290
 - 16.7. Complex mechanisms 293
 - 16.7.1. 's and *of* in English 295

16.8. Internal structure of an NP which includes possession	296
16.9. Constructions which establish a possessive relationship	298
16.9.1. Using 'have'	299
16.9.2. Using a copula construction	302
16.9.3. Using intransitive verb 'exist'	303
16.9.4. Other techniques	304
16.10. Summary	305
16.11. What to investigate	306
Appendix Possessive constructions in Fijian	307
Sources and notes	310
17. Relative Clause Constructions	313
17.1. The canonical relative clause construction	314
17.2. The common argument (CA)	317
17.2.1. The nature of the CA	318
17.2.2. Allowed functions for CA in MC and in RC	319
17.2.3. Possibilities for realization of CA	326
17.3. The relative clause (RC)	338
17.3.1. Marking of a relative clause	338
17.3.2. Structure of a relative clause	348
17.3.3. Functions of a relative clause	350
17.3.4. Meanings of relative clauses	352
17.4. How to recognize a relative clause construction	354
17.5. Non-canonical constructions	356
17.5.1. The co-relative construction	356
17.5.2. 'Adjoined relative clauses'	358
17.5.3. Condensed (or fused) relative clauses	359
17.5.4. Relative clauses marked by 'to'	360
17.6. Lines of diachronic development	361
17.7. Summary	365
17.8. What to investigate	366
Sources and notes	368
18. Complement Clauses and Complementation Strategies	370
18.1. Distinguishing complement clauses	371
18.2. Grammatical criteria	375
18.3. Grammatical parameters	381
18.4. Types and meanings	389
18.5. Semantic types of verbs and varieties of complement clause	394
18.5.1. Primary-B semantic types	395
18.5.2. Secondary semantic types	399
18.6. Complementation strategies	405
18.6.1. Serial verb construction strategy	406
18.6.2. Relative clause strategy	407

18.6.3. Nominalization strategy	408
18.6.4. Complementation strategies involving linked clauses	409
18.7. Summary	413
18.8. What to investigate	415
Sources and notes	418

Glossary 422

References 433

Combined indices for Volume 1 and Volume 2

Author Index 464

Language Index 472

Subject Index 480

List of tables and figures

Tables

- 11.1. Meanings for *sun*, *moon*, and *star* when used as noun and as verb 47
- 11.2. Meanings for *mother*, *father*, and *baby* when used as noun and as verb 47
- 11.3. Verbs derived from nouns 48
- 11.4. Nouns derived from verbs 48
- 12.1. Summary of the semantic content of small and smallish adjective classes in eight languages 75
- 12.2. How often word classes occur in functional slots in Fijian 108
- 12.3. Contrasting possibilities for the three word classes in Fijian 112
- 13.1. Sample semantic types of the Verb class in English, and their roles 127
- 13.2. Illustration of the semantic parameters underlying transitivity 142
- 14.1. Outline of the semantic relations shown in copula constructions for English 159
- 14.2. Clause types 161
- 14.3. Examples of types of Identity relation in English 171
- 15.1. Forms of 2nd person pronouns in languages with different number systems 191
- 15.2. Contrastive properties of free and bound pronouns 212
- 15.3. Forms of nominal and adverbial demonstratives, illustrated for 'this' and 'here' 231
- 15.4. Anaphora and cataphora possibilities for English 249
- 15.5. Anaphoric and cataphoric functions of nominal demonstratives 251
- 16.1. Different marking of possessive constructions according to the nature of the possessor 272
- 16.2. Languages with two sets of nouns for D function 279
- 16.3. Languages with three sets of nouns for D function 281
- 16.4. Languages with four sets of nouns for D function 282
- 16.5. Gender in and on NPs involving whole-part possession in Jarawara 298
- 16.6. Possessive constructions in Boumaa Fijian 308
- 17.1. Possible functions of the common argument in the two clauses of a relative clause construction 321

Figure

- 13.1. The nominal hierarchy 138

How to read this book

This book is, of course, designed to be read from first to last page. But other strategies are possible.

Chapter 1 outlines the approach followed throughout, and should be consulted first. The lengthy Chapter 3 provides an overview of grammatical structures and systems found across the languages of the world. Many of these topics are dealt with in more detail in the chapters of Volume 2 and of the projected Volume 3. Ideally, Chapter 3 should be studied at an early stage, although experienced linguists may choose to skim it.

The remaining chapters of Volume 1, and those of Volume 2, could be read in any order. However, recurrent themes are developed across chapters and maximal benefit will be obtained by reading the chapters in the order in which they were written.

Preface

For more than four decades I have been doing linguistics in the true sense of the word—undertaking immersion fieldwork, writing grammars, compiling lexicons. I've studied, in fair detail, more than two hundred published grammars, and consulted several hundred more. I have worked—by inductive generalization—on a number of topics in typological theory, and have read everything I could lay my hands on that is relevant to this endeavour. However, despite having been learning, learning all along the way, I feel that I know only a fraction of what I would like to know.

This book is a distillation of what I have learned thus far—the most satisfactory and profitable way to work, and what pitfalls to avoid. In short, how best to obtain reliable and satisfactory results which have scientific validity. Volume 1 sets the scene, with chapters on aspects of methodology. Volumes 2 and 3 then deal in fair detail with each of a number of grammatical topics.

The reader will find opinions expressed straightforwardly, without demur. Some of the things that are said may go against certain of the current 'fashions'. I do not expect others to agree with everything I say. But all the points made here have validity, and are worthy of serious consideration.

The languages I know best are those that I have worked on myself and published on—the Australian languages Dyirbal (1972), Yidiñ (1977a, 1977b, 1991b), and Warrgamay (1981), plus Boumaa Fijian (1988a), Jarawara from Brazil (2004a), and English (1991a, 2005). If some point can be illustrated from one of these languages then I do so, rather than using data from another language which I know less well. This applies especially to the general discussions in Volume 1. For points which do not occur in these languages, and for further exemplification of points that do, information from many other grammars is used.

Sources are sometimes included in the text but more usually in notes at the end of a chapter. It has not been thought necessary to quote sources for well-described languages such as Latin, French, German, Estonian, Turkish, Hebrew, Mandarin Chinese, Quechua, Swahili, Thai, and the like. Specific references are sometimes not given for the languages I have worked on. If, say, an example is taken from Jarawara, the interested reader can easily consult my comprehensive grammar of that language (Dixon 2004a) to see how the matter under discussion fits into the overall linguistic system of the language. Sources are provided for information from other languages. There is a glossary of technical terms, included at the end of each volume.

There is today a fashion in linguistics—and no doubt in other disciplines as well—of what can be called ‘quotationitis’. That is, attempting to cite every single thing published on or around a topic, irrespective of its quality or direct relevance. Not unusually, quotations are provided from several sources which are contradictory in assumptions and import, without attention being drawn to this. I have used citations sparingly; these only reflect a small proportion of the grammars and general works which I have studied. The present work is conceived of as being like a well-organized garden; I have tried to avoid it degenerating into an impenetrable jungle.

At several places I mention the number of languages currently spoken across the world. The habit has arisen of quoting a figure of well over six thousand, which is the number of ‘language names’ listed in *Ethnologue* (Gordon 2005). This is put out by a missionary body with the main purpose of indicating where there is considered to be need for translation of the Christian Bible. The volume is uneven in scope and reliability, particularly as regards what is a language and what is a dialect (decisions on these questions frequently relate to policies concerning translation teams, and decisions may change as policies change). More than two hundred languages are listed for Australia (many labelled ‘nearly extinct’ or even ‘extinct’), but sixty would be an optimistic estimate for the number which are still actually spoken (or else well remembered). The actual number of distinct languages currently in use across the world is no more than four thousand, quite likely a fair number fewer.

This book has been envisaged, planned, and written in close collaboration with my colleague Alexandra Y. Aikhenvald. We have discussed every topic, often many times. I have benefited from her grammars of Warekena (1998), Tariana (2003), and Manambu (2008a), and from her typological studies (particularly 2000, 2004b). I am the one who has written the book (and Aikhenvald would not necessarily agree with every single word in it) but the ideas, analyses, and generalizations are in very many instances our joint work.

Nick Enfield carefully read every chapter and provided the most useful comments, corrections, and suggestions. And I owe a considerable debt to the several score students and colleagues whose grammatical descriptions I have assisted with over the years, having learnt from each of them.

These volumes have been brought to fruition through the help and encouragement of John Davey, linguistics editor *sans pareil*. Of the several publishers I have worked with over almost five decades, Oxford University Press is, in every department, the most efficient and caring. John Davey exudes an enthusiasm which makes one feel valued and wanted, and works in a friendly and unobtrusive way to assist each author in realizing their potential.

And so, I cast my pebble upon the beach.

Abbreviations and conventions, for Volumes 1 and 2

Some abbreviations are used through the book (for example, A, S, and O), others only in chapters where a particular topic is being discussed (for example, RC for relative clause).

There are abbreviations employed in interlinear glossing of examples, such as **ERG** for ergative and **CLASS** for classifier. However, where an example is short, with plenty of room on the line, a full label **ERGATIVE** or **CLASSIFIER** is written out. It would be pedantic (and otiose) to insist on always employing **ERG** and **CL** when there is no spatial limitation which requires abbreviation. My aim, through the volumes, has been to try to be as reader-friendly as circumstances permit.

-	affix boundary
=	clitic boundary
'	stress (or accent)
1	1st person
2	2nd person
3	3rd person
A	transitive subject
ABS	absolutive
ACC	accusative
AN	animate
ART	article
AUX	auxiliary
CA	common argument (shared by main and relative clauses in a relative clause construction)
CC	copula complement
CL, CLASS	classifier
CoCl	complement clause
COMP	complement clause marker
COMPL	completive

CONTIN	continuous
COP	copula
CS	copula subject
CTV	complement-taking verb (Chapter 18)
D	possessed (Chapter 16)
D	specific description in copula construction (§14.4)
DEC	declarative
DEM	demonstrative
DIM	diminutive
du, DU	dual
E	extension to core
ERG	ergative
exc	exclusive
F	focal clause (§3.11)
FEM, F, f, fem	feminine
FIN	finite
FUT	future
G	general description in copula construction (§14.4)
GEN	genitive
IMM	immediate
IMPERV	imperfective
inc	inclusive
INTERROG	interrogative
INTR	intransitive
LOC	locative
MASC, M, m, masc	masculine
MC	main clause
Mf	marker attached to focal clause (§3.11)
min	minimal
Ms	marker attached to supporting clause (§3.11)
NEG	negation
NOM	nominative
NON.FIN	non-finite

NP	noun phrase
nsg	non-singular
O	transitive object
∅	zero
PART	particle
PERF	perfect
PERFV	perfective
pl, PL	plural
POSS	possessive
PRED	predicate marker
PREP	preposition
PRES	present
R	possessor (Chapter 16)
R	specific referent in copula construction (§14.4)
REDUP	reduplicated
REL	relative clause (marker)
REP	reported
S	intransitive subject
S	supporting clause (§3.11)
Sa	‘active’ S, marked like A
sg	singular
So	‘stative’ S, marked like O
SUBORD	subordinate
SVC	serial verb construction
TAM	tense, aspect, and modality
TR	transitive
VCC	verbless clause complement
VCS	verbless clause subject
VP	verb phrase

This page intentionally left blank

Grammatical Word and Phonological Word

10.1 Types of word

The two parts of grammar—syntax and morphology—are linked together through the unit ‘word’. As described in §3.13, on morphology, the basis of a word is a lexical root (or roots) to which may be applied various morphological processes: compounding, reduplication, shift of stress or change of tone, internal change, subtraction, and—most common of all—affixation. Once all relevant morphological processes have applied, we have produced a word. The discipline of syntax deals with the organization and interrelation of grammatical elements. A clause will have a predicate, which selects a number of arguments. Corresponding to predicate will be a verb phrase, which typically has a verb as head. An argument may be realized by a bound pronoun and/or by a noun phrase, which typically has a noun as head (and optionally also adjective(s) as modifiers). Thus, words are created by morphological processes, and—alone or in combination—provide realizations for syntactic elements. As stated in §3.1, ‘word’ is the central unit, the intersection of syntax and morphology. But ‘word’ is not infrequently defined, at least in part, on phonological criteria.

What is ‘word’? P. H. Matthews commences the section ‘What are words?’ in the second edition of his seminal textbook *Morphology* (1991: 208) with: ‘there have been many definitions of the word, and if any had been successful I would have given it long ago, instead of dodging the issue until now.’

Some of the definitions which have been suggested are simple and appealing. These include Sapir’s (1921: 34) ‘one of the smallest, completely satisfying bits of isolated “meaning” into which the sentence resolves itself’ and Žirmunskij’s (1966: 66): ‘the word is the most concise unit of language, which is independent in meaning and form.’ But each of these is essentially vague; they do not provide definite criteria for deciding ‘what is a word’ in a given language.

Sweet (1875/6: 474) suggests: ‘we may, therefore, define a word as an ultimate or indecomposable sentence.’ That is, anything which is a word can make up a

complete sentence. Sweet offers as examples of this (from English) *Come!* and *Up?* (meaning ‘Shall we go up?’). However, he is then concerned over what to do with forms like English *the* and *a*, which he terms ‘half-words’.

Bloomfield (1933: 178) pursues a similar line in his definition: ‘a word, then, is a free form which does not consist entirely of (two or more) lesser free forms; in brief, a word is a *minimum free form*’ (his italics). This is probably the most oft-quoted definition of word but is, in fact, scarcely workable. As Matthews (1991: 210) points out, ‘Latin *et* “and” would normally be called a word, and so would English *my* or *the*. But are these words that could occur on their own?’ They could do so in a kind of citation (‘Did you mean *et* or *aut*?’ ‘*Et.*’) but so too could a part of a word. Matthews recalls having heard a dialogue: ‘(A) “Did you say *révisé* or *dévisé*?” (B) “*Re.*”’

Many linguists have suggested more detailed definitions, for ‘word’ in general or just in a particular language. Newman (1967: 182–3) begins his perceptive study of words and word classes in Yokuts (California) with lists of phonological and grammatical criteria, stating ‘morphological criteria serve to supplement the phonological features for delimiting the unit word.’ And Wells (1947: 99) states: ‘because of their insufficiency, the phonemic criteria of a word must be supplemented, for every or nearly every language, by criteria of the second kind ... the grammatical.’

But why should we expect phonological criteria and grammatical criteria to coincide, and identify an identical unit? Attempts to combine phonological and grammatical factors do, not unnaturally, lead to conflicts and ambiguities. Wells rightly states—working in terms of a single unit ‘word’—‘in fact, the word is most solid as a unit in those languages where phonemic and grammatical criteria reinforce each other.’

Recent work has shown that best practice is not to try to combine criteria of different types, but to apply them separately and then compare the results. That is, we should:

- (a) Recognize ‘phonological word’, determined on entirely phonological principles.
- (b) Recognize ‘grammatical word’, determined on exclusively grammatical (that is, morphological and syntactic) principles.
- (c) Compare the two units. In some languages, grammatical word and phonological word may coincide. (These will be the languages Wells refers to as: ‘where phonemic and grammatical criteria reinforce each other.’) In other languages, grammatical and phonological word will coincide in most cases, but with a number of instances where one grammatical word may consist of more than one phonological word, and/or vice versa.

In languages where grammatical and phonological word do not always coincide, they will do so in *the majority* of cases. Indeed, this is the justification for including ‘word’ in the name of each type of unit.

A methodological analogy may help explain the procedure which should be followed. As pointed out in §4.7, if in doubt about some putative contrast it is most sensible to adopt a cautious approach. When one begins to transcribe a previously undescribed language, a narrow transcription is most appropriate, noting considerable phonetic detail. The linguist observes both a dorso-velar voiceless stop [k] and a dorso-uvular voiceless stop [q]; these are recorded as such. Further work may show that these are variants of one phoneme, which can then be conveniently written as /k/. Or it may turn out that there are two contrasting phonemes, /k/ and /q/. A narrow transcription records all the phonetic details which may, as the analysis proceeds, be found to be phonologically contrastive. If the linguist had begun with a broad transcription, writing both [k] and [q] as ‘k’ (on the principles that they would be unlikely to be found to be contrastive), then—if they *were* found to be contrastive—every word transcribed with a ‘k’ would have to be reassessed to see whether it is in fact dorso-velar or dorso-uvular.

Similarly with types of words. It may be that after phonological word has been recognized on phonological criteria, and grammatical word on grammatical criteria, the two units will be found to coincide. All well and good. But in many languages they *do not quite* coincide. A linguist who has just tried to recognize a single unit ‘word’, mingling phonological and grammatical criteria, will find himself in trouble.

§§10.3–4 survey criteria for recognizing phonological and grammatical word. Before that it will be useful to examine ways in which the word *word* is used.

10.2 What is a word?

Many of the labels employed in linguistic work are entirely technical, being scarcely used in everyday language—*morpheme*, *phoneme*, *clause*. *Word* is an exception, occurring in the everyday lexicon of English (and similarly for corresponding terms in other European languages). The current meaning of *word* is, however, a recent development. In Old English its primary meaning was (a) for referring to speech, as contrasted with act or thought. There was a second sense, which may then just have been emerging: (b) what occurs between spaces in written language. In the development to modern English (b) has become the major sense—the one used in this book—with sense (a) still surviving mainly in fixed phrases such as *the spoken word*, *the written word*, *the Word of God*, *she always has the last word*, *a word of warning/advice/caution*,

Can I have a quick word with you? Similar remarks apply to corresponding terms in some other European languages, e.g. *mot* in French and *slovo* in Russian. German also has a noun, *Wort*, with these two senses, but there are here two plural forms—*Worte* for ‘speech’ and *Wörter* for ‘what is written between spaces’.

Over on the other side of the world, *vosa* in Fijian is a verb meaning ‘speak, talk’ and also a noun, with several related senses: ‘language’, ‘talk, speech’, and ‘word’. It is likely that we have here a similar line of semantic development to *word* in English.

The vast majority of languages spoken by small tribal groups (with from a few hundred to a few thousand speakers) have a lexeme meaning ‘(proper) name’ but lack a word with the meaning ‘word’. This applies to many languages from Australia, Amazonia, and New Guinea.

It has occasionally been suggested that linguistic analysis does not need a unit ‘word’ (or ‘sentence’ or ‘clause’ or ‘phrase’), the only necessary elements being ‘utterance’ and ‘morpheme’ (Harris 1946). Such an idea has not met with general acceptance. And Hockett (1944: 255) maintains that ‘there are no words in Chinese’. This seems unlikely and, in fact, the leading grammarian of Chinese, Yuen-Ren Chao, reaches a quite different conclusion. He recognizes a ‘syntactic unit’ in Chinese which satisfies our criteria for grammatical word—it has fixed internal structure but ‘unlimited versatility’ in syntactic constructions; in addition, one may pause at a word boundary, etc. (Chao 1946, 1968: 136–93).

The label ‘word’ can be applied in a number of different ways. Consider the following examples, from English and Latin, of the root or underlying form of a lexeme and its inflected forms, as used in a sentence.

root or underlying form	inflected forms	
(a) <i>look</i>	<i>look</i>	present, non-3sg subject
	<i>looks</i>	present, 3sg subject
	<i>looked</i>	past, and past participle
	<i>looking</i>	present participle
(b) <i>take</i>	<i>take</i>	present, non-3sg subject
	<i>takes</i>	present, 3sg subject
	<i>took</i>	past
	<i>taken</i>	past participle
	<i>taking</i>	present participle
(c) <i>lup-</i> ‘wolf’	<i>lupus</i>	nominative sg
	<i>lupō</i>	dative/ablative sg
	<i>lupī</i>	genitive sg, nominative pl
	etc.	

The term ‘word’ is sometimes used in reference to the root or underlying form, and sometimes in reference to the inflected forms. That is, we hear on the one hand things like ‘*look, looks, looked, and looking* are forms of the same word’, and on the other hand things like ‘the lexeme *look* is realized as word-forms *look, looks, looked, and looking*’.

Bally (1950: 287–9) is so concerned about this ambiguity of usage that he recommends abandoning the label ‘mot’ in French (and ‘word’ in English) and instead employing ‘sémantème’ for the root or underlying form and ‘molécule syntaxique’ for inflected forms. Lyons (1968: 197) prefers a different course. While recognizing that in classical grammar ‘word’ was used to mean ‘sémantème’ he notes that modern usage tends to employ ‘word’ as a label for ‘molécule syntaxique’ and suggests standardizing on this.

I have followed Lyons’s suggestion, of using ‘lexeme’ as the label for ‘root or underlying form’ and ‘(grammatical) word’ for ‘inflected form of a lexeme’. Note that Lyons uses *italics* for words and SMALL CAPITALS for lexemes—thus, the word *looked* is the past tense form of the lexeme LOOK.

Lyons’s convention is useful from another viewpoint, for dealing with lexemes that involve two words. These include phrasal verbs in English such as MAKE UP, as in *I made the story up* and *I made it up*. As discussed under (a) in §1.10, the words of this lexeme are mapped onto two non-contiguous syntactic slots—an inflected form of *make* goes into the verb slot while *up* follows the object NP. That is, the lexeme MAKE UP consists of two words, each of which has its own syntactic behaviour. If we had decided on ‘word’ as the label for lexeme, there would then be need for a separate notion of ‘syntactic word’. We would have had to say that the (lexical) word *make up* consists of two syntactic words, *make* and *up*. This is avoided by describing MAKE UP as a lexeme that consists of two (grammatical) words, an inflected form of *make* and the preposition *up*. (Similar remarks apply to phenomena such as separable preverbs in German and Hungarian.)

In summary, LOOK and TAKE are lexical elements. After applying the morphological process of affixation to root *look*, we get grammatical word *looked* (past tense form of this verb). After applying the morphological process of internal change to root *take*, we get grammatical word *took* (past tense form of this verb).

To a literate non-linguist, a ‘word’ is an orthographic unit—what comes between spaces in writing. (But note that not all writing conventions include gaps between words—for instance, the earliest writing system for Greek, and the modern system for Thai and Mandarin Chinese.) And writing conventions are unlikely to be absolutely consistent. In English, for instance, the custom is to write *cannot* as one word but the analogous *must not* as two. There appears to be no reason for this; it is just a convention of the language community.

The Bantu languages of southern Africa have a complex and agglutinative verb structure. Van Wyk (1967: 230) describes different conventions used in these languages for writing word divisions:

- (a) disjunctivism—‘according to which relatively simple, and, therefore, relatively short, linguistic units are written and regarded as words’; and
- (b) conjunctivism—‘according to which simple units are joined to form long words with complex morphological structures’.

He exemplifies with the Northern Sotho sentence ‘we shall skin it with his knife’. The two ways of writing this are:

- (a) *re tlo e bua ka thipa ya gagwe*, according to the disjunctive system; and
- (b) *retloebua kathipa yagagwe*, according to the conjunctive system.

Here *re-* is the 1pl subject prefix, *-tlo-* is the future prefix, *-e-* is a 3sg object prefix, *-bua* is the verb root ‘to skin’, *ka-* is an instrumental prefix, *thipa* is the noun ‘knife’, *ya-* is a class 9 prefix (agreeing with the class 9 noun ‘knife’), and *gagwe* is ‘his’.

In fact different orthographic strategies have been adopted for different Bantu languages. Northern Sotho, Southern Sotho, and Tswana are written disjunctively while Zulu and Xhosa are written conjunctively. There is no inherent grammatical difference between these languages; it is just that different writing conventions are followed. In the conjunctive system spaces are written between grammatical words (which may be long); in the disjunctive system spaces are written between morphemes within grammatical words. This may have been influenced by the fact that some of the prefixes are bound pronouns and case-type markers, corresponding to free pronouns and prepositions in languages such as English and Dutch—the languages of the Europeans who helped devise these writing systems—which are there written as separate words.

The orthographic conventions used for a language tend to reflect what the language was like at the time when an orthography was first adopted. For example, *knee* was pronounced with an initial *k* when English was first written. A language may undergo considerable changes, few of which get incorporated into the orthography. French, for instance, has shifted from a mildly synthetic structure to one that is now highly synthetic. A sentence such as *je ne l’ai pas vu* ‘I have not seen it’ can be considered a single word, on both grammatical and phonological criteria. But the language is—as a reflection of its history—written disjunctively, with the consequence that speakers will say that the sentence consists of five or six words (see Vendryes 1925: 87–8). This is one of the reasons why linguists have found it harder to decide ‘what is a word’

for French than for many other languages. (This point is further pursued by Matthews 2002.)

One most interesting question is: when phonological word and grammatical word fail to coincide, do speakers write a word break between phonological words, or between grammatical words? We return to this at the end of §10.6.

Units ‘phonological word’ and ‘grammatical word’ can, without doubt, be recognized for all languages. In highly synthetic languages, a grammatical word may include a number of phonological words (often, as rhythmic units).

We can now discuss typical criteria for ‘phonological word’, and then those for ‘grammatical word’, in a given language.

10.3 Phonological word

It is clear that there is no single criterion which can serve to define a unit ‘phonological word’ in every language. Rather there is a range of types of criteria such that every language which has a unit ‘phonological word’ (which is probably every language in the world) utilizes a selection of these.

We can offer the following definition:

A **phonological word** is a phonological unit larger than the syllable (in some languages it may minimally be just one syllable) which has at least one (and generally more than one) phonological defining property chosen from the following areas:

- (a) *Segmental features*—internal syllabic and segmental structure; phonetic realizations in terms of this; word boundary phenomena; pause phenomena.
- (b) *Prosodic features*—stress (or accent) and/or tone assignment; prosodic features such as nasalization, retroflexion, vowel harmony.
- (c) *Phonological rules*—some rules apply only within a phonological word; others (external sandhi rules) apply specifically across a phonological word boundary.

Note that there is likely to be a close interaction between these types of features. For example, many phonological rules, under (c), operate in terms of stress assignment within a word, under (b); the appearance of certain phonemes at certain positions within a phonological word, under (a), may be a consequence of the operation of certain phonological rules, under (c).

We can now briefly discuss these types of criteria, one at a time.

(a) Segmental features

In some Australian languages, for example, a root or suffix may have one or more syllables but every phonological word must involve at least two syllables. In Walmatjari (Hudson 1978: 37–43) a disyllabic verb root may take a zero tense-mood suffix, e.g. *luwa-∅* ‘hit!’ (the allomorph of imperative for the conjugation to which this verb root belongs is zero), whereas a monosyllabic root must take a suffix that is at least one syllable in extent, e.g. *ya-nta* ‘go!’ (here the imperative allomorph is *-nta*). In the Mbyá variety of Guaraní (Tupí-Guaraní branch of Tupí family) a monosyllabic root, when used without affixes, is obligatorily reduplicated in order to satisfy the requirement that each word have at least two syllables, e.g. root *hũ* ‘black’ becomes *hũʔhũ* as a complete word (Guedes 1991: 44, 49). In other languages each word must have at least two moras; thus, if a word is monosyllabic it must include a long vowel or a diphthong—this happens in Warekena (Arawak family; Aikhenvald 1998: 409) and in Fijian (Dixon 1988a: 25).

Looking now at segmental restrictions, there are languages in which a word-medial syllable may begin with a lateral but a word-initial syllable may not, e.g. the Western Australian language Yingkarta (Dench 1998: 15). One of the most common restrictions is that a word may not commence with *r* (as for the Arawak language Tariana, see Aikhenvald 2003: 29). And there are languages in which a word-medial syllable may end in a consonant but every phonological word must be vowel-final, e.g. the Pitjantjatjara dialect of the Western Desert language of Australia (Dixon 1980: 209). Phonotactic possibilities sometimes vary for words of different types; for example, nouns and verbs may show different phonotactic possibilities.

Trubetzkoy (1969: 275) notes that in some languages ‘certain distinctive oppositions’ occur only in initial or final position: ‘This is true, for example, for the aspirated occlusives of the Scottish-Gaelic dialect of Barra Island, the aspirated and recursive consonants of East Bengali, the recursive occlusives and emphatic palatalised consonants of Chechen.’

There are often different possibilities for sequences of phoneme types within a phonological word and across word boundaries; for example, a sequence of two vowels may only occur between words. In some Australian languages each word begins with a single consonant and ends with a vowel or a single consonant so that there can be a sequence of at most two consonants across a word boundary; however, within a phonological word there can be a sequence of three consonants (for example, *bulmbun* ‘mourning’ in Yidiñ). In contrast, Zoque allows syllables beginning in CCC and CCCC only in word-initial position (Wonderly 1951: 116). In Estonian ‘only the first syllable of a word may begin with a vowel; every non-first syllable begins with a one mora

consonant... If a word ending in a vowel is followed by a word beginning with a vowel, the occurrence of the sequence serves as a boundary marker, since only the first syllable of a word may begin with a vowel' (Lehiste 1962: 179–80).

The realization of vowel clusters between consecutive syllables may vary depending on whether the syllables belong to the same or different phonological words. In Fijian, for instance, certain vowel sequences are pronounced as diphthongs within a phonological word (e.g. the /oi/ in *boica* 'smell') but the same sequence across a phonological word boundary is pronounced as two distinct vowels e.g. *ilo.ilo* 'glass' (where '·' indicates a phonological word boundary within a grammatical word, here the boundary of an inherent reduplication).

Quite apart from the possible positioning of phonemes within a word, their phonetic realization often depends on position in a word. For example 'in Japanese, where "g" initially is realised as the voiced obstruent *g*, and medially as a nasal *ŋ*, *g* is a positive and *ŋ* a negative non-phonemic boundary signal' (Trubetzkoy 1969: 292). Similarly, the operation of certain phonological rules—see (c) below—can signal boundaries. For instance, in the Papuan language Yimas 'the final nasal plus stop cluster simplification rule only applies at the end of words' while 'initial semivowel formation only applies at the beginning of words' (Foley 1991: 80).

For the Arawak language Bare, Aikhenvald (1996) states that aspirated consonants are only found in word-initial position (most of them come from phonological rules which only apply at this place in the word, e.g. *me-haba* '3pl-fingernail' → *m^heba* 'their fingernails'). Thus, the presence of an aspirated consonant marks the beginning of a phonological word in Bare. And the presence of a nasalized vowel marks the end of a word, since this is the only structural slot in which nasalized vowels occur. In the Australian language Arrernte (Henderson 2002), the realization of vowels at word boundaries constitutes a criterion for the recognition of these boundaries. Meillet (1964: 137–40) has a useful discussion of processes applying at the ends of words in Indo-European languages (and see also Meillet 1970: 43–9). Trubetzkoy (1969: 273–97) provides an incisive discussion of boundary signals, mostly relating to the phonological word.

In some languages, words have special final features when followed by a pause. For example, in Warekena the occurrence of an *-hV* indicates that this must be the end of a phonological word that is followed by a pause (Aikhenvald 1996: 503; 1998: 411). In Semitic languages, such as Biblical Hebrew and Classical Arabic, words have distinctive forms when followed by a pause—see, for example, Gray (1934: 28–9) and Dresher (1994). The occurrence of pausal forms is never likely to constitute a necessary and sufficient

criterion for recognizing a phonological word, but can be a useful concomitant feature.

(b) Prosodic features

In very many—but not quite all—languages, stress (or accent) provides one criterion for phonological word (see Vendryes 1925: 55–6). Many languages have fixed stress—on the first or last or penultimate or antepenultimate syllable (or mora) of a phonological word. It should then be possible to ascertain the position of word boundaries from the location of stress. (See the examples given in Bloomfield 1933: 182 and Trubetzkoy 1969: 277–8.) The placement of stress may be linked to the segmental properties of phonemes; for example, in Latin stress falls on the penultimate syllable if it is long and on the antepenultimate if the penultimate is short.

In languages with contrastive stress there will generally be just one syllable with primary stress per word—see Weinreich (1954) on Yiddish, and Joseph and Philippaki-Warbuton (1987: 242–3) on Modern Greek. Although here phonological word boundaries cannot be deduced from the position of stress, one can tell from the number of stressed syllables in an utterance how many phonological words it contains (and one can deduce that a word boundary must lie somewhere between two stressed syllables).

However, in some languages stress placement may depend on a combination of morphological and phonological factors. In such cases stress may not be a useful criterion for phonological word.

A tonal system may relate to the syllable or to the phonological word—the latter applies in Lhasa Tibetan (see Sprigg 1955) and to the Papuan language Kewa (Franklin 1971; Franklin and Franklin 1978), for example.

A suprasegmental prosody such as nasalization or retroflexion will have a syntagmatic extent, and this may be a phonological word. For example, Allen (1957) provides a prosodic account of aspiration in nominals for Hāṛautī (Rajasthani) in terms of the unit ‘word’. Among his conclusions is: ‘a breathy transition is never followed or preceded by another breathy transition within the same word.’ Robins (1957) describes vowel nasality in Sundanese (Austronesian, Indonesia) as having prosodic extent. A nasal consonant engenders nasalization of a following vowel and of all subsequent vowels if separated from it only by a glottal stop or *h*; this continues until a word boundary is reached. (Robins points out that this applies to all nominal words except for loans and onomatopoeias.) In §7.5, there was a discussion of the prosody of retroflexion in Sanskrit, which applies up to the end of a word.

Vowel harmony is a prosody which operates over a certain syntagmatic extent, and this is often the phonological word. In Turkish, for instance, the

vowels in certain types of word must either be all front or all back (Bloomfield 1933: 181; Waterson 1956). Trubetzkoy (1969: 285) mentions an associated phenomenon (found in Kazakh and a number of other Turkic languages) which he calls ‘synharmonism’—a word can contain only front vowels and palatalized consonants or only back vowels and velarized consonants. This is also found in North-Eastern Neo-Aramaic (Jastrow 1997: 352–3).

(c) Phonological rules

In many languages the optimum analysis involves recognizing underlying forms for roots and affixes and then a number of phonological rules which apply to generate the surface forms. Each rule applies over a certain syntagmatic extent. Many rules apply just within the phonological word while some apply across a phonological word boundary.

We can first look at rules that only apply within a phonological word. In Hungarian, for instance, a rule of palatalization assimilates dentals *d*, *t*, *l*, or *n* to a following semi-vowel *j*, yielding the corresponding palatal sound, and the rule applies just within a phonological word (Kenesei, Vago, and Fenyvesi 1998: 438, 440; Nespore and Vogel 1986: 123–4).

At the end of §3.13, there was an account of phonological rules that apply in the Australian language Yidiñ after morphological processes of affixation have operated. Extending the discussion, we can note that some trisyllabic nominals are assigned an underlying form ending with a morphophoneme, e.g. *gajarrA* ‘brown possum’ (see also §4.9, and Dixon 1977a: 42–98; 1977b). There are the following rules that apply within a phonological word:

- (i) If a phonological word has an odd number of syllables then the penultimate vowel is lengthened.
- (ii) If a morphophoneme *A* is the last segment of a phonological word, it is omitted; otherwise it is realized as *a*.

We can compare what happens to *gajarrA* with zero suffix (for absolutive case) and with suffix *-gu* (for purposive case).

(1)	underlying form	<i>gajarrA</i>	<i>gajarrAgu</i>
	rule (i)	<i>gaja:rrA</i>	—
	rule (ii)	<i>gaja:rr</i>	<i>gajarragu</i>

A root plus monosyllabic suffix (such as purposive *-gu*) forms one phonological word. But a disyllabic suffix always commences a separate phonological word. For example, *gajarrA* ‘brown possum’ plus privative suffix *-gimbal* ‘without’ gives *gajarrA.gimbal*, a single grammatical word that consists of two phonological words (again using ‘.’ for a phonological word boundary within a grammatical word). To this can be added purposive

suffix *-gu*, which is part of the same phonological word as *-gimbal*. Rules (i) and (ii) then apply separately to the two phonological words within this grammatical word.

- | | | |
|-----|-----------------|---------------------------|
| (2) | underlying form | <i>gajarrA.gimbalgu</i> |
| | rule (i) | <i>gaja:rrA.gimba:lgu</i> |
| | rule (ii) | <i>gaja:rr.gimba:lgu</i> |

If *gajarrAgimbalgu* had been one phonological word, it would consist of an even number of syllables. Rule (i) would not apply and the surface form would be **gajarragimbalgu*; the occurring form is, in fact, *gaja:rr.gimba:lgu*.

In some languages the phonological rules that apply within a phonological word relate to stress or tone, and are thus an extension of (b).

Then there are some languages in which a special set of ‘(external) sandhi rules’ apply across word boundaries. In these languages word boundaries may be recognized partly by the operation of the sandhi rules. Allen (1972) is a detailed account of sandhi in Sanskrit. Mutation in Celtic languages is a phenomenon of the same general type (see, for example, Gregor 1980: 149–57; Ball 1993: 9–10). (Rice 1990 has a useful discussion of types of phonological rules and the syntagmatic domains over which they apply.)

Different types of criteria are relevant for defining phonological word in different languages. And the relative importance and weighting of criteria differ from language to language. For example, in some languages a rule of vowel harmony may constitute a necessary and sufficient condition for recognizing phonological words, whereas in others it may be sufficient but not necessary (for example, in Turkish not all words are subject to vowel harmony; see Bazell 1953: 67–8).

Sign languages employ a different medium of expression from their spoken cousins. Nonetheless, criteria similar to those discussed above have been enunciated. See Sandler’s (1999, 2000) discussion of phonological word in Israeli sign language, including phonological rules which operate within and across phonological words. (There is a summary of this in Zeshan 2002.)

10.4 Grammatical word

For phonological word we could offer only a number of *types of* criteria, no one of which applies in every language. In the case of grammatical word it is possible to put forward general criteria, although tempered by a number of caveats. We will discuss eight criteria, (a)–(h), one or two at a time. Note that (a)–(d) are the main criteria for identifying a grammatical word.

A grammatical word

- (a) has as its base one or more lexical roots to which morphological processes (compounding, reduplication, shift of stress, change of tone, internal change, subtraction, affixation) have applied; and
- (b) has a conventionalized coherence and meaning.

Looking first at criterion (b), this indicates that the speakers of a language think of a word as having a certain unity and character. That is, they may talk about a word (but are unlikely to talk about a morpheme). Confronted with a word like *untruthfulness*, people may talk in various ways about *true* or *truth* or *untruth* or *truthfulness* or *untruthfulness*, etc., but scarcely of *-th* or *-ness* (although they may possibly talk about the suffix *-ful*, since it is homonymous with the word *full* which has some semantic similarities, or about *un-*, since this has a clear meaning, of negation). And it must be noted that, while the meaning of a word is related to the meanings of its parts, it is often not exactly inferable from them. As pointed out in §3.13, *blackbird* refers to a particular species of bird that is black—the European common thrush, *Turdus merula*—not to any black bird. The noun *action* is a nominalization from *act* but has a shifted meaning—not every instance of ‘acting’ could be described as an ‘action’ (e.g. ‘She died midway through the act of repenting’ or ‘He acted the fool’ wouldn’t normally be).

Turning now to criterion (a), when the grammatical process involves stress shift, tone change, or change in internal form, the grammatical word is simply the product of the process applied to an underlying root; for example, grammatical word *took*, the past tense form of the verb, is the result of internal change on underlying root *take*. The recognition of a root form is, of course, an analytic decision by the linguist. In the case of subtraction, omission of some part of the root creates a grammatical form. As illustrated under (5) in §3.13, in Samoan imperative form *silā* ‘see!’ is derived from root *silaf* by deletion of the final consonant of the root.

There are various types of reduplication—see (2) in §3.13 and also §6.4. When a process of partial reduplication applies, the resulting form is easily recognized as one grammatical word; for instance *jo-joko* ‘push a bit’ from *joko* ‘push’ in Jarawara. Full reduplication may pose greater problems, especially in a language which allows repetition of a word in discourse. For instance, in Manambu (Ndu family, New Guinea; Aikhenvald 2008a) full reduplication of a verb indicates intensity and/or completeness. Thus *kabəl* ‘surround’, *kabəl.kabəl* ‘surround fully’; the reduplicated form is one grammatical word consisting of two phonological words. It is also possible, in narrative, to repeat the verb *kabəl* (which is one grammatical word and one phonological word) several times, the meaning then being ‘surround many things’; or ‘surround

a thing more than once (at different times)'. In repetition there can be two or three or more occurrences of the same word, and the repetitions could be separated by adverbs; for reduplication there can only be two occurrences of *kabəl*, and nothing can come between them. Each verb in Manambu requires an inflectional suffix; just one inflection goes at the end of reduplicated form *kabəl.kabəl*; in contrast, in a string of repeated verbs, inflection will be required on each.

When the morphological process involves compounding or affixation—in a largely agglutinative language—is it necessary to distinguish between a sequence of forms which make up one grammatical word and the same forms (or homonyms of them) which function as separate words in syntactic construction. Consider:

(3) The boys rest less than they should

(4) The boys are in a restless mood

In (4), *restless* is one word, verb *rest* plus suffix *-less* (which derives an adjectival stem). The two components of this grammatical word must occur together, in this order, with nothing intervening between them. (This grammatical word is also one phonological word.) But in (3) verb *rest* and adverb *less* are separate grammatical words which just happen to occur together in this sentence. One could add a temporal adverb after *rest* (as in *The boys rest after lunch less than they should*) or a modifier before *less* (as in *The boys rest a little less than they should*) or both (*The boys rest after lunch a little less than they should*).

This leads to further criteria:

When a **grammatical word** involves compounding or affixation, its component grammatical elements

- (c) always occur together, rather than scattered through the clause (the criterion of cohesiveness); and
- (d) generally occur in a fixed order.

We can illustrate these criteria for Dyirbal, where there are two forms with similar meanings, *bulayi* 'two' and *jarran* 'two, each of two, a pair'. One could say either of:

(5) ban yibi bulayi bani-ñu
 DETERMINER(fem) woman two come-PAST
 The two women came

(6) ban yibi jarran bani-ñu
 DETERMINER(fem) woman two come-PAST
 The two women (a pair of women) came

Dyirbal is a language with remarkably free word order. In (5) the four forms *ban*, *yibi*, *bulayi*, and *baniñu* can be permuted and occur in any order (e.g. *yibi ban baniñu bulayi*). However in (6) *jarran* must follow *yibi*; here we can only permute *ban*, *yibi-plus-jarran*, and *baniñu*. This shows that *bulayi* is a separate grammatical word, the adjective ‘two’, while *-jarran* is a nominal suffix, with dual meaning.

In many—but not all—languages, it is useful to distinguish between derivational and inflectional processes; see §3.13 and §6.3. Derivational processes, which are all optional, apply first. These may simply add semantic elements (adjective *un-happy* derived from adjective *happy*) or they may change word class (for example, noun *sens-ation* from verb *sense*). After all derivational processes have applied, a choice is made from the inflectional system appropriate for the word class of the derived stem. If no derivation had applied, the verb *sense* would choose one of the inflectional endings available for verbs in English: *-ed*, *-ing*, *-s*, or zero. The derived form *sensation* must choose from the inflectional system of number, which applies for count nouns in English: *-s* or zero.

In a language where derivational and inflectional processes may fruitfully be distinguished, we may have a further criterion:

- (e) There will be just one inflectional affix per word.

In Latin each word in an NP must show the appropriate inflection for number and case. The same applies in Dyirbal, but just for case. Harking back to (5–6) suppose that we have NPs:

- (7) *ban yibi bulayi* ‘two women’
 (8) *ban yibi-jarran* ‘a pair of women’

Now, when dative case *-gu* is added to these two NPs we get:

- (9) *bagun yibi-gu bulayi-gu* ‘for two women’
 (10) *bagun yibi-jarran-gu* ‘for a pair of women’

The dative form of the determiner *ba-n* is *ba-gu-n*, with the dative suffix *-gu* coming between root *ba-* and feminine suffix *-n*. The point to note is that in (9) noun *yibi* ‘woman’ and adjective *bulayi* ‘two’ are separate words and each takes the dative suffix *-gu*. But in (10) *yibi-jarran* is one word and it takes a single token of *-gu*, after the dual suffix *-jarran*.

In a language where inflections do not go onto every word of an NP (but only, say, onto the head, or only onto the last word or the first word) this criterion would have to be modified but could still be applicable. In a language such as Turkish or Hungarian, where number and case are separate,

obligatory suffixes, the criterion would have to be modified in a further way but again could still be applicable. The criterion may also apply with respect to inflections on verbs. As discussed earlier, it also applies for reduplication: there will be just one inflection for a reduplicated word (rather than an inflection for each reduplicand).

Concerning criterion (d), it is in fact sometimes possible for affixes to occur in alternative ordering within a word, but there must then be some semantic difference—that is, a change in (d) affects (b), the coherence and meaning of the word. Dyirbal has a rich array of derivational affixes to nouns including the dual suffix *-jarran*—as in (6), (8), and (10)—and *-gabun* ‘another’. As mentioned under (3) in §5.3, these can occur in either order with, of course, a meaning difference. Thus *yibi-jarran-gabun* is ‘another two women’ (where there have been a number of pairs of women and here is another pair) and *yibi-gabun-jarran* is ‘two other women’ (where there have been a number of women and here are two more). (See Dixon 1972: 232–3 where a further example is given.) Nedjalkov (1992) illustrates alternative orderings of the affixes ‘want to’ and ‘begin’ (‘begin to want’ versus ‘want to begin’) in Evenki (Tungus family, Russia).

Matthews (1991: 213) provides a nifty pair of examples from English of the varied order of application of derivational processes. In (11), suffixal processes *-al* and *-ize* apply before *-ation* and in (12) they apply after it:

- | | | |
|------|---|-----------------------------|
| (11) | noun root | nation |
| add | <i>-al</i> , deriving an adjective stem | nation- <i>al</i> |
| add | <i>-ize</i> , deriving a verb stem | nation- <i>al-ize</i> |
| add | <i>-ation</i> , deriving a noun stem | nation- <i>al-ize-ation</i> |
| (12) | verb root | sense |
| add | <i>-ation</i> , deriving a noun stem | sense- <i>ation</i> |
| add | <i>-al</i> , deriving an adjective stem | sense- <i>ation-al</i> |
| add | <i>-ize</i> , deriving a verb stem | sense- <i>ation-al-ize</i> |

Once all derivational processes have applied, the resulting stem takes the inflection appropriate to its word class. *Nationalization* is a derived noun and can take plural suffix *-s*; *sensationalize* is a derived verb and takes one of the inflectional suffixes available for verbs in English, *-ed*, *-ing*, *-s*, or zero.

Matthews (1991: 213) suggests a further criterion. Whereas syntactic processes are often recursive (e.g. a relative clause within a relative clause, or just saying something like *very very very good*) we find that:

- (f) Morphological processes involved in the formation of words *tend* to be non-recursive. That is, as a rule, one element will not appear twice in a word.

But, as Matthews points out, this is not a hard-and-fast principle. It does apply without exception for some languages (Latin being an example). In others, it applies most of the time. Turkish, for instance, may have a causative derivation applying twice within a given word, so that two instances of the causative suffix occur in sequence (although with slightly different forms). In Dyirbal an intransitive verb (e.g. *ñinay*- ‘sit’) can take the comitative derivational suffix *-mal-*, producing transitive stem *ñinay-mal-* ‘sit together with’. This can then be made intransitive by adding the reflexive suffix *-rriy-* and then a second token of comitative *-mal-* may be added to this, giving *ñinay-ma-rri-mal-* ‘two (people) sit with (a third)’—see Dixon (1972: 98, 246–7). And one can say things like *re-rediscover* in English, although these are highly marked.

We also find some instances of a single grammatical category being marked twice in a word. In Yiddish, plural is generally marked just at the end of a word, as in *hant* ‘hand’, diminutive *hant-l*, diminutive plural *hent-lex* (note that there is also vowel shift here). There is, however, a class of nouns where plural is marked twice, by suffix *-im* to the root and by the plural form, *-lex*, of the diminutive suffix. (Note that these are hybrid forms, including the Hebrew marker *-im* together with plural diminutive *-lex* of Germanic origin.) Thus we get (Bochner 1984: 414–15):

po ¹ yer	‘peasant’	po ¹ yer-l’	diminutive
po ¹ yer-im	plural	po ¹ yer-im-lex	diminutive plural

Aikhenvald (1999a, 2003: 176–9, 253–4) gives examples of both plural and gender being marked twice—both within a noun and within a verb—in Tariana.

We have said that a grammatical word is centred on a root or else on a combination of roots (a compound stem). In some languages there can be a variant type of grammatical word, with no root at all (or perhaps with a zero root). In the Amazonian language Jarawara, the 1sg pronoun prefix *o-* can attach to the feminine declarative suffix *-ke*, to form *o-ke*, which is both one grammatical word and one phonological word. And some verbal suffixes may be added to an auxiliary root, *-na-*, but cause the auxiliary to drop if it also bears a prefix; thus, underlying *o-na-bisa* ‘1sg-AUXILIARY-ALSO’ becomes *o-bisa*, one (phonological and grammatical) word which consists just of prefix *o-* and suffix *-bisa* (Dixon 2002b, 2004a: 31, 129–30).

In §10.3 we discussed boundary phenomena, characteristic features of the beginning and end of phonological words in particular languages. Similar features can be recognized for grammatical words. Van Wyk (1968: 554) mentions: ‘in Northern Sotho, for example, the negative morpheme *ga-* only appears on initial boundaries of verbs and the relative morpheme on the final boundaries of verbs.’ Thus, the negative prefix *ga-* always marks the beginning of a grammatical word in Northern Sotho. Similarly, in English past tense suffix

-ed (with allomorphs /-t/, /-d/, and /-ɪd/) marks the end of a verb. These are language-particular criteria which can be of great help to a linguist working on a previously undescribed language.

We can now look at two of the most quoted criteria for word, concerning the placement of pauses and the ability of words to make up complete utterances.

Bloomfield (1933: 180) and Lyons (1968: 202) lay stress on the criterion of pausing or interruptability:

(g) A speaker may pause between words but not within a word.

Bloomfield (1933: 180) exemplifies this with: 'one can say *black—I should say, bluish-black—birds*, but one cannot similarly interrupt the compound word *blackbirds*.'

This criterion should, however, be treated with caution. First, it is at best a tendency. In a synthetic language one certainly tends to pause more often between words than within words but it is by no means unheard of to pause between morphemes within a word—one does hear things like *it's very un- <pause> suitable*.

The second caveat is the most important. Pausing appears in most cases (although perhaps not in all) to be related not to grammatical word but to phonological word. In English, for instance, there are just a few examples of two grammatical words making up one phonological word, e.g. *don't, won't, he'll*. One would not pause between the grammatical words *do-* and *-n't* in the middle of the phonological word *don't* (one could of course pause between the *do* and *not* of *do not*, since these are distinct phonological words).

The places where expletives may be inserted, as a matter of emphasis, are closely related to (but not necessarily identical to) the places where a speaker may pause. Expletives are normally positioned at word boundaries (at positions which are the boundary for grammatical word and also for phonological word). But there are exceptions—for instance the sergeant-major's protest that *I won't have no more insu bloody bordination from you lot* or such things as *Cinda bloody rella* and *fan fucking tastic*. McCarthy (1982) shows that in English expletives may only be positioned immediately before a stressed syllable. What was one unit now becomes two phonological words (and the expletive is a further word). Each of these new phonological words is stressed on its first syllable; this is in keeping with the fact that most phonological words in English are stressed on the first syllable.

In highly synthetic languages—especially those with an agglutinative profile—a grammatical word may be rather long, but typically involves several phonological words. Pauses are often possible at phonological word boundaries within a grammatical word; see Russell 1999 on Cree (Algonquian,

Montana). Associated with pause is the phenomenon of ‘self-repair’. If a speaker realizes that they have made a mistake in the middle of an utterance, they are likely to pause. The mistake will have to be corrected and the utterance resumed. The interesting question is how far (if at all) one has to go back, in this process of repair. Woodbury (2002: 96–7) shows how in Cup’ik Eskimo (an extremely synthetic language) ‘if a pause or speech error occurs in the middle of a phonological word, the speaker will go all the way back to the beginning of the [phonological] word and start again’.

We can now turn to the criterion of isolatability—Sweet’s ‘ultimate or indecomposable sentence’ and Bloomfield’s ‘minimum free form’:

(h) A word may constitute a complete utterance, all by itself.

When this criterion is examined it is seen to apply not to grammatical word and not to phonological word. Rather it applies to a combination of these—to a unit which is both a grammatical word and a phonological word. Or to something which is a grammatical word consisting of a whole number of phonological words; or to something which is a phonological word consisting of a whole number of grammatical words. Woodbury (2002: 86) states that every grammatical word in Cup’ik may stand alone as a complete utterance, except for most clitics (which are one grammatical word, but not a separate phonological word).

That is, a grammatical word which is just part of a phonological word may not make up a complete utterance (e.g. *n’t* from English *don’t*). Nor may a phonological word which is part of a grammatical word (e.g. *gimba:lgu* from Yidiñ *gaja:rr.gimba:lgu* in (2) of §10.3).

Even then, criterion (h) has no more than limited applicability—to only *some* words in *some* languages, depending on the conventions for discourse organization and on other factors. Note also that, in certain speech situations, part of a word may make up a complete utterance. Matthews’s example of an utterance consisting just of ‘Re’ was mentioned in §10.1. And I have heard an airline clerk ask a passenger whether they would like a smoking or non-smoking seat, the answer being just ‘Non’.

In summary, (a–d) are the main criteria for defining a grammatical word, with caveats mentioned above. Criterion (e), distribution of inflections and (f), non-recursiveness, do apply well in certain languages. The principle of pausing/interruptability, (g), is only a tendency—which may apply more to phonological than to grammatical words—but can be a useful support for the other criteria. And (h), isolatability, is again a tendency which can be of use when it is realized that it only applies to a unit which consists of a whole number (one or more) of grammatical words and also a whole number (one or more) of phonological words.

10.5 Clitics

The unit ‘clitic’ was described, and its properties contrasted with those of ‘affix’, in §5.4. This term is typically used of something which is a grammatical word, but not a phonological word in its own right. For example, it may not have sufficient bulk to satisfy the requirements of a phonological word—a clitic may only have one mora whereas a phonological word requires at least two. A clitic is attached to a host phonological word, as a sort of optional extra. There are some items that may have the form either of a clitic or of a full phonological word. For example, *the* in English is generally a proclitic [ðə=] but can, when used contrastively, be accorded a full vowel which is stressed, [ði] (as in ‘Is that *the* man you saw yesterday?’).

Typically, a clitic is added—as an additional, unstressed syllable (or smaller unit)—to a fully articulated phonological word after stress placement and other phonological rules have applied. Consider an example from Yidiñ of verb root *warrŋgi-* ‘do all around’, past tense inflection *-ju*, and the clitic with meaning ‘now’ which has form *=la* after a vowel and *=ala* after a consonant. (It is useful to use ‘=’ for a clitic boundary.) Recall rule (i) from §10.3, which states that if in Yidiñ a phonological word has an odd number of syllables then the penultimate vowel is lengthened. A further rule, (iii)—mentioned at the end of §3.13—omits the final *-u* of past tense *-ju* from a word with an odd number of syllables. We get the following derivation (Dixon 1977a: 237):

underlying form	<i>warrŋgiju</i>	<i>=(a)la</i>
rule (i) applies to an odd-syllabled form	<i>warrŋgi:ju</i>	<i>=(a)la</i>
rule (iii) applies to an odd-syllabled form	<i>warrŋgi:j</i>	<i>=(a)la</i>
the clitic attaches	<i>warrŋgi:j=ala</i>	

If the clitic were attached to the underlying form *warrŋgiju* we would have *warrŋgijula* which has four syllables, and rules (i) and (iii) would then not apply. But these rules do apply to *warrŋgiju*, showing that *=(a)la* is added to the phonological word as the very last step in word-building, after all other rules.

Clitics may sometimes form part of a host phonological word for purposes of assignment of prosodic features (such as stress and vowel harmony) and for the application of phonological rules. And there are instances of two clitics, which occur in sequence, combining to form one phonological word. In Boumaa Fijian, for example, the preposition *i=* ‘to’ is generally a proclitic to a following noun and so is the common article *a=*, as in

<i>i=vanúa</i>	<i>a=vanúa</i>
‘to land’ (as opposed to ‘to sea’)	‘the land’

Note that primary stress goes on the syllable containing the second mora from the end of a word and secondary stress on the syllable containing the fourth mora from the end; here the clitics *i=* and *a=* bear no stress (although they are in the fourth mora from the end), showing that they are attached to the phonological word *vanúa* after the stress rule has applied.

However, when the preposition and article are used together (the article then has allomorph *na*) they make up a phonological word, which has penultimate stress:

í=na vanúa
 ‘to the [place on] land’

(See Dixon 1988a: 116, 29.) Similar ‘clitic-only’ words are reported for Tariana by Aikhenvald (2002b, 2003: 55) and for Cup’ik by Woodbury (2002).

Aikhenvald (2002b) presents a comprehensive typology of fifteen parameters in terms of which clitics vary, with discussion and exemplification of each. Just a few will be mentioned here. Some clitics attach to the first word (or the first constituent)—or to the last word—of a sentence and have scope over the whole sentence (for instance, such a clitic may mark a polar question). Others attach just to a word or phrase of a particular grammatical type, and have scope over this (for instance, with emphatic effect). Clitics may have different phonological form from roots and affixes. And clitics may combine together in special ways.

Most clitics either add on to the end of the host word (these are ‘enclitics’) or to the beginning (‘proclitics’). Some can be enclitic in one construction type and proclitic in another, as in Italian, where clitics precede an indicative verb (for example *me=lo=dici* ‘you tell me it’) but follow an imperative (*di-mme=lo* ‘tell me it!’).

In almost every case, a proclitic will precede prefixes and an enclitic will follow suffixes (as in the Yidiñ example just given). It is sometimes said that Portuguese constitutes an exception to this principle, in that a pronominal clitic can intervene between verb root and future tense suffix. However, the facts are as follows. Future tense marking in present-day Portuguese has developed from a periphrastic form involving the verbal infinitive plus an inflected form of the verb ‘have’. The form of the ‘have’ auxiliary first developed into an enclitic *=ei* which can follow an object pronominal enclitic added to the infinitive form of a verb (such as *procurar* ‘look for’), for example *procurá=lo=ei* ‘I will look for it’. An alternative is to place the object pronoun before the verb (‘it’ is then just *o*), giving *eu* (‘I’) *o* (‘it’) *procurar-ei* ‘I will look for it’. In this construction the future tense (plus 1sg subject) form *-ei* has evolved further, to be a suffix to the verb. The important point is that in *procurá=lo=ei* the *=ei* is a clitic, not a suffix. (Verbs of the form

procurá=lo=ei are still freely used in the Portuguese spoken in Portugal, but in Brazil they are confined to the written register and to a formal spoken style which deliberately reflects the conventions of writing. See Prista 1966: 60–1.)

10.6 Relationship between grammatical and phonological words

Rather few linguists, in writing grammars of languages, have clearly distinguished between phonological and grammatical words. Often, the unit word is taken for granted, with no justification or criteria offered. Sometimes criteria are offered but they may mix grammatical and phonological characteristics without clear discussion of whether these always define the same unit. However, there are sufficient clear descriptions for us to be able to recognize each of three simple types of relationship between the two kinds of word: (a) the units coincide; (b) a phonological word may consist of one or more grammatical words; and (c) a grammatical word may consist of one or more phonological words. We discuss these first, before looking at more complex relationships, in (d).

(a) Phonological and grammatical word coincide

Newman (1967) clearly distinguishes phonological and grammatical criteria in Yokuts, implying that these converge on a single unit ‘word’. A similar conclusion is explicitly stated by Czaykowska-Higgins (1998) for Moses-Columbia Salish (see the discussion at the end of this section).

A considerable search of grammars has found almost none which provide explicit criteria for phonological word and for grammatical word and state that these coincide. It may be that grammars tend only to mention instances where the two units do not coincide; or that in those languages which have been investigated from this point of view the two units never exactly coincide. More work is needed on the topic.

(b) Phonological word consists of (usually) one or (sometimes) more than one grammatical word

Many languages have clitics, which are grammatical words that do not constitute a phonological word on their own but must be attached to a phonological word primarily associated with some other grammatical word, e.g. *-n’t* as in English *mustn’t*. In Dyirbal there is a clitic *-ma* (marking a clause as a polar interrogative) which is a grammatical word that attaches—as an enclitic—to the end of the first phonological word of the sentence. For example, the interrogative version of sentence (5) in §10.4 would be *ban=ma yibi bulayi bani-nū* ‘Did the two women come?’

In Jarawara, a non-inflecting word is followed by auxiliary verb *na* to which prefixes and suffixes are attached. For example:

- (13) *ámo ó-na-hára*
 sleep 1sgS-AUXILIARY-IMMEDIATE.PAST.EYEWITNESS(fem)
 I just slept

Similar to Fijian, stress goes on the syllable including the penultimate mora, and every alternate mora before that.

Jarawara has several score suffixes but all are optional (a distinction between derivational and inflectional affixes is not useful in this language). 3sg subject is shown by a zero prefix. Thus, auxiliary *na* can occur without any affix. A phonological word must include at least two moras, and so *na* is too short to constitute a phonological word on its own. What happens is that this grammatical root attaches to the uninflecting verb which precedes:

- (14) *amó=na*
 He/she is sleeping

The interesting fact is that *na* is added to *amo* before the stress assignment rule applies. Stress goes on the *o* of *amona* (the penultimate mora of the whole word) and not on the *a* of *amo* as it does in (13).

Other examples of one phonological word consisting of more than one grammatical word are found in Tariana (Aikhenvald 2002b, 2003: 53–60), Cup'ik (Woodbury 2002), and Arrernte (Henderson 2002). Nespor and Vogel (1986) provide useful discussion of how phonological word and grammatical word boundaries do not coincide, in a number of languages. However, in no case do they provide full criteria for phonological word and grammatical word in a given language.

(c) Grammatical word consists of (usually) one or (sometimes) more than one phonological word

In Yidiñ we may find one grammatical word consisting of two phonological words; this applies both to nouns, illustrated in (2) above, and to verbs. Foley (1991: 80–7) reports a similar situation in the Papuan language Yimas.

There are a number of types of grammatical construction which typically fall under this heading. A compound is by definition one grammatical word but in many languages the components are separate phonological words. For nominal compounds in Yimas, Foley (1991: 86) notes, ‘each of the nouns in these compounds constitute a phonological word in themselves, as shown by the individual primary stresses. Yet they form one grammatical word in that there is only one inflection for number.’ Similar remarks apply for compounds in Fijian (Dixon 1988a: 22), and in Jarawara (Dixon 2004a: 29–30). Nespor and Vogel (1986: 120) state that in Turkish ‘additional evidence that the two

members of a compound do not form a single phonological word is provided by vowel harmony’.

These languages are different from English. Bloomfield’s definition of word as a ‘minimum free form’ appears to encounter difficulties with compounds such as *blackbird* since *black* and *bird* are themselves minimum free forms. He is able to argue that *bird* in *blackbird* is not the same as *bird* in *bláck bírd* since it does not bear major stress. This argument works for English. It would not be applicable to Yimas, Fijian, Jarawara, or Turkish—languages in which a compound is one grammatical but two phonological words.

In some languages with verb serialization, the verbs involved are effectively compounded together—see Foley (1991: 84–5) on Yimas. This is another typical instance where a grammatical word (the serialized verb compound) may consist of several phonological words (the individual verbs involved).

The other typical example of a grammatical word consisting of two phonological words involves full reduplication. A reduplicated form is one grammatical word (if it were not it would simply be repetition) but in many languages the reduplication boundary is also a phonological word boundary. We saw under (a) in §10.2 how a sequence of *o*-plus-*i* forms a diphthong within a phonological word in Fijian but in an inherent reduplication like *ilo.ilo* ‘glass’ each vowel is pronounced as a separate syllable. (Stress rules support this analysis—see Dixon 1988a: 24.) Similar remarks apply to Jarawara (Dixon 2004a: 30). In the Australian language Warrgamay a long vowel may only occur in the initial syllable of a phonological word. The only grammatical words with two long vowels are *ji:ji:* ‘bird (generic)’ and *bi:lbil:* ‘pee wee (*Grallina cyanoleuca*, species of bird)’ words with inherent reduplication (Dixon 1981: 17). This shows that in Warrgamay, as in many other languages, a reduplication boundary is also a phonological word boundary within a grammatical word.

(d) More complex relationships between grammatical and phonological word

There are some languages where one type of word does not necessarily consist of a *whole number* of instances of the other type. In Fijian the derivational prefix *i-* is added to a verb and derives a noun; for example, *sele* ‘to cut, slice’ → *i-sele* ‘knife’. The unusual feature is that *i-* coheres with a preceding common article *a* to form one phonological word with it:



The grammatical words are *a* and *i-sele*, but the phonological words are *ai* (pronounced as a diphthong, which only happens within a phonological word) and *séle*. Thus the grammatical word *i-sele* consists of one full phonological word (*sele*) and a part of another (*i* from *ai*) while the phonological word *ai* consists of one full grammatical word (the article *a*) and a part of another (the derivational prefix *i-* from the noun *i-sele*).

The following complex NP comes from near the end of a text, as the narrator states that he has been telling:

- (16) $\overline{a}i-$ $tú'u-tú'u$ \overline{se} *i-tàlanóa*
 ARTICLE+DERIVATION REDUPLICATED-tell OR DERIVATION-recount
 a report or a story

It involves the disjunction of two derived nouns—*i-tu'u.tu'u* ‘report’, from verb *tu'u* ‘tell (a piece of news)’, with reduplication, and *i-talanoa* ‘story’, from verb *talanoa* ‘tell (a story)’. This is one NP and the article *a* comes only once, at the beginning. The interesting point is that the *i-* of *i-tu'u.tu'u* forms a phonological word with the article, *a* (producing diphthong *ai*), whereas the second *i-*, having no article to attach to, is simply a prefix to *tàlanóa*.

Early missionaries—and later linguists—in Fiji found it hard to decide where to write the word boundary in a phrase like (15). There are three possibilities:

- (17) (i) *ai sele* (ii) *a i sele* (iii) *a ise le*

Hazlewood (1850), in his grammar, opted for (i). Churchward (1941) criticized this and preferred (ii). Then Milner (1956) went to the other extreme and used (iii). In fact there is merit in each of these alternatives: (i) shows the phonological word, (iii) the grammatical word, while (ii) simultaneously recognizes both kinds of word boundary. (See also Dixon 1988a: 21–31; 1988b.)

Another example of one type of word consisting of other than a whole number of instances of the other type of word concerns Arrernte, as described by Henderson (2002: 113–14). It relates to the VC(C) syllable structure which Henderson posits for this language.

Literate native speakers—and some linguists—think of a word as what comes between spaces in writing. Indeed, in his *Phonemics*, Pike (1947: 89) defines ‘word’ as ‘the smallest unit arrived at for some particular language as the most convenient type of grammatical entity to separate by spaces; in general, it constitutes one of those units of a particular language which actually or potentially may be pronounced by itself.’ Pike here implies that the ideal orthographic convention is to write spaces between grammatical words. The first part of his definition is circular—spaces are written around a grammatical word and a

grammatical word is what is felt to be appropriately written between spaces; that is, no independent criterion for 'grammatical word' is provided. The second part of his definition is essentially Bloomfield's 'minimum free form', discussed under (h) in §10.4.

An important question is: if there is a difference between phonological word and grammatical word, where do people prefer to insert a space—between grammatical words or between phonological words? In order to provide a fully informed answer to this we would need an array of studies for individual languages, which is not at present available. But some preliminary remarks may be offered.

In many cases people will place word boundaries around the larger unit. Thus, if a phonological word involves two grammatical words they will write spaces around the phonological word (for example, *mustn't* in English) and not between the grammatical words within the phonological word. And if a grammatical word consists of two phonological words they will write spaces before and after the grammatical word and not between the two constituent phonological words (this applies to reduplication and compounding in many languages).

But what of case (d), in Fijian, where there is no 'whole number of units' inclusion between the two kinds of word? Well, most spontaneous written material (and the Bible translation) in Fijian works in terms of alternative (i) in (17). Similarly, when speakers dictate material they say 'ái -pause -séle' and stoutly maintain that *ái* is one word and *séle* another. When the narrator of (16) was assisting in its transcription, he said *ái* <pause> *tú'u.tú'u* <pause> *sée* <pause> *i-tàlanóa*. This shows that in this instance it is the phonological word which determines word spaces (and that this is the unit which has 'psychological validity'—see §10.8).

We have discussed phonological words and grammatical words as if they were quite separate units, and then investigated the types of relationship between them. In fact, the two kinds of word are always closely intertwined. Each type of morpheme in a language is likely to have its own accentual potentiality (for example, some affixes may bear inherent stress while others lack this), so that the way in which the components of a grammatical word are combined defines its phonological status.

Phonological words of different compositions may show varying prosodic properties. In Modern Greek, if a long phonological word consists of more than one grammatical word it has an obligatory secondary stress, whereas the inclusion of a secondary stress is always optional in a long phonological word which consists of just one grammatical word (Joseph and Philippaki-Warbuton 1987: 243).

Czaykowska-Higgins (1998) presents an illuminating discussion of words in Moses-Columbia Salish, showing that although phonological word and morphological word coincide in extent, their internal structures—in terms of phonological and grammatical bracketing—differ. For example, reduplication is a grammatical process of suffixation applying to a ‘morphological root’ whereas, in terms of phonological processes, the reduplicated portion forms an inherent part of the ‘phonological root’.

For each language we can recognize a hierarchy of grammatical units; this is, typically: morpheme, grammatical word, phrase, clause, sentence. There must also be a hierarchy of phonological units; this is, typically: phoneme, syllable, foot (in some languages), phonological word, intonation group, utterance. (An alternative phonological hierarchy is suggested by Nespor and Vogel 1986 and repeated in Hall and Kleinhenz 1999: 9: syllable, foot, phonological word, phonological phrase, intonational phrase, phonological utterance.) The way in which the hierarchies relate varies from language to language. The place at which the two hierarchies are most likely to converge concerns grammatical word and phonological word—these may wholly coincide or else often coincide, for a given language. That is why it is appropriate to use the term ‘word’ for units on both hierarchies.

10.7 Interjections

An ‘interjection’ can be described as a conventionalized cry, typically indicating the speaker’s emotional response to something that has happened to them, or something which they have observed or become aware of. Interjections extend over a wide range of emotions. A partial list covering several dialects of English is:

- types of surprise—(i) *oh*; (ii) *my God*; *goodness gracious me*; *stone the crows*; *good grief*
- admiration—(i) *wow*
- joy—(i) *yippee*
- understanding, recognition—(i) *aha*
- delight—(i) *ooh*
- relief—(i) *phew*
- grief—(i) *oh dear*
- impatience—(i) *tut-tut*
- disbelief or dissatisfaction—(i) *humph*
- comment on some impediment—(i) *oops*; *oops-a-daisy*
- pain—(i) *ouch*; *ow*
- dislike for something in the environment—(i) *ugh*; *yuk*

- dislike for what has happened to speaker—(ii) *bother*; *damn*; *bloody hell*; *shit*; *fuck*
- dislike for what someone (say, John) has done, to the detriment of the speaker: (ii) *damn your eyes* (, *John*); *damn you* (, *John*); *fuck you* (, *John*)
- desire for quiet so that the speaker can hear something—(i) *ssh* (also written *sh* or *shh*)

An interjection stands outside the lexical and grammatical systems of the language. It can make up a complete utterance. Or an interjection can be in apposition to a following sentence, which describes the reason for the emotional response. For example: *Wow, that was a great goal*; or *Yippee, we've won*; or *Oh dear, I've broken it*; or *Tut-tut, you should know better than that*; or *My God, it's Lazarus come back*.

In terms of their forms, interjections fall into two broad classes, as indicated on the list above. Class (ii), many of which are multi-word items, utilize regular lexemes and grammatical elements of the language, although often with special meanings—for example, the interjection *Shit!* carries no expectation of defecation. And they may involve non-standard constructions—as in *Goodness gracious me!* They may even be productive. For example, if I stub my toe on a protruding chair-leg, I may exclaim *Damn that chair-leg!*

Class (i) has been recognized by grammarians (since Roman times) as a special word class of ‘interjections’—*aha*, *wow*, *phew*, *yuk*, and so on in English, with a corresponding set of conventionalized exclamations in every other language. There is now an obvious question to address—are these interjections each a phonological word, or each a grammatical word, or both, or neither?

Looking first at ‘grammatical word’—generally, no morphological processes can apply to an interjection. It can make up a complete utterance, and one may always pause after it. It does have a conventionalized meaning. An interjection is certainly not a normal type of grammatical word, within the linguistic system of the language.

It is a feature of interjections—in very many languages—that they have special phonetic and phonological characteristics. This may involve consonant and vowel sounds not found elsewhere in the language, unusual phonotactics, or special prosodies (of stress or tone or vowel lengthening, etc.). For example, the English interjection *phew* generally begins with a voiceless bilabial (rather than labio-dental) fricative, followed by a close front rounded vowel, [ɸy]. The interjection written as *ugh* can consist of a close back unrounded vowel followed by a voiceless dorso-velar fricative, [ux]. The interjection indicating disbelief or dissatisfaction—written as *humph*—typically consists of a sequence of voiced and voiceless bilabial nasals, [mmm] or [mm̥m] (Jones

1956). The interjection written as *ssh* or *sh* or *shh* is simply a lengthened sibilant [ʃ:].

Typically, interjections include sounds which fall outside the regular system of the language (although these do occur as phonemes in other languages). Jarawara has /s/ as a phoneme but not /ʃ/; however [ʃ] occurs in the interjection *fee* ‘oh no’, marking unpleasant surprise (Dixon 2004a: 389, where interjections involving further unusual sounds are listed). And in English the interjection indicating impatience involves a repeated alveolar click (a sound which occurs as a phoneme in some southern African and one Australian language).

The writing system for a language is devised to reflect the regular phonological system. Interjections may fall outside this, such that it can be difficult to render them in the orthography. For example, *humph* scarcely does justice to [m̥m̥m̥] or [m̥m̥m̥] and *tut-tut* is a poor representation of a sequence of alveolar clicks. Sometimes, literate speakers may adopt a spelling pronunciation, and actually say [tʌt-tʌt] in place of the alveolar clicks.

If—in a language like English—interjections were to be treated as ‘phonological words’, a number of additional members would have to be added to the vowel and consonant systems, each occurring in a single interjection. This would be an ad hoc and undesirable measure. It is best to consider interjections as lying a little outside the lexical, grammatical, and phonological structure of the language. It is thus not appropriate to enquire whether they constitute a ‘grammatical word’, or a ‘phonological word’.

Interjections can, of course, be quoted within direct speech, as in *She read his letter and exclaimed* ‘[ɑːhɑː], *now I understand why he acted as he did*’. And they can be made into verbs, as in *He’s always* [ʃ:-]ing *someone*. However, when an interjection with unusual phonetics is made into a verb, it is most likely to be converted into a normal phonological form—*He’s always shushing someone*.

There are a number of other items which are sometimes listed by grammarians within the class of interjections. All are conventionalized signs, but none are spontaneous exclamations indicating the mental states of the speaker. These include:

- (a) Short responses to a question, ‘yes’ and ‘no’. Forms of this nature are found in many—but by no means in all—languages. In a number of languages ‘no’ has the same form as clausal negator ‘not’; in others it differs.
- (b) Calls to attract attention, such as *Hey!* and *Hi!* in English. (There are often special terms for attracting the attention of various kinds of animal.)
- (c) Greetings. Every language community has a set of greetings, used when meeting someone on the road, arriving at a house, or departing. In

most cases these are conventionalized expressions from within the language, such as ‘Where are you going?’, ‘I’ve arrived’, ‘Please may I be allowed to depart’, ‘You stay, I will go’. English, and some other languages, use special greeting expressions. *Goodbye* developed out of *God be with you!* and *Hello* is said by the *OED* to have developed from the emphatic imperative of a verb ‘to fetch’ (especially used in hailing a ferryman). Be that as it may, *goodbye* and *hello* are now not felt to have any internal structure, or any function other than as greetings.

Forms of types (a), (b), and (c) do sometimes have unusual phonological form. In Dyirbal, for instance, there are only two invariable monosyllabic words—*ŋa* ‘yes’ and *ŋu* ‘alright’. The fact that all of (a)–(c) are conventionalized signs can be regarded as justification for extending the class of ‘interjections’ to include them.

(Onomatopoeic expressions—such as *miaow* as imitation of a noise made by a domestic cat—and ideophones have sometimes been classed as interjections. This is misguided. As pointed out in §8.3, ideophones may have rather different phonology from the rest of the lexicon; but they have regular function and certainly each is a grammatical word.)

10.8 The social and mental status of words

Although it is likely that all languages have words (as we have characterized ‘word’ in this chapter), their social role differs widely.

In English and other European languages (with an established tradition of writing) the word is the unit of the language about which people talk and argue. A quite different kind of unit may fulfil this role in other languages. Chao (1968: 136) explains that in Chinese a unit called *tzyh* (nowadays written *zì*) is the ‘sociological’ unit of the language, meaning by this ‘that type of unit, intermediate in size between a phoneme and a sentence, which the general, nonlinguistic public is conscious of, talks about, has an everyday term for, and is practically concerned with in various ways. It is the kind of thing which a child learns to say, which a teacher teaches children to read and write in school, which a writer is paid for so much per thousand, which a clerk in a telegraph office counts and charges so much per, the kind of thing one makes slips of the tongue on, and for the right or wrong use of which one is praised or blamed. Thus it has all the social features of the common small change of every day speech which one would call a “word” in English.’ Chao (1946: 4) mentions that *tzyh* is translated as ‘word’ by ‘most of those who speak in English on Chinese’, a footnote adding ‘such as Sinologists, missionaries, and Chinese students studying abroad’. But in fact *tzyh* is not a ‘word’ on any of the

accepted definitions; it is a character. As mentioned in §10.2, Chao provides criteria for a 'syntactic unit' in Chinese (called *cí*, see Packard 2000: 14–20) which satisfies our criteria for grammatical word (it consists of one or more *tzyh*) but states that it 'plays no role in the Chinaman of the street's conception of the subunits of the Chinese language' (1968: 138).

That is, Chinese does have 'word' but this unit has no social status for the language community. In much the same way that speakers of English and other languages talk about words, speakers of Chinese talk about *tzyh* 'characters', which roughly corresponds to the grammatical morpheme and/or phonological syllable. This social difference is undoubtedly related, at least in part, to the different writing systems employed by the Chinese and the English.

Many topics remain to be investigated in connection with the unit 'word'. Two of the most important are:

(a) Sapir (1921: 33–4; 1933) talks of the 'psychological reality' of 'word'. Does this relate to a phonological word or to a grammatical word or to both? What does it imply—that 'word' is a cognitive unit? What are the consequences of this? Does the difference between Chinese and English just described relate to a cognitive difference between the language communities?

Work is also needed on the role of 'word' in language acquisition. Studies of how children learn a language appear seldom to first establish what types of words the language has, and then to study how children acquire units of the various types.

Although phrases *can* be borrowed, the most common type of loan form is a word. But is it phonological word or grammatical word that is the basis for loans? Preliminary work by Aikhenvald (2002a: 28, 271) suggests that, at least in Amazonia, phonological word is the unit which is borrowed.

(b) Some societies plainly operate in terms of words. All over the world there is tabooing, and it is generally words that are tabooed (see Rosetti 1947: 43). In indigenous communities of Australia, when an important person dies their name may be tabooed for a while, and so too any common noun or verb which is phonologically similar to it. For example, in 1975 a man called Djaayila died at Yirrkala (in north-east Arnhem Land); his name was tabooed and so too was the verb *djaal*- 'to want, to be desirous of' (Dixon 1980: 28). Note that it is the lexeme which is here being tabooed, in the form of all words based on it. The whole question of tabooing deserves detailed study, on a cross-linguistic basis, from the point of view of the language units concerned.

Bits of language may be endowed with mystical or religious properties, but these are seldom just words. Rather, one has magical spells, pious incantations,

and the like. Even a ‘password’ is more often a phrase than a single word. It may indeed be forbidden to pronounce a certain name (e.g. ‘Jehovah’) but—as noted in §10.2—whereas it is likely that every language has a term ‘(proper) name’, only a minority of languages have a term ‘word’.

Further work is needed on which units of language are assigned mystical properties, and which types of unit are tabooed.

10.9 Summary

We have found that although many types of definition have been suggested for ‘word’, there has often been lack of a clear distinction between lexeme and word form, and/or between phonological and grammatical criteria. It is suggested that different sorts of criteria should be kept strictly apart—phonological criteria define phonological word, which is a unit in the phonological hierarchy, while grammatical criteria define grammatical word, which is a unit in the grammatical hierarchy. In some languages grammatical and phonological words coincide so that we have a single unit functioning in both hierarchies. Many languages have clitics which are grammatical words but not independent phonological words and here we have one phonological word consisting of a number of grammatical words. Other languages can have one grammatical word consisting of several phonological words (especially in compounding or reduplication). And there are examples of a grammatical word consisting of one-and-a-bit phonological words (and vice versa). Note that in all these languages phonological and grammatical word do coincide *most* of the time; it is only in a minority of cases that their borders differ.

There is no one criterion that characterizes a phonological word in every language. In §10.3 we defined a phonological word as a phonological unit larger than the syllable with at least one (and generally more than one) defining property from the following areas: segmental features, including word boundary phenomena; prosodic features such as stress and/or tone assignment and vowel harmony; the domain of application of phonological rules.

A grammatical word was defined in §10.4 as what results from applying morphological processes to a lexical root (or roots); a grammatical word has conventionalized coherence and meaning. If the components are separable, they must occur together, generally in a fixed order. Other useful criteria, in certain languages, can be the distribution of inflections, and the non-recursiveness of morphological processes.

The possibility of pause may relate to a phonological word more often than to a grammatical word. Bloomfield’s criterion of ‘minimum free form’ appears to apply to something which consists of one or more grammatical words and also of one or more phonological words (that is, not to a phonological word

which is part of a grammatical word, nor to a grammatical word which is part of a phonological word). In similar fashion, in writing there is a tendency to place a word space where there is both a phonological word boundary and also a grammatical word boundary.

10.10 What to investigate

The following questions should be addressed when working on the grammar of a language.

- Phonological word
 1. What are the criteria for phonological word in this language (in terms of those discussed in §10.3)?
 2. What are the other units in the phonological hierarchy; how does phonological word fit into the hierarchy?
- Grammatical word
 1. What are the criteria for grammatical word in this language (in terms of those discussed in §10.4)?
 2. What are the other units in the grammatical hierarchy; how does grammatical word fit into the hierarchy?
 3. What are the prosodic possibilities of units that make up a grammatical word?
- Does the language have (one or more kinds of) clitics? These should be described with respect to the parameters outlined in §10.5 (and expounded and discussed in more detail in Aikhenvald 2002b):
 - (a) The direction in which they attach to a host—proclitics before the host and enclitics after it.
 - (b) The type of host they attach to, for example—the first word in a clause, the last word in an NP, any noun.
 - (c) Their scope—a clitic marking negation or a polar question may have scope over an entire clause, while one marking emphasis or ‘also’ may have scope over a phrase or perhaps just over a word.
 - (d) Whether or not they form part of a host phonological word for assignment of prosodic features (stress, vowel harmony, etc.) and for the application of phonological rules.
- At what sort of boundaries can pauses occur? (This—and other—questions are especially important for highly synthetic languages.)
- Which types of words can make up a complete utterance?

- What is the relationship between phonological word and grammatical word?
- What is the social (and mental) status of ‘word’? Do people talk in terms of ‘words’ (in everyday discourse and/or in a fieldwork context)? If there is any tabooing in the language community, what is it that gets tabooed? (Lexemes, or something else?)
- Is there any information available on the role of phonological word and/or grammatical word in child language acquisition?

Sources and notes

This chapter is a revision (and shortening) of Dixon and Aikhenvald (2002b). The original includes a fuller history of how linguists have attempted to define ‘word’. The other chapters in Dixon and Aikhenvald (2002a)—most of which are explicitly referred to in the present chapter—provide excellent accounts of ‘phonological word’ and ‘grammatical word’ in a variety of languages.

10.1. Some of the definitions suggested for word are horrifying in their complexity and clearly infringe the principle that a definition should not be more difficult to understand than the term it purports to define. There are useful surveys of definitions of ‘word’ in Rosetti (1947), Weinreich (1954), and Ullmann (1957).

Krámský devotes a whole monograph to discussing ‘word’. He surveys past definitions and then comes up with his own (1969: 67): ‘the word is the smallest independent unit of language referring to a certain extra-linguistic reality or to a relation of such realities and characterised by certain formal features (acoustic, morphemic) either actually (as an independent component of the context) or potentially (as a unit of the lexical plan).’ Another extreme example is that by Longacre (1964: 101), a definition conceived within the formal framework of tagmemics: ‘a class of syntagmemes of a comparatively low hierarchical order, ranking below such syntagmemes as the phrase and the clause and above such syntagmemes as the stem (as well as above roots which have no external structure and are therefore not syntagmemes). It may be of greatly varied structure... Words tend to be rigidly ordered linear sequences containing tagmemes which (aside from those manifested by stems) are manifested by closed classes of morphemes unexpandable into morpheme sequences and giving only stereotyped bits of information.’

10.3. Nespor and Vogel (1986) provide some useful discussion of phonological word, but the reader should be warned that there are a good number of errors and inaccuracies in their account. For instance, on page 34 they refer to ‘Yidij, a language spoken in Central Australia’ and on page 134 to ‘Yidij, an Australian

language spoken in northeast Queensland'. The same language is referred to and in fact exactly the same data is presented on the two pages; but it is presented in a misleading manner. Nespor and Vogel say that underlying *gumari-daga-ju* becomes *gumá:ridagá:ju* after a rule of penultimate lengthening has occurred. In fact this is an intermediate stage in derivation, not an occurring form. The surface form is (after further rule application) *gumá:rdagá:p* (Dixon 1977a: 91; 1977b: 28).

10.4. Concerning the discussion under (g), note that Zwicky and Pullum (1983) suggest that *don't* and *won't* could each be analysed as one grammatical word.

'Grammatical word' forms the interface between morphology and syntax. Morphology deals with the composition of words while syntax deals with the combination of words. One could imagine slightly different words being required as ideal units for these two purposes. That is, there could be a 'morphological word' and a 'syntactic word' which would perhaps generally coincide but might not always do so. I am not aware of this sort of distinction having been fully justified for any language; but it is certainly a possibility. (Rankin et al. 2002 put forward the idea that the term 'syntactic word' could perhaps be used—in Siouan languages—for a type of word incorporating a relative clause, the whole constituting one phonological word.)

The possibility of a unit 'syntactic word' is mentioned by Di Sciullo and Williams (1987) without, however, the formulation of any explicit cross-linguistic or language-specific criteria. This question is also aired in Gak (1990). Dai (1998) establishes separate units 'syntactic word', 'phonological word', and 'morphological word' in Chinese. He suggests that a compound is one syntactic word and also one morphological word but that it may have different syntactic and morphological structures.

A number of other types of 'word' have been suggested. For example, Packard (2000: 7–14) lists: orthographic word, sociological word, lexical word, semantic word, phonological word, morphological word, syntactic word, and psycholinguistic word.

Where one may pause in natural speech is undoubtedly related to (but not necessarily identical to) where people do pause when dictating. Firth (1957: 5) suggests that one way of discovering the words of a language is 'by slow dictation, using any feeling for word-units the native [speaker] may have'. Sapir (1921: 33–4) is more definite, stating: 'no more convincing test could be desired than this, that the naive Indian, quite unaccustomed to the concept of the written word, has nevertheless no serious difficulty in dictating a text to a linguistic student word by word.' However, Bloomfield (1933: 178) puts forward a contrary opinion: 'people who have not learned to read and write, have some difficulty when, by any chance, they are called upon to

make word divisions.’ (See further discussion of this in Dixon and Aikhenvald 2002b: 12.)

There is a considerable literature on ‘self-repair’ (mostly in English), commencing most notably with Schegloff, Jefferson, and Sacks (1977).

10.6. In Fijian, a phonological word may, as illustrated here, consist of one-and-a-bit grammatical words. It may also consist of two-and-a-bit grammatical words—see Dixon (1988a: 21–2; 1988b).

In some languages, for which it is a useful unit, ‘mora’ may be included in the phonological hierarchy, between ‘phoneme’ and ‘syllable’.

10.7. Fuller discussions of interjections include Sweet (1891: 151–2, 443–4) and Ameka (1992b, 1994). The papers in Ameka (1992a) provide a varied view of interjections across a range of languages. Lists of interjections which include sounds outside the regular phonological system are provided by Evans (1995: 396–7) for the Australian language Kayardild, and by Aikhenvald (2008a) for the New Guinean language Manambu.

Distinguishing Noun and Verb

Do all languages have distinct classes of noun and verb (let alone of adjective, discussed in the next chapter)? Some grammarians have suggested that they do not. Writing of Fijian, Milner (1956: 10) states that the language has ‘a great many words which can be used either as verbs or as nouns. All such words will be called bases.’ Every kind of lexeme is referred to just as a ‘base’ throughout his grammar, terms ‘noun’ and ‘verb’ not being used again. Bloomfield (1917: 146), in his analysis of Tagalog, distinguishes between ‘particles’—which ‘either express the syntactic relations between full words or act as attributes of full words’—and ‘full words’—which ‘act not only as attributes, but also as subject or predicate, and any full word may, in principle, be used in any of these three functions’. The terms ‘noun’ and ‘verb’ nowhere occur in Bloomfield’s account. Writing of Nootka (Wakashan family, British Columbia), Swadesh (1938: 78) says: ‘normal words [words other than particles] do not fall into classes like noun, verb, adjective, preposition, but all sorts of ideas find their expression in the same general type of word, which is predicative or non-predicative according to its paradigmatic ending.’ Frachtenberg (1922a: 318) says of Hanis (Coosan family, Oregon): ‘all stems seem to be neutral, and their nominal or verbal character depends chiefly upon the suffixes with which they are used.’ However—unlike Milner and Bloomfield—Frachtenberg then refers to ‘noun’ and ‘verb’ throughout the grammar.

Should such opinions be taken to imply that there are languages for which classes of noun and verb cannot—or should not—be recognized? The answer is ‘no’. Sapir’s (1921: 119) discussion of the question concludes with: ‘no language wholly fails to distinguish noun and verb, though in particular cases the nature of the distinction may be an elusive one.’ In his authoritative article ‘Parts of speech systems’, Schachter (1985: 6–7)—himself an expert on Tagalog—states: ‘the distinction between “nouns” and “verbs” is one of the few apparently universal parts-of-speech distinctions. While the universality of even this distinction has sometimes been questioned, it now seems that the alleged counter-examples have been based on incomplete data, and that there are no languages at all which cannot be

said to show a noun–verb distinction when all relevant facts are taken into account.’

In some languages the distinction between noun class and verb class has many manifestations, in others only a few. The fact that criteria for the distinction may in a number of languages be relatively difficult to discern (Sapir’s ‘elusive’) does not mean that they are impossible to find. In a nutshell—people who say that in language X there is no distinction between noun and verb simply haven’t looked hard enough. Kinkade (1983), Jelinek and Demers (1994), and others have asserted that languages of the Salish family ‘lack a noun/verb contrast’, whereas detailed justification for such a contrast is provided by van Eijk and Hess (1986), Hébert (1983), and half a dozen others (see Sources and notes at the end of this chapter). Similarly for Tagalog and for the Wakashan languages (details are provided later in the chapter). The appendix to Chapter 12 provides a goodly array of criteria for distinguishing noun, verb, and also adjective in Fijian. These classes *must* be recognized in Fijian for a full characterization of the language.

It is amply apparent that there is no set of criteria which will serve to recognize noun and verb classes across all languages. However, it is likely that speakers of every language do operate in terms of underlying classes of noun and verb. A noun has prototypical function in an NP, which is an argument of a predicate, and a verb has prototypical function as head of a predicate. But the functions of nouns and verbs may overlap. The class of nouns always includes terms referring to concrete objects, and the verb class always includes lexemes which describe actions. But the full semantic scope of each class varies from language to language; to a limited extent, a concept which is realized through a noun in one language may be coded as a verb in another.

This chapter describes the characteristic grammatical profiles of the two word classes. It examines how properties which generally apply to a noun may also be applicable to a verb, but with considerable limitations. And vice versa—if a noun enters into a function which typically applies to a verb, there will be restrictions associated with this. First, though, we need to reiterate some basic principles concerning the nature of noun and verb, and of a predicate and its arguments.

11.1 Preliminaries

In every language, word classes must be recognized on the basis of grammatical criteria internal to that language. The nature of the criteria is likely to depend on the structural profile of the language. The exemplification of this for Latin and English, provided in §1.8, can profitably be repeated here.

For Latin, we recognize three lexical word classes, with the following properties:

- class A, inflects for case and number
- class B, inflects for case, number, and gender
- class C, inflects for tense, aspect, mood, person, and number

For English, we also recognize three lexical word classes and here the criteria are:

- class X, takes suffix *-ing*
- class Y, may be immediately preceded by an article and need not be followed by another word
- class Z, may be immediately preceded by an article and is then followed by another word (either one from class Y or another word from class Z)

The lexemes belonging to each of these classes show a certain range of meaning. They also have typical behaviour in filling functional slots within a clause. It is because of a measure of similarity of meaning and function that we may identify word classes between languages, and use the same label for them.

Cross-linguistically, word classes ‘noun’ and ‘verb’ can be characterized as follows.

Noun

- *Function*. Can always occur in an NP, which is an argument of a predicate. In some languages it has a secondary function as head of predicate.
- *Semantics*. The class of nouns always includes words referring to concrete objects (and their parts), such as ‘tree’, ‘stone’, ‘star’, ‘woman’, ‘foot’, ‘water’, ‘axe’.

Verb

- *Function*. Can always occur as head of a predicate. In some languages it has a secondary function in an NP.
- *Semantics*. The class of verb always includes words referring to actions, such as ‘jump’, ‘sit’, ‘burn’, ‘eat’, ‘laugh’, ‘talk’, ‘see’.

On this basis, classes A and Y are both identified as ‘noun’ and C and X as ‘verb’ (and B and Z as ‘adjective’, discussed in the next chapter). Note that the criteria employed for recognizing lexical word classes are different for the two languages. Latin has a rich morphology but no strict ordering of words within a phrase and within a clause. English has rather little morphology but fairly strict rules of ordering.

As emphasized in §1.8, it is important to bear in mind that, although word classes have similar semantic content between languages, the full ranges of meanings they cover are never identical. The central members—as exemplified above—are likely to correspond (although there is no guarantee that every single one will). But there can be considerable variation among non-central members. Kin relations are expressed just by nouns in some languages but by a mixture of verbs and nouns in others. The idea of needing to eat may be expressed through noun (‘hunger’) or verb (‘be hungry’) or adjective (‘hungry’), depending on the language. In essence, it is *not* possible to decide which class a word belongs to in a given language solely on the basis of its meaning. If this were the case then the word for ‘needing to eat’, or for a male parent, would be in the same word class for every language, which they are not.

Another point which needs to be repeated—from (c) in §3.2—is the critical difference between a word being used as head of an intransitive predicate—as *widen* is in (1)—and as the copula complement argument within a copula construction, as *wide* is in (2).

(1) [The river]_S [widens]_{INTRANSITIVE.PREDICATE} below the waterfall

(2) [The river]_{CS} [is]_{COPULA.PREDICATE} [wide]_{CC} below the waterfall

There has arisen the bad habit of referring to *wide* in (2) as a ‘nominal predicate’. It is not a predicate if ‘predicate’ is being used as a grammatical label. The copula verb *is*—which carries tense—functions as predicate in (2); it takes two arguments, copula subject *the river* and copula complement *wide*.

In these examples, intransitive verb *widen* is derived from adjective *wide* through the addition of suffix *-en*. But *-en* may only be added to a form ending in one of a limited set of stops and fricatives (*p, t, k, d, f, s, j, θ*). It cannot be added to *narrow*, so that corresponding to (1–2) we get:

(3) [The river]_S [narrows]_{INTRANSITIVE.PREDICATE} above the waterfall

(4) [The river]_{CS} [is]_{COPULA.PREDICATE} [narrow]_{CC} above the waterfall

People who talk of ‘nominal predicate’ would say that *narrow* has predicate function in both (3) and (4), which fails to recognize a critical grammatical distinction.

The English lexeme *comb* is a noun, which can be used as head of an NP in S or A or O or CS function, or in copula complement function, as in:

(5) [This]_{CS} [is]_{COPULA.PREDICATE} [a comb]_{CC}

A secondary function of *comb* is as head of a transitive predicate, as in

(6) [The actress]_S [combed]_{TRANSITIVE.PREDICATE} [her hair]_O

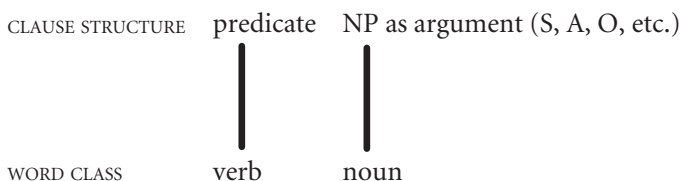
In (5), *comb* is head of an NP which is an argument of the predicate (here a copula). This must be clearly distinguished from its secondary function as transitive predicate, in (6).

Throughout this book, the term ‘predicate’ covers intransitive predicate, transitive predicate, and the copula of a copula construction. It does *not* include copula complement.

11.2 Major functions

The easiest distinction between what is a noun and what is a verb comes in languages where a noun can only occur in an NP and a verb only as head of a predicate. Repeating the diagrams from §3.5, this can be shown as:

I Canonical scheme

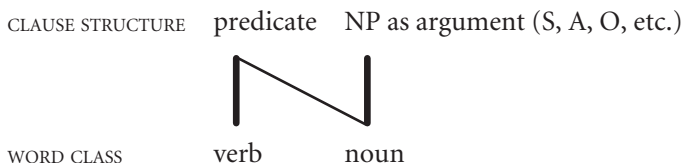


Very many languages show this scheme—Latin, German, and other Indo-European languages, Dyirbal and other Australian languages, plus languages from many other families and areas.

Scheme I is likely to be encountered in languages where all noun roots, or all verb roots, or both, are bound. It is especially likely when there is fusion between root and affix, as in Latin and German.

We can now look at variants on the canonical system in I. These are: II, when a noun can also be head of a predicate; III, when a verb can also be head of an NP; and IV, when both apply.

II Noun can also be head of predicate



The main function of a noun (shown by the thick line) is as in scheme I, within an NP which is a predicate argument. But it may have a secondary function as well (shown by a thin line), as head of a predicate. Generally, nouns are restricted to intransitive predicates.

There are a fair number of languages with scheme II, where either a verb or a noun may be head of a predicate (but a verb may not be head of an NP). For example, in Nenets (Samoyed branch of Uralic, Siberia; Hajdú 1963: 67) ‘the substantive [noun] can also be the predicate of the sentence, and in this function it can take verbal person suffixes’ (examples were given in §3.5). In Mao Naga (Tibeto-Burman, India; Giridhar 1994: 262) ‘the verb ... is defined as a word which can potentially take mood markers [he uses this term to cover both mood and modality, see §11.6]. Tense and aspect are structural properties not of the verb but of the predicate, which can be either a noun or a verb.’ For Korku (Munda branch of Austro-Asiatic, India; Nagaraja 1999: 30) we read: ‘while some nouns and almost all adjectives can function as verbs [i.e. as intransitive predicates] by taking tense markers, etc., the verbs themselves cannot be used conversely.’

When a language does not accord with scheme I, there can be a tendency to deny that there is a distinction between noun and verb classes. Frachtenberg’s (1922a: 318) assessment of Hanis (Coos) was mentioned before: ‘all stems seem to be neutral, and their nominal or verbal character depends chiefly upon the suffixes with which they are used.’ However, perusal of his grammar suggests that this language most probably employs scheme II.

A cursory examination of Fijian suggests that the head of an intransitive predicate can be a verb, a noun, an adjective, or a pronoun. (The head of a transitive predicate may only be a verb bearing a transitive suffix). Closer study shows the correct analysis to be: an NP can function as head of an intransitive predicate. That is, the predicate head may consist of noun plus adjective, of noun plus demonstrative, of possessive pronoun plus noun, etc.; minimally, it can be just a noun. An NP in argument function must commence with an article: with common article *a* ~ *na* if the NP head is a common noun or with proper article *o* if the head is a proper noun (the name of a person or place) or a pronoun. When an NP functions as head of an intransitive predicate, a common article is omitted but a proper article retained (Dixon 1988a: 65).

III Verb can also be head of NP as predicate argument

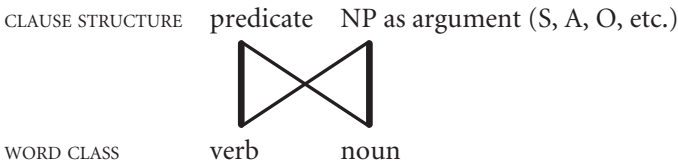
CLAUSE STRUCTURE predicate NP as argument (S, A, O, etc.)



Scheme III is found in Chinese (described in §3.5). It appears that it is less common than scheme II. This correlates with the fact that, cross-linguistically,

derivational processes for forming noun stems from verb roots are more frequent than processes for deriving verb stems from noun roots. Since it is quite common for a noun to be used as head of a predicate, there is less need for a verbalization process which creates verbs out of nouns. And since it is relatively uncommon for a verb to function as head of an NP, there is consequently more need for explicit processes of nominalization, which derive a noun from an underlying verb.

IV Noun can also be head of predicate, and verb can also be head of NP as predicate argument



This is what has been reported for languages of the Salish family, and for Nootka and other languages of the Wakashan family (these families are located in the same region, spanning south-west Canada and north-west USA). The illustrative sentences in Nootka (from Swadesh 1938: 78; Jacobsen 1979: 85) can be repeated here from (3–4) in §3.5:

- (7) [ʔi·h-ma']_{INTRANSITIVE.PREDICATE} [qo'ʔas-ʔi']_s
 be.large-3sg.INDICATIVE man-ARTICLE
 The man is large
- (8) [qo'ʔas-ma]_{INTRANSITIVE.PREDICATE} [ʔi·h-ʔi]_s
 man-3sg.INDICATIVE be.large-ARTICLE
 The large one is a man

One of the languages most often cited as having no distinction between noun and verb classes is Mundari (Munda branch of Austro-Asiatic, India). This is based entirely on the work of Hoffmann (1903, 1905); careful study shows that Hoffmann's assertion was unfounded. One can, speculatively, reconstruct the train of thought of this German missionary (and no doubt of some others who have denied a distinction between word classes in a variety of languages). Languages such as German and Latin—with which Hoffmann would have been familiar before coming to India—adhere to Scheme I, in which a noun can only function in an NP and a verb only in a predicate. He could have assumed that this is the universal behaviour of nouns and verbs. Mundari certainly doesn't fit the scheme. Ergo, it does not distinguish between noun and verb classes.

In a convincing study, ‘Mundari: the myth of a language without word classes’, Evans and Osada (2005) put forward the criterion of ‘bidirectionality’. For a language to qualify for scheme IV, *every* noun must have secondary function as head of an intransitive predicate, and *every* verb must have secondary function as head of an NP. They find (p. 376) that ‘though we can use some verbs freely as arguments, the vast majority must effectively be converted into headless clauses before being placed in an argument slot’. There are also semantic differences between a given lexeme as used in NP and in predicate functions, suggesting that lexical derivation (with zero realization) is at work here, rather than just functional freedom. (This is further discussed in §11.4.)

Hoffmann appears to have moderated these views in later years. His sixteen-volume *Encyclopaedia Mundarica* (1930–79) provides information on syntactic function for each lexeme. From examination of a sample of 3,824 items (about 10 per cent of those in the *Encyclopaedia*), Evans and Osada find that 772 (20 per cent) are given only as nouns, 1,099 (29 per cent) just as verbs, with only 1,953 (51 per cent) being shown as having two functions. A smaller sample of 105 lexemes is used for comparison with English. Of the 41 lexemes which are basically verbs in Mundari, only 27 (66 per cent) may be used as NP head, and of the 64 basic nouns, only 47 (73 per cent) can function as predicate head. Figures for corresponding lexemes in English are 67 per cent and 65 per cent respectively. That is, as good a case can be made out for English having scheme IV as for Mundari. But in fact English, as discussed in §11.4, shows a quite different profile.

Eskimo—varieties of which are nowadays called Cup’ik, Yupik, Inuktitut, Inupiatun, Kalaallisut—has also been a victim of the ‘having no word classes’ accusation; Sadock (1999) provides a perceptive account. The excellent early grammar by Kleinschmidt (1851) ‘is organized around the difference between nouns and verbs’. But then Thalbitzer (1911: 1006) denied this, stating, ‘the lines of demarcation between classes of words are vague’. Sadock provides a plausible explanation. Thalbitzer was interested in the psychology of language, and to further this interest decided to study a ‘primitive language’, Eskimo being chosen for this purpose. A primitive language (on a priori judgement) requires primitive qualities, such as the failure to distinguish nouns and verbs. Other Eskimologists have noted a ‘sharp formal contrast’ between noun and verb classes. Sadock concludes that ‘it is hard to imagine a poorer choice of a language group to accuse of not having fundamental part-of-speech distinctions than Eskimo’.

Sadock states that Eskimo does have ‘a fair number of homonymous roots and stems that (on one meaning) accept case-endings and (on another meaning) mood signs’. Describing Samoan (Oceanic branch of Austronesian),

Mosel and Hovdhaugen (1992: 77) say that ‘many, perhaps the majority of, roots can be found in the function of verb phrase [predicate] and noun phrase nuclei’. On this basis, they maintain that ‘the categorisation of words into nouns and verbs is not given a priori in the lexicon. It is only their actual occurrence in a particular syntactic environment which gives them the status of a verb or a noun.’ One infers from this that many (possibly the majority of) roots *are* classifiable as nouns or as verbs. Mosel and Hovdhaugen then say that some of these indeterminate items ‘predominantly function as verbs, whereas others are more likely to be found in the function of nouns’. And, a little later, they describe how ‘Samoan has some affixes which derive nouns and verbs (i.e. words which are nearly exclusively used as nouns or verbs)’.

A number of other Oceanic languages are like Samoan. For the North-East Ambae language (spoken on Vanuatu), Hyslop (2001: 72, 91) describes two open classes—of verbs (including adjectives as a subclass) and of nominals. And she states that a minority of lexemes can function as NP head or as predicate head, so that ‘it is difficult to state categorically whether they belong essentially in the class of nouns or verbs’. It is not the case in North-East Ambae, or in Samoan, or in Eskimo, that a distinction between noun and verb classes cannot be made. The great majority of lexemes have primary function either within an NP or in a predicate. In Samoan and in North-East Ambae, it appears that there are a limited set of lexemes which can occur with equal facility in both functional slots. This can be described in one of a number of ways—by saying that there is a degree of overlap between noun and verb classes, or that there are a limited set of noun/verb homonyms.

The most touted candidates for scheme IV are languages from the neighbouring Wakashan and Salish families. One would want to be fully convinced that Evans and Osada’s criterion of ‘bidirectionality’ holds. In these—as in all other—languages, there are some lexemes which occur most often as head of an NP. Is it the case that *every single one* of these may also have secondary function as head of a predicate? And there are lexemes which are most frequently found as head of a predicate; can *each* of these also function as head of an NP? Whatever the results of such a searching enquiry, there are in these languages other criteria for distinguishing classes of noun and verb. When a noun is head of a predicate it often has properties much more restricted than those for a verb as predicate head. Similarly, when a verb is head of an NP, the possibilities for NP structure are likely to be different from when a noun is NP head. These matters will be discussed in §11.4. But first we need to look at English, which exhibits a rather different character from those in schemes I–IV.

In English, many nouns can be used—without any obvious derivational process having applied—as predicate; for example, *boss*, *market*, and *knot*. And many verbs can be used—without any obvious derivational process having applied—as head of an NP; for example, *witness*, *love*, and *concern*. But this is not a productive process, applying to all nouns, or to all verbs (or even to all of a certain semantic type). Typically, out of a certain set of nouns, only that with the most generic meaning may also be used as a verb; for example, there is a verb *to house*, but no verbs relating to nouns *cottage*, *bungalow*, or *apartment* (Dixon 2005: 57–8).

(9) [e dredre]_{INTRANSITIVE.PREDICATE} [a tuuraga]_s
 3sgS laugh[**VERB**] ARTICLE chief[**NOUN**]
 The chief is laughing

- Noun *tuuraga* 'chief' has the same function in the predicate of (10) as verb *dredre* 'laugh' does in the predicate of (9). Identical markers of person and number of subject, of tense, and of many other semantic features, have the same possibilities for inclusion in both predicates.

In Fijian, *every* noun can be head of an intransitive predicate, if it would be semantically plausible for it to be so. In English, only *some* nouns have this possibility, and it is hard to predict which they would be. In addition, the meanings of ‘nouns used as verbs’ vary dramatically. Compare the meanings associated with three names of celestial bodies, all of which can be used as nouns and as verbs, as set out in Table 11.1. *Star* can be used as a verb in senses

TABLE 11.1. Meanings for *sun*, *moon*, and *star* when used as noun and as verb

	AS NOUN	AS VERB
<i>sun</i>	'ball of fire in the sky by day'	transitive verb, generally used in reflexive construction: <i>sun oneself</i> 'lie in a place where the sun shines on one'
<i>moon</i>	'large illuminated object in the sky at night'	intransitive verb, typically followed by <i>around</i> or <i>away</i> : 'spend time doing nothing in particular'
<i>star</i>	(i) 'small point of light in the sky at night' (ii) 'an object with four, five or six points in a regular pattern' (iii) 'person who plays a leading role in a performance, etc.'	(ii) 'place a star or asterisk on or against an item' (iii) intransitive verb: 'play a leading role in a performance, etc.'

(ii) and (iii). *Sun*, as a verb, refers to being *in* the sun(shine). The semantics of *moon* as a verb is scarcely directly relatable to its meaning as a noun. It will be seen that the meaning of each of the verbs is essentially idiosyncratic, and cannot be predicted from the meaning of the corresponding noun.

Only a small number of nouns referring to kin relations and age groups may be used as verbs. And for those that can, the meaning of the verb is once again not predictable from the meaning of the noun. This is illustrated in Table 11.2.

All three verbs are transitive. Verb *father* relates exactly to the meaning of the noun: if Y fathered X then Y is X's father. The A argument of verb *mother* (X) is *not* the mother of X, but treats X as one would expect a mother to. For

TABLE 11.2. Meanings for *mother*, *father*, and *baby* when used as noun and as verb

AS NOUN	AS VERB
<i>mother</i> (of X) 'be the woman who gives birth to X'	transitive verb: <i>mother</i> (X) 'treat X (who can be a child or an adult) in the way that a mother would treat her child'
<i>father</i> (of X) 'be the person who made X's mother pregnant with X, through sexual intercourse'	transitive verb: <i>father</i> (X) 'make X's mother pregnant with X, through sexual intercourse'
<i>baby</i> 'a very young child who cannot yet do things for itself'	transitive verb: <i>baby</i> (X), 'look after X in the way that one would look after a young child'

verb *baby* (X), it is the O argument, X, that is treated like a baby, although X is not in fact a baby.

These examples indicate that it is not adequate to say that *sun*, *moon*, *star*, *mother*, *father*, and *baby* is each a single lexeme which can function as a noun or as a verb. In each case there are *two* lexemes, one a noun and the other a verb; there is generally a semantic link between them, but this is different in each instance. (For *moon*, the semantic link is nebulous.)

English has a good many morphological processes. For quite a few of these there are a number of alternative realizations, one of which is zero. In §3.13 there was illustration of plural marking on nouns—by /-z/ in *dog-s* /dɒgz/, by /-s/ in *cat-s* /kæts/, by /-əz/ in *horse-s* /hɔ:səz/, and zero in *sheep* /ʃi:p/.

There are a fair number of word-class-changing derivational processes, for which one allomorph is zero. These are illustrated in Tables 11.3 and 11.4.

For each pair in these two tables, the lexemes in the left-hand column have similar meanings, and so do their derivations in the right-hand column. For the first member of each pair there is plainly a derivational process, realized by the addition of suffix *-ize* in (a–c), *-er* in (d), *-red* in (e), and *-ion* in (f).

TABLE 11.3. Verbs derived from nouns

	UNDERLYING NOUN	DERIVED TRANSITIVE VERB
(a)	<i>patron</i> <i>boss</i>	<i>patron-ize</i> ‘act like a patron towards’ <i>boss</i> ‘act like a boss towards’
(b)	<i>hospital</i> <i>market</i>	<i>hospital-ize</i> ‘put into a hospital’ <i>market</i> ‘take to/put on market’
(c)	<i>liquid</i> <i>knot</i>	<i>liquid-ize</i> ‘make become liquid’ <i>knot</i> ‘tie into a knot’

TABLE 11.4. Nouns derived from verbs

	UNDERLYING VERB	DERIVED NOUN
(d)	<i>observe</i> <i>witness</i>	<i>observ-er</i> ‘someone who observes’ <i>witness</i> ‘someone who witnesses’
(e)	<i>hate</i> <i>love</i>	<i>hat-red</i> ‘the feeling Y has for X when Y hates X’ <i>love</i> ‘the feeling Y has for X when Y loves X’
(f)	<i>exhaust</i> <i>concern</i>	<i>exhaust-ion</i> ‘what one feels when someone or something has exhausted one’ <i>concern</i> ‘what one feels when something concerns one’ (extended to refer to ‘some situation which does or could cause concern’)

The same analysis should be provided for the second member of each pair, and here we have zero derivation.

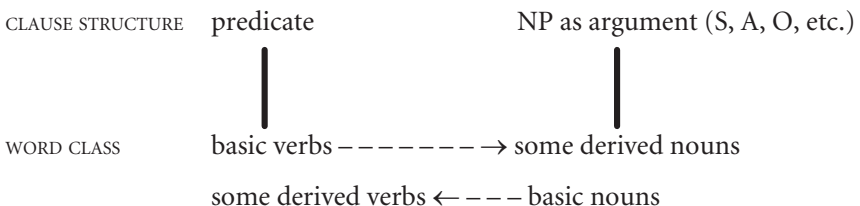
We have seen that some nouns in English appear to function in predicate slot. The appropriate analysis is to say that verbs can be derived from a number of nouns, the derivation having zero marking in some instances and being shown by an affix in others. Similarly for verbs. We should not say that there are verbs which can function as NP head. Rather, from a number of verbs can be formed a derived noun, often with zero marking but other times shown by a derivational affix, or by internal change or stress shift (for example, verb *think* → noun *thought* and verb *im'port* → noun *'import*). This analysis is justified by:

- (i) The correspondence between zero-derived and non-zero-derived forms, as shown in Tables 11.3 and 11.4 (many further examples could be provided).
- (ii) The unpredictable difference in meaning between noun and derived verb, and between verb and derived noun, as shown in all four tables. This indicates that in each instance verb and noun constitute distinct lexemes.
- (iii) The fact that there are corresponding verbs for only *some* basic noun lexemes, and vice versa.

Correlating with this is the fact that the majority of zero-derived verbs are transitive, whereas in languages—of schemes II and IV—where a noun can feature as predicate head, this is almost always confined to an intransitive predicate.

Corresponding to the diagrams for schemes I–IV, we can represent the situation in English by:

V Canonical scheme, with considerable derivations between noun and verb at the lexical level



From the brief account of Mundari provided by Evans and Osada (2005), it appears that this language may be like English, involving zero derivation to create new lexemes rather than multifunctionality. A critical consideration is

the lack of semantic correspondence between instances of a single form in two functions. For example:

FORM	MEANING AS NOUN	MEANING AS VERB
<i>dasi</i>	'male servant'	'work as a servant'
<i>hoRo</i>	'person, Munda person'	'speak Munda language'
<i>buru</i>	'mountain'	'heap up'

Note that Mundari does have a copula construction, which would be used to render 'My son is a Munda person' and, presumably, 'That is a mountain'.

Korku, another language from the Munda subgroup, was in §11.2 tentatively identified as having scheme II. But Nagaraja (1999: 30) states that '*some nouns ... can function as verbs*' (my italics). It would be instructive to explore whether Korku, like Mundari, is of the English type with scheme V. In fact, many languages for which it has been claimed that 'all nouns can be head of a predicate' and/or 'all verbs can be head of an NP' should be investigated from this point of view, particular attention being paid to semantic content.

11.4 Structures of noun phrases with different heads

There is always a range of possibilities for what can be the head of what is conveniently referred to as 'noun phrase' (NP). In every language, the fullest set of structural possibilities is when the head is a common noun.

Languages do vary in this, but the most frequent situation is for there to be one or more adjectives modifying the common noun as head, plus a number or quantifier ('many', 'few', 'all'), a demonstrative or article, a relative clause, and often a modifying time or place phrase ('from yesterday', 'in the garden'). Generally, all are optional, save the article (or a determiner) in some languages. Sometimes there may be a second common noun modifying the head; this will typically refer to sex ('man child'), material ('metal door'), or purpose ('dog biscuits'). If a language has a set of classifiers or generic nouns, one of these may accompany the common noun as head. And in some languages a part noun may occur with a common noun as head, sometimes agreeing with it in gender (for example, 'mother foot', exemplified in §5.6).

An NP head may be modified by a possessor NP—an NP within an NP—as in the English phrase [*[the bald man from across the road]'s two fierce dogs which you detest*]. One needs to enquire what can be the head of an NP in possessor function.

Alternatives to a common noun as head are likely to include proper noun (name of person or place), free pronoun, demonstrative, and interrogative. Although there is considerable variation from language to language, for each of these the structural possibilities of the NP are likely to be restricted,

compared to those when a common noun is head. For example, a proper noun or pronoun may not be modifiable by a number or quantifier, or a possessor NP.

The critical question for the present enquiry is: when a verb functions as head of an NP, does the NP have the same structural possibilities as when a common noun is head? Note that the possibilities have to be *precisely the same* for a claim that there is no distinction between noun and verb classes to be substantiated.

For every language which has been thoroughly examined from this point of view, the answer is 'no'. In a perceptive and detailed study of the Wakashan language Makah, Jacobsen (1979: 121) states that a verb can only be head of an NP when it takes the 'article' *-^oiq* (this could be said to function as a nominalizer), whereas a common noun can be NP head either with or without *-^oiq*. This would be sufficient criterion for distinguishing classes of nouns and verbs in Makah (and for other Wakashan languages, where the same property holds). Support is provided by Nakayama (2001: 49–50), who reports that in Nootka (here called Nuuchahnulth), 'nominals', but not verbs, 'can be modified with expressions of property concepts, quantity or quantifiers'. And there are a number of further corroborating criteria, which will be mentioned later.

Describing the two Salish languages Lillooet and Lushootseed, van Eijk and Hess (1986) point out that when a verb is head of an NP it *cannot* take possessive affixes, whereas a noun *can*. This, again, is sufficient evidence for a distinction between noun class and verb class. In Fijian, a verb can be head of what appears to be an NP (functioning as argument of the predicate), but it *must* then take a possessor modifier (coding the underlying subject of the verb), whereas the statement of a possessor is *optional* for a common noun as NP head. And only a common noun, not a verb, may be modified by a number or quantifier. (In fact, what appears to be an NP with verb as head in Fijian is better analysed as a type of complement clause. See the appendix to Chapter 12.)

In the case of every language for which it has been suggested that verbs can function as NP head, the optional and obligatory elements in NP structure must be fully investigated—including whether a verb can be head of an NP in possessive function, or whether a verb as NP head can or must or cannot be modified by a possessor element.

11.5 Properties of predicates with different heads

There is more variation between languages—and overall less structural complexity—concerning what fills the predicate slot than there is for NPs which fill an argument slot. In many languages, a single lexeme functions

as predicate head, accompanied by a good number of grammatical elements (realized by affixes or other morphological processes, or by clitics, or by small grammatical words). There may be a lexical adverb as modifier and this may be sensitive to what is functioning as predicate head; for example Nakayama (2001: 49–50) shows that, for Nuuchahnulth (Nootka), verbs, but not ‘nominals’, can be modified ‘by qualifying expressions like... “almost” or... “barely”’.

Some languages have a ‘serial verb construction’, where two or more verbs jointly fill the predicate slot; others—such as English—lack anything of this nature. For languages which show both a serial verb construction and nouns having secondary function within a predicate, it would be important to enquire how nouns feature in serial verb constructions. Little attention has so far been given to this question.

It is the predicate which determines the argument structure of a clause. All languages have transitive verbs which, as predicate head, select A and O core arguments. As mentioned before, when a noun has secondary function as head of a predicate, it is generally restricted to an intransitive predicate, with one core argument in S function. This applies to Fijian, and to Nootka (Jacobsen 1979).

It also applies to Tagalog. Schachter and Otnes (1972: 62–79) show that if predicate head is ‘nominal’ or ‘adjectival’, there is a single argument which they dub ‘topic’. However, with verb as predicate head there can be up to three core arguments—‘actor’ plus ‘object’ or ‘directional’ or both. One argument must be put in ‘focus’, the choice of argument being shown by a focus marker on the verb (the focus marker can be prefix, suffix, or infix). The vital point here is that the focus system applies only to verbs, not to nouns (or adjectives) when functioning as predicate head. These would be sufficient criteria for distinguishing noun and verb classes. In fact, there is a further criterion, described in the next section.

When one means of realization for a core argument is a bound pronoun, this is generally attached to the predicate head, whether it be verb or noun, as in (7–8) in §11.2. Indeed, this is likely to be a way of telling that a noun is being used in predicate head function.

11.6 Grammatical categories associated with verbs

There are a number of grammatical categories typically associated with a verb when it is head of a predicate.

- (a) **Tense** and/or **aspect** and/or **modality** (TAM, to be discussed in Volume 3). As pointed out in §3.15, these are simultaneously properties

of the clause and of the predicate. They are often marked by morphological processes applying to the verb which is head of the predicate. Alternatively, they may be realized by clitics or grammatical words either within the predicate or elsewhere in the clause (for example by an enclitic to the first phrasal constituent of the clause). In a minority of languages, markers of tense—but, much less often, of aspect or modality (and never of mood)—may also be applied, sparingly, to an NP; for example ‘my old house (house-PAST)’ or ‘her husband-to-be (husband-FUTURE)’; see Nordlinger and Sadler (2004).

- (b) **Mood.** This is a category associated with the sentence, through its main or focal clause (see §3.11). Since a clause consists of one predicate and a variable number of (or no) NP arguments, it is convenient to mark mood on the verb, as head of the predicate. Imperative mood is frequently marked in this way, but declarative and interrogative less often; see (b) in §3.2.
- (c) **Evidentiality.** This is also a category associated with the clause and, as with mood, is typically marked on a verb which is predicate head.

The critical question here is—if a language can have a noun as predicate head, which of the categories may it take from those that are associated with a verb as predicate head? In some languages, such as Fijian, the answer is ‘all’. For others it is ‘some’.

- *Tagalog.* Schachter (1985: 13) states that there is a fairly ‘clear-cut’ distinction between noun and verb classes in that ‘only verbs are inflectable for aspect’. Schachter and Otanes (1972: 66, 361–8) recognize three aspects which they call ‘perfective’, ‘imperfective’, and ‘contemplated’.
- *Wakashan languages.* For Makah, Jacobsen (1979: 140) says ‘nouns may be predicates in their own right, but are rather restricted in the inflectional categories they may take, being limited to the durative aspect, and apparently not occurring in the future tense’. That is, nouns as predicate heads do not occur in iterative or repetitive aspects, or in future tense.
- *Salish languages.* Information on a variety of languages from this family indicates that the TAM categories associated with verbs are generally not all available for nouns in predicate head slot. Describing Kalispel, Vogt (1940: 23) explains: ‘the verbs are characterized by the category of aspect, expressed by affixes unknown in the noun.’ (He adds: ‘in the great majority of cases it is possible by purely formal criteria to assign a given form to the verb-class or the noun-class.’) For Lillooet and Lushootseed, van Eijk and Hess (1986: 322) state: ‘verbs (i.e. stems that do not take possessive affixes) can be subject to aspectual operations. Nouns do not allow these operations.’

Similar restrictions on nouns in predicate slot are found in languages for which the noun/verb distinction has never been doubted. For Tariana (Arawak family, Brazil; Aikhenvald 2004a: 105): ‘a noun in the predicate slot cannot take *aktionsart* enclitics ... neither can it take intentional or apprehensive moods, or be used in a positive or negative command.’ In §11.2, we quoted Mao Naga as having aspect and tense marked on both noun and verb as predicate head, whereas only a verb may take ‘mood’ markers. (Giridhar 1994: 305–58 identifies thirty ‘moods’, which are what we would call modalities, plus the imperative mood.)

Sapir’s (1921: 119) dictum that the distinction between noun and verb class applies in every language although sometimes ‘the nature of the distinction may be an elusive one’ was mentioned at the beginning of this chapter, together with Schachter’s (1985: 6–7) claim that ‘there are no languages at all which cannot be said to show a noun–verb distinction when all relevant facts are taken into account’. Tongan is a language for which—like its close relative Samoan and fairly close relative Fijian—it has long been acknowledged that basic word classes can be clearly distinguished (see Churchward 1953). Broschart (1997: 153) appears to have begun with the idea that there is no distinction between noun and verb, and demonstrates functional commonality between the classes in Tongan. He appears not to have looked carefully for criteria which would serve to distinguish the classes. Broschart’s cavalier approach is demonstrated when he admits, ‘there are no doubt some structural differences between the “nominal” and “verbal” predications, but these differences are slight’. No doubt they may be slight but, if Broschart had teased them out, no doubt they might prove to be criterial.

One type of grammatical category generally associated with verbs is valency-changing derivations. Since nouns can in most instances only occur in intransitive predicates, valency-reducing derivations (such as passive, antipassive, reflexive, and reciprocal) are unlikely to apply to them. What of valency-increasing derivations—such as causative and applicative—which apply to intransitive verbs? Little attention has so far been paid to this question, in discussion of languages in which nouns can be predicate head; it may well turn up further criteria.

11.7 Grammatical categories associated with nouns

A number of categories are typically associated with noun or NP; they have varying grammatical status.

- (a) If there is a **noun class** or **gender** system, then each noun is likely to have a fixed value from the system; this is an inherent property of the

- noun. Gender or noun class may be marked in the form of the noun, but often it is shown only by agreement with modifiers within the NP (adjective, inalienably possessed part noun, demonstrative, etc.) and/or by agreement in the predicate.
- (b) When there is a set of **classifiers**, then an integral feature of each noun is the classifier(s) it may occur with.
 - (c) As pointed out in §1.10, **number** is a referential property of a complete NP. It may be shown by a morphological process applying to the head noun, and often also by agreement processes on modifiers. In some languages, including English, the number system only applies to some nouns, which are called ‘countable’. Whether it is classed as countable or non-countable is an inherent property of a noun.
 - (d) **Definiteness** is a discourse property of an NP, indicating that a unique referent has been identified.
 - (e) **Case** has quite different grammatical status, marking the function an NP has in its clause. It may be shown by a grammatical word or clitic at the end of the NP, or by a morphological process which applies either to every word, or just to one (or more) word(s)—it may be the last word, or the first, or the head. (In some languages some terms from the case system have secondary function. Added to a predicate they may mark varieties of clause linkage, or add aspectual or modal meanings to the clause as a whole; see Aikhenvald 2008b.)

The question now is—when a verb functions as NP head, which of the properties may it have; out of those associated with a noun as NP head? Generally, the categories associated with noun and verb as NP head do vary.

As mentioned in §11.4, in Wakashan languages a verb may be NP head only when accompanied by the ‘article’ *-^oiq*, which indicates definiteness (the *-^oiq* is optional when a noun is NP head). In Fijian, a noun as NP head may select from a set of classifiers whereas for a verb in that slot no choice is involved (see the appendix to Chapter 12).

Possession, properly a matter of NP structure, was mentioned in §11.4—how, for example, a verb as NP head cannot take a possessor modifier in Salish languages, but must do so in Fijian.

11.8 Further criteria

There are a number of other properties which may be used to distinguish word classes.

- (i) *Morphological process of reduplication*. In many languages, nouns and verbs show different patterns of reduplication, in terms of both form and

meaning. In Dyirbal reduplication of a noun or adjective involves copying the full form of the word and indicates plural reference; for example, *mulari* 'initiated man', *mulari-mulari* 'initiated men'. For a verb, reduplication involves repeating just the first two syllables before the root, indicating 'do to excess'; for example *miyanda-* 'laugh', *miya-miyanda-* 'laugh more than is appropriate'.

This may sometimes assist in distinguishing noun and verb classes in languages where each has multiple syntactic functions. In Fijian, verbs may be freely reduplicated with the meaning 'do something several times, or over a long period'. There is no productive reduplication for nouns. In at least some Salish languages, one type of aspectual operation, associated only with verbs, repeats the consonant after the stressed vowel, and expresses an 'ongoing process'; for example *pála?* 'to be one', *pállá?* 'to get together' (van Eijk and Hess 1986: 322).

(ii) *Word-class-changing derivations*. If there are derivational processes which change word class, this presupposes the existence of distinct word classes. Very many languages have a process for deriving a noun stem from a verb root or stem, and a considerable number have a process for deriving a verb stem from a noun root or stem.

In Fijian, for example, prefix *i-*, when added to a verb (sometimes reduplicated, other times not), derives a noun which describes an instrument, or the place of an activity, or the result of the activity, or the activity itself. Examples include *cula* 'pierce', *i-cula* 'needle'; *moce* 'sleep', *i-moce-moce* 'bed'; *vola* 'write', *i-vola* 'letter, book' (further examples were given under (d) in §10.6). And in Tagalog prefix *taga-*, added to a verb root, creates a noun meaning 'the person employed or delegated to perform the action'; for example *sulat* 'write', *taga-sulat* 'writer' (Kroeger 1998: 16; Schachter and Otnes 1972: 105–6).

The ways in which roots combine to form compounds always relates to their word class membership and is likely to constitute a further criterion for word class recognition.

(iii) *Special reference stems*. These are reported for the Wakashan language Nootka and there they are a most valuable criterion. As quoted at the beginning of this chapter, Swadesh (1938: 78) wrote of Nootka 'normal words do not fall into classes like noun, verb, adjective, preposition'. However, later in the same paper (pp. 98–100) he notes that there are seven sets of 'special reference stems', and that each lexeme selects just one set; each set involves a pronominal-like 'indirect reference stem' (used for anaphora), a 'relative stem' (marking the argument shared between a relative clause and its main clause), and an 'interrogative stem' (used in questions). Lexemes divide into seven classes, according to which set of special reference stems they relate

to. Swadesh's labels for the classes, and summaries of their memberships, are:

1. 'Entity—referring to species of flora and fauna and supernatural beings, age and other classes of people and other beings, body parts, a group of classes of objects according to shape, and other entities.'
2. 'State—express quality, condition, color, size, position, mental state or attitude, conditions of the weather, and other notions.'
3. 'Action—express movement and various other activities; some are implicitly transitive, some intransitive.'
4. 'Location.' A 'virtually complete' list of twelve items is given, including 'there', 'out to sea', and 'on the left'.
5. 'Time.' A list of nine items is given, including 'right away', 'again', and 'never'.
6. 'Quantity.' (a) numeral, (b) amount.
7. 'Indication (demonstrative notions)' including 'this', 'that', 'some', 'different', and 'another'.

The choice of special reference stem for each lexeme constitutes an internal grammatical criterion for recognizing word classes. Swadesh's labels, and descriptions of the semantic content of each class, clearly show that the first three should be labelled 'noun', 'adjective', and 'verb'.

There are other grammatical properties which may be relevant for word class identification. For instance, negation is often marked differently in a predicate and in an NP (as in English—predicate *not arriving* and NP *non-arrival*). It would be instructive to investigate whether the same form of negator is used for a noun as for a verb as predicate head, and for a verb as for a noun as NP head, in languages which show such multiple functions. As a further example, there may be processes creating 'diminutive' and/or 'augmentative' forms from roots of a certain word class or classes.

As briefly mentioned in §§7.4–5, in some languages there may be different phonological profiles for words belonging to noun and verb classes, providing a further defining feature.

11.9 Summary

We can conclude—after a fairly thorough investigation—that every human language has distinct classes of noun and verb, recognizable on morphological and/or syntactic criteria. In some languages these are easily recognizable, in others the clues are more subtle. As an analogy, for some species of birds their sex is easily perceived due to different colouring, whereas for others the sex is

hard to tell. But, for each species, the sex of a bird is equally important. Similarly for languages: irrespective of how obvious they may be on the surface, noun and verb classes of lexemes are of fundamental importance in each language. The major function of a noun is always to be head of an NP which is a predicate argument, and the major function of a verb is always to be head of a predicate. The class of nouns invariably includes lexemes with concrete reference (what else there is varies from language to language) and the class of verbs always includes lexemes which refer to movement, giving, actions, and talking.

As said several times before, word classes must be recognized on an individual basis in each language, in terms of internal grammatical criteria in that language. In a language where each word class is associated with a fair number of morphological processes, these will furnish sufficient criteria (as in Latin). For a language with limited morphology (such as Vietnamese), occurrence of a lexeme with small grammatical words will provide criteria—nouns with classifiers and demonstratives, verbs with particles expressing tense, aspect, and mood.

In addition to its primary function in an NP, in some languages a noun can have secondary function as head of a (usually intransitive) predicate. And, in addition to its primary function as predicate head, a verb may have secondary function in the head slot of an NP. Each word class occurs much more frequently in primary than in secondary function (as shown by the thick and thin lines in the diagrams representing schemes II–IV in §11.2). We have seen that when a verb is NP head it is likely to have more restricted properties than a noun as NP head, with respect to the structure of the phrase and/or the grammatical categories associated with the head. The same applies to a noun as predicate head, compared to a verb in that slot.

As mentioned in §1.11, in each language the class of nouns has, at the least, several thousand members. In the great majority of languages, the verb class has, at the least, many hundreds of members. Just a few languages have a more limited inventory of verbs, each with a rather general meaning. But they enter into very many combinations (typically including serial verb constructions)—many of which have idiosyncratic meanings—so that overall the language's inventory of verbal complexes is considerable.

There are a number of languages for which a great song-and-dance has been made in the literature concerning their lacking a noun/verb distinction (many of these are secondary sources, quoting from other secondary sources). As described in §§11.2–3, Evans and Osada have shown that in Mundari it is far from the case that all lexemes are multifunctional. Indeed, the optimal analysis is probably that there is no multifunctionality per se, but instead Mundari is like English in making extensive use of zero derivation, often involving an idiosyncratic semantic shift.

For other languages touted as not distinguishing nouns and verbs, a number of adequate criteria were found:

- *Tagalog*. Only verbs as predicate heads may take multiple core arguments, with a ‘focus’ system (§11.5). Only verbs accept aspect modification (§11.6). And there are derivations which form a noun stem from a verb root (§11.8).
- *Nootka and other Wakashan languages*. A verb can only be NP head if it takes an ‘article’, whereas this is optional for a noun. Only nouns, not verbs, ‘can be modified with expressions of property concepts, quantity or quantifiers’ (§11.4). Within a predicate, verbs but not nouns can be modified by qualifying expressions like ‘almost’ or ‘barely’ (§11.5). And a noun as predicate head has much more restricted aspect and tense properties than a verb (§11.6). Finally, the sets of ‘special reference stems’ clearly relate to—and establish—the lexical and grammatical word classes of the language (§11.8).
- *Salish languages*. Unlike a noun, a verb as NP head cannot be marked for possessor (§11.4). A noun as predicate head generally cannot take the full set of TAM categories which are available for a verb (§11.5) and, related to this, aspectual reduplication is limited to verbs (§11.8).

In addition, it remains to be investigated—for Tagalog, for Wakashan languages, and for Salish languages—whether *every* single noun may have secondary function as head of a predicate, and whether *each* verb, without exception, has secondary function as head of an NP.

The many distinguishing properties of noun and verb (and adjective) in Fijian are set out in the appendix to Chapter 12.

11.10 What to investigate

Lexical word classes—noun, verb, adjective, and sometimes also adverb—and grammatical word classes—pronoun, demonstrative, etc.—must be recognized on internal grammatical criteria within the language (see the illustration in §1.8, repeated in §11.1). Note that the meaning of a lexeme cannot be used as a criterion for which word class it should belong to. However, after the classes have been established, their semantic content should be studied.

When reading a new grammar, the first thing I look for is a statement of the word classes—the grammatical criteria for recognizing them, whether they are open or closed, whether lexical or grammatical, and an outline of the semantic content of each. This must come at the beginning of the grammar, since it is a prerequisite for everything that follows.

The primary function of a noun will always be as head of an NP which is a predicate argument, and the primary function of a verb will always be as head of a predicate. There may also be secondary functions:

- A limited set of nouns may function as modifier within an NP; these generally indicate sex, or material, or purpose, or parts (§11.4).
- A noun, or a complete NP, may function as head of a predicate. Check which types of predicate (whether just intransitive, or of either transitivity), and what the properties of the predicate are with a noun as head (§11.5–6).
- One needs to check whether a noun can be modifier within a predicate. This is rare (which is why it was not mentioned earlier) but not impossible.
- A verb may be modifier within a predicate, for example, in an asymmetrical serial verb construction. This function should be fully investigated.
- A verb may function as head of an NP. Check the structural possibilities for such an NP, and grammatical categories associated with a verb as head, compared to those when a noun is head (§11.4).
- A verb is unlikely to be able to directly modify the head of an NP (it will generally only do so within a relative clause construction, or as a derived adjective). But this possibility should be checked out.

Sources and notes

A useful historical discussion of the literature concerning the absence or presence of noun and verb classes in Wakashan and Salish languages is in Jacobsen (1979: 84–108). Mithun (1999: 56–67) surveys work on North American languages; see also Mithun (2000) for confirmation that a noun/verb distinction can be recognized for Cayuga and other Iroquoian languages. People writing on North American languages seem particularly prone to say that there is no noun/verb distinction—apparently because there is not a distinction of the same kind as in European languages—but then to work in terms of ‘noun’ and ‘verb’; among others, see Hoijer (1933: 23) on Tonkawa (isolate, Texas), and Andrade (1933: 236) on Quileute (Chimakuan family, state of Washington, geographically next to the Wakashan and Salish families).

In the course of careful study of the literature on this topic, the following generalization has emerged. People who deny a noun/verb distinction tend to be those who have just looked at limited aspects of a given language, whereas those who have studied the language in depth, and (in most cases) published a full grammar of it, most often do recognize distinct word classes.

11.2. Concerning scheme II, Launey (1994) coins the term ‘omnipredicative’ for a language in which any lexeme can function as head of a predicate (he illustrates this from Classical Nahuatl). Use of this term would tend to ignore the fact that different word classes are likely to have rather different structural and inflectional possibilities when used as head of a predicate.

Some adjustments have been made to the information included in Evans and Osada (2005). A number of their percentages were wrongly calculated; for example 1953/3824 is given as 52 per cent when in fact it is 51 per cent (more accurately 51.072).

Adopting the stance that a language lacks nouns and verbs makes it difficult to employ normal terminology when discussing typological topics. For instance, Mosel (2004) continues to deny that Samoan has regular word classes; as a consequence, she cannot talk of ‘noun incorporation’ but has to substitute the label ‘juxtapositional constructions’ (this is, as she puts it, a construction ‘which corresponds to noun incorporation . . . in other languages’).

11.3. Clark and Clark (1979) provide an admirable account of ‘when nouns surface as verbs’ in English, examining the semantic and pragmatic factors involved. There is discussion of nouns derived from verbs in English in Dixon (2005: 322–52). Although *witness* first came into English as a verb, as it is used in the language today speakers tend to perceive the noun sense as primary.

The term ‘derivation’ refers to any morphological process which forms a stem from a root or stem. It can either change word class or maintain the same word class membership. It would be useful to have a generally accepted term for ‘word-class-changing derivations’. Sweet (1891: 38–40) did coin the term ‘conversion’ for this. However, in recent years Sweet’s term has been reinterpreted as describing only word class changing by a derivation which has zero realization. Indeed Bauer and Huddleston (2002: 1640) insist that the term ‘conversion’ should be used *instead of* ‘zero derivation’ or ‘zero affixation’.

For *Tukang Besi* (Austronesian, Indonesia), Donohue (1999: 88) suggests a continuum, ranging from purely verbal, to weakly verbal, then pre-categorial, followed by weakly nominal, and finally purely nominal.

11.4. Another study of Nootka which supports Jacobsen’s and Nakayama’s conclusions is Davidson (2002).

Much has been published on the question of whether Salish languages make a distinction between noun and verb classes. Among other publications which support this distinction—repeating the ‘possessive affix’ criterion and/or presenting a variety of further criteria—are Vogt (1940), Kuipers (1967, 1974), Hébert (1983), Nater (1984), Galloway (1993: 237–8), van Eijk (1997), Haag (1998), Watanabe (2003: 66–75), Kroeber (1999: 33–6).

The Adjective Class

It has occasionally been suggested that some languages lack a distinction between classes of noun and verb. In the last chapter we showed that, in Schachter's (1985: 6–7) words, 'it now seems that the alleged counter-examples have been based on incomplete data' and that distinct noun and verb classes are always recognizable 'when all relevant facts are taken into account'.

For a greater number of languages, it has been suggested that it is impossible or inappropriate to identify an adjective class. (Indeed, I was one of those who promulgated this view, in Dixon 1977c/1982.) But detailed examination (over the past thirty years) of languages for which this claim had been made suggests that, once again, when all relevant facts are taken into account an adjective class can be (and should be) recognized for every language, distinct from noun and verb classes.

§12.1 sets out the parameters of variation for adjective classes; §12.2 explains the reasons for recognizing an adjective class and surveys its history in grammatical work. Criteria for recognition are outlined in §12.3, and then §12.4 discusses the semantic content of adjective classes. §12.5 discusses types of adjective class and features that are likely to serve for distinguishing it from the verb class in some languages, and from the noun class in others. Later sections examine languages with restricted functional possibilities for adjectives, those with two adjective classes, correlation with other grammatical parameters (and resulting correlative change), and the question of semantic overlapping between the major word classes.

12.1 Parameters of variation

Cross-linguistically, adjective classes differ in (I) their size and productivity, and (II) whether their grammatical properties are similar to those of nouns, or of verbs, or both, or neither.

I. Size. In every language, the class of nouns has several thousand members. The verb class generally has at least several hundred, but there are languages which have only a few score or so verbs (and many complex expressions

involving them)—this was illustrated for the Australian language Yawuru in §1.11. Adjective classes present a rather different picture. We can distinguish:

- (i) Languages with a large, open class of adjectives which include hundreds of members. New items may be added to the class, by derivations from within the language, and by loans from without. In a given language, there are generally fewer monomorphemic adjectives than verbs. Typically, a higher proportion of adjectives than of nouns and verbs will be derived forms (see Givón 1970: 816).
- (ii) Languages with a small closed adjective class, to which new members may *not* be added. An adjective class may be exceedingly small, as in Yimas (Lower Sepik family, New Guinea) for which just three adjectives are identified by Foley (1991: 93): *kpa* ‘big’, *yua* ‘good’, and *ma* ‘other’. Watters (2002: 111) reports just three adjectives of native origin for Kham (Tibeto-Burman, Nepal); they are *gehppa* ‘big’, *zimza* ‘small’, and *twi:za* ‘short’. Other languages may have eight or ten or twelve adjectives, or forty or fifty. The semantic content of small adjective classes is summarized in §12.4.

II. Grammatical properties. There is a rough division into four types of adjective class:

- (a) Adjectives have similar grammatical properties to those of verbs, as in Mandarin Chinese, Thai, Vietnamese, and Korean. Typically, verb and adjective may both function as head of an intransitive predicate, taking similar marking for some or all of tense, aspect, modality, and mood. It is often the case that verbs and adjectives may only modify a noun—which is head of an NP—through a relative clause construction.
- (b) Adjectives have similar grammatical properties to those of nouns, as in Latin, Spanish, Finnish, Hungarian, Igbo, Quechua, and Dyirbal. Typically, both noun and adjective may be restricted to occurrence in an NP (that is, they cannot be used in a predicate). An NP may include noun, or noun plus adjective, or just adjective. Adjectives may take the same inflectional processes as nouns, for instance relating to gender and number.
- (c) Adjectives combine some of the grammatical properties of nouns with some of those of verbs (as in Berber languages from North Africa, Tariana from Amazonia, Nunggubuyu from north Australia, and Takelma from Oregon). For example, they may be able to occur in an NP, then inflecting like a noun, and also as head of an intransitive predicate, then inflecting like a verb.

- (d) Adjectives have grammatical properties different from those of nouns and from those of verbs (as in English, Tunica from Louisiana, and Mam and Teribe from Central America). An adjective cannot be the sole lexeme in an NP, neither can it function as intransitive predicate. It takes none of the morphological processes available to nouns and to verbs but instead has categories of its own.

There are fewer languages of types (c) and (d) than of types (a) and (b). It is for languages of the two major types that justification for recognizing an adjective class has been denied.

For languages of type (b), where adjectives show similarities to nouns, grammars not infrequently maintain that adjectives *are* nouns. For example, Bright (1957: 56), writing on Karok (isolate, north-west California), has 'adjectives' as one of four subclasses of noun. And for languages of type (a), where adjectives share many properties with verbs, it has often been said that adjectives *are* verbs. An example here is Swanton (1911a: 270) writing on Haida (isolate, Queen Charlotte Islands, off the coast of British Columbia): 'adjectives may always be used as verb-stems and so belong to this category.' For the Kolokuma dialect of Ijo (Kwa family, southern Nigeria), six subtypes of verb are set up by Williamson (1965: 34), one of them being 'adjective verbs'. Indeed, many grammars make no mention at all of adjectives, silently subsuming them into the noun or verb class. One has to study a grammar most carefully in order to work out in which class adjectival concepts have been included. It appears to be the noun class for Hanis (Coosan family, Oregon), described by Frachtenberg (1922a: 318), and also for Siuslawan (isolate, southern Oregon) also described by Frachtenberg (1922b). And it can be inferred that adjectival concepts are included within the verb class for Tlingit (Na-Dene family, south-east Alaska) in Swanton's (1911b) grammar, and for Sioux (Siouan-Catawba family) in Boas and Swanton's (1911) study.

We noted, in Chapter 11, that the distinction between noun and verb classes may have many surface manifestations in some languages, only a few in others. Nevertheless, it is equally valid and useful in all instances. Similarly for an adjective class. It has been suggested that adjectives cannot be distinguished from nouns in Turkish. However, as Lewis (2000: 50) puts it: 'the dividing line between noun and adjective is a thin one, but it is still worth drawing.' Similarly for a language such as Thai, which has been said to express adjectival concepts through verbs. It is true that both verb and adjective can occur as head of an intransitive predicate; nevertheless, a distinction between verb and adjective classes can be established, and does play an important role in the grammar. For example, only an adjective may occur in unmodified form as the parameter in a comparative construction; adjective and verb show different

possibilities for direct modification of a noun; and the two word classes behave differently with respect to reduplication (see Post 2008, who provides a list of ten properties in which adjectives differ from verbs).

The previous chapter commenced with a quotation from Milner (1956: 10), maintaining that there are no lexical classes at all in Fijian; in his grammar, every kind of lexeme was combined together in a class of ‘bases’. But we showed in §11.8 that classes of noun and verb must be distinguished for proper statement of reduplication processes and for word-class-changing derivations. In addition, an adjective class must be recognized for proper description of NP structure, further word-class-changing derivations, the conditions for occurrence of certain predicate modifiers, choice of classifier, and so on. The appendix to this chapter provides an account of the necessity for recognizing classes of noun, verb, and adjective in Fijian.

In these and all other cases an adjective class *can* be recognized, and it *is* useful. We need now to explain why this is so.

12.2 Why recognize an adjective class?

As stated in §1.1, linguistics has two interwoven components—description and theory. The description of a language is framed in terms of basic linguistic theory, choosing from the available roster those categories and construction types which are relevant and useful for the language under study. The theory itself is made up of interrelated inductive generalizations based on good-quality descriptions. As each new language is described, it will throw up significant features which lead to the refinement, revision, or extension of part of the theory.

The recognition of an ‘adjective class’ in the grammar of a particular language (as for every other category) is justified on two grounds—(a) its usefulness and explanatory power within that grammar, and (b) its relation to the general typological theory.

(a) **Utility in description.** Unlike many formal theories, basic linguistic theory does not consist of a list of components which every grammar *must* include. What it does, instead, is provide a range of theoretical tools and a pool of conceptual categories, each of which *may* be utilized in the grammar of a particular language if it fulfils a useful role there in description and explanation. As an illustration, for some languages it is not appropriate to recognize anything which could be felicitously called a complement clause construction. A distinction between derivational and inflectional processes is most helpful in some instances, but there are quite a few languages for which the distinction is simply inapplicable. Basic linguistic theory does *not require*

that distinct classes of noun and verb be recognized for each language. Our conclusion in the previous chapter that it is appropriate to identify these two classes in every language is based on empirical investigation, rather than it being a theoretical postulate. It is the same with the adjective class.

The inventory of categories and construction types recognized in basic linguistic theory provides the fieldworker with an idea of what to look for in a new language. For example, a few decades ago little had been published on the grammatical category of evidentiality. As a consequence, for a number of languages with a system of evidentiality distinctions the category was not recognized. Now that evidentiality is fairly well understood and well described (its typical content, and patterns of cross-linguistic variation), a student working on a previously undescribed language will be on the lookout for the category. And if it occurs they will be in a position to describe it accurately, and within a cross-linguistic perspective.

There is never just one point of justification for an analytic decision in linguistics. It is always the case that a number of criteria come together—and reinforce each other—to define a category. This category will then play a role in explanation. This is as true for the adjective class as for any other feature of a grammar.

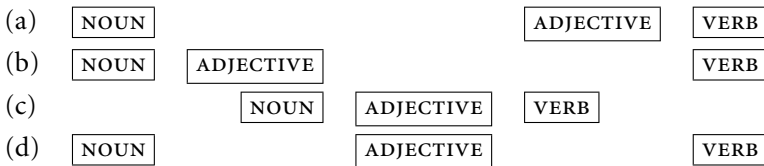
Saeed (1999: 104–9) recognizes a smallish class of about forty-two adjectives in Somali (Cushitic branch of Afro-Asiatic, Somalia, Ethiopia, and Kenya). Criteria for recognizing adjective as a class distinct from noun include: adjectives ‘do not occur with suffixed determiners’ and ‘they do not have inherent number and gender’. They differ from verbs in that they ‘may mark plural agreement with a nominal head by reduplication, which does not occur with this function in verbs’. Having established an adjective class, it has further properties within the grammar. For example, only adjectives can occur as complements of the copula verb *yahay* ‘be’, being positioned between satellite clitics and the verb. Adjectives may then fuse with the present tense form *yahay* (for example *wanaagasán* ‘good’ plus *yahay* ‘is’ gives *wanaagaányay*). In addition, adjectives enter into comparative constructions with adposition *ká* ‘from’ as marker of the standard of comparison (see §3.23).

For every language which has been thoroughly studied from this point of view, once an adjective class has been recognized, it does play a significant role in the grammar. That is, there is never just one property which serves to identify this or any other word class; there are always—at the least—several.

(b) Role in theoretical generalization. As said before, every new grammatical description is likely to provide feedback into the make-up of basic linguistic theory. But, in order to achieve this, the description must be framed within the general theoretical matrix. In §6.1 ‘Requirement for consistent

analysis', three languages were contrasted. It was shown that, under one method of analysis, an adjective class would be recognized for all three languages, whereas, if the same criteria were applied in a different order, only language 3 would be assigned a distinct adjective class. The discussion there can now be extended.

Consider four languages whose adjective classes show different grammatical properties—types (a)–(d) described under II in §12.1. They can be diagrammed, with grammatical similarities and differences between word classes modelled by spatial distance:



In language (d), adjectives have grammatical properties different from those of nouns and verbs, so that a distinct adjective class must be established. However, alternative analyses are available for languages (a) and (b). In (a), the adjective class has similar grammatical properties to the verb class. We could either:

- (i) Say that adjectives constitute a subclass of a combined verb-adjective class (conveniently called just 'verb class'). Or:
- (ii) Say that adjectives are a separate class, noting that their grammatical properties are similar to those of verbs.

In language (b), the adjective class has similar grammatical properties to the noun class. The same two analyses are available, *mutatis mutandis*. Either:

- (i) Say that adjectives constitute a subclass of a combined noun-adjective class (conveniently called just 'noun class'). Or:
- (ii) Say that adjectives are a separate class, noting that their grammatical properties are similar to those of nouns.

If one were interested only in the description of a single language, either alternative would be equally good, the difference being pretty much terminological. But if the linguist is interested in relating their grammar to a general theoretical framework, then the consequence of choosing one alternative over the other is immense. Under analyses (ii) all of languages (d), (a), and (b) have a major word class 'adjective'. As will be shown below, the three adjective classes will have similar functional properties and semantic content.

Were analysis (i) to be followed, only language (d) would have a major word class ‘adjective’. It would be possible—but both complex and unnecessary—to try to relate the adjective class in (d) to a subclass of verbs in language (a) and to a subclass of nouns in (b). Analysis (i) would greatly impede the task of comparing languages and working towards a simple and elegant general theory of language structure.

And what about language (c), where adjectives share significant grammatical properties with both verbs and nouns? If analysis (i) were extended to language (c), we would have to say that adjectives are simultaneously a subclass of noun and of verb. There would just be two major word classes, noun and verb, with overlapping identity. Again, the approach followed in analysis (i) has complex and unnecessary consequences.

For every language that has been closely examined, an adjective class can be recognized, although for languages of types (a) and (b) there is an alternative analysis as ‘subclass of verbs’ or ‘subclass of nouns’. If the linguist is interested not only in description of their language but also in the continued refinement of the general grammatical framework in terms of which grammars are written, then analysis (ii) is the alternative to follow.

It is interesting to briefly survey how ‘adjective class’ has been treated in the past. Both the ancient grammar of Sanskrit by Pāṇini and the early grammars of Greek and Latin—which began the western tradition—failed to make any distinction between noun and adjective. It was only at about 1300 CE, in the scholastic grammar of Thomas of Erfurt, that the criterion of gender was invoked—each noun has one inherent gender, whereas an adjective has no gender in itself but may show any of the genders, by agreement with the noun it relates to. On the basis of the European languages they knew, it became the accepted doctrine among linguists that adjectives are a class with similar morphology to nouns, differing from nouns in terms of gender possibilities. Indeed, it appears that Jespersen (1924: 72) considered this to be the only criterion. Since Finnish has no genders he inferred that in this language adjectives could not be distinguished from nouns. There are, in fact, a fair number of other relevant criteria in Finnish—only nouns (not adjectives) take possessive suffixes, and only adjectives (not nouns) take comparative and superlative suffixes.

Australian languages are like the languages of Europe in that adjectives have very similar morphological possibilities to nouns. Some languages have noun classes (similar to genders) and this is accepted as a viable criterion. But for languages without this aid, it is often said that there is no separate class of adjectives. It is instructive to consider the implications of this position. If a language has a category of gender, then it will have a class of adjectives. If

it loses gender, then presumably it loses adjectives as a separate word class. If it then redevelops gender marking, it will regain an adjective class. Such a scenario is surely unacceptable.

In a classic study, Alpher (1991: 22–6) investigates the basis for recognizing a class of adjectives in Yir-Yoront, an Australian language which lacks noun classes/genders. There is no obvious clear-cut criterion to distinguish adjectives from nouns, the two types of word having virtually the same morphological and syntactic properties. Alpher is, however, able to suggest five fairly subtle properties in which nouns and adjectives differ. One he labels ‘grading’: ‘Both “nouns” and “adjectives” occur with postposed *morr* “real, actual, very”’. With common nouns, *morr* has the sense “actual present-day”, as in *kay morr* “the present-day (steel) axe”, or “real and not imaginary”, as in *warrchuwrr morr* “real woman (not one in a dream)”. With “adjectives” susceptible of grading, however, *morr* means “very”: *karntl morr* “very big”, *wil morr* “very bitter”. Such adjectives, moreover, can be quantified with adpositions like *mangl* “a little”, as in *mangl-karntl* “a little bit big”, *wil-mangl* “a little bit bitter”; common nouns lack this possibility.’

The modern discipline of linguistics has been centred on the study of European languages, and is generally undertaken by speakers of European languages. There has, as a consequence, arisen the idea that if a language has an adjective class, then it should be similar to the adjective class in European languages; that is, functioning directly as the modifier of a noun in an NP, acting as copula complement, and showing morphological categories similar to those of nouns (number, case, etc.), quite different from the categories applying to verbs (tense, aspect, mood, etc.).

This has undoubtedly played a role in the failure to recognize an adjective class for languages in which adjectives show a rather different profile, functioning as head of an intransitive predicate (rather than as copula complement), and having some of the same morphological properties as verbs. There is an oft-repeated tradition of saying that in Chinese ‘all adjectives are verbs’ (see, among many others, Hockett 1958: 223; Lyons 1968: 324–5; Li and Thompson 1981: 141; Schachter 1985: 18). This lacks insight. In an important study, Xu (1988) demonstrates a range of criteria for recognizing adjectives to be a separate word class in Chinese. For example, adjectives and verbs show different syntax when modifying a noun within an NP, have different aspectual possibilities when functioning as intransitive predicate, take different derivational possibilities. In addition, reduplication has different semantic implications for the two word classes; see examples (8–9) in §12.5.1.

Even when a linguist does provide criteria for distinguishing adjectives from verbs (in a language where adjectives can function as intransitive predicate), there is often an unwillingness to use the label ‘adjectives’, simply because

these adjectives are so different in grammatical properties from the familiar kind of adjective occurring in European languages. A term like ‘descriptive verb’ may be used instead (for example, Seki 1990, 2000, on Kamaiurá, Tupí-Guaraní branch of Tupí family, Brazil). Oceanic languages typically have an adjective class similar in grammatical properties to the verb class. Buse (1965), writing on Rarotongan, called them ‘statives’, and this label (or ‘stative verbs’) has become institutionalized in Oceanic linguistics.

In summary, there has been a tendency for linguists working on a language of type (a)—where adjectives have similar grammatical properties to verbs—to opt for analysis (i), treating adjectives as a subclass of verbs (or else simply saying that all adjectival concepts are realized as verbs), whereas linguists working on a language of type (b)—where adjectives have similar grammatical properties to nouns—appear to be more open to recognizing ‘adjective’ as a distinct word class. Anyone maintaining that adjectives are a type of verb in Chinese should be consistent and also say that adjectives are a type of noun in Spanish—analysis (i) in each instance. But, as stated above, if the grammar is to be oriented towards the general typological theory which aims to model human language ability, then alternative (ii) should be followed in each case. Every language is then accorded an adjective class (on a par with noun and verb classes), the properties of which will now be discussed.

12.3 Criteria for recognition

As outlined in §3.6, there are two major semantic tasks for an adjective to perform: (A) state a property, and (B) further specify the referent of a noun. In some languages an adjective has an additional function: (C) serving as the parameter in a comparative construction. And in a number of languages we also find (D): an adjective may function like an adverb in further specification of the reference of a verb. These will be discussed in turn.

(A) Make a statement that something has a certain property. There are two syntactic techniques for coding this:

(A-i) In many languages this is achieved by placing an adjective in copula complement slot, as in the English example:

- (1) [The chief]_{COPULA.SUBJECT} [is]_{COPULA.PREDICATE}
[tall]_{COPULA.COMPLEMENT}

As discussed at (f) in §2.5, and at (c) in §3.2, the predicate in this clause is simply the copula verb *is*. The adjective *tall* is *not* part of the predicate, but rather the copula complement, an argument of the predicate (in the same way that *the chief* is an argument, the copula subject).

(A-ii) There are many languages in which an adjective may function as head of an intransitive predicate, and this is then the way in which statement of a property is achieved. It can be exemplified from Fijian:

- (2) [E balavu]_{INTRANSITIVE.PREDICATE} [a tuuraga]_S
 3sgS tall ARTICLE chief
 The chief is tall

In Fijian, the head of an intransitive predicate can be a verb or a noun—as in (9–10) of §11.3—or an adjective—as in (2) here—or a pronoun or a complete NP. In each case the predicate takes identical markers for person and number of subject, for tense, and various semantic modifiers.

(B) As a specification that helps identify the referent of the head noun in an NP. This is shown by the adjective functioning as a modifier within an NP, as in (3) from English and (4) from Fijian. In each example, the modifying adjective is underlined.

- (3) [The tall chief]_S [laughed]_{INTRANSITIVE.PREDICATE}
- (4) [E aa dredre]_{INTRANSITIVE.PREDICATE} [a tuuraga
3sgS PAST laugh ARTICLE chief
balavu]_S
tall
The tall chief laughed

However, the ways in which an adjective may be used to modify a noun vary; they are discussed in §§12.5–6.

In most languages all adjectives have functions (A) and (B). In some, just a few adjectives may be confined to one of these functions. (For examples and discussion of adjectives in English which can occur only as modifier or only as copula complement, see Bolinger 1967; Johansson and Lysvåg 1987: 93–6.) There are also languages in which the entire class of adjectives only has function (B); and there may well be others where it only has function (A). These cases are discussed in §12.6.

(C) Some—but by no means all—languages have a comparative construction. Adjectives may always function as the ‘parameter of comparison’ (and sometimes they are the only words which may be the parameter). Illustration can again be provided from English, in (5), and from Fijian, in (6).

- (5) [Suva]_S [is]_{COP,PREDICATE} [more beautiful]_{CC} [than Nadi]_{STANDARD}

type combine techniques (A-ii) and (A-i)—allowing adjectives to occur both as intransitive predicate and as a copula complement.

- (d) The grammatical properties of adjectives are different from those of nouns and verbs. Typically, in such languages, an adjective cannot function as intransitive predicate and it cannot be sole lexeme within an NP. When it modifies a noun within an NP, it does not repeat any grammatical specifications from the noun. An adjective can function on its own as copula complement, whereas a noun may not be able to. (All of these properties apply for adjectives in English.)

12.4 The semantic content of adjective classes

An adjective class is recognized by virtue of its showing property (A) or (B), usually both, and often also (C) and/or (D), as set out in the previous section. And also by the meanings it covers—the semantic types included within the class.

The idea of dividing the lexicon of every language into a number of semantic types was explained in §1.11 (and applied in §1.9). Semantic types which relate to the adjective class fall into three sets.

SET A. There are four core semantic types, which are typically associated with both large and small adjective classes.

1. DIMENSION—‘big’, ‘small’, ‘long’, ‘tall’, ‘short’, ‘wide’, ‘deep’, etc.
2. AGE—‘new’, ‘young’, ‘old’, etc.
3. VALUE—‘good’, ‘bad’, ‘lovely’, ‘atrocious’, ‘perfect’, ‘proper(/real)’, etc. (And also concepts such as ‘odd’, ‘strange’, ‘curious’, ‘necessary’, ‘crucial’, ‘important’, ‘lucky’.)
4. COLOUR—‘black’, ‘white’, ‘red’, etc.

SET B. Three semantic types are typically associated with medium-sized and large adjective classes.

5. PHYSICAL PROPERTY—‘hard’, ‘soft’, ‘heavy’, ‘wet’, ‘rough’, ‘strong’, ‘clean’, ‘hot’, ‘sour’, etc. And a subclass referring to corporeal properties, e.g. ‘well’, ‘sick’, ‘tired’, ‘dead’, ‘absent’.
6. HUMAN PROPENSITY—‘jealous’, ‘happy’, ‘kind’, ‘clever’, ‘generous’, ‘cruel’, ‘proud’, ‘ashamed’, ‘eager’, etc.
7. SPEED—‘fast’, ‘quick’, ‘slow’, etc.

SET C. A number of other semantic types are associated with large adjective classes in some languages. These include:

8. DIFFICULTY—‘easy’, ‘difficult’, ‘tough’, ‘hard’, ‘simple’, etc.
9. SIMILARITY—‘like’, ‘unlike’, ‘similar’, ‘different(/strange)’, ‘other’, etc.
10. QUALIFICATION—‘definite’, ‘true’, ‘probable’, ‘possible’, ‘likely’, ‘usual’, ‘normal’, ‘common’, ‘correct’, ‘appropriate’, ‘sensible’, etc.
11. QUANTIFICATION—‘all(/whole)’, ‘many’, ‘some’, ‘few’, ‘only’, ‘enough’, etc.
12. POSITION—‘high’, ‘low’, ‘near’, ‘far/distant’, ‘right’, ‘left(/strange)’, ‘northern’, etc.
13. CARDINAL NUMBERS. (In some languages these constitute a separate word class.) And ‘first’, ‘last’ (together with other ordinal numbers).

Small adjective classes tend to have all (or almost all) their members from the core semantic types. A classic example comes from Igbo, where we find an antonymic pair from each of the types in Set A:

DIMENSION	kuku	‘large’	natà	‘small’
AGE	ọhụrụ	‘new’	oncè	‘old’
VALUE	ọma	‘good’	ọjọọ	‘bad’
COLOUR	ọjaí	‘black, dark’	ọca	‘white, light’

Table 12.1 summarizes the number of adjectives in each of the semantic types from Sets A and B for eight languages whose classes range in size from seven to about eighty-four members. In addition, two tiny classes were mentioned in §12.1. That in Kham consists just of three DIMENSION terms, ‘big’, ‘small’, and ‘short’, while the class in Yimas has one DIMENSION adjective, ‘big’, one from the VALUE type, ‘good’, and ‘other’ from the SIMILARITY type.

As an adjective class gets larger than the eight in Igbo, it is likely to include more words from the four core types (for example, ‘long’, ‘short’, ‘red’) and also some PHYSICAL PROPERTY items (for example, ‘raw, green, unripe’, ‘heavy’, ‘light’, ‘sharp’, ‘hot’). Only when an adjective class is much bigger (with at least a couple of score members) is it likely to include terms referring to HUMAN PROPENSITIES (for example, ‘happy’, ‘jealous’, ‘clever’).

Not every small adjective class is as symmetrical as that in Igbo. Indeed, the main members of a semantic type may belong to different word classes. In Yoruba (Kwa family, Nigeria), for instance, there are three adjectives with a meaning similar to ‘good’ but only a verb ‘be bad’ (this language has a small class of about fifteen adjectives: George Madugu 1976). In Jarawara there is an adjective ‘bad’ but only a verb ‘be good’. In Hausa, ‘bad’ is an adjective but ‘good’ can only be rendered in terms of the noun *kyau* ‘that which delights the eye’.

One interesting feature is that, in Table 12.1, the small adjective classes from Papuan and South American languages lack COLOUR terms, but small classes in African languages include them. It does seem that COLOUR terms have a

TABLE 12.1. Summary of the semantic content of small and smallish adjective classes in eight languages

	TOTAL NUMBER	Set A				Set B			Set C
		DIMENSION	AGE	VALUE	COLOUR	PHYSICAL PROPERTY	HUMAN PROPENSITY	SPEED	
Kamula (Papuan area)	c7	2	2	2	—	1	—	—	—
Igbo (Benue-Congo family, Nigeria)	c8	2	2	2	2	—	—	—	—
Hausa (Chadic branch of Afro-Asiatic, Nigeria)	c12	5	2	1	3	1	—	—	—
Jarawara (Arawá family, Brazil)	c14	4	2	3	—	3	—	—	2
Sare (or Kapriman) (Sepik Hill family, Papuan area)	c23	4	4	2	2	11	—	—	—
Somali (Cushitic branch of Afro-Asiatic, Somalia, Ethiopia, and Kenya)	c42	3	1	3	3	15	6	—	11
Akan (Kwa family, Ghana)	c50	10	6	8	3	15	5	3	—
Northern Subanen (Austronesian, Philippines)	c84	18	3	11	10	19	8	6	9

particular cultural saliency for languages spoken in Africa. Indeed, Bing (1991) argues that a small adjective class is in the process of evolution in Krahn/Wobé (Kru family, Liberia), consisting in the first place just of three COLOUR terms.

The most frequently occurring PHYSICAL PROPERTY adjectives typically mean something like 'raw, alive, uncooked'. For example, the sole PHYSICAL PROPERTY adjective in Kamula is *pesekalo* 'raw, alive', and that in Hausa is *danya* 'fresh, raw, unripe'. Jarawara has two terms with this meaning, both used of a fruit—*kini* 'small, immature, not yet reached its full size' and *tati* 'full-sized but not yet ripe and ready to eat'. (The third PHYSICAL PROPERTY adjective in Jarawara is *hinita* 'empty, alone'.)

Lexemes from the semantic types in Set C occur sparsely in small adjective classes, more frequently in larger classes. The three-member class in Yimas includes *ma* 'other' from the SIMILARITY type. Jarawara has *one* 'other' and *hinima* 'all and only'. The classes in Somali and Northern Subanen each include between one and four lexemes from the DIFFICULTY, SIMILARITY, QUANTIFICATION, and POSITION types. Interestingly, the middle-sized class in Akan is composed just of lexemes from semantic types in Sets A and B.

It is interesting to enquire how, in a language with just a small adjective class, the other typical adjectival concepts are coded. The following tendencies have been noted:

- PHYSICAL PROPERTY terms, if not in the adjective class, are generally in the verb class;
- HUMAN PROPENSITY terms, if not in the adjective class, may be in either the noun class or the verb class;
- SPEED terms tend to be in the adjective class if PHYSICAL PROPERTY terms are in this class, and in the adverb class if PHYSICAL PROPERTY terms are in the verb class.

In languages with large adjective classes there may be differences of various kinds between the core and peripheral types. For example, Blackwell (2000) studied how children acquire syntactic functions for adjectives from seven semantic types in English, and found that terms from the DIMENSION, AGE, VALUE, COLOUR, and SPEED types tend to be used first in modifier function, while those from the PHYSICAL PROPERTY and HUMAN PROPENSITY types tend to be used first in copula complement function.

12.5 Distinguishing types of adjective class

The following sections deal in turn with adjective classes which show grammatical properties of types (a)–(d) from §12.1, with special attention to criteria

for distinguishing adjectives from verbs in type (a) and adjectives from nouns in type (b). A major contribution to this investigation is provided by the chapters in *Adjective classes: A cross-linguistic typology* (Dixon and Aikhenvald 2004), which emanated from an International Workshop on ‘Adjective classes’ organized by Alexandra Aikhenvald and myself in August 2002.

For example, previous writers had stated that adjectives cannot be distinguished from nouns in Totonac (Mexico); Levy (2004) provides ample argumentation and exemplification as to why this is untrue. She establishes an adjective class with thirteen monomorphemic (and many derived) members—only adjectives may modify a noun, and may enter into comparative constructions.

Korean is like Chinese in that it has often been suggested that adjectives are indistinguishable from verbs in this language, but Sohn (2004) demonstrates manifold differences. A member of the (large and open) adjective class functions as intransitive predicate like a verb but may not occur with certain moods, has different marking for indicative, may not take certain conjunctive suffixes, and may take the intensifier *-ti*. The suffix *-élan/-ala* has imperative meaning with a verb, but exclamatory function with an adjective. In addition, adjectives can form adverbs, and can occur in comparative and superlative, among many other criterial differences.

12.5.1 *Distinguishing between adjective and verb classes*

Where both adjectives and verbs can fill the intransitive predicate slot—in a language of type (a)—criteria for distinguishing the two word classes include: (1) different possibilities within the predicate slot; (2) different transitivity possibilities; (3) different possibilities as modifier within an NP; (4) different possibilities in comparative constructions; (5) different possibilities for forming adverbs (that is, modifiers to verbs). We can discuss these one at a time. (There is a fuller list in §12.11.)

1. Different possibilities within the predicate slot

In some languages exactly the same morphological processes and syntactic modifiers may apply to a verb and an adjective within a predicate. However, in many languages the possibilities vary.

Most typically, an adjective is far more restricted than a verb when it occurs as predicate head. For example, in the Iroquoian language Cherokee (Feeling 1975), a verb as predicate head allows three types of prefix and two varieties of suffix. In contrast, an adjective as predicate head allows only pronominal prefixes:

(7) Predicate structure in Cherokee

verb					
±initial prefixes	adjective		adjective		
	+pronominal prefix	±reflexive prefix	+verb/ adjective root	±non-final suffixes	±final suffix
8 orders, including negative, 'again', 'since'				13 orders, including reversive, repetitive, completed, tense/aspect, interrogative	pre-incipient, future, infinitive, tense, etc.

Note that only those positions which are obligatory for verbs (indicated by '+') are found with adjectives.

Another language in which adjectives have more limited possibilities than verbs is Temiar (Aslian branch of Austro-Asiatic; Benjamin 1976: 184); only verbs (not adjectives) may take the modal affix *-m-* and form causatives.

In other languages, verbs allow some modifiers which adjectives lack, and adjectives permit some which verbs lack. For example:

- In Vietnamese (Nguyễn 1987: 791), only adjectives can be preceded by *rất* 'very' and *khát* 'rather', and only verbs can occur with the exhortative particle *hãy*.
- In Chamorro (Austronesian, Guam; Topping 1973: 231), only verbs can take a modifier of manner, and only adjectives may take an intensifier.
- In Kamaiurá (Seki 2000: 64), adjectives differ from verbs in that (a) verbs but not adjectives can occur in the circumstantial mode; (b) in indicative, exhortative, and imperative moods, adjectives take pronominal proclitics while verbs take pronominal prefixes; (c) the gerund is marked by *-ram* on a verb but by *-m* on an adjective.

Typically, adjectives show fewer possibilities for mood than do verbs, particularly for imperative (and its subtypes such as hortative).

Adjectives may have wider possibilities than verbs. As shown in the appendix to this chapter, the pre-head predicate modifier *rui* 'more than a usual amount' in Fijian is allowed when the predicate head is an adjective, not when it is a verb.

Another recurrent criterion concerns reduplication possibilities. In Chinese (Xu 1988), a verb when reduplicated carries the meaning 'do a little bit', for example:

- (8) dòng 'to move' dòngdòng 'to move a little'

In contrast, when an adjective is reduplicated, the semantic effect is 'intensification of the quality', as in:

- (9) hóng 'red' hónghóng 'vividly red'

In Qiang (Tibeto-Burman, China; LaPolla and Huang 2004), reduplication usually signifies reciprocity for verbs but either plurality or intensification or both for adjectives (depending on the formal nature of the reduplication). In Mupun (Chadic branch of Afro-Asiatic, Nigeria; Frajzyngier 1993: 63–73), both verbs and adjectives may reduplicate, which serves as a process of nominalization. But whereas a reduplicated verb just forms an abstract noun (e.g. *rán* 'write', *ránrán* 'writing'), when an adjective is reduplicated it adds a sense of intensity (e.g. *móol* 'thick', *məmóol* 'great thickness'). (And see the note on methodology concerning the semantics of reduplication, under (2) in §12.5.2.)

Adjectives may also differ from verbs in possibilities for derivation. In Mandarin Chinese (Xu 1988), different sets of derivational suffixes apply to verbs (e.g. agentive nominalizer *-zhě*) and to adjectives (e.g. verbalizer *-huà*).

2. Different possibilities for transitivity

In Fijian, almost every verb can be used either intransitively (then not bearing a suffix) or else transitively (with a transitive suffix). For some verbs the intransitive subject (S) relates to the transitive subject (A), and for others S relates to the transitive object (O). For example (full details are in Dixon 1988a: 200–19):

- | | | | | | |
|------|------------|---------------|--------------|----------------------|------------|
| (10) | | INTRANSITIVE | | TRANSITIVE | |
| | type S = A | <i>la'o</i> | 'go' | <i>la'o-va</i> | 'go for' |
| | | <i>dredre</i> | 'laugh' | <i>dredre-va'ina</i> | 'laugh at' |
| | type S = O | <i>cori</i> | 'be tied' | <i>cori-ta</i> | 'tie' |
| | | <i>rogo</i> | 'be audible' | <i>rogo-ca</i> | 'hear' |

Unlike verbs, adjectives do not take a transitive suffix; that is, adjectives only occur in intransitive—not in transitive—predicates. (There are a few verbs which are only used intransitively, including *gaadee* 'stroll' and *bona* 'stink'. These are distinguished from adjectives by other tests, e.g. their non-occurrence with *rui* 'more than a usual amount'.)

It is interesting to study the allocation of adjectival concepts into word classes in Fijian. Words from the DIMENSION, AGE, VALUE, COLOUR, PHYSICAL PROPERTY, and SPEED types are adjectives, but HUMAN PROPENSITY items are placed in the verb class. It is not hard to see why this should be so.

Most adjectives in English just describe a property of some thing (for example, 'big', 'new', 'heavy', 'sharp'). However, HUMAN PROPENSITY adjectives describe an attitude on the part of one participant towards someone or something else. When they function as copula complement, this second argument may be shown by an optional prepositional phrase; for example 'happy (about)', 'clever (at)', 'jealous (of)', 'afraid (of)', 'brave (at)', 'angry (at/about)'.

These ideas are coded in Fijian by verbs, each of which can be used intransitively (with no suffix) or transitively (with a suffix); they are all of type S = A. The O of the transitive verbs corresponds to the prepositional argument in English. For example:

- | | | |
|------|---------------------|--------------------------------|
| (11) | INTRANSITIVE | TRANSITIVE |
| | maarau 'be happy' | maarau-ta'ina 'be happy about' |
| | vu'u 'be clever' | vu'u-ta'ina 'be clever at' |
| | vuuvuu 'be jealous' | vuuvuu-ta'ina 'be jealous of' |
| | rere 'be afraid' | rere-va'ina 'be afraid of' |
| | dou 'be brave' | dou-va'ina 'be brave at' |

Now some verbs in Fijian may choose between two transitive suffixes, which bring different participants into the second core argument slot. For example:

- | | | | |
|------|--------------|-------------------------|--|
| (12) | INTRANSITIVE | TRANSITIVE ₁ | TRANSITIVE ₂ |
| | dabe 'sit' | dabe-ca 'sit on' | dabe-va 'sit (waiting) for' |
| | vana 'shoot' | vana-a 'shoot at' | vana-ta'ina 'shoot with
(e.g. a gun)' |

A few of the verbs relating to the HUMAN PROPENSITY semantic type can also make a choice of transitive suffix, effectively corresponding to a choice of preposition in English. For example:

- | | | | |
|------|-----------------|---------------------------------------|--|
| (13) | INTRANSITIVE | TRANSITIVE ₁ | TRANSITIVE ₂ |
| | pu'u 'be angry' | pu'u-ca 'be angry at
(e.g. child)' | pu'u-ca'ina 'be angry about
(e.g. child's behaviour)' |

3. Different possibilities as modifier within an NP

There are a number of ways in which adjectives may differ from verbs in the modification of a head noun within an NP. The most straightforward difference is that only an adjective can directly modify a noun, not a verb. This appears in Kamaiurá (Seki 2000: 70, 117), in Tigak (Austronesian, Papua New Guinea; Beaumont 1980: 85), and in Papatlantla Totonac (Levy 2004).

In some languages with a verb-like adjective class, both verb and adjective can modify a noun through a process of nominalization, but there may be

A number of languages have adjectives and verbs modifying a noun through a relative clause construction. In Mojave (Yuman, California; Schachter 1985: 19), a relativizing particle is obligatory with a verb, when modifying a noun, but optional with an adjective. In Èdo (Kwa family, Nigeria; Qm̧ŗuyi 1986), both adjective and verb require a relative marker when in modifying function, but there is phonological reduction of the relative marker only in the case of adjectives. In Bororo (Macro-Jê grouping, Brazil; Crowell 1979), a relative clause which has a verb as predicate must precede the noun it modifies, whereas if the head of the relative clause predicate is an adjective the relative clause may either precede or follow the noun.

(14a) wì(l) klī
child little
a little child

(14b) wì(l) mà: klī
child RELATIVE.MARKER little
the little child

(15) (a) *óbi ójòó* ‘hard-heartedness, meanness’ as an inherent character trait (literally ‘heart bad’)
 (b) *úzó dí njó* ‘road which is bad’, which can, after all, be repaired

There may be other kinds of restriction on a verb in modifying function, which do not apply to an adjective. In Chemehuevi (Uto-Aztec; Press 1979: 58), verbs must co-occur with a demonstrative when modifying a noun;

adjectives need not. In *Tukang Besi* (Austronesian, Indonesia; Donohue 1999: 144, 303–7), adjectives can modify a noun directly but verbs require a subject focus marker. In *Mupun* (Frajzyngier 1993: 69), both adjectives and verbs may only modify a noun together with the relative clause marker *de*; but whereas verbs require a subject to be stated within the relative clause (this is underlined in (16a)), adjectives do not, as in (16b).

- (16) (a) n-dem ngwe [de wu cii]
 1sg-like man RELATIVE.MARKER 3m refuse
 I like a man who refuses
- (b) n-dem ngwe [de ci]
 1sg-like man RELATIVE.MARKER different
 I like a different man

As described in the appendix to this chapter, there are different possibilities for the classifier in a ‘clausal NP’ in Fijian, depending on whether its predicate is headed by an adjective or a verb.

4. Different possibilities in comparative constructions

Not all languages have a comparative construction (types of comparative construction were illustrated in (5–6) above). In some of the languages that do, the ‘parameter of comparison’ can only be an adjective, but in others there are wider possibilities. In *Èdo*, for example, both adjectives and verbs may occur in comparative constructions (Ọmọruyí 1986). However, in some languages only adjectives can be compared, and this furnishes a criterion for distinguishing between adjective and verb classes; such a property applies to *Toba-Batak* (Austronesian, Indonesia; Nababan 1981: 7), *Korean* (Sohn 2004), *North-East Ambae* (Austronesian, Vanuatu; Hyslop 2004), *Qiang* (LaPolla and Huang 2004), and *Lao* (Enfield 2004).

5. Different possibilities for forming adverbs

In Fijian, for example, adverbs can be formed from adjectives (but generally not from verbs) by means of the prefix *va'a-*; for example, *va'a-levu* ‘greatly’ from *levu* ‘big’ and *va'a-dodonu* ‘correctly’ from *dodonu* ‘correct’. In Japanese, too, it is mainly adjectives which may function as adverbs, this being one of the properties which link the two adjective classes into one macro-class (see the discussion in §12.7).

There are other properties which recur. For example, adjectives typically behave in a special way within Serial Verb Constructions; this applies for *Tariana* (Aikhenvald 2004a), *North-East Ambae* (Hyslop 2004), *Semelai* (Aslian branch of Austro-Asiatic, Malaysia; Kruspe 2004b), and *Qiang*

(LaPolla and Huang 2004). In many languages, some intensifiers ('very') and quantifiers ('much') typically apply to adjectives and not to verbs.

The discussion in this subsection has been of languages where adjectives function as intransitive predicate, rather than as copula complement. Not all languages have a copula construction. One might expect a correlation: languages in which adjectives can be intransitive predicate might be thought likely to lack a copula construction, with languages for which adjectives cannot function as intransitive predicate being likely to have a copula construction. From examination of a range of languages, it appears that there is in fact *no such* correlation. That is, whether or not a language has a copula construction is quite independent of whether or not adjectives can be intransitive predicates.

Languages with verb-like adjectives differ with respect to the possibilities for using an adjective in the copula complement slot. In Mupun (Frajzyngier 1993), a copula complement can only be an NP (e.g. 'this man is the chief'), not an adjective. In Chinese (Xu 1988), an adjective can occur as copula complement only when in nominalized form, as in (17a). This has a rather different meaning from a clause in which the adjective is intransitive predicate, as in (17b).

- (17) (a) [táng]_{COPULA.SUBJECT} [shì]_{COPULA}
 sugar is
 [tián-de]_{COPULA.COMPLEMENT}
 sweet-NOMINALIZER
 Sugar is a sweet thing
- (b) [táng]_S [tián]_{INTRANSITIVE.PREDICATE}
 sugar sweet
 Sugar is sweet

The nominalizer can be omitted from a sentence like (17a) in marked circumstances, when it is in emphatic or contrastive function.

In some languages where adjectives may function both as head of an intransitive predicate and also as modifier in an NP, there is a definite preference for employing them in the former function. For example, a Korean will be more likely to say 'Men are numerous' than 'There are many men' (Ramstedt 1939: 35). And Kimball (1991: 484) reports that in the Muskogean language Koasati there is a preference for saying, literally 'The willow is long-, green-, and many-leaved', rather than (as in English) 'The willow has many long green leaves'.

(Interestingly, in languages where an adjective can function as copula complement or modifier in an NP, no preferences have been reported with respect to one of these syntactic possibilities.)

12.5.2 Distinguishing between adjective and noun classes

There are a number of kinds of criteria for distinguishing adjectives from nouns, where these share grammatical properties in a language of type (b): (1) the internal syntax of NPs; (2) morphological possibilities; (3) the comparative construction; and (4) adverbial use. These will be discussed one at a time. (There is a fuller list in 12.11.)

1. The internal syntax of NPs

The prototypical NP has a noun as head and one (or, sometimes, several) adjectives as modifiers. Where this scheme is closely adhered to there is no difficulty in distinguishing between nouns and adjectives; this applies in Hua (Papuan region, Haiman 1980: 268–9), in Basque (Saltarelli 1988: 144), in Upper Necaxa Totonac (Beck 2000), and in Papatla Totonac (Levy 2004).

However, there are some languages in which a noun may also function as modifier. Generally, the possibilities for noun modifiers are rather limited. It may be that an NP can include no more than one noun modifier, but several adjective modifiers. And whereas every, or almost every, adjective is likely to function as modifier within an NP, only a limited set of nouns may have this function. Jarawara is typical in that the only nouns used as modifiers are those referring to sex (*fana* ‘woman, female’ and *maki* ‘man, male’) and to material (such as *jati* ‘stone’, *awa* ‘wood’). In Tariana just human nouns may function as modifier.

In some languages a noun can be modifier only under particular grammatical conditions. In Bilin (Cushitic branch of Afro-Asiatic, Ethiopia; Palmer 1967: 206), for example, a modifying noun must be in genitive form. In Igbo, when a noun is modified by another noun or by a number, these form an ‘associative construction’ (with tone change); this does not apply when a noun is modified by an adjective (Welmers and Welmers 1969).

The other variation on the prototypical pattern is for an adjective to make up a complete NP. In some languages this can be described as the adjective becoming head of the NP, but in most instances it is better treated as an NP whose head noun has been omitted (under certain discourse conditions), which consists just of a modifier. In languages with gender, the ellipsed noun is likely to determine the gender of the modifier adjective. The possibilities for ellipsis can depend on some characteristic of the head noun; for example, in Modern Standard Arabic, only a noun with human reference can be omitted.

Generally, when an adjective occurs without a noun in an NP, it may not receive any syntactic modification. That is, an NP may consist of a noun plus one or more adjectives; or it may just consist of an adjective; this applies, for example, in Amele (Papuan region; Roberts 1987: 155).

A further criterion, in some languages, lies in the existence of a pre-modifier ‘very’, which can apply to adjectives but not to nouns. This applies in Buriat (Mongolian, Mongolia; Poppe 1960) and in Quechua (Cole 1982: 99), among many other languages.

2. Morphological possibilities

One of the most useful criteria for distinguishing between nouns and adjectives is gender or noun classes. In Latin, for instance, each noun belongs to just one of the three genders, while an adjective can be in any gender, agreeing with the noun it is modifying. A similar criterion is given by Sokolov (1967: 43) for Avestan and by Fortune (1942: 55–6) for the Papuan language Arapesh; and see the discussion of Russian in Corbett (2004).

However, this criterion is not always watertight. In Dyirbal a noun is generally accompanied by a noun marker, a determiner-like element which indicates location/visibility, agrees with the noun in case, and marks the noun class of the noun (this is not shown on the noun itself). Most nouns relate to just one noun class, while most adjectives can occur with a noun marker of any class (in agreement with the gender of the noun they modify). Compare (noting that in fact the words in an NP can occur in any order):

- (18) (a) *bayi yara* ‘man’
 (b) *balan yibi* ‘woman’
 (c) *balam mirrañ* ‘black bean’
 (d) *bala diban* ‘stone’
- (19) (a) *bayi (yara) midi* ‘small (man)’
 (b) *balan (yibi) midi* ‘small (woman)’
 (c) *balam (mirrañ) midi* ‘small (black bean)’
 (d) *bala (diban) midi* ‘small (stone)’

The noun marker ‘there’ (shown by initial *ba-*), in absolutive case, has four forms, masculine *bayi*, feminine *balan*, edible *balam*, and neuter *bala* (see Dixon 1972 for full details). Now a head noun can be omitted from an NP (under discourse conditions). Thus, while the noun *yara* ‘man’ can only occur with *bayi*, *yibi* ‘woman’ only with *balan*, etc., an adjective such as *midi* ‘small’ can occur with all four noun markers, as in (19a–d).

However, there is a handful of ‘hybrid’ nouns that can take either masculine or feminine markers; these include *bayi/balan jaja* ‘male/female baby’ and *bayi/balan bimu* ‘father’s elder brother/sister’. And while adjectives such as *midi* ‘small’ can modify any noun, there are adjectives which—by virtue

of their meaning—may only modify a noun which has human reference; for example *wugija* ‘generous, always sharing things’ and *jilbay* ‘experienced/expert at some task’. There are thus a few nouns which can occur with either masculine or feminine noun marker, and a few adjectives which are restricted to masculine and feminine markers. That is, while noun class co-occurrence is a pretty good criterion for distinguishing nouns and adjectives in Dyrbal, it is not perfect. Other criteria need to be brought in to deal with words like *jaja*, *bimu*, *wugija*, and *jilbay*.

In some languages only certain adjectives may take gender or noun class marking. This applies in Swahili, where the adjective class has two subclasses. One subclass consists of about fifty native roots which take the concordial prefix of the noun they modify; the other subclass involves a score or so of borrowed adjectives (mostly from Arabic) which do not take the prefixes. However, the subclasses are linked by all their members sharing other grammatical properties.

In Jarawara, some inalienably possessed nouns and some adjectives show a gender distinction. However, the rules for gender agreement within an NP are different for the two word classes. Following a non-singular 1st or 2nd person as head of an NP, plus a possessed noun, a further possessed noun will be in masculine form while an adjective will show feminine form (this is illustrated by examples (34–6) in Dixon 2004c).

The gender/noun class distinction spans morphology and syntax. A similar criterion is provided by classifiers; generally, an adjective may occur with a larger set of classifiers than may a noun (potentially, an adjective may occur with all classifiers, while a noun may be limited to one, or to just a few). (See Dixon 1977a: 122 on the Australian language Yidiñ.)

A survey of the literature shows a number of different kinds of morphological differences between nouns and adjectives. Only a noun may take possessive affixes—in Finnish and in Hungarian, and also in the Papuan language Alamblak (Bruce 1984: 74 provides a most useful table of the various morphological differences between adjectives and the other word classes in Alamblak).

Typically, adjectives will accept only a subset of the affixes available to nouns. Arnott (1970: 78–130) states that in Fula (Atlantic family) an adjective takes all noun class suffixes but a noun will only take some, whereas nouns take all the remaining nominal suffixes, while adjectives just accept a selection of them (for example, singular *-wo* and plural *-be* are confined to nouns).

In Maasai (Chari-Nile subgroup of Nilotic, Kenya; Tucker and Mpaayei 1955: 3–13), a noun—as head of an NP—inflects for gender and number, while an adjective—as modifier—inflects only for number. But if the head noun is omitted, so that the NP consists just of an adjective, then that inflects for

gender and number, like a noun. The principle appears to be that number marking goes on every word in an NP, but gender marking just onto one word (a head noun, if present, otherwise an adjective).

Another distinguishing feature is when a given grammatical form has different allomorphs when used with nouns and with adjectives. For example:

- In Awa Pit (Barbacoan family, Ecuador/Colombia; Curnow 1997a: 91) the ‘collective action’ suffix has allomorph *-tuzpa* with a noun, and *-tuz* on an adjective which makes up a full NP.
- In Venda (Bantu, South Africa; Poulos 1990: 121), both adjectives and nouns take noun class prefixes but with some differences of form; for example, class 15 is shown by *hu-* on an adjective but by *u-* on a noun.

In some languages a given suffix may be used on both noun and adjective, but with a difference of meaning. For the Australian language Bandjalang, Crowley (1978: 30) describes how the suffix *-bu* means ‘still’ with an adjective (for example *mirij-bu* ‘still alive’) but ‘along’ with a noun (for example *balun-bu* ‘along the river’).

Reduplication is another grammatical process which may have different semantic effect with nouns and with adjectives. In the Australian language Emmi (Ford 1998: 140), reduplication of a noun indicates plurality (for example, *perre* ‘grub’, *perreperre* ‘grubs’) while reduplication of an adjective indicates intensity (for example, *duk* ‘big’, *dukduk* ‘very big’).

A note on methodology is in order here. It might be suggested that the semantic effect of reduplication is a consequence of the semantic nature of a lexeme, not of its grammatical word class. On this principle, lexemes referring to ‘properties’ would be marked for intensity, and not for plurality, whatever word class they belonged to. That this is untrue is shown by comparing the semantic effect of reduplicating nouns and adjectives in Emmi and in Dyirbal, where reduplication indicates plurality on both nouns and adjectives:

(20)	REDUPLICATION OF NOUN
Emmi	plural, e.g. <i>perreperre</i> ‘grubs’
Dyirbal	plural, e.g. <i>jambunjambun</i> ‘grubs’
	REDUPLICATION OF ADJECTIVE
Emmi	intensity, e.g. <i>dukduk</i> ‘very big’
Dyirbal	plural, e.g. <i>bulganbulgan</i> ‘many big (things)’

This shows that the semantic effect of reduplication does not here operate on the basis of the semantics of the lexeme involved, but rather upon its word class, with different languages having varying specifications for their word classes.

The placement of case marking seldom provides a criterion for distinguishing adjectives from nouns. Case indicates the function of an NP in its clause. Each language has its own rule for the assignment of case to the words within an NP—it may go onto every word, or just the last word, or just the first word, or just the head. Whether or not an adjective bears case will depend on these rules, and on the position of the adjective in the NP. For example, in Bilin, case goes onto the last word in an NP. If there is an adjective modifier (which follows the head noun), this will take case; only if there is no modifier to a noun will case attach to the noun (Palmer 1967). However, in Buriat, case goes onto the head word. If an adjective functions as modifier, it takes no case affix; if an adjective makes up a whole NP, then it does take case (Poppe 1960: 76).

Note, though, that the case system for adjectives may differ in size from that for nouns. Nichols (1994: 95–9) states that in Ingush (North-East Caucasian) nouns may select from eight cases but adjectives just from two—nominative (corresponding to nominative on nouns) and oblique (corresponding to genitive, dative, ergative, instrumental, locative, comparative, and allative on nouns). Estonian is similar to Ingush in this feature.

3. Comparative construction

In some languages—for example, Finnish, Hungarian, Russian, and Papantla Totonac (Levy 2004)—only an adjective can occur as the parameter of comparison in a comparative construction, and this serves to distinguish adjectives from nouns. However, in other languages nouns and adjectives share this property and it is thus not a relevant criterion; this applies for Portuguese, for Sanskrit (Bhat 1994: 181–2) and also for Dyrbal (Dixon 1972: 226–8).

4. Use as adverbs

In Tariana, in Mandarin Chinese, and in Buriat (Poppe 1960), only adjectives—not nouns—also have adverbial function.

A further distinguishing property is given by Johanson (2006) in his discussion of the adjective classes in South Siberian and other Turkic languages—nouns answer the questions ‘who?’ and ‘what?’, whereas adjectives answer questions ‘what kind of?’ and ‘how?’. This would apply for many other languages.

12.5.3 *Adjectives grammatically similar to both verbs and nouns*

The last two sections have discussed languages in which adjectives have similar grammatical possibilities to verbs, and languages in which they have similar possibilities to nouns. What more natural than for a language to combine these

features—for an adjective to inflect like a noun when occurring in an NP, and to inflect like a verb when functioning as predicate head, type (c) from §12.1?. In fact, a rather small number of languages appear to be of this type. We can present a number of well-documented examples.

1. In languages from the Berber subgroup of Afro-Asiatic, adjectives—like other lexemes—have triconsonantal roots, e.g. *m - l - l* ‘white’. An adjective will inflect for gender and number, like a noun, when in an NP and as copula complement. It will inflect for tense and for person and number of the subject, like a verb, when functioning as head of an intransitive predicate. (See, for example, Aspinion 1953; Sadiqi 1986.)
2. In the Australian language Nunggubuyu (Heath 1984: 152), an adjective can function as modifier in an NP; it then takes a noun class prefix and case and number suffixes, like a noun. An adjective may also take an intransitive subject prefix, just like a verb; it must then be functioning as an intransitive predicate. But note that, as in many languages, an adjective has more limited morphological possibilities than a verb in this slot; it can only take tense and aspect suffixes if the inchoative derivation suffix is first added. It is likely that in Nunggubuyu adjectives are just beginning to take on grammatical properties similar to those of verbs; see §12.8 below.
3. In Tariana, an adjective can modify a noun and then agrees with it in number and classifier. It may also function as predicate head and may then take tense/evidentiality, aspect, mood (other than imperative), and most other suffixes that are available for a verb.

There is an explanation for the rich syntactic possibilities available to adjectives in Tariana. Their functioning as intransitive predicate is an inherited property, shared with other Arawak languages. Their functioning as copula complement is a property which has been borrowed from East Tucanoan languages, as one aspect of the large-scale diffusion of grammatical patterns that characterizes the Vaupés linguistic area (to which Tariana and East Tucanoan languages belong). See examples (8a–9) in §14.2 and Aikhenvald (2004a: 153–74).

4. Sapir’s grammar of Takelma (Takelman family, Oregon) was written at a time when many grammars of North American languages silently included adjectival lexemes in either the verb or the noun class. He comments (Sapir 1922: 255–6):

Adjectives can not in Takelma without further ado be classed as nouns or verbs, as they have certain characteristics that mark them off more or less clearly from both; such are their distinctly adjectival suffixes and their peculiar method of forming the plural. In some respects they closely approach the verb, as in

the fact that they are frequently preceded by body-part prefixes... They differ, however, from verbal forms in that they can not be predicatively used... nor provided with the pronominal suffixes peculiar to the verb; a first or second person relation is brought about by the use of appropriate forms of the copula *ei-* BE. They agree with the noun and pronoun in being frequently followed by the distinctively denominative suffix *-t'a*... and in the fact that, when forming part of a descriptive noun, they may take the personal endings peculiar to the noun... It thus appears that the adjective occupies a position midway between the noun and the verb, yet with characteristics peculiar to itself.

This appears to be another language of type (c).

There are hints in the literature of further languages of this type, but insufficient information to check them out in detail. For instance, Swadesh (1946: 320–1) says of Chitimacha (isolate, Louisiana): ‘very much like certain kinds of verbs is the adjective, part of whose inflection coincides with that of the verb, but which has two additional forms called the substantival singular and plural. Moreover, it is precisely the substantival forms which are the most commonly used.’

In other languages, adjectives may be most similar to one of nouns and verbs, but have some properties in common with the other. In Upper Necaxa Totonac, adjectives have grammatical properties similar to those of nouns. However, an adjective as copula complement may be modified by *tunká* ‘very’; nouns do not take *tunká*, but intransitive state verbs (e.g. ‘be ashamed’) do (Beck 2000: 233–4). In the Australian language Emmi (Ford 1998: 139–40), adjectives inflect like nouns but are negated, like verbs, by the particle *way* (nouns, in contrast, are negated by the negative copula *piya*).

12.5.4 Adjectives grammatically different from both verbs and nouns

In a further set of languages, the morphological and syntactic properties of adjectives differ from those of verbs and of nouns, type (d) from §12.1. I will mention just four examples of this.

1. English. Only nouns may take a plural suffix; only verbs may take tense-aspect suffixes; only adjectives may take comparative and superlative marking, shown either by affixes (*-er*, *-est*) or by pre-modifiers (*more*, *most*). Generally, an adjective cannot occur as head of an NP (while a noun can), nor as predicate (while a verb can). Only an adjective can occur alone as copula complement, as in *John is tall*; a noun requires an article or other determiner in this slot, as in *John is a doctor/my son*.
2. Teribe (Chibchan family, Panama and Costa Rica; Quesada 2000). Verbs take aspect, modality, and mood suffixes, and nouns take plural marking;

none of these is available to adjectives. An adjective may modify a noun in an NP, may occur in a comparative construction, and may be complement in a verbless (copula-type) construction.

3. Mam (Mayan family, Guatemala and Mexico). England (2004) describes an adjective class which, like those in other Mayan languages, has about fifty monomorphemic members. All types of words may occur as intransitive predicate (although non-verbs have limited TAM possibilities). Properties of adjectives which distinguish them from nouns and from verbs include: taking certain derivational suffixes, forming a comparative, modifying a noun, being used adverbially, and functioning as complement of an existential predicate.
4. Tunica (isolate, Louisiana; Haas 1941). Verbs take a variety of morphological processes for aspect and modality, and also mark person and number of the subject argument. Nouns take article-like prefixes. Adjectives constitute a quite distinct word class, not undergoing any morphological processes. They may function as copula complement and may modify a noun within an NP.

12.6 Languages with restricted functional possibilities for adjectives

As described in §12.3, in the great majority of languages adjectives have two canonical functions:

- (A) In a statement that something has a certain property, coded through the adjective functioning either as intransitive predicate (A-ii) or as copula complement (A-i).
- (B) As a specification that helps focus on the referent of the head noun in an NP, the adjective functioning as modifier to the head.

In a fair number of languages, adjectives can have one or both of two further properties:

- (C) As the parameter of comparison in a comparative construction.
- (D) As modifier to a verb, in adverbial function.

There are some languages whose adjectives do not have both (A) and (B) functions. They can be divided into three classes.

Class (1), adjectives which just function as modifier within an NP, and lack function (A). This applies to Malayalam (Dravidian, South India; Asher and Kumari 1997: 350–5; Variar 1979: 24–36); to Hua (Haiman 1980:

268–9); to Yoruba (George Madugu 1976); and to Dagbani (Gur family, Ghana; Olawsky 1999 and p.c.). In Yoruba, for example, one simply cannot say ‘Olu is good’ or ‘Ibadan is large’; a copula complement must be an NP, including a head noun, as in ‘Olu is a good girl’, ‘Ibadan is a large city’ (George Madugu 1976: 93).

Class (2), adjectives which just function as copula complement, and lack function (B). A number of languages from the northern branch of the Carib family are of this type, including Hixkaryana (Derbyshire 1979: 81; 1985: 10–15, 27–8) and Tiriyó (Meira 1999: 334–6). These languages have a word class whose members may:

- function as copula complement (like nouns, and unlike verbs);
- have adverbial function, as modifier to a verb.

They cannot directly modify a head noun in an NP but must first be nominalized (just as a verb must be).

It was remarked in §12.2 that, as a consequence of the Eurocentrism of much linguistic work, there is sometimes a reluctance to use the term ‘adjective’ for a class of words which does not have similar grammatical properties to nouns (as adjectives do in European languages). From this viewpoint, words which cannot function as modifier within an NP (except in nominalized form) may appear un-adjective-like. As a consequence, Derbyshire (1979, 1985) prefers not to use the label ‘adjective’ for the class of words just described, in Hixkaryana and other North Carib languages. However, Derbyshire’s preferred label, ‘adverbs’, is scarcely appropriate; an adverb cannot normally occur as copula complement. This word class in Carib languages is certainly not a typical adjective class, since it does not have function (B), but it is no more untypical than those languages—in class (1)—whose adjectives only occur as nominal modifier and lack function (A).

The semantic content of the non-prototypical adjective classes in languages of classes (1) and (2) does accord with the scheme outlined in §1.11. Yoruba has about twenty-five members in its adjective classes, Malayalam has nine, Hua has four, and Dagbani has about seventy. They are:

- Yoruba: five DIMENSION, four AGE, four VALUE, three COLOUR, five PHYSICAL PROPERTY, four HUMAN PROPENSITY.
- Malayalam: four DIMENSION, three AGE, one VALUE, plus ‘humble’.
- Hua: two DIMENSION (‘big’, ‘little’), two PHYSICAL PROPERTY (‘raw, false’, ‘wild, not tame’).
- Dagbani: five or more in each of DIMENSION, AGE, VALUE, COLOUR; about twenty in PHYSICAL PROPERTY, etc.

For class (2), about thirty adjectives are reported for Hixkaryana and about forty-four for Tiriyó. These include terms from DIMENSION, VALUE, PHYSICAL PROPERTY, HUMAN PROPENSITY, and SPEED (AGE terms are nouns while COLOUR are derived adjectives). But the adjective class in North Carib languages also includes terms for QUANTIFICATION ('all') and NUMBER ('one', 'two') together with items relating to PLACE ('hither', 'thither', 'beyond', 'this side of') and TIME ('later', 'soon', 'now', 'yesterday'). It is perhaps not surprising that the Carib adjective class, which functions only as copula complement and as adverb, should include words of place and time which are typically coded as adverbs in other languages.

Class (3), adjectives which only function as intransitive predicates, and lack function (B). In some of the languages with verb-like adjectives that can function as intransitive predicate, both adjectives and verbs may modify a noun through a relative clause construction. As mentioned under (3) in §12.5.1, it is often the case that a relative clause marker is obligatory with a verb but optional with an adjective; adjectives could thus be said to have function (B), modifying a noun directly; these would thus be prototypical adjective classes in terms of their syntactic functions.

Edo might be a candidate for class (3), since a relative marker is required with adjectives as well as with verbs. However, as pointed out in §12.5.1, there is phonological reduction of the relative marker only in the case of adjectives.

12.7 Languages with two adjective classes

In §12.3, the criteria for recognizing an adjective class were set out as: a word class distinct from noun and verb, including words from the prototypical adjective semantic types, and functioning either as intransitive predicate or as copula complement; and/or modifying a noun in an NP. It is possible for there to be two word classes which satisfy these criteria; that is, a language might have two adjective classes. I will mention three well-documented instances of this.

1. Macushi (information from Abbott 1991 and p.c.) has an adjective₁ class similar to that described in §12.6 for the related North Carib languages Hixkaryana and Tiriyó. Unlike its relatives, Macushi has a second small class, adjective₂, whose members may modify a noun in an NP (or make up a full NP, with the head noun ellipsed). They may not function as modifiers to the verb (that is, as adverbs), and can only be copula complement when the denominalizer *pe* is included. In summary:

(21)		ADJECTIVE ₁ CLASS	ADJECTIVE ₂ CLASS
	can modify noun	only with nominalizer	✓
	can make up whole NP	—	✓
	can be copula complement	✓	only with denominalizer
	can function as adverb	✓	—

According to Abbott (1991: 88, 129–30), each class is rather small. The reported members are:

(22)		ADJECTIVE ₁ CLASS	ADJECTIVE ₂ CLASS
	DIMENSION	'big', 'deep'	'long', 'fat'
	VALUE		'good', 'bad'
	PHYSICAL PROPERTY	'hard', 'well'	'hot', 'cold'
	SPEED		'fast'
	QUANTIFICATION	'all', 'few', 'many'	
	NUMBER	'two'	
	PLACE	'here', 'there'	
	TIME	'today', 'yesterday', 'long ago', 'later', 'regularly', 'afternoon'	

It will be seen that two of the recurrent semantic types for adjective classes, DIMENSION and PHYSICAL PROPERTY, have members in both classes.

2. Japanese has two adjective classes, each of which is quite large. Their grammatical properties can be summarized as follows (based on Backhouse 1984; Takeuchi 1999: 81–2, and the full discussion in Backhouse 2004).

- What are called inflected adjectives may function as intransitive predicates, like verbs. They take most of the inflections available to verbs, although with allomorph *-i* for present tense as against *-ru* on verbs. Adjectives differ from verbs in not taking imperative and hortative suffixes, and in not combining with auxiliaries to mark aspect, benefaction, etc. Like verbs, they may modify nouns.
- The class of uninflected adjectives is like nouns in not taking any inflections, and in functioning as copula complement. These adjectives cannot function as intransitive predicate (without a verbalizing suffix being added), and they may only modify a noun if the marker *na* or *no* is also included.

The properties just listed indicate the differences between the two adjective classes. They do, however, share important syntactic properties: for example,

members of both classes may be modified by an intensifier, and they may also function as adverbs. Some of their major grammatical properties can be tabulated:

(23)	inflected adjectives (verb-like)	uninflected adjectives (noun-like)
can be intransitive predicate	✓	only with derivational suffix
can be complement of copula <i>da</i>	—	✓
can modify noun	✓	needs <i>na</i> or <i>no</i>
can be modified by intensifiers	✓	✓
can function as adverb	✓	✓
may accept verbalizing suffix - <i>sugiru</i> 'too'	✓	✓

Looking now at semantic types, AGE, COLOUR, and SPEED terms are all inflected adjectives. Most DIMENSION and PHYSICAL PROPERTY terms are also inflected adjectives, although some belong to the uninflected class, while VALUE terms are divided between the two classes. HUMAN PROPENSITY terms are predominantly in the uninflected class, although inflected items are not uncommon.

In terms of size and composition, the inflected adjective class has about 700 members (some lexically complex); all the lexically simple members are native roots. The uninflected adjective class has more than three times as many members, some native forms and some loans from Chinese and from European languages; new forms can be added to this class (but not to the inflected class).

3. Manange (Tibeto-Burman, Nepal) also has two adjective classes. What Genetti and Hildebrandt (2004)—in their detailed and instructive account—call 'verb-like adjectives' (a class with about fifty-seven monomorphemic members) may function as intransitive predicate but lack most of the morphological processes available to verbs. Their 'adjectives' (a class with about thirty members, which could be referred to as 'noun-like adjectives') may occur as copula complement, like a noun, but cannot be NP head and have distinct phonotactics. The two adjective classes share properties; for example, both may directly modify a preceding noun in an NP.

The COLOUR, SPEED, and QUANTIFICATION semantic types consist only of (noun-like) adjectives, HUMAN PROPENSITY and DIFFICULTY involve only verb-like adjectives, while DIMENSION, AGE, VALUE, PHYSICAL PROPERTY, and POSITION include members from both classes.

12.8 Correlations with other grammatical parameters

It is interesting to enquire whether there is any correlation between the type of adjective class found in a language and other grammatical parameters. I have observed one, relating to ways in which the functions of core arguments are shown within a clause. We can distinguish:

- Predicate-marking, where bound pronouns within the predicate provide information about the syntactic functions of core arguments.
- NP-marking, where there is case marking (or something similar) on NPs which are the (whole or partial) realizations of core arguments identifying syntactic functions.

(Following Nichols 1986, the terms ‘head-marking’ and ‘dependent-marking’ have recently come to be used for what I called ‘Predicate-marking’ and ‘NP-marking’. As mentioned in §5.6, the term ‘head’ is typically—and appropriately—used with respect to a phrase. It is not really advisable to extend the meaning of this term so that a predicate is called the ‘head’ of a clause. For this reason the terms ‘Predicate-marking’ and ‘NP-marking’ are preferred to the labels introduced by Nichols, when referring to a clause.)

Surveying the languages of the world, there is a striking quantitative correlation:

- (24) • Adjective classes of type (a)—with grammatical properties similar to those of verbs—tend to be found in languages with Predicate-marking and in languages with neither Predicate- nor NP-marking
- Adjective classes of type (b)—with grammatical properties similar to those of nouns—tend to be found in languages with NP-marking

Tentative examples of the correlation include:

- *Type (a) and basically Predicate-marking*
 - Many of the languages of North America (including most languages in Na-Dene, Algonquian-Ritwan, Salish, Siouan, Iroquoian, Muskogean, Tsimshian, Zuni)
 - Some languages from South America (including the Arawak family)
 - Most Austronesian languages (excluding those in the Philippines)
 - Ainu
- *Type (a) and neither predicate-marking nor NP-marking*
 - Most languages from South-East and East Asia (including Sinitic, Tibeto-Burman, Tai-Kadai, and some Austro-Asiatic)

- *Type (b) and basically NP-marking:*
 - Most of the languages of Europe, North Africa, North and West Asia, and North India (Indo-European, Basque, Uralic, Turkic, North-East Caucasian, Afro-asiatic, Burushaski)
 - Most of the languages of Australia
 - Most of the languages of the Philippines
 - Some languages from North America (including Yokuts, Sahaptin, Sierra Miwok, Tarascan)
 - Some languages from South America (including Quechua)

It should be emphasized that this is very much a first run-through of the data. Detailed study of the adjective classes in individual languages is required. There may, indeed, be languages of more than one type within a single genetic or areal grouping. Surveying Nilo-Saharan languages, Dimmendaal (2000: 218–19) notes that—in accordance with (24)—‘adjectives tend to pattern with nouns in dependent-marking [i.e. NP-marking] languages and with verbs in head-marking [i.e. Predicate-marking] languages’.

There are a number of exceptions to the generalization in (24), some of these being of particular interest. In each instance, an historical explanation can be provided, in terms of the generalizations.

1. It is clear that, at an earlier stage, Australian languages were entirely NP-marking; in keeping with this, adjective classes are almost all of type (b), with grammatical properties similar to those of the noun class. (In fact, fairly subtle criteria have to be applied, in most languages, to distinguish between adjectives and nouns; see the discussion of Alpher’s criteria in §12.2.)

In recent times, bound pronouns have evolved over a good deal of the continent. In most of the languages in which they occur, these are clitics attached to the verb or a verbal auxiliary, and they are not always obligatory. However, languages over a continuous area in the central north have developed obligatory pronominal prefixes to verbs, a clear Predicate-marking strategy. As a consequence they have lost or are in the process of losing NP-marking from NPs (see Dixon 2002a for full details).

Interestingly, a couple of these Predicate-marking languages appear to be assigning more verb-like properties to their adjective class. It was mentioned, in §12.5.3, that in Emmi adjectives are negated like verbs, differently from nouns. And that in Nunggubuyu an adjective may take subject pronominal prefixes, like an intransitive verb, showing that it is functioning as head of an intransitive predicate.

The shift of a language from an NP-marking to a Predicate-marking profile is well attested. Bound pronouns develop from what were free forms, and are obligatorily included in each predicate, with the old NP-marking dropping

out of use. It may be that the shift from (b) a ‘noun-like’ to (a) a ‘verb-like’ adjective class—in order to re-establish the correlation in (24)—tends to follow the shift from NP-marking to Predicate-marking, but operating at a slower pace.

2. Japanese is an NP-marking language; as discussed in §12.7, there are two adjective classes, one verb-like and one noun-like. This suggests a rather speculative historical scenario:

STAGE ONE. Japanese lacked NP-marking. It probably also lacked Predicate-marking, showing syntactic function by the order of phrasal constituents within a clause. There was a single class of adjectives (the present inflected class), similar to verbs in their grammatical behaviour. Japanese thus conformed with the correlation in (24).

STAGE TWO. The language developed NP-marking. In association with this, it developed a second class of adjectives (the present uninflected class), which are ‘noun-like’.

The following points can be adduced to support this scenario:

- The marking of the function of NPs in a clause is by syntactic particles, rather than by case suffixes. This is a little surprising, since Japanese is a fairly synthetic language with verbs taking a variety of suffixes. It is consistent with NP-marking having been introduced rather recently. Indeed, Shibatani (1990: 333–57) states that although the topic-marking particle *wa* is present in the earliest records (from about the eighth century CE), the particles *ga*, marking subject, and *o*, marking object, developed fairly recently from other grammatical elements (the object marker evolving before the subject marker).
- The verb-like inflected adjective class appears to be archaic, being restricted to native lexemes; although large, it does not accept loans. It includes all AGE, COLOUR, and SPEED items, and most from the DIMENSION and PHYSICAL PROPERTY types (the types that are typically associated with an adjective class).
- The noun-like uninflected class is now bigger than the inflected class and is growing; it accepts all kinds of loans. This class includes some DIMENSION and PHYSICAL PROPERTY terms, and most of the HUMAN PROPENSITY adjectives.

This scenario is speculative but not implausible. It suggests that, as with the Australian languages Emmi and Nunggubuyu, once a language shifts its profile with respect to Predicate-marking and NP-marking, then there will be a tendency to reorient the grammatical properties of adjectives in

accordance with the correlation in (24). In Nunggubuyu, the adjective class has had its grammatical possibilities extended so that it may now accept subject pronominal prefixes, like an intransitive verb (it does not yet directly accept tense and other verbal suffixes; this would be the next step). In Japanese, a new adjective class has been established, which is steadily increasing in size.

Korean is a clear exception to the correlation in (24), being NP-marking and having just one adjective class, which is ‘verb-like’. The speculative scenario just suggested for Japanese could be extended to Korean—supposing that the language originally lacked both Predicate-marking and NP-marking, and had a class of ‘verb-like’ adjectives, but then developed NP-marking. Syntactic function is, as in Japanese, shown by particles following an NP, and some of these are thought to have developed recently. (For example, Sohn 1999: 30 mentions that the subject particle *ka*—which is now a conditioned allomorph of the earlier subject marker *i*—first appeared in the literature in 1572 CE and may possibly have been a borrowing from the Japanese subject particle *ga*.)

It may be that extending the Japanese scenario to also apply to Korean is transcending speculation in the direction of fantasy. Like other results in linguistic typology, (24) is a statistical correlation, not a hard-and-fast rule. There are exceptions to it. Besides Korean, these include Southern Paiute (Uto-Aztecan; Sapir 1930–1), which is also NP-marking and shows ‘verb-like’ adjectives. It is, of course, worthwhile according detailed examination to these and other exceptions, to see whether there is an explanation (along historical or other lines). But it is unlikely that every exception will be provided with an explanation, and it is an error to try forcibly to provide one. Some languages just do have a typologically unusual combination of properties, in some area of the grammar.

12.9 Semantic overlapping between word classes

It was mentioned in §1.8 that while each word class has a similar semantic core across languages, there are a number of non-core concepts whose word class membership varies between languages; for example ‘hunger/hungry/be hungry’ can be a noun, an adjective, or a verb.

We have seen that some languages have a small closed adjective class. However, most languages exhibit a large, open class. For such languages, it is interesting to compare the semantic membership of the three open classes—noun, adjective, and verb. This will now be attempted for Dyirbal, contrasting it with English.

The adjective class in Dyirbal is large and open, and it is like the noun class in its grammatical properties (very similar to European languages). The semantic contents of the adjective classes in Dyirbal and English are similar; that is, most adjectives in Dyirbal correspond to adjectives in English, and vice versa.

However, there are a number of small semantic fields for which Dyirbal has adjectives while English has verbs. Some of these are exemplified in (25).

(25)	VERB IN ENGLISH	ADJECTIVE IN DYIRBAL
(a)	divide split crack smash tear fold	ñarri 'divided up' yagi 'split' gajala 'cracked' muñi 'smashed up' gini 'torn' wujun 'folded'
(b)	gather heap muster	balmбу 'gathered together' gurruñ 'heaped up' guwurr 'mustered'
(c)	lean	yulgarra 'leaning'
(d)	marry	julbun 'married'

It will be seen that the terms in (a) deal with related concepts, to do with changing the form of an object; those in (b) have to do with getting together several things. Note that, just as adjectives can be derived from verbs in English (*cracked*, *torn*, *leaning*, etc.), so can verbs be derived from adjectives in Dyirbal—adding *-bi-l* to form an intransitive and *-ma-l* to form a transitive stem (for example, *yagi-bi-l* 'be split', *balmбу-ma-l* 'gather together').

There are thus some differences in the ways in which languages divide up semantic space into word classes. Corresponding to adjectives *happy*, *clever*, and *jealous* in English, Fijian has S = A type ambitransitive verbs *maarau* (*-ta'ina*) 'be happy (about)', *vu'u(-ta'ina)* 'be clever (at)', and *vuuvuu(-ta'ina)* 'be jealous (of)'. Corresponding to verbs *divide* and *marry* in English, Dyirbal has adjectives *ñarri* 'divided up' and *julbun* 'married'.

It is now relevant to ask whether there is any semantic overlap between word classes in individual languages, and—if so—whether different languages show similar or different overlaps. That is, we can enquire whether a given concept may be coded (1) by both verb and noun; (2) by both adjective and noun; or (3) by both verb and adjective.

Looking at English, there are many instances of (1) and (2), but few of (3). Some examples of verb/noun and of adjective/noun overlap are given in (26).

(26)	VERB	NOUN	ADJECTIVE	NOUN
(a)	hit	blow	big, small	size
	go	journey	new, young, old	age
	happen	event	fast, slow	speed
(b)	arrive	arrival	long, short	length
	think	thought	accurate	accuracy
	announce	announcement	happy	happiness

The pairs in (a) are non-cognate between word classes. Those in (b) are some of the many examples of nouns derived from verbs and from adjectives (a different form of derivation appears in each example).

There are very few instances, in English, of verb and adjective with similar meanings (where the forms are not related through a productive derivation). One example consists of *fear* and *afraid*, as in:

- (27) (a) I fear having to enter the lion's cage
 (b) I was afraid when I entered the lion's cage.

The adjective *afraid* is generally used to refer to the feeling one gets when one is actually in contact with something scary, while the verb *fear* tends to be used for a general feeling in connection with something that might happen.

In fact, this particular verb/adjective overlap recurs in a number of languages. For instance, the Australian language Yidiñ has a transitive verb *yarrga-n* and an adjective *munu* with very similar meanings to *fear* and *afraid* in English (Dixon 1991b: 240, 272). However, not all languages have such an overlap. Describing the Oceanic language Mokilese, Harrison (1976: 150) mentions that there is just one lexeme, *mijik*, corresponding to both 'fear' and 'afraid' in English. (Following the Oceanic tradition—see §12.2—Harrison says that this belongs to the class of stative verbs; however, applying the criteria set forth in the present study, the label 'adjective' is appropriate.)

English thus has considerable semantic overlap between the verb and noun classes, and between the adjective and noun classes, but very little between verb and adjective classes. Turning now to Dyirbal, we find exactly the opposite situation. Here there is no overlap at all between verb and noun classes, or between adjective and noun classes. Basically, Dyirbal does not have abstract nouns such as 'journey', 'event', 'thought', 'size', 'happiness', or 'colour'. One simply has to use the appropriate verb or adjective.

However, Dyirbal does have considerable semantic overlap between the classes of verb and adjective. A sample of these is set out in (28).

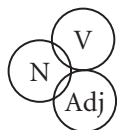
(28)	VERB	ADJECTIVE
(i)	<i>ñaju-l</i> 'cook'	<i>ñamu</i> 'cooked'
(ii)	<i>dadi-l</i> 'cover'	<i>ŋulguñ</i> 'covered'
(iii)	<i>gulba-l</i> 'block'	<i>gumun</i> 'blocked'
(iv)	<i>wanda-l</i> 'hang'	<i>burrgaligan</i> 'hanging'
(v)	<i>baŋganda-y</i> 'be sick'	<i>wulmba</i> 'sick'

There is in fact a clear difference of meaning in each case, with the verb referring to an action, or getting into a state, or being in a state that varies with time, and the non-cognate adjective referring to either a state that is the result of an activity, or a state that is semi-permanent. There is a slightly different semantic contrast for each verb/adjective pair. Taking them one at a time:

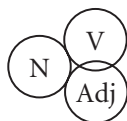
- (i) The transitive verb *ñaju-l* refers to the act of cooking; its participle *ñajuŋu* can describe something being cooked a bit or a lot, not enough or too much. In contrast, the non-cognate adjective *ñamu* means 'cooked to perfection, ready to eat'.
- (ii) The transitive verb *dadi-l* refers to any sort of act of covering; its participle *dadiŋu* can describe a blanket over just half a sleeping person. In contrast, the adjective *ŋulguñ* means 'properly covered, covered all over'.
- (iii) The transitive verb *gulba-l* can refer to any kind of blocking; its participle *gulbaŋu* can be used to describe a temporary obstruction across a path. In contrast, the non-cognate adjective *gumun* refers to something permanently blocked; for example, a road that has been closed off for good, or a road that simply stops at a certain place, never having been constructed any further.
- (iv) The transitive verb *wanda-l* is used to describe hanging something up; the participle *wandaŋu* can be used of a basket or bucket which has been hung from a hook. The adjective *burrgaligan* refers to something hanging down; for example, long hair on a person, or bark hanging off a tree (it can also be used to describe something that has been hung up).
- (v) The intransitive verb *baŋganda-y* is used to describe feeling sick or ill (or just weary); the participle *baŋgandaŋu* refers to someone who is under the weather at present, but is expected to get better. In contrast, the adjective *wulmba* refers to someone who is truly sick and is expected to die. (Death is believed to be caused by sorcery, so that using *wulmba* of a person is saying that a sorcerer has done something to them which will result in their death.)

The kind of overlap between these three major word classes in English and in Dyirbal can be shown diagrammatically:

Type (a) English



Type (b) Dyirbal



It appears that most of the languages of Europe are basically of type (a), like English. It is interesting to enquire what other languages are of type (b), like Dyirbal. Data is hard to come by, since very few linguists provide a detailed description of the semantic characteristics of word classes, let alone investigate the possibility of semantic overlap between classes. There are just a few hints available. For Zuni (isolate, New Mexico), Newman (1968: 66) provides the following examples of overlapping between the verb class and what should probably be recognized as the adjective class:

- (29) hemmo 'to boil' woleya 'to be boiled'
 ʔahha 'to buy' ʔayyo 'to be bought'
 šalu 'to stretch' pilla 'to be stretched'

Sohn (2004) offers some illuminating remarks on the semantic overlap between word classes in Korean. He shows that there is overlap between verb and adjective classes involving just native lexemes, whereas noun/verb and noun/adjective overlaps often involve one native and one loan item (the latter from the Sino-Korean stratum of vocabulary).

Overall, one would expect semantic overlap between word classes to be found most commonly in languages which maintain a strict correspondence between word class and functional slot. Dyirbal is of this type—a noun can only function as head of an NP (in predicate argument function) and a verb only as head of a predicate. A language with more fluid class-slot correspondences (such as Nootka, briefly discussed in Chapter 11 above) might be less likely to feature semantic overlap between its word classes.

12.10 Summary

The label 'adjective class' is here used for a word class that:

- is grammatically distinct from noun class and verb class;
- (A) functions either as intransitive predicate or as copula complement and/or (B) modifies a noun in an NP;

- includes words from some or all of the prototypical adjective semantic types—DIMENSION, AGE, VALUE, and COLOUR.

In some languages two separate adjective classes can be recognized. Note that although the prototypical adjective class combines functions (A) and (B), in some languages the class has only one of these functions.

There are two broad parameters of variation for an adjective class—size and grammatical properties. Some languages have a small closed class (with anything from a handful to a few score members) whereas others have a large open class (to which new derivations and/or new loans may be added). And adjectives may be similar in their grammatical properties to nouns, or to verbs, or to both, or to neither.

It can sometimes be a tricky matter finding criteria to distinguish ‘verb-like’ adjectives from verbs, or ‘noun-like’ adjectives from nouns. I believe that for every language which is studied in detail, such criteria can be found. Criteria are not always of the definitive ‘yes-or-no’ variety; as Alpher (1991) shows for the Australian language Yir-Yoront (see §12.2), a collection of statistical tendencies can combine to provide a satisfactory grammatical characterization of the adjective class (as opposed to noun and verb classes).

In §12.8, a tentative correlation was established—adjective classes which have grammatical properties similar to nouns tend to be found in languages with NP-marking, while adjective classes with grammatical properties similar to those of verbs tend to be found in languages with Predicate-marking or with neither NP-marking nor Predicate-marking. I suggested, with some supporting exemplification, that if a language shifts its Predicate-marking/NP-marking profile, then the orientation of its adjective class is likely slowly to change, to re-establish the correlation.

§12.9 looked briefly at the kinds of semantic overlap between the three major word classes. We saw that English has considerable verb/noun and adjective/noun but rather little verb/adjective overlap, while Dyirbal is almost exactly the reverse, with considerable verb/adjective but no verb/noun or adjective/noun semantic overlap.

12.11 What to investigate

Word classes should be distinguished on language-internal grammatical criteria. The label ‘adjective’ is appropriate for that class which (A) functions either as intransitive predicate or as copula complement and/or (B) modifies a noun in an NP. It is extremely likely that this class will include lexemes from the semantic types DIMENSION, AGE, VALUE, and COLOUR (and, if the class has more than a few dozen members, also from PHYSICAL PROPERTY, HUMAN PROPENSITY, and SPEED).

The linguist should investigate the size of the adjective class—whether it is small and closed, or large and open to the addition of new members. The morphological and syntactic properties of the adjective class should be studied—whether these are similar to the grammatical properties of nouns, or to those of verbs, or to both, or to neither.

There may be two adjective classes; typically one is grammatically similar to verbs and the other to nouns, but they are linked together through a number of shared properties. An appropriate topic for study, at an advanced stage of analysis, concerns the semantic overlap between noun, adjective, and verb classes (see the discussion in §12.9).

Some of the recurrent criteria for distinguishing adjectives from verbs and from nouns were discussed and exemplified in §§12.5.1–2. A fuller list of features to examine can now be provided. (These can, of course, only be investigated after the full set of morphological and syntactic properties of nouns and of verbs has been established.)

(i) **Useful criteria when adjectives are grammatically similar to verbs**

1. Can an adjective function (like an intransitive verb can) as head of an intransitive predicate? And can a noun or pronoun (or any other kind of word) or NP also be head?
 - (a) Do the same morphological processes apply to an adjective as to an intransitive verb in this slot? (TAM marking, pronominal marking, etc.) Or just some of them?
 - (b) If the same processes apply, are they realized in the same way on verbs and on adjectives?
 - (c) If a verb may take an auxiliary, may an adjective do so too?
2. Are there any affixes or particles which apply to both verbs and adjectives but have a different meaning with the two word classes?
3. Can an adjective be used in imperative mood, in the same way that a verb can?
4. Do adjectives behave like verbs with respect to derivational processes (for example, causative)?
5. What are the transitivity values of predicates that accept adjectives and verbs as heads? (Generally, adjectives only occur in intransitive predicates. There may be several subclasses of verbs, each having distinct transitivity potential.)
6. Does reduplication apply to adjectives but not to verbs (or vice versa)? If it applies to both, does it have the same or different form and/or meaning with the two word classes?
7. If there are nominalization process(es), do they apply in the same way to verbs and to adjectives?

8. Can an adjective modify a noun (and also pronoun?) in an NP?

(a) Directly? (preceding or following the head or either?)

(b) Indirectly, with relative clause or other marker?

If a verb may also be modifier, how do its possibilities and marking differ from those of an adjective modifier?

9. Can an adjective be head of an NP? Can a verb also function as head of an NP?

If so, does this apply for a simple NP, or an NP which is a nominalized clause?

10. Does an adjective have the same possibilities as a verb for functioning as head of the predicate in a subordinate clause?

(ii) Useful criteria when adjectives are grammatically similar to nouns

11. When an adjective is modifier to a head noun

(a) Does it agree with the head in gender/number/definiteness/etc.?

(b) If it shows a category in common with the head noun in the NP, does this category have the same realization on noun and on adjective?

12. Is gender/noun class (or classifier) choice a criterion for distinguishing adjectives from nouns? (A noun may have just one inherent gender/noun class, whereas an adjective may take any gender/noun class, agreeing with the noun it is modifying.)

13. Do adjectives differ from nouns in number marking? In system or in realization? For example, in Berber languages, nouns have irregular number marking but adjectives are all regular.

14. Is there an affix or particle that applies to both nouns and adjectives but has a different meaning with the two word classes?

15. Does reduplication apply to adjectives but not to nouns (or vice versa)? If it applies to both, does it have the same or different form and/or meaning with the two word classes?

16. Can an adjective make up an NP all by itself? If so, is this best described as an elliptical NP, with head omitted, consisting just of a modifier, or as an NP for which the adjective is head?

17. If an adjective is NP head, can it be modified in the same way as can a noun when it is NP head?

18. If the language has case marking, investigate the rules for its realization on an NP; for example, on last/first/head word, on every word, on words of certain types. Does case go onto adjectives according to this general rule? (For example, if case goes on the last word of the NP, it attaches to an adjective just when it is the last word.) Or does a special case rule apply for adjectives?

- (a) If both nouns and adjectives show case, does the same case system apply for the two word classes?
- 19. Do adjectives function in the same way as nouns for possession? For example, do they take bound possessive pronouns or other markers of possession?
 - (a) Can an adjective be (i) possessor, (ii) possessed, within a possessive construction?
- 20. If a noun can also be a modifier to an NP head:
 - (a) Which nouns may be modifier?
 - (b) Do noun modifier and adjective modifier behave in the same way? For example, if an adjective modifier may be further modified by an adverb (or 'very'), may a noun modifier also be?
 - (c) If a noun may have several adjective modifiers, may it have several noun modifiers?
- 21. Can an adjective modifier be used anaphorically, for a full NP? And can a noun modifier be?
- 22. Are there alternative constructions for linking a head noun with an adjective modifier? Does each of them also apply for head noun and noun modifier?

For example, in Hausa, 'Adj₁ Genitive-Copula Noun₂' can also be expressed as 'Noun₂Adj₁', but 'Noun₁ Genitive-Copula Noun₂' *cannot* be expressed as 'Noun₂ Noun₁' (Parsons 1960).
- 23. If there are verbalization process(es) do they apply in the same way to nouns and to adjectives?
- 24. Can an adjective function as copula complement?
 - (a) In bare form?
 - (b) Or only as part of an NP? Or only when nominalized?

Compare with the possibilities for a noun in a copula complement.

(iii) Useful general criteria

- 25. Can an adjective have manner adverbial function, modifying a verb?
 - (a) In bare form? (And can noun and/or verb also have this function?)
 - (b) In derived form? (Does the derivation apply to any other word type?)

If only some adjectives have this property, which semantic types do they belong to?
- 26. Can an adjective function as 'parameter of comparison' in a comparative construction (if the language has such a construction type)? What else can function in this slot? Verb? Noun?
- 27. Can some or all adjectives be modified by an intensifier with a meaning like 'very' or a quantifier such as 'much'? If so, does this property also apply to nouns and/or verbs?

- 28. How is an adjective negated? In the same way as a verb? Or as a noun?
- 29. Does an adjective have any affixes/take any particles etc. which all other word classes lack, e.g. comparative, superlative; or augmentative, diminutive?
- 30. Do adjectives lack any properties that all other word classes have? For example, in Tamil and Telugu (from the Dravidian family) the adjective class is the only word class whose members do not accept any clitics.

Appendix Distinguishing Noun, Verb, and Adjective in Fijian

At the beginning of Chapter 11, Milner (1956: 10) was quoted as maintaining that there are no word classes (no noun, no verb, no adjective) in Fijian, just a set of lexical ‘bases’. In his high-quality doctoral dissertation, Arms (1974: 7–17) presented a set of criteria for distinguishing the three major word classes. Nevertheless, Schütz (1985: 95–7), while distinguishing noun and verb, makes no mention whatsoever of adjective (it appears that adjectives were silently incorporated into the verb class by Schütz).

Classes of noun, verb, and adjective can be recognized without difficulty, and indeed must be recognized for an adequate grammatical description. Table 12.2 summarizes the occurrences of word classes in structural slots. After a brief mention of word-class-changing derivations, there is then discussion of basic clause structure, predicate structure, and NP structure, comparing the properties of noun, verb, and adjective in each.

Functions of lexical word classes

- A noun frequently occurs in its primary function, as head of an NP, and occasionally in its secondary function, as head of an intransitive predicate. A small number of nouns (referring to material, such as ‘stone’) may be modifier within an NP (shown by ** in Table 12.2).
- An adjective has two primary functions, as head of an intransitive predicate and as modifier within an NP; it may also be used as head of an NP.
- Most verbs may be used as head of a transitive predicate (and then take a transitive suffix) or of an intransitive predicate (without a suffix). The * in Table 12.2 indicates that about one-fifth of verbs are only used transitively, and a very small

TABLE 12.2. How often word classes occur in functional slots in Fijian

	VERB	ADJECTIVE	NOUN
HEAD OF INTRANSITIVE PREDICATE	frequently*	frequently	occasionally
HEAD OF TRANSITIVE PREDICATE	frequently*	—	—
HEAD OF NP	occasionally	occasionally	frequently
MODIFIER IN NP	—	frequently	—**

number are only used intransitively. Verbs also have secondary function as head of an NP.

Word-class-changing derivations

The following derivations are word class specific.

- prefix *dau-* added to a verb, derives an adjective ‘habitually, often’; for example:
 - verb *pu'u* ‘be angry’; adjective *dau-pu'u* ‘habitually/often angry’
 - verb *buta'o* ‘steal’, adjective *dau-buta'o* ‘habitually stealing’
 - verb *qito* ‘play games’, adjective *dau-qito* ‘habitually playing games’
- prefix *i-* added to a verb, derives a noun with meaning ‘instrument’, ‘place of activity’, ‘result of activity’, ‘activity’—see examples under (ii) in §11.8.

The class of adverbs consists entirely of forms derived from adjectives through prefix *va'a-*. For example: adjective *levu* ‘big, great’; adverb *va'a-levu* ‘to a great extent’; *kaukaua* ‘strong, hard, powerful’; *va'a-kaukaua* ‘do strongly’.

Basic clause structure

The only obligatory constituent of a clause is the predicate, which includes a subject pronoun and, if transitive, an object pronoun. For example:

- (1) *au la'o* ‘I am going’
- (2) *e la'o* ‘he/she is going’
- (3) *e pu'u* ‘he/she is angry’

The subject argument may be expanded by an NP which follows the verb, as in

- (4) [*e pu'u*]_{INTRANSITIVE.PREDICATE} [*a gone*]_S ‘the child is angry’

Here the NP which expands on the 3sg pronoun *e* within the predicate has common noun *gone* ‘child’ as head, preceded by the ‘common article’ *a*.

In a transitive clause, the verb bears a transitive suffix (see §12.5.1):

- (5) *e pu'u-ca* ‘he/she is angry with him/her’

The transitive suffix here has underlying form *-ci*, and becomes *-ca* to indicate a 3sg O argument. Subject and/or object can be expanded by a post-predicate NP:

- (6) [*e pu'u-ca*]_{TRANSITIVE. PREDICATE} [*a gone*] [*a tuuraga*]

Here the two NPs have as their heads *gone* ‘child’ and *tuuraga* ‘chief’. The order of NPs after the predicate is variable, so that (6) could mean either ‘the chief was angry with the child’ or ‘the child was angry with the chief’. The first interpretation is most likely, on pragmatic grounds.

Predicate structure

A predicate consists of:

- (a) subject marker (obligatory), such as 1sg *au*, 3sg *e*
- (b) tense and/or aspect marker (optional)
- (c) pre-head modifiers (optional), such as *via* ‘want to’, *rui* ‘to a high degree’
- (d) HEAD (obligatory), plus suffix if transitive
- (e) object pronoun (if transitive)
- (f) post-head modifiers and/or adverbs (optional)

Pre-head modifier *rui* ‘to a high degree’ may only be used if the head is an adjective, or if an adverb (derived from an adjective by prefix *va'a-*) is in slot (f). One can say:

- (7) *e rui levu* ‘he/she/it is very big’

but not:

- *(8) *e rui pu'u*

Since *pu'u* ‘be angry’ is a verb, it cannot in itself be modified by *rui*. However, if it is modified by the adverb *va'a-levu* ‘to a great extent’ (derived from adjective *levu* ‘big, great’), then *rui* is allowable:

- (9) *e rui pu'u va'a-levu* ‘he/she is very angry’

Prefix *dau-* derives an adjective from a verb, and *dau-pu'u* may be modified by *rui*:

- (10) *e rui dau-pu'u* ‘he/she is habitually very angry’

In addition, only an adjective (not a verb or noun as predicate head) may follow *sega soti* ‘not very’, and only an adjective can be used in a comparative construction.

The primary function of a verb is to be head of a predicate. One of the primary functions of an adjective is to be head of an intransitive predicate. A secondary function of a noun is as head of an intransitive predicate. In fact, a complete (and quite complex) NP can function as head of an intransitive predicate—‘tall man’ or ‘that woman’ or ‘the last instance of your eating people’ or ‘you’; see §11.2.

Noun phrase structure

The structure of an NP with a common noun as head is, basically:

- (a) number modifier (optional)
- (b) article (obligatory)—*a* if head is a common noun, *o* if it is a proper name or a pronoun
- (c) possessor element (optional)
- (d) HEAD (obligatory)
- (e) adjective as modifier and/or ‘material’ noun as modifier (optional)
- (f) one NP, expanding bound pronoun in possessor element (optional)

The possessor element has two components: a classifier followed by a pronominal suffix. For example:

- (11) a o-na da'ai 'his/her gun' (which he/she owns)
 (12) a 'e-na da'ai 'his/her gun' (which will be used to kill him/her)

The 3sg possessive pronoun *-na* 'his/her' is attached to a classifier element, here *o-* 'owned' or *'e-* 'relating to, but not owned' (the other possibilities are 'edible' and 'drinkable').

The 3sg pronominal suffix *-na* in (11) and (12) can be expanded by a further NP after the head of this NP, as in

- (13) a o-na da'ai [a tuuraga] 'the chief's gun' (which he owns)

This is, literally, 'his/her gun, the chief'.

It appears that a verb or adjective can be head of an NP. The differences from an NP with a common noun as head are:

- there can be no number modifier, in slot (a)
- the possessor element, in slot (c), is obligatory
- there may be more than one NP in slot (f).

In addition, the classifier options within the possessor element are different. Whereas a common noun as NP head has four classifier possibilities, a verb just has one, *o-*. For an adjective there are two, exemplified in:

- (14) a 'e-na kaukaua 'his her/its (inherent) strength'
 (15) a o-na kaukaua 'his/her (acquired) power'

In fact, what appears to be an NP with verb or adjective as head is better regarded as a type of complement clause (termed a 'clausal NP' in Dixon 1988a). From the simple clause

- (4) [e pu'u]_{INTRANSITIVE.PREDICATE} [a gone]_S 'the child is angry'

can be derived a 'clausal NP', as in

- (16) [au rai-ca] [a o-na pu'u [a gone]_S]_O 'I saw the child being angry'

The clause in (4) is made into a complement clause (in O function to transitive verb *rai-ca* 'see') in (16) by means of:

- placing article *a* at the beginning
- replacing the subject pronoun by the corresponding possessor element, here 3sg *e* by *o-na*
- retaining everything else in the clause

There may be one post-predicate NP, expanding the underlying subject, as in (16). Or, for a clausal NP based on a transitive clause—such as (6)—there could be two

NPs, one expanding the underlying subject and the other the object. (And there could be peripheral NPs as well—marked by prepositions—in the underlying clause, carried over into a clausal NP.)

A noun derived from a verb by prefix *i-* (such as *i-vola* ‘book’) behaves exactly like an underived noun—it may be modified by a number, the possessor element is optional, etc. In contrast, the underlying verb (*vola* ‘write’) can only be head of a clausal NP—no number modifier, obligatory possessor element, etc.

Reduplication

- Reduplication of the first two syllables of a root applies freely to verbs with meaning ‘do several times, do over a long period’; for instance *buta.buta* ‘steal on a number of occasions’.
- Colour adjectives have a reduplicated form when modifying a noun that does not refer to an animal or ‘stone’ or ‘earth’. For example *a pusi loa* ‘black cat’ but *a nu’u loa.loa* ‘black sand’.
- There is no productive reduplication for nouns (or for other adjectives).

Summary

The contrastive properties of noun, verb, and adjective in Fijian provide ample justification for recognizing them as three word classes. These properties comprise: their major functional possibilities, in Table 12.2; derivational processes, set out above, and the properties just enunciated, which are summarized in Table 12.3.

It will be seen that a grammar which does not recognize the three open word classes—such as Milner (1956) and Schütz (1985)—is essentially superficial, and cannot adequately describe the intricacies of Fijian grammar.

TABLE 12.3. Contrasting possibilities for the three word classes in Fijian

	VERB	ADJECTIVE	NOUN
WHEN HEAD OF INTRANSITIVE PREDICATE			
can take modifier <i>rui</i> ‘to a high degree’	no	yes	no
can follow <i>sega soti</i> ‘not very’	no	yes	no
can enter into comparative construction	no	yes	no
WHEN HEAD OF NP			
can be modified by number	no	no	yes
modification by possessor element	obligatory	obligatory	optional
there is a choice of classifier available	no	yes	yes
head may be followed by how many NPs	two (or more)	one	one
PRODUCTIVE REDUPLICATION	yes	(only colour adjectives)	no

Sources and notes

This chapter is based on Dixon (1977c) and Dixon (2004b). It repeats some discussion and examples from them, but also includes a fair amount of new material. For reasons of space, not all of the earlier materials could be included here. The interested reader may wish to consult Dixon (1977c) for discussion of the semantic content and grammatical properties of the adjective class in English, for further details of languages with small adjective classes, and for discussion of ‘adjectival’ and ‘verbal’ languages.

It should be noted that some instances of what I call an ‘adjective class’ are not accorded this label in the grammars from which I take the data. Nevertheless, they should each be labelled ‘adjective class’ according to the criteria used in this study—a word class distinct from noun and verb, including words from the prototypical adjective semantic types, and functioning either as intransitive predicate or as copula complement; and/or modifying a noun in an NP.

12.1. Enfield (2004) discusses adjectives in Lao, which is either a dialect of the same language as Thai or a very closely related language. He shows that adjectives share basic properties with verbs, but also exhibit crucial differences: only adjectives may feature in comparative constructions, may take intensifiers *khanaat5* ‘extent’ and *teep5* ‘rather’, and may undergo a type of reduplication with the meaning ‘is’. Both verbs and adjectives may be preceded by modifier *jaak5*; this indicates ‘want’ with verbs but may signify ‘somewhat’ with adjectives. Enfield has ‘adjective’ and ‘state verb’ as subdivisions of ‘stative verb’ which itself is a primary subclass of ‘verb’; but note that he mentions rather more points of difference between adjectives and stative verbs—and thus between adjectives and all other verbs—than points of similarity.

12.2. The authoritative text on evidentiality is Aikhenvald (2004b). One of a number of grammars of Australian languages to deny that an adjective class can be recognized is Eades (1979) on Gumbaynggir. An example of a grammar of an Austronesian language which uses the label ‘stative verbs’ for adjectives is Hyslop (2001) on North-East Ambae.

12.3. In some languages a noun may be modified by more than one adjective. There is generally a preferred order in which the semantic types will occur. In English, where adjectives precede the noun, the unmarked order is VALUE, DIMENSION, PHYSICAL PROPERTY, SPEED, HUMAN PROPENSITY, AGE, COLOUR (Dixon 1982: 24–5). In languages where adjectives follow the noun, the ordering is roughly the reverse of this. That is, a term referring to COLOUR, a fairly fixed property, tends to occur nearest to the head noun, and one referring to VALUE, which is a subjective judgement, tends to appear furthest out. A full

cross-linguistic study of adjective ordering lies outside the scope of the present chapter. (There is discussion of adjective order for Qiang in LaPolla and Huang 2004.)

There may, of course, be further syntactic patterns available to adjectives in individual languages. A comprehensive study of the syntactic possibilities open to adjectives in English will be found in Ferris (1993).

12.4. It is not uncommon to find that basic colour terms ('white', 'red', etc.) have different grammatical properties from non-basic colour terms ('silver', 'indigo', etc.).

Sources for information in Table 12.1 are as follows. Kamula—Routenaa (n.d.); Igbo—Welmers and Welmers (1968, 1969), Welmers (1973); Hausa—Abraham (1959), Bargery (1934), and Migeod (1914) (also see Dixon 1982: 4, 39); Jarawara—Dixon (2004a, 2004c); Sare—Sumbuk (1999); Somali—Saeed (1999); Akan—Osam (1999); Northern Subanen—Daguman (2004).

Some languages allow a given adjective to either precede or follow the head noun, with a difference in meaning. For example, in French *un curieux homme* is 'a curious/strange man' while *un homme curieux* is 'a curious/inquisitive man' (in English the adjective *curious* is ambiguous between the 'strange' and 'inquisitive' senses). See Jespersen (1924: 168–9) on English, and Waugh (1977: 182–3) on French.

12.7. From Rose's (2003) account, it seems that Emerillon (Tupí family, French Guiana) also has two adjective classes. The verb-like class relates to DIMENSION, COLOUR, and VALUE lexemes while the noun-like class includes PHYSICAL PROPERTY and HUMAN PROPENSITY items.

12.8. Wetzer (1992, 1996) and Stassen (1997) put forward a rather different kind of correlation, suggesting that languages with 'nouny adjectivals' tend to show a tense system, while languages with 'verby adjectivals' tend to lack such a system (where tense is defined as, minimally, a distinction between past and non-past). There appears to be a degree of statistical support for this generalization, although there are a considerable number of exceptions. A much more fine-grained study is needed, with greater attention to the varying grammatical properties of adjective classes in individual languages.

12.9. Interestingly, the best example I can find of semantic overlap between the verb and adjective classes in English involves *afraid*, one of the set of adjectives that can only occur in copula complement function (not as a modifier).

Transitivity

Transitivity is a many-layered phenomenon:

- A. Each clause has a transitivity value, which determines the number of core arguments it requires. §13.1 discusses clausal transitivity, the functional roles which define arguments, and the ways in which they are identified in surface structure.
- B. There is generally some convention for marking core arguments, in order that a listener may recognize them. Canonical systems of marking are described in §13.2. (Non-canonical marking is discussed in §13.6.)
- C. Verbs divide into transitivity classes, depending on the transitivity types of clauses they may appear in. This is the subject of §13.3. The parameters of variation described under C can be combined with those in A and/or with those in B. Such possible combinations are outlined in §13.4.

Like everything else in grammar, transitivity has a semantic foundation. §13.5 explains how this works. It outlines the semantic basis for recognizing core argument functions and how semantic roles of diverse nature (associated with different verb types) may be mapped onto the same syntactic function. Semantic considerations may fulfil a further role—in some languages there can be alternative markings for one or more core arguments. That is, besides the canonical marking convention there may be non-canonical marking, depending on the nature of the referent of an (A, S, or O) core argument, or on the nature of the reference of the verb which functions as predicate head. Such non-canonical marking is the topic of §13.6.

The chapter closes with two appendices. One explains that labels ‘ergative’ and ‘accusative’ have a conceptual value in describing an association of S and O, or of S and A, in argument-marking conventions. It is not appropriate to extend these labels to other kinds of recurrent association—of a quite different kind—between S and O, and between S and A. The other appendix surveys the number of disparate ways in which terms ‘unergative’ and ‘unaccusative’ have been utilized, leading to such confusion that the terms are best avoided.

One point to be stressed—and always kept in mind—is that *transitivity is a syntactic matter*. When a clause is said to have a certain transitivity value, and when a verb is said to show certain transitivity possibilities, these are syntactic—not semantic—specifications. As described in this chapter, there are semantic parameters which underlie several aspects of transitivity. But it makes little sense to say, for example, that a given verb is ‘semantically transitive’ or ‘semantically intransitive’. It is more appropriate to describe it as having a semantic profile which is consistent with a certain transitivity profile at the syntactic level.

13.1 Clausal transitivity

The structure of a clause was outlined under (c) in §3.2. Leaving aside copula and verbless constructions—which are dealt with in Chapter 14—each clause involves a predicate and a number of core arguments, which must be either stated or inferable from the context of discourse.

The two major clause structures, across the languages of the world, are intransitive, with one core argument, and transitive, with two:

CLAUSE TYPE	PREDICATE	CORE ARGUMENTS
intransitive	intransitive	S (intransitive subject)
transitive	transitive	A (transitive subject) and O (transitive object)

There may also be peripheral arguments, which are optional, and can generally be included in either clause type. They cover things like instrument (for example, ‘with a stick’), beneficiary (‘for the child’), time (‘in the afternoon’), and place (‘under the tree’).

The only obligatory argument in an intransitive clause is identified as being in S function. Allocating functions A and O to the two core arguments in a transitive clause has a semantic basis. Briefly, that argument whose referent is most likely to be relevant to the success of the activity is identified as A. (An A argument most often has animate reference, and it is then that argument which could initiate or control the activity.) And that argument whose referent is most likely to be saliently affected by the activity will be in O function. There is fuller discussion of this, in §13.5.1–2, in relation to the association between syntactic functions and the semantic roles of the various semantic types of verbs.

In some languages, there are two further transitivity types: extended intransitive, which involves core arguments S and E (standing for ‘extension’ to core), and extended transitive, with three core arguments, A, O, and E (this is often referred to as ‘ditransitive’). It will be useful to repeat here the example sentences from Tongan presented in §3.2:

- (1) intransitive na'e 'alu ['a e fefiné]_S
 PAST go ABSOLUTIVE ART woman
 The woman (S) went
- (2) extended na'e sio ['a e fefiné]_S [ki he
 intransitive PAST see ABSOLUTIVE ART woman DATIVE ART
 tangatá]_E
 man
 The woman (S) saw the man (E)
- (3) transitive na'e taa'i ['a e tangatá]_O ['e
 PAST hit ABSOLUTIVE ART man ERGATIVE
 he fefiné]_A
 ART woman
 The woman (A) hit the man (O)
- (4) extended na'e 'oange ['a e tohi]_O ['e he
 transitive PAST give ABSOLUTIVE ART book ERGATIVE ART
 fefiné]_A [ki he tangatá]_E
 woman DATIVE ART man
 The woman (A) gave a book (O) to the man (E)

Tongan has an absolutive–ergative case system: S and O functions are shown by absolutive case (marked by particle 'a, where ' indicates a glottal stop) and A function by ergative case (particle 'e). Dative is shown by *ki*. Note that the noun phrases may occur in any order after the predicate, their functions being shown by initial case particles.

Argument profiles for the four clause types are (with case assignment in Tongan):

CLAUSE TYPE/PREDICATE	CORE ARGUMENTS		
intransitive	S (absolutive)		
extended intransitive	S (absolutive)	E (dative)	
transitive	A (ergative)	O (absolutive)	
extended transitive	A (ergative)	O (absolutive)	E (dative)

It will be seen that both extended intransitive and plain transitive involve two core arguments. They are identified as S and E in (2) and as O and A in (3). This is because the Perceiver ('the woman') in (2) shows the same grammatical properties as the S argument ('the woman') in the plain intransitive sentence, (1), and different properties from the A argument ('the woman') in the plain transitive (3). In addition, 'the woman' in (2) is marked by absolutive preposition 'a, just like the S argument 'the woman' in (1) (and like the O arguments

in (3) and (4)). Note also that the E arguments in (2) and in (4) are marked in the same way, by dative preposition *ki*.

As stated in §1.10, a grammar should properly examine the underlying structures and systems of a language, rather than looking just at surface realizations. Criteria for recognizing A, S, and O (plus E in languages which have this as a core relation) relate to the functional roles which these core argument types play in the grammar. These involve the following:

- (a) Their role in constraints involved in the formation of coordinate and subordinate constructions. Complex sentences may require that linked clauses share an argument which is in certain specified functions in each clause. For example, coordination may work in terms of an S/A pivot in some languages and an S/O pivot in others, referred to as ‘syntactic accusativity’ and ‘syntactic ergativity’ respectively. Or else, the occurrence of a common argument in one of a number of specified functions in main clause and in relative clause may be a requirement for a well-formed relative clause construction; see Chapter 17. Similarly for complement clause constructions, as described in Chapter 18.
- (b) Almost every language has some valency-changing derivations, which are framed in terms of core arguments. From an intransitive clause, with core argument in S function, can be derived a causative construction, where the erstwhile S argument goes into O function, and/or an applicative construction, where S becomes A. From a transitive clause, with core arguments A and O, passive derives an intransitive construction in which the O argument is now in S function, and antipassive derives an intransitive with A having taken on S function. (The original A in the first instance and O in the second are now placed in peripheral function, and may be omitted.)
- (c) In a language where reflexive and/or reciprocal constructions involve a reflexive or reciprocal pronoun (with the transitivity value of the clause remaining unaffected), it is always the A or S argument which is grammatical ‘controller’. That is, the A or S argument is fully stated while the reflexive or reciprocal pronoun is placed in O or E or a peripheral function. For example, in English we get: *John*_A *cut himself*_O, *Mary*_S *talks a lot [about herself]*_{PERIPHERAL}, [*The boys*]_A *hit [each other]*_O, and [*The girls*]_A *told stories*_O [*to each other*]_{PERIPHERAL}.

Almost every language has some surface grammatical mechanism(s) for marking core and peripheral arguments so that they may be recognized—and the discourse understood—by listeners. As outlined under (b) in §3.9—and in §6.1—argument functions may be identified:

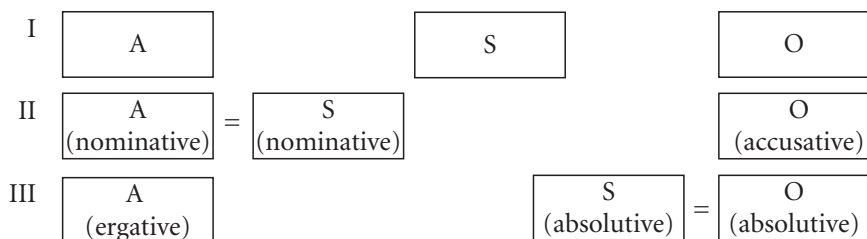
- (i) through marking on an NP which provides realization of an argument—by choice from a system of case inflections, or by adpositions.
- (ii) by the form of a bound pronoun which realizes an argument; this may attach either to the predicate or to some other constituent of the clause.
- (iii) by constituent order, as in English.

A very few languages (Thai is one example) essentially lack all of (i)–(iii) and rely on the pragmatics of the situation of utterance for identification of which argument is in which syntactic function.

It is vitally important to realize that the syntactic arguments of a predicate have significance because of their functions within the grammar, *not* because of the way in which they are marked. In one language, the subject (in A or S function) may be the first constituent of a clause. This does provide a ready means of identification. However, the import of the arguments relates to the role of subject relations in clause combining, in valency-changing derivations, and so on. Being placed at the beginning of a clause is not a function, but rather a surface marker of these functions, whose significance lies in the way they operate within the underlying structure of the grammar.

13.2 Marking of core arguments

Arguments in A and O function occur in a transitive and those in S function in an intransitive clause. Since S never co-occurs with A or O, it need not be marked differently from them, and very seldom is. There are two recurrent patterns—S marked like A and S marked like O. The possibilities can be diagrammed:



Row I shows A, S, and O all being marked differently. It is extremely unusual to find such a 'tripartite' system applying right across a grammar, although it can be part of a mixed marking system (see §13.5.4). The most common system, II, has A and S marked in the same way (nominative case) and O marked differently (accusative). Less common—but still found in about one-quarter of the world's languages—is III, where S and O are marked in the same

way (absolutive case) and A differently (ergative). Many languages combine marking types I and II (or I, II, and III), according to various semantic and syntactic parameters; these are discussed in §13.5.4.

Absolutive is always the functionally unmarked member of an absolutive–ergative contrast, and if either member is formally unmarked it will be absolutive (that is, it has zero realization or at least a zero allomorph). Nominative is almost always the functionally unmarked term in a nominative–accusative system, and may also be formally unmarked. However, there are some languages in which accusative is both functionally and formally unmarked; these include Yuman languages from California and a number of languages from North and East Africa (see Dixon 1994: 63–7; Aikhenvald 1995; and König 2008: 138–203).

In schemes I, II, and III, argument S is marked in the same way for all intransitive verbs. In II, S is marked in the same way as A, that argument of a transitive clause whose referent is most likely to be relevant to the success of the activity; this is typically an argument with animate reference which can initiate or control the activity. In III, S is marked in the same way as O, that argument of a transitive clause whose referent is not likely to initiate or control (it may be significantly affected by the activity).

An alternative scheme is where S is marked like A (shown as Sa) for some verbs in intransitive clauses and like O (So) for others:

$$\text{IV} \quad \boxed{\text{A}} = \boxed{\text{Sa}} \quad \boxed{\text{So}} = \boxed{\text{O}}$$

Typically, an intransitive verb with an S argument which acts volitionally will be marked like A (Sa) while an S argument whose referent essentially lacks control over the activity is marked like O (So). Such a language is said to have ‘split-S’ marking.

Mali (Baining family, New Guinea; Stebbins forthcoming and personal communication) shows syntactic function by constituent order. An NP realizing the A argument precedes the predicate, while that realizing O follows it, as in (5).

- (5) ngu_A ser ngeet_O
 1sg weave 3.NEUTER
 I weave it

In an intransitive clause Sa is like A in preceding the predicate, as in (6), while So is like O in following it, as in (7).

- (6) ngu_{Sa} namas (7) irees ngo_{So}
 1sg lie.down be.sad 1sg
 I lie down I am sad

A further group of languages shows a variant of this pattern. The S argument of an intransitive verb may be marked *either* like A (Sa) *or* like O (So) depending on the specific meaning of the verb in an instance of use. This is called a ‘fluid-S’ system. In Bats (or Tsova Tush, North-East Caucasian family) the 1sg pronoun has form *as* for A function and *so* for O function within a transitive clause. Either *as* or *so* could be used for the S of the intransitive verb *wože* ‘fall’, depending on whether or not the activity was the fault of the referent of the S argument:

(8) *as wože* ‘I (Sa) fell’ [it was my own fault that I fell down.]

(9) *so wože* ‘I (So) fell’ [no implication that it was my fault.]

In an important piece of fieldwork, Holisky (1987) checked 303 intransitive verbs with native speakers to see whether Sa or So or both were acceptable, and if so which was preferred. She found:

- 31 verbs are acceptable only with So—these refer to states or activities that cannot be controlled; for example ‘tremble’, ‘be hungry’, and ‘be ripe, grow up’.
- 78 verbs are acceptable only with Sa—they refer to activities that are naturally controlled, such as ‘walk, wander’, ‘talk’, and ‘think’.
- The remaining 194 verbs were judged acceptable with either Sa or So:
 - For some of these So is preferred since there is unlikely to be control; for example ‘die’, ‘burn’, and ‘become old’.
 - For some Sa is preferred since normally there is control; these include ‘wash’, ‘laugh out once’, and ‘begin’.
 - For the rest, Sa and So were considered equally likely; these include ‘lose weight’, ‘get lost’, ‘be late’, ‘get drunk’, and ‘slip (when So)/slide (when Sa)’.

The label ‘active/stative’ has been applied to split-S systems, with verbs taking Sa called ‘active’ and those taking So called ‘stative’. And the same label has been used of fluid-S systems. As pointed out under (d) in §2.5, this can lead to confusion. More specific designations would be needed, such as ‘split active-stative’ and ‘fluid active-stative’. (Although ‘split-S’ and ‘fluid-S’ are shorter and easier.)

13.2.1 Recognizing cases

There are two alternative ways of recognizing cases—according to form or according to function. These can be illustrated for Dyrbal.

Dyrbal has a case affix which has the following allomorphic variants:

- *-ŋgu* after a disyllable form ending in a vowel;
- *-gu* after a form of three or more syllables which ends in a vowel;
- a homorganic stop plus *-u* after a nasal or *ɣ* (*-bu* after *m*, *-du* after *n*, etc.);
- *-ru* after a form ending in a liquid (with the liquid dropped).

This suffix marks the A argument of a transitive clause, and also a peripheral argument in instrumental function, which can occur in a transitive or intransitive clause. The two functions are illustrated in:

- (10) *guda-ø*_O *yara-ŋgu*_A *balga-n* *yugu-ŋgu*_{PERIPHERAL}
 dog-ABSOLUTIVE man-ERGATIVE hit-PAST stick-INSTRUMENTAL
 The man hit the dog with a stick

The two methods for recognizing cases can be illustrated from these data:

(I.) If form is taken as primary we would have to say that there is a single case, with allomorphs *-ŋgu*, *-gu*, etc., which has two functions. It could be called ‘ergative/instrumental’ or just ‘ergative’ (focusing on the major function) with a note that it is also used to mark an instrument.

There could then be a temptation to say that, in a transitive clause such as (10), the A and the instrument comprise a single NP, a sort of composite argument. Words can occur in any order within a sentence in Dyirbal, so that (10) could be rearranged, and analysed, as:

- (10') *guda-ø*_O [*yara-ŋgu yugu-ŋgu*]_A *balga-n*
 Literally: ‘[man (with) stick]_A hit dog_O’

Instrumental constituents are most often found in transitive clauses, such as (10), but can occur in intransitives. This poses a problem for such an analysis.

(II.) If function is taken as primary, we must recognize two cases: ergative (marking a core function) and instrumental (a peripheral function), which happen to receive the same marking. In fact, ergative and instrumental arguments do behave quite differently under various syntactic derivations.

Antipassive is marked by derivational suffix *-ŋa-* on the verb (between root and tense inflection); the original A NP goes into S function, O goes into a peripheral function, marked by dative, and the instrumental NP remains unchanged:

- (11) *yara-ø*_S *balgal-ŋa-ñu* *guda-gu*_{PERIPHERAL}
 man-ABSOLUTIVE hit-ANTIPASSIVE-PAST dog-DATIVE
 *yugu-ŋgu*_{PERIPHERAL}
 stick-INSTRUMENTAL
 The man hit the dog with a stick

Here ‘man’ which was originally marked by ergative *-ŋgu* is now in absolutive case (with zero realization) while ‘stick’ retains its instrumental suffix *-ŋgu*. The opposite applies in an applicative derivation:

- (12) *yugu-∅*_O *yara-ŋgu*_A *balgal-ma-n*
 stick-ABSOLUTIVE man-ERGATIVE hit-APPLICATIVE-PAST
 *guda-gu*_{PERIPHERAL}
 dog-DATIVE

The man uses a stick to hit the dog (lit. The man hits-with a stick to the dog)

An applicative construction is marked by derivational suffix *-ma-* on the verb. The instrumental role now moves into O function, while the original O is placed in peripheral function, marked by dative. Here ‘stick’, which bore instrumental suffix *-ŋgu* in (10), is now marked by absolutive case (with zero realization), while ‘man’ retains its ergative case marking *-ŋgu*.

Alternative II assigns case in terms of underlying function—ergative and instrumental are distinct cases, since these arguments behave quite differently within the grammar. They just happen to be marked in the same way, by suffixes with identical allomorphy. Alternative II is plainly to be preferred over I, an analysis relating solely to surface structure.

‘Traditional grammar’ analyses surface structure, following alternative I. In primers of Latin it is said that accusative is ‘the object case’ but it is then reported that accusative case may, amongst other things, also mark ‘time how long’—as in ‘I slept the whole night’, where ‘whole’ and ‘night’ are in accusative form—and ‘space how far’—as in ‘They crawled three miles’, where ‘three’ and ‘miles’ are in accusative form. Within alternative II, which assigns case on a functional basis, two cases would be recognized here, having identical surface marking—accusative (for O function) and a case describing extent in time and in space.

(The ‘tradition’ of pursuing analysis in terms of surface structure led to recognition of genitive—which marks relationship within an NP—as having grammatical status similar to cases such as nominative, accusative, and dative—which mark relationships within a clause. This error of analysis was discussed under (f) in §1.10.)

13.3 Transitivity classes of verbs

Verbs can be arranged in classes according to which transitivity types of clause they may occur in:

- (i) Strictly intransitive verbs—may occur only in intransitive clauses; for example, *go* and *chat* in English.
- (ii) Strictly transitive verbs—may occur only in transitive clauses; for example, *recognize* and *promote* in English.
- (iii) Ambitransitive of type $S = A$. These verbs can occur in either an intransitive or a transitive clause, with the S of the intransitive corresponding to the A of the transitive. For example *knit* in English—one can say either *She_S is knitting* or *She_A is knitting [a scarf]_O*.
- (iv) Ambitransitive of type $S = O$. These verbs can occur in either an intransitive or a transitive clause, with S corresponding to O . For example, *trip* in English—one can say either *He_S tripped* or *She_A tripped him_O*.

It is of the highest importance when recognizing ambitransitive verbs (also called 'labile verbs') to specify whether of type $S = A$ or of type $S = O$.

In some languages, all verbs have a fixed transitivity—each is either strictly transitive or strictly intransitive. Latin and Dyirbal are essentially of this type. For such languages, one expects to encounter a number of valency-changing processes, deriving transitive stems from intransitive roots and/or deriving intransitive stems from transitive roots.

Many languages have a fair number of ambitransitives, of either or both varieties. There is, in fact, often a preponderance of one type. In Manambu (Ndu family, New Guinea) over 80 per cent of verb roots are ambitransitives of type $S = A$, with smaller numbers of strictly intransitives, strictly transitives, and $S = O$ ambitransitives (plus a few extended intransitives and extended transitives). Similarly in Tariana (Arawak family, Brazil), where well over half the verbs are $S = A$ ambitransitives. In contrast, Jarawara (Arawa family, Brazil) has, on my dictionary count, 28 per cent of its verbs $S = O$ ambitransitive as against just 3 per cent $S = A$ ambitransitive (in addition to 52 per cent strictly intransitive and 17 per cent strictly transitive). Some of the ambitransitives occur much more often in a transitive clause and rather seldom in an intransitive one; others show the opposite profile.

There may be a small number of verbs which are both $S = A$ and $S = O$ ambitransitive. Just two roots of this type have been uncovered in Jarawara.

- (a) *afi -na-* appears to be basically intransitive with the meaning 'bathe, take a bath'. As a transitive verb with $S = O$ orientation it means 'give a bath to, bathe (e.g. a baby)'. As a transitive verb of $S = A$ type it can mean 'jump into water' and the O NP is always *faha* 'water'.
- (b) *-awa-* is basically a transitive verb 'see, look at, look for', as in (13). When used intransitively it can either be of $S = O$ type, with meaning 'be visible', as in (14), or of $S = A$ type, with meaning 'look', as in (15).

Note that in the S = A intransitive sense the S NP must include the possessed noun *noki* (feminine)/*noko* (masculine) ‘eye’.

- (13) *jifo*_O *o-wa*
 firewood (feminine) 1sgA-look.for:feminine
 I look for firewood
- (14) *abarikos*_S *awe* *ama-ka*
 moon(masculine) be.visible:masculine EXTENT-DEC:masculine
 The moon is visible
- (15) *nokos*_S *awe-himata-mona-ka*
 eye(masc) look-FAR.PAST:noneyewitness:masc-REP:masc-DEC:masc
 He is said to have been looking (lit. his eye is said to have been looking)

Note that ambiguity is avoided between the two transitive constructions relating to *afi -na-* ‘bathe’ since for the S = A type noun *faha* ‘water’ must fill the O argument slot. And similarly for the two intransitive constructions relating to *-awa-* ‘see’, since the S = A type must include the inalienably possessed noun *noki/noko* ‘eye’ within its S NP.

For Lango (Nilotic, Uganda), Noonan (1992: 125) identifies just one verb which can be both S = O and S = A ambitransitive. Interestingly, it has the same meaning as one of the two verbs of this type in Jarawara. Corresponding to transitive *nènnò* ‘see (something)’, there is S = A ambitransitive *nénò* ‘see’ and also S = O ambitransitive *nèn* ‘be visible’. And David Watters (personal communication) points out that the Tibeto-Burman language Kham has just two verbs which are both S = A and S = O ambitransitive, ‘see’ and ‘hear’. If the basically transitive verb ‘see’ is used in an intransitive clause it is taken as S = O (‘appear, be visible’) unless the S NP includes noun ‘eye’ in which case it is S = A, literally ‘X’s eye sees’ (typically used in the negative, ‘X’s eye doesn’t see’ indicating that X is blind). Similarly for ‘hear’ and ‘ear’.

Just a few basically transitive verbs in English may have ambitransitive correspondents of both S = A and S = O varieties. For example, one can say [*The chef*]_A *is cooking* [*the soup*]_O. Or, focusing just on the general activity the chef is involved in, [*The chef*]_S *is cooking*. Or, describing what is happening to the soup, [*The soup*]_S *is cooking*.

13.4 More complex types

There tends to be a limit on how much detail the grammar of any particular language can tolerate. That is, it is not very likely that a language would combine all of the complicated mechanisms outlined above. But we do find some combinations, and others are certainly possible.

- Both split-S marking and ambitransitive verb type. Tariana has both of these properties. Its ambitransitive verbs are actually of types $S_a = A$ and $S_o = O$. Warekena (another language from the Arawak family) has a subclass of verbs of type $S_a = A$, and also a few verbs with $S_o = A$ (for example, ‘like’ and ‘fear’), and some with $S_a = O$ (including ‘break’).
- Ambitransitive verbs combined with extended intransitive and/or extended transitive clause types. Consider the following sentences in Samoan (Cook 1978: 61):

(16) *saa manatu-a [le teine]_O [e le tama]_A*
 PAST remember ARTICLE girl ERGATIVE ARTICLE boy
 The boy remembered the girl

(17) *saa manatu [le tama]_S [i le teine]_E*
 PAST think ARTICLE boy ABOUT ARTICLE girl
 The boy thought about the girl

Sentence (16) consists of a transitive clause, with the O argument unmarked and A indicated by ergative preposition *e*. The verb is specified as transitive by suffix *-a*. In (17) the same verb is used intransitively (with no suffix); the S argument being unmarked with the E argument indicated by preposition *i*. The important point is that the verb *manatu(-a)* is essentially ambitransitive, with $A = S$ and $O = E$. Note the meaning difference—‘remember’ versus ‘think about’—relating to the difference between transitive and extended intransitive construction types.

- A language could combine: (A) extended intransitive and/or extended transitive clause types; (B) split-S (or even fluid-S) marking; and (C) ambitransitive verb types. We await reports of such an interwoven system.

13.5 The semantic bases

The semantic parameters which underlie (a) the allocation of semantic roles to syntactic arguments, (b) the surface marking of core arguments, and (c) the transitivity classes of verbs, are intertwined. Building on these are the semantic parameters relating to non-canonical marking of core arguments, discussed in §13.6.

13.5.1 Identifying A

When one stops to think of it, the semantic reasons for identification of a transitive subject argument (A) are scarcely crystal clear. What is there in common between ‘Mary’ in *Mary cut the bread* (employing a knife), *Mary*

lent her car to John (she let him take it for the weekend), *Mary told John a joke in French* (since she knows he understands that language), *Mary considered the red skirt* (whether she should buy it or not), *Mary saw the collision* (it just happened in front of her, and she just wished she had been somewhere else at the time), and *Mary hates Mozart* (his music always makes her feel depressed). Is there anything in common between the verbs in these six sentences, or to Mary's six roles? Basically, rather little.

As mentioned in §1.9 and §1.11, a number of classes of lexemes, called 'semantic types', can be recognized, each associated with a major word class in individual languages. Those relating to the adjective class were discussed in Chapter 12. Both large and small adjective classes are generally associated with the DIMENSION, AGE, VALUE, and COLOUR types. In languages with a largish adjective class, this is likely to include lexemes from the PHYSICAL PROPERTY and HUMAN PROPENSITY types. When the adjective class is small, PHYSICAL PROPERTY typically goes into the verb and HUMAN PROPENSITY into the noun or verb class. In every language, the verb class has associated with it semantic types which include AFFECT, GIVING, SPEAKING, THINKING, and ATTENTION.

Each of the semantic types associated with the verb class has a number of semantic roles, describing the participants involved in the activity, state, or property described by its verbs. Some of the main semantic types for the verb class in English were illustrated in Table 3.1—under (b) in §3.3—repeated here as Table 13.1.

Repeating some of the discussion from §3.3, in English it is the leftmost of the roles in each row of Table 3.1 which is placed in A function. We need to enquire what the principle is underlying this association between semantic roles and syntactic functions. What is there in common between a person wielding an implement, for AFFECT, someone who—permanently or temporarily—transfers their possession of something, in GIVING, a person whose mouth produces an utterance, for SPEAKING, someone whose mind focuses on something, for THINKING, a person who—either purposely or

TABLE 13.1. Sample semantic types of the Verb class in English, and their roles

<i>Semantic type</i>	<i>Roles</i>				
AFFECT (for example, <i>hit, cut</i>)	Agent	Target	Manip		
GIVING (for example, <i>give, lend</i>)	Donor	Gift	Recipient		
SPEAKING (for example, <i>speak, tell</i>)	Speaker	Addressee	Message	Medium	
THINKING (for example, <i>consider</i>)	Cogitator	Thought			
ATTENTION (for example, <i>see, hear</i>)	Perceiver	Impression			
LIKING (for example, <i>like, love, hate</i>)	Experiencer	Stimulus			

involuntarily—receives a sense impression, for ATTENTION, and someone who experiences a certain internal feeling about something, in LIKING?

The principle appears to be as follows: ‘that role which is most likely to be related to the success of the activity is placed in A function.’ This is plainly the Agent for an AFFECT verb, the Donor for GIVING, the Speaker for SPEAKING, the Cogitator for THINKING, and the Perceiver for ATTENTION. Most often, the referent of the semantic role will be human, or at least animate. This role then equates with ‘the participant who could control and/or initiate the activity, state, or property, if anyone could’.

In many instances, the activity referred to by a verb may or may not be volitional. John could cut his finger on purpose or by accident. *Mary saw the circus* could describe her paying the admission fee and watching the show, or it could describe her—by chance—noticing a string of circus trucks driving down the road. Many verbs have the possibility for an A with inanimate reference, as in *The rain ruined the garden party* or *The fallen tree blocked the road*. (‘The rain’ and ‘the fallen tree’ are those roles which determine ‘ruining’ and ‘blocking’ respectively.) But many such verbs may, alternatively, have a human controller as A, as in *The butcher ruined the garden party* (through swearing at the vicar for buying sausages from a rival vendor) or *The protesters blocked the road* (so that politicians could not get through to the summit meeting).

In some languages, every transitive verb has an A argument whose referents can include animate beings with the capability of initiation and/or control. But there are languages which include just a few verbs whose A argument may *only* have an inanimate referent. For example, the dictionary I compiled for Yidiñ, an Australian language, has just four transitive verbs of this type:

- *wigi-l*, (rich food: A) makes (a person: O) feel sick.
- *manja-n*, (something: A) fills up (a part of a person’s body: O). For example: tiredness (A) fills my shins (O); vegetables (A) have filled my stomach (O).
- *jaja-l*, (sacred water: A) has adverse reaction to (someone: O). If some person angers the rainbow serpent (a powerful spirit), he may turn a sacred waterhole (which he controls) against that person, causing them to become sick and perhaps die.
- *yama-l*, (something: A) makes (a person: O) feel cold. The A is likely to be a cold wind, or cold weather.

Note that in each instance it is the referent of the A argument which is responsible for the state or activity described by the verb. The A always has inanimate reference, but for each verb the O argument refers to a person or a part of a person’s body.

Many languages have some LIKING verbs which are transitive, with the Experiencer role as A and Stimulus as O. Many fewer have verbs like those of the ANNOYING type in English, with role/function linkings reversed. In some languages ideas of 'liking' and 'pleasing' are expressed by extended intransitive verbs. For example, Portuguese has extended intransitive verb *gostar* and Spanish has extended intransitive *gustar*, which are cognate. A surprising difference is that in Portuguese the Experiencer is in S function for *gostar*, as in (18), while Spanish has the Stimulus as S for *gustar*, shown in (19).

- In summary we find that, across almost all languages, the semantic roles mapped onto A function are the Agent for a verb of AFFECT, the Donor for GIVING, the Speaker for SPEAKING, and the Cogitator for THINKING. For the ATTENTION type, the Perceiver role is generally mapped onto the A argument of a transitive verb, or the S argument of an extended intransitive (see §13.6.1). But for LIKING, there is a great deal of variation across languages. Appropriate verbs may be transitive, with Experiencer as A and Stimulus as O, or vice versa (or both of these), or intransitive with either of the roles in S function. In some

languages the idea of ‘liking’ may only be expressed by an adjective, or only by a noun.

13.5.2 Identifying O

For the verbs of a semantic type which has just two semantic roles, that which is not mapped onto A syntactic function will correspond to O. In English, this is the Thought for THINKING verbs, the Impression for ATTENTION, and the Stimulus for LIKING, as illustrated in:

(20) Mary_{COGITATOR:A} considered [the red skirt]_{THOUGHT:O}

(21) Mary_{PERCEIVER:A} saw [the collision]_{IMPRESSION:O}

(22) Mary_{EXPERIENCER:A} hates Mozart_{STIMULUS:O}

For verbs from these semantic types, the semantic role which relates to O syntactic function is defined negatively, as: ‘that participant which is *not* most likely to be related to the success of the activity’.

But, as shown in Table 13.1, some semantic types have more than three semantic roles. In (6)–(7) of §3.2, we illustrated how, for some AFFECT verbs in English, either the Target or the Manip (thing manipulated) role can be in O function:

(23) John_{AGENT:A} hit [the vase]_{TARGET:O} ([with a stick]_{MANIP})

(24) John_{AGENT:A} hit [a stick]_{MANIP:O} [against the vase]_{TARGET:E}

The principle here appears to be: ‘that role whose referent is most likely to be significantly affected will be placed in O function’. Hearing (23), one would infer that in all likelihood the stick is strong and the vase weak, so that this action may result in the vase being damaged. In contrast, (24) would be used if a brittle stick were struck against a sturdy vase so that it would be most likely that the stick should be damaged.

For verbs of SPEAKING, there are three roles in addition to that of Speaker (which is always in A function). The Medium role can be in O function (but then Addressee and Message would not be likely to be included), as in *Mary speaks Bengali* and *John talks rubbish*. The more frequently stated roles are Addressee and Message, either of which may—in appropriate circumstances—be placed in O function. Both alternatives are possible with verb *tell* in English:

(25) Mary_{SPEAKER:A} told John_{ADDRESSEE:O} [a joke]_{MESSAGE}

(26) Mary_{SPEAKER:A} told [that joke about the bishop]_{MESSAGE:O} [to some of her friends]_{ADDRESSEE}

That (non-A) semantic role which has definite, specific, individuated reference is likely to be placed in O function—*John* and *that joke about the bishop*

respectively. The non-O role has, in each sentence, vaguer denotation—a *joke* and *some of her friends*.

There are many verbs of SPEAKING in English. Some are like *tell* in occurring in alternate syntactic frames but others lack this freedom. Consider:

- (27) John_{SPEAKER:A} informed [the police]_{ADDRESSEE:O} ([of/about the accident]_{MESSAGE})
- (28) John_{SPEAKER:A} reported [the accident]_{MESSAGE:O} ([to the police]_{ADDRESSEE})

The meaning of *inform* focuses on the Addressee and that of *report* on the Message; these are the roles which *must be* in O function for the two verbs.

The range of syntactic frames available cross-linguistically for GIVING verbs is discussed in §13.5.3.

Every language has verbs of AFFECT and of SPEAKING with two or more semantic roles (other than that role which is in A function). For an AFFECT verb, either the Target or the Manip may be in focus for particular tokens of an activity. Grammars have different ways of handling this. As illustrated in (23–4), the English verb *hit* may occur in alternate syntactic frames—one with the Target role in O function and the other with Manip as O. In contrast, the verb *balga-l* ‘hit with a long rigid implement, held in the hand’ in Dyirbal is restricted to one syntactic frame, with Target in O function and Manip as a peripheral argument, marked by instrumental suffix *-ŋgu*. This is illustrated in (10) of §13.2.1. And, as explained there, Dyirbal has an applicative derivation, which moves the Manip role into O function; this is exemplified in (12). Thus, English and Dyirbal attain the same ends by quite different means—by alternate syntactic frames, and by employing a syntactic derivation.

Sentences (25–6) illustrated the two syntactic frames for *tell* in English, one with the Addressee and the other with the Message as O. Dyirbal here adopts a different strategy—it has two verbs ‘tell’, with different expectations for role/function associations. The more general of the two, *buwa-y* ‘tell’, expects to have Addressee as O argument, with Message marked by dative case. Then there is *jinga-y* ‘recount some particular piece of news, or story’, which generally has Message as O with Addressee indicated by dative.

In Fijian, the great majority of words can be used intransitively, with no suffix, or in a transitive clause, then taking a suffix. (They divide into S = O and S = A classes, discussed in §13.5.5.) Just a few transitive verbs may occur in alternate syntactic frames. For example *'eli-a* ‘dig’ has as O argument either what is dug in (such as ‘the ground’) or what is dug for (this could be ‘yams’). The role not marked as O is placed in a peripheral phrase.

What is particularly striking about Fijian is that a fair number of verbs have a choice of two transitive suffixes, relating to different semantic roles in O function. The verb *qoli* ‘catch fish with a net’ can be used intransitively, or transitively with suffix *-va*, for which the Target (the fish) is in O function—as in (29)—or transitively with suffix *-va'ina*, for which the Manip (the net) is in O slot—as in (30).

- (29) au aa qoli-va [a i'a]_{TARGET:O} [i-na lawa
 1sgA PAST catch-TR ART fish PREP-ART net
 yai]_{MANIP}
 THIS

I caught fish with this net

- (30) au aa qoli-va'ina [a lawa yai]_{MANIP:O} [i-na
 1sgA PAST catch-TR ART net THIS PREP-ART
 i'a]_{TARGET}
 fish

I fished with this net for fish

Other verbs with two transitive suffixes, each marking a different semantic role in O function, include (with sample O arguments included in each instance):

VERB	FIRST TRANSITIVE SUFFIX	SECOND TRANSITIVE SUFFIX
vana ‘shoot’	-a ‘at (O: animal)’	-ta'ina ‘with (O: gun)’
pu'u ‘be angry’	-ca ‘at/with (O: naughty child)’	-ca'ina ‘about (O: someone’s habits)’
vuunau ‘advise’	-ca O: person advised	-ta'ina ‘about (O: advice given)’
miimi ‘urinate’	-ca ‘on (O: some surface)’	-ca'ina O: what is passed (e.g. blood)

Verbs *ve'a(-ca/-ca'ina)* ‘defecate’ and *kaasivi(-ta/-ta'ina)* ‘spit’ behave in the same way as *miimi(-ca/-ca'ina)* ‘urinate’.

Jarawara has a single form for each verb but there are often alternative syntactic frames with different roles corresponding to O function. For *mii* ‘defecate, shit’ the O argument can either be what is expelled from the body, as in (31), or what it is expelled onto, as in (32).

- (31) inamatewe_A ama_O mii
 child blood(feminine) shit
 na-ka
 AUXILIARY-DECLARATIVE:masculine
 The boy child shat blood

- (32) inamatewe_A [mesa mese]_O mii
 child table(feminine) top.of shit
 na-ka
 AUXILIARY-DEC:masculine
 The boy child shat on top of the table

There is not likely to be any danger of ambiguity—‘the top of the table’ can only be what is shat on, not what is expelled from the body, in (32), and in (31) ‘blood’ would be understood as that which is expelled. Two other verbs in Jarawara behave in exactly the same way—*soo -na-* ‘urinate, pee’ and *saa -na-* ‘vomit’.

Jarawara allows wide freedom for association of semantic roles with syntactic function O. Textual examination of syntactic possibilities for the verb *tisa -na-* ‘shoot (with an arrow or slingshot)’ reveals that the A is always the hunter, but the O argument can be any of the other semantic roles involved in the activity. It is most frequently the animal or fish that is shot at, as in:

- (33) aba_O mee otaa tisa na otaa-ke
 fish(m) 3plO 1pl.excA shoot AUXILIARY 1pl.exc-DECLARATIVE:f
 We shot lots of fish

(Note that in Jarawara the unmarked gender, feminine (f), is always used to cross-reference pronouns, as with 1st person plural exclusive form *otaa* in (33), 3rd plural *mee* in (34), and 1st plural inclusive *ee* in (35), all in A function.)

Alternatively, the arrow that is used in the action can be placed in O function, as in:

- (34) faja wati_O mee tisa
 THEN arrow(m) 3nsgA shoot
 ne-mete-mone-ke fahi
 AUX-FAR.PAST.NON-eyewitness:f-REPORTED:f-DECLARATIVE:f THERE
 They are then said to have shot off arrows there

And there is one example where the O argument is ‘water’:

- (35) faha_O ee tisa ne-ne
 water(f) 1pl.incA shoot AUXILIARY-IRREALIS:f
 We could shoot into the water (to try to catch fish)

In summary, it will be seen that there is relatively little variation concerning which semantic roles are associated with syntactic function A but, in contrast, considerable freedom in relation to O.

13.5.3 Ways of expressing 'giving'

In some languages we find a small number of verbs for which three semantic roles must be stated (or understood). That is, they occur in an extended transitive syntactic frame. *Give* in English is like this. One can say either of

(36) John_{DONOR:A} gave [his old coat]_{GIFT:O} [to a beggar]_{RECIPIENT:E}

(37) John_{DONOR:A} gave [the winner]_{RECIPIENT:O} [a prize]_{GIFT:E}

It is not possible to omit anything from these constructions. That is, none of the following are judged as acceptable (in a neutral context): **John gave his old coat*, **John gave to a beggar*, **John gave the winner*, **John gave a prize*. The Gift and Recipient must be specified (although the latter could be shown by an adverb, as in *John gave his old coat away* or *John gave away his old coat*.)

(As in many instances, one or more roles may not be overtly expressed if they can be inferred from the context. If a charity person rattles a collecting box under your nose, you could say just *I've already given today*, which will be clearly understood as *I've already given some money to your charity today*.)

As in earlier examples (23–6), the role which is placed in O function in (36) and (37) is likely to have specific and individuated reference—*his old coat* in (36) and *the winner* in (37). When there is a choice of role-function association, it is likely to be determined by saliency of referents, and by the discourse and extra-discourse context of the clause.

Many verbs from the GIVING type in English can occur in the two syntactic frames illustrated in (36) and (37); they include *lend*, *sell*, *pay*, *owe*, and *bequeath*. But the meanings of some verbs from this semantic type are such that in most circumstances the Gift must be in O function (not the Recipient); for example, *donate*, *contribute*, and *deliver*.

We can briefly survey the different grammatical and lexical techniques which languages employ for dealing with 'give'.

A. A single lexeme which can be used in two syntactic frames, one with Gift and the other with Recipient in O function, as illustrated for English in (36–7). In Dyirbal, the verb *wuga-l* 'give' occurs in three frames:

DONOR	GIFT	RECIPIENT
(a) A function	peripheral argument with instrumental marking	O function
(b) A function	O function	peripheral argument with dative marking
(c) A function	O function	possessive modifier (with genitive marking) to head of O NP

Banjarese (Austronesian, Indonesia; Sari 1984) also behaves in this way.

- B. One lexeme which has a single syntactic frame. There are two possibilities:
- Bi. Gift role is always in O function. The Recipient is marked by dative case, or an adposition, or something similar. This applies to Burmese, Russian, Polish, Hindi, Telugu, Abkhaz (North-West Caucasian), as well as Austronesian languages such as Acehnese, Fijian, and Paamese, and Australian languages such as Warlpiri and Kalkatungu.
- Bii. Recipient role is always in O function, the Gift being a peripheral argument. This is less common than pattern Bi, but it is found in a fair number of languages, including the Uto-Aztecan language Huichol, the Austronesian language Tawala (from New Guinea), and the Australian language Nakkara. Typically, the O argument takes accusative case and the third core argument will often take dative case.

The two possibilities are:

SEMANTIC ROLE	Bi	GIFT	RECIPIENT	Bii	GIFT	RECIPIENT
SYNTACTIC FUNCTION		O	indirect object		indirect object	O
MARKING		accusative	dative, etc.		dative, etc.	accusative

It is important to distinguish semantic role, syntactic function, and marking of syntactic function, as shown in the diagram here. One sometimes hears scheme (ii) being described as ‘the object is in dative case’, identifying Gift as O because Gift is assigned to O function in many languages, shown as scheme (i). In (i) it is the Gift which has the same grammatical status as O in a simple transitive construction, and in (ii) it is Recipient which has this status.

- C. Two lexemes, one with Gift and the other with Recipient in O function. This can be exemplified from Jarawara:

- (38) jimawa_{GIFT:O} taa ti-na-hi
 knife give 2sgA-AUXILIARY-IMPERATIVE:fem
 [owa ni.jaa]_{RECIPIENT:PERIPHERAL}
 1sg POSTPOSITION
 You give the knife to me!
- (39) owa_{RECIPIENT:O} teekawa-hi
 1sg 2sgA:give-IMPERATIVE:fem
 [jimawa jaa]_{GIFT:PERIPHERAL}
 knife POSTPOSITION
 You give the knife to me!

In (38) the Gift role, *jimawa* ‘knife’, is in O function with Recipient, *owa* ‘me’, marked by postposition *ni.jaa*. The situation is reversed in (39), with Recipient as O and Gift as a peripheral argument, shown by postposition *jaa*. Note that the verb in (38) is non-inflecting *taa*, which is accompanied by an auxiliary, *-na-*, to which are added the 2sg A prefix *ti-* and immediate positive imperative suffix *-hi* (in unmarked feminine form, as determined by the pronoun in A function). The verb in (39) is inflecting *-kakawa-*, taking the same prefix and suffix; the first two syllables of *ti-kakawa-hi* (2sgA-give-IMPERATIVE) reduce to *tiwa-* and then to *tee-*.

- D. Some languages lack an extended transitive clause type and require a combination of two verbs to achieve what other languages do with one. In Koiari (New Guinea; Dutton 1996: 19) one can only express ‘Give me the hammer!’ through a biclausal construction:

- (40) hama_O mime da_O momi!
 hammer get 1sg give
 Get the hammer and give it to me! (lit: get hammer give me!)

Or there may be a serial verb construction, which has two verbs combined in one predicate. For example, in Kana (Benue-Congo family, Nigeria; Ikoro 1996a: 254) one finds:

- (41) Bārìlè āā sú zǐá nè ɲwíí
 name PROGRESSIVE take food give child
 Bārìlè is giving food to the child (lit: Bārìlè takes food gives child)

In many languages, ‘tell’—as exemplified in (25–6) of §13.5.2—and ‘show’ have similar syntax to ‘give’. That is, the Message for ‘tell’ and ‘what is shown’ for ‘show’ behave like Gift, while Addressee and ‘person to whom shown’ are like Recipient.

What syntactic function a given argument is in can generally be recognized by surface coding; for example, place in constituent order, or case marking. In English, S and A arguments precede the predicate and O follows, without any marking. Other arguments are recognized through being introduced by a preposition. In (36), *John gave his old coat to a beggar*, it is clear that *his old coat* is in O and *a beggar* is in E function. In (37), *John gave the winner a prize*, we see that *the winner* is in O function. But what about *a prize*, which follows *the winner* but is not marked by a preposition? This has been referred to as a ‘second object’. But does it exhibit any of the defining properties of an O argument?

One important criterion concerns passivization. For the great majority of transitive verbs in English, the underlying O argument can be put into S

function within a passive construction. One can say *The winner was given a prize (by John)*, confirming the O status of *the winner*. But *a prize* from (37) cannot become A argument in a passive. That is, for most speakers it is not permissible to say **A prize was given the winner (by John)*. It appears that *a prize* in (37) is a type of non-O argument which, unusually, is not marked by a preposition.

A number of Bantu languages allow two NPs to follow the verb, neither bearing any case or adpositional marking. For example, one could refer to Gift and the other to Recipient, and these may occur in either order. It turns out that the NP which immediately follows the verb is passivizable and cliticizable, and is thus in O function. That is, constructions in these languages are just like (36) and (37) in English, if no *to* were included in (36). Hock (1985) describes how, in Sanskrit, two constituents in a clause may both receive accusative case. However, only one is in O function; this *must* be marked with accusative and it may be passivized. For the other constituent, another case (dative, ablative, instrumental, etc.) can be substituted for accusative, and passivization is not possible.

Many other tests may be invoked to distinguish an O from other arguments, within a specific language. These include: being realizable by a bound pronoun, and having a role in rules for coreferential NP omission.

13.5.4 Split systems of marking

Ways of marking core arguments were discussed in §13.2—accusative system, ergative system, and so on. Many languages combine several types of system, the split being semantically conditioned. Relevant parameters can be: (i) the referents of core arguments; (ii) tense and/or aspect of the clause; and (iii) the syntactic status of the clause—whether main or subordinate clause (and which kind of subordination). Item (i) will be discussed in some detail, with (ii) and (iii) being briefly summarized at the end of this section.

(i) Referents of core arguments

It is usual to provide a special mark for some unusual feature. A few decades ago, people would talk about *male nurse* and *female professor*. Most nurses were women and the expected referent of *nurse* was female. For a man to be a nurse was sufficiently unusual for the term *male nurse* to be employed. Similarly, in England the great majority of professors were male; if a woman did attain this high academic rank, the label *female professor* might be used. This principle, of providing specific marking for something which is unusual, extends into grammar.

Split systems of case marking are conditioned by the ‘nominal hierarchy’, shown in Figure 13.1. (It will be seen that this combines parameters of person,

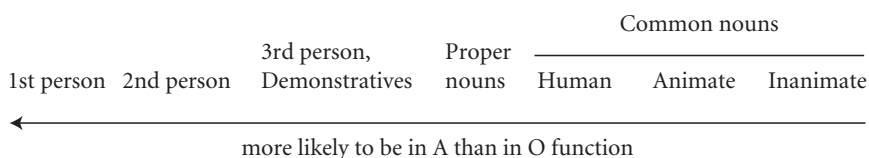
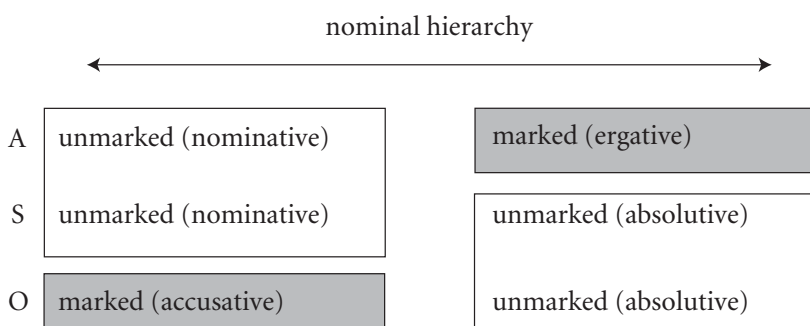


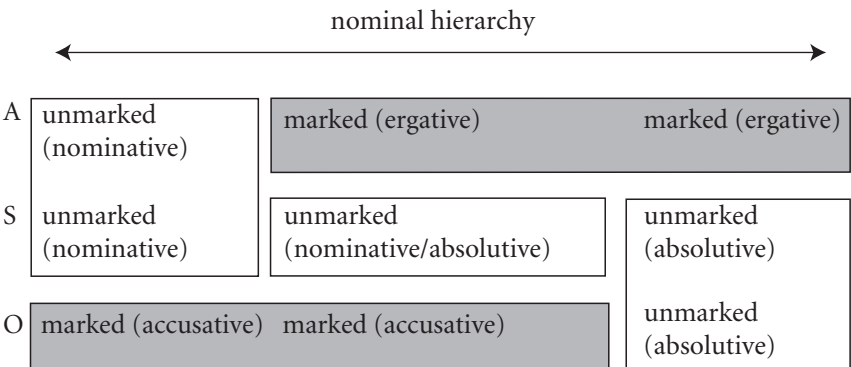
FIGURE 13.1. The nominal hierarchy

humanness and animacy, proper and common nouns, etc.) The basic principle is that the further to the left of the hierarchy the referent of a core argument appears, the more likely it is to be in A function rather than in O function. That is, an A argument is more likely to be animate than inanimate, more likely to be human than animate, more likely to be a proper than a common noun, and so on. A 1st or 2nd person pronoun is very likely to be in A function and—within speech act participants—a speaker is more likely to conceive of themselves doing something to an addressee than vice versa. (There are occasional exceptions to the hierarchy; in a few languages 2nd person comes to the left of 1st; details are in Dixon 1994: 83–97.)

What arises from this is the following. For participants at the left end of the nominal hierarchy, A is the default syntactic function which is likely to be unmarked. O is the marked function, and is likely to attract a non-zero case (accusative). At the right end of the hierarchy, O is likely to be the default syntactic function, and here A will be marked (by ergative case). Accusative marking—which extends in from the left—and ergative marking—which extends in from the right—must at least meet in the middle. This gives a (nominative–)accusative system for the left-hand portion of the hierarchy and an (absolutive–)ergative system at the right-hand end:



Or, accusative and ergative marking may overlap in the middle of the hierarchy:



We now have a tripartite system in the middle (with all of A, S, and O marked differently) flanked by an accusative system on the left and an ergative one on the right. This can be illustrated from the Australian language Yidiñ:

A	∅	ERG	ERG	ERG	ERG
S	∅	∅	∅	∅	∅
O	ACC	ACC	(ACC)	∅	∅
	1st and 2nd person pronouns	human deictics and interrogatives	inanimate deictics, proper names, kin terms	inanimate interrogatives	common nouns and adjectives

1st and 2nd person pronouns have a nominative-accusative system; ACCusative is $-(\text{:})\tilde{n}$ and nominative receives zero marking (note that this language lacks 3rd person pronouns). On the right, an absolutive-ergative system applies for inanimate interrogatives, common nouns, and adjectives: ERGative has forms $-\eta gu \sim -lu$, etc., and absolutive has zero marking. Human deictics and interrogatives show tripartite marking. (For inanimate deictics, proper names, and kin terms, the ACCusative suffix is optional.)

It should be noted that the nominal hierarchy is not an absolute scale, but rather a general scheme which has minor variations as it applies to individual languages. For instance, in Dyirbal proper names (which can take accusative marking) are to the left of demonstratives (which cannot).

A participant from the left-hand end of the nominal hierarchy is likely to be the initiator or controller of an activity, and thus to be in A syntactic function. Similarly, a participant from the right-hand end is more likely to be saliently affected by an activity, and thus to be in O function. All this applies

to transitive clauses. An intransitive clause has a single core argument, in S function, shown as consistently marked in the diagrams just presented. But, as described in §13.2, some languages have split-S marking—either So, with S marked like O in a transitive clause, or Sa, with S marked like A.

The semantic basis for the Sa/So split varies from language to language. Study of languages with this kind of system—in particular, those described in Mithun's (1991) excellent survey—reveals two basic semantic schemes:

	Sa	So	example languages
(a)	volitional	non-volitional	Guaraní, Lakota
(b)	non-affected	affected	Central Pomo, Caddoan

Basically, in languages of type (a), if the referent of the S argument can act volitionally Sa marking is employed, otherwise So. In contrast, in type (b) languages it is So which is positively specified, as an argument whose referent is affected; otherwise Sa marking is used. The difference between the two types can be illustrated:

	volitional, not affected, e.g. 'go', 'swim'	not volitional, not affected, e.g. 'be tall', 'be strong'	affected, not volitional, e.g. 'be cold', 'be angry'
(a)	Sa	So	So
(b)	Sa	Sa	So

In the great majority of instances, placement of verbs in Sa and So classes can be explained in terms of general semantic principles, such as those briefly outlined above. There are, however, always a number of exceptions—for example, in a language with scheme (a) some verbs which describe non-volitional acts may nevertheless be coded as Sa. There can be any of a variety of reasons for this. One may relate to borrowing a verb from another dialect, maintaining its Sa or So marking, but shifting the meaning in such a way that volitionality or affectedness is altered.

Or else diachronic change could be responsible. For Haida (isolate; Queen Charlotte Islands, Canada), Enrico (2003: vol. 1, 95) notes that the verb *raganjuu* originally meant 'breathe', a volitional activity, and was marked as Sa. Over time, the meaning has shifted to 'be hanging on, still breathing, not dead yet'. In the modern language, this verb fails tests for volitionality—for example, it cannot occur in imperative mood, nor with 'try'. Yet it is still marked as Sa, a historic relic of past meaning. An example in the opposite direction concerns *gadas* which originally described a non-volitional activity 'come off, ricochet

off' and was naturally So. The meaning has now changed to the volitional 'leave to live elsewhere', but So marking has been retained.

Then there are fluid-S systems where some intransitive verbs can have their sole core argument marked either like A or like O, with a difference of meaning, as exemplified for Bats in §13.2. It appears that the principle here is: use Sa if the activity is being controlled or is the 'fault' of the referent of the S argument, and use So elsewhere. But its interpretation may well be culture-based. For example, Holisky (1987: 115) reports on her work with Bats: 'When I constructed the first person form for the verb "get poor" ... using [Sa] marking, my consultant did not say *categorically* that it wasn't possible. She said it isn't possible, because you would never want to be poor.'

(ii) Tense and/or aspect of the clause

Suppose that a stretch of text describes one participant (A) doing something to another participant (O). This can be viewed in either of two ways:

- (a) From the point of view of the O—how it is affected by a series of actions. This can be the basis for an (absolutive-)ergative system of marking.
- (b) From the point of view of the A—how this participant plans, initiates, and controls the actions. This can be the basis for a (nominative-)accusative marking system.

For something that is completed (perfect aspect) or has already happened (past tense), either of the two viewpoints could be invoked. For something which has not yet eventuated (in the future) only (b) is likely to be plausible. A person may plan to undertake a series of actions—viewpoint (b)—but it is much less likely for it to be anticipated that something could be the O of a series of actions—viewpoint (a).

We would thus predict that when a language has an ergative/accusative split conditioned by tense or aspect, an ergative system—under viewpoint (a)—would apply for perfect aspect or past tense—contrasting with an accusative system—under viewpoint (b) elsewhere. This is what is found in a number of Iranian and Indo-Aryan languages, in Classical Armenian, in some Mayan languages from Central America, and in Burushaski (an isolate, spoken on the Karakoram range in northern Kashmir). See Dixon (1994: 97–101) for fuller discussion, examples, and references.

(iii) Syntactic status of clause

Another kind of split involves a main clause showing one variety of marking system, with a different system applying in some type(s) of subordinate clause. How this split works depends on the kind of subordinate clause involved.

We can again invoke viewpoints (a) and (b). A purposive construction (someone doing something in order to facilitate doing something else) is necessarily of the accusative type, (b). If the main clause differs, it will have to be the one to employ an ergative system. This is what is found in the Nilotic language Pãri (Andersen 1988).

A relative clause typically describes something that has happened, and can represent viewpoint (a). This kind of split will be expected to show an ergative system in the relative clause as opposed to an accusative system in the main clause. Further discussion, exemplification, and references are in Dixon (1994: 101–4).

13.5.5 Transitivity classes

All languages have a class of transitive verbs and a class of intransitive verbs, but their memberships do not fully correspond. For instance, in English *laugh* is a strictly intransitive verb, but the corresponding lexeme in the Australian language Guugu Yimidhirr is the strictly transitive *diingal* ‘laugh at’. We can investigate which semantic parameters help to determine whether a given verbal concept is likely to be expressed by a transitive or by an intransitive verb. Basically, these are: (i) whether the verb has two or more syntactic roles; (ii) whether there is volitional control; (iii) whether or not the verb describes an action; and (iv) whether the referent of one role is saliently affected. Table 13.2 illustrates five verbal concepts in terms of these four semantic parameters.

The four parameters can now be briefly discussed in turn.

(i) Number of semantic roles

There are very few verbal concepts for which only one semantic role can be recognized (leaving aside general location and time specifications, which can apply to virtually every verb). One example is ‘hiccup’, describing an involuntary bodily gesture. Otherwise, a concept which is coded as a strictly

TABLE 13.2. Illustration of the semantic parameters underlying transitivity

	(i) two or more roles	(ii) volitional	(iii) describing action	(iv) one role saliently affected
‘hit’	✓	✓	✓	✓
‘touch’	✓	✓	✓	—
‘follow’	✓	✓	✓	—
‘praise’	✓	✓	(✓)	—
‘like’	✓	—	—	—
‘hiccup’	—	—	—	—

intransitive verb in one language may be transitive (or ambitransitive) in another.

English has a fair number of intransitive verbs, Fijian very few. We can illustrate how Fijian supplies a second core argument for a selection of verbs which are strictly intransitive in English. All are, in fact, ambitransitive of type S = A in Fijian:

ENGLISH	FIJIAN INTRANSITIVE	FIJIAN TRANSITIVE	IDENTITY OF O
<i>arrive</i>	<i>yaco</i> 'arrive, happen'	<i>yaco-va</i>	place arrived at
<i>creep</i>	<i>yaqa</i> 'creep (like a crab), walk with body bent down'	<i>yaqa-va</i>	place crept/walked towards (e.g. doorway)
<i>go</i>	<i>la'o</i> 'go'	<i>la'o-va</i>	EITHER: thing gone for (e.g. coconut); OR: distance travelled (e.g. three miles)
<i>sneeze</i>	<i>suru</i> 'sneeze'	<i>suru-ta</i>	loved one thinking of you, believed to be the reason one sneezes

It will be seen that the semantic roles mapped onto the second core argument (to make the verb transitive) can be culture-specific, as with Fijian *suru-ta* 'sneeze because of'.

In essence, just about any verbal concept *could* be allocated two semantic roles (relating to A and O syntactic functions), for the verb to fall into the transitive class. The exceptions are involuntary gestures like 'hiccup' (and there are very few of them). But even for 'hiccup' it might be that some society imputes a cultural design to the activity, which could yield a semantic role in O function (as we saw for 'sneeze' in Fijian).

(ii) Volitional control

It is sometimes said that if volitional control is involved in an activity, it is likely to be represented by a transitive (rather than an intransitive) verb. (See, for example, Hopper and Thompson 1980.) This is not entirely true. There are many intransitive verbs where the referent of the S argument exercises control—for example, 'walk', 'run', 'crouch' (these are among the verbs likely to be classed as Sa in a split-S language). Indeed, 'walk', 'run', and 'crouch' *must* involve volitional control. Prototypically transitive verbs such as 'hit', 'cut', and 'touch' generally also do so, but not invariably—someone may hit, cut, or touch a person (or animal or thing) accidentally.

Indeed, for some transitive/intransitive pairs—with corresponding meanings—it is the intransitive member which requires volition. English distinguishes between *hear*, a transitive verb, and *listen*, which is basically intransitive. If any noise eventuates close to a non-deaf person, they cannot help but hear it; there is no choice or control or volition involved. In contrast, intransitive verb *listen* implies a degree of concentration—that is, volition is involved, and control. One can use an intransitive construction in imperative mood: *You listen!*

What could be said is not that volition involves an activity being coded through a transitive verb, but that lack of volition makes an activity more likely to be coded by an intransitive verb (the So class of split-S languages). This is, however, no more than a tendency.

(iii) Whether or not the verb describes an action

As outlined in §13.1, some languages have two construction types which each feature two core arguments:

- (a) SIMPLE TRANSITIVE core arguments A and O
- (b) EXTENDED INTRANSITIVE core arguments S (same as the sole core argument in a simple intransitive clause) and E (same as the third core argument in an extended transitive)

Construction (a) is plainly ‘more transitive’ than (b), which is a special type of intransitive. As mentioned under (c) in §3.2, verbs from semantic types ATTENTION (‘see’, ‘hear’, ‘smell’, etc.) and LIKING (‘love’, ‘hate’, etc.) typically occur in an extended intransitive frame, in languages which have this construction type. These have in common that they do not describe actions, and—harking back to parameter (ii)—that they do not have to involve volitional control. That is, scoring a ‘no’ for these two parameters makes a two-role verb more likely to be classed as extended intransitive.

(iv) One role being saliently affected

This again ties in with ATTENTION and LIKING verbs being coded as extended intransitive in languages which exhibit this construction type. In *Mary saw/heard/smelt John* and in *Mary loves/hates John*, the referent of the O argument, *John*, may not be aware of this, let alone affected by it.

Other verbs where no participant is saliently affected may be transitive in one language, and intransitive in another. For example, *enter* is used in transitive clauses in English but in the Australian language Yidiñ the corresponding verb *bila-n* ‘enter’ is strictly intransitive, ‘what is entered’ being shown by a peripheral NP marked by dative or locative case. *Follow* is ambitransitive of type S = A in English but Jarawara has only a strictly intransitive

verb *joto -na-* ‘follow’. Whereas in English one might say *The women followed the men*, a Jarawara would prefer to express it by ‘The men went first and the women followed.’

Other verbal concepts which do not refer to any participant being affected are likely to be coded as intransitive in some languages. For example, ‘know’, ‘imagine’, ‘remember’, ‘imitate’, ‘obey’. The effect of touching someone or something is much less than the effect of hitting or cutting, and so ‘touch’ might well be rendered as an intransitive verb.

These four parameters are at best indicators. Many languages have robust transitive verbs which do in fact register ‘no’ for all of (ii)–(iv). For example, *imply* in English. This can include a volitional participant—as in *John implied that we would have to go*—but it need not. For example, [*That George Bush invaded Iraq*]_A *implies* [*that he has an insatiable greed for oil*]_O.

Many languages have ambitransitive verbs (occurring in transitive and intransitive clause types) of both S = A and S = O varieties. What is the basis for this split orientation? We noted in §13.3 that it is not uncommon to find a preponderance of one variety of ambitransitives. Thus, Tariana and Manambu each have many verbs of the S = A type and only a few S = O, with this being reversed in Jarawara. It will be most instructive to examine a language in which the two subclasses of ambitransitives are of approximately equal size.

For Boumaa Fijian, I studied about 460 verbs (probably rather more than one-third of the total verb class—see Arms 1974: 125). About four-fifths of these may be used in both transitive and intransitive clause types; of these 53 per cent are S = A and the remainder S = O ambitransitives. The underlying semantic principles are fairly clear and can be briefly summarized. (Similar principles apply for the semantic explanation of S = A and S = O ambitransitives in other languages.)

(i) For verbs of MOTION, REST, and GIVING the guiding principles are:

- Verbs where the motion or action of the referent of the A argument is seen as the most significant aspect of the situation described are S = A ambitransitive.
- Verbs where the effect on the referent of the O argument is seen as the most significant aspect are S = O ambitransitive.

Verbs referring to different modes of motion are S = A; for example *lade* (-*va*) ‘jump (for/over)’, *yaqa*(-*va*) ‘creep (to)’, *ada*(-*va*) ‘run (for)’, and *bale* (-*ta*) ‘fall (on)’. Verbs for direction of motion are also of S = A type, including *la'o*(-*va*) ‘go (for)’, *viro*(-*ca*) ‘return (to)’. But the general verb *yaavala*(-*ta*) ‘put in motion (e.g. start engine); be in motion’ is S = O.

If there is some change in the orientation of the referent of the O argument, then the verb will be in the S = O class. This covers lexemes such as *vu'i(-ca)* 'turn', *lo'i(-a)* 'bend (at a joint)', *lobi(-a)* 'fold', *tobe(-a)* 'plait', and *cori(-ta)* 'tie, tether'.

Verbs of 'pulling' and 'pushing' are particularly interesting. General verbs of this set are of the S = A type since it is the activity of the referent of the A argument which is focused upon—*dree(-ta)* 'pull' and *bili(-ga)* 'push'. But more specific verbs such as *cavu(-ta)* 'pull up' and *beti(-a)* 'pluck (fruit)' are S = O type since it is the effect on the referent of the O argument which is seen as most significant (a fruit is attached to a branch, and then detached from it).

At first sight, verbs of carrying appear to be randomly split between the two ambitransitive types. However, more careful examination reveals that those for which the referent of the O argument must be human are of S = O type—*'eve(-ta)* 'carry (baby) on hip', *vava(-a)* 'carry (baby) on back'—and those for which it must be inanimate or can be anything are S = A type—*drewe(-ta)* 'carry (e.g. a bag) on back', *roqo(-a)* 'carry (anything) in arms', *qumi(-a)* 'clench (anything) in fist'. If the O of the carrying verb must be human, then this is the argument which is focused on, and the verb is S = O; otherwise it is the person doing the carrying who is the most significant participant, and the verb is S = A.

(ii) There are many verbs describing some activity which affects the argument of the O NP and they are uniformly S = O. This covers verbs of peeling, breaking and dividing, rubbing and painting, covering, cutting and piercing, and lighting and extinguishing a fire.

(iii) Verbs of SPEAKING are either S = A ambitransitive or strictly transitive. Those from the THINKING semantic type are S = A (e.g. *vuli(-ca)* 'learn'). As noted under (2) in §12.5.1, the HUMAN PROPENSITY semantic type is in Fijian associated with the verb class, and its members are all S = A ambitransitives; for example *maarau(-ta'ina)* 'be happy (about)' and *'idacala(-ta'ina)* 'be shocked or amazed (by)'.

(iv) Looking at the ATTENTION type, we find that the two most common verbs have different syntactic profiles *rai(-ca)* 'see, look at' is S = A, while *rogo(-ca)* 'hear, listen to (tr); be audible (intr)' is an S = O ambitransitive. The explanation for this is that 'being visible' and 'being audible' are rather different. Just by existing, a thing or a person will be visible—in 'I saw this child' the significant thing is not the child being there but the fact that the speaker saw them. So *rai(-ca)* is naturally an S = A type verb. However, people and things do not continuously emit noise. And once a noise is made, it is impossible for anyone in the vicinity not to hear it. In 'I hear this child', the fact that the child produces a noise is regarded as the most significant thing, hence *rogo(-ca)* is an S = O ambitransitive.

This asymmetry between ‘see’ and ‘hear’ is not confined to Fijian. It was noted in §13.3 that in Jarawara *-awa-* ‘see, look at, look for’ is one of only two verbs identified as being both $S = A$ and $S = O$ ambitransitive. However *-mita-* ‘sense by hearing, taste, smell, or touch (tr); be audible (intr)’ is—like *rogo(-ca)* ‘hear, listen to (tr); be audible (intr)’ in Fijian—in the class of $S = O$ ambitransitives, no doubt for the same reason.

English has some $S = O$ ambitransitives (typically referring to an activity in which the referent of the S/O argument is saliently affected; for example *break*, *trip*, *melt*) and also a fair number of verbs of the $S = A$ type. The latter could be viewed as omitting an O argument, and it is instructive to investigate when this may occur. Three different circumstances can be identified.

(a) The O NP for a verb in the semantic types **ATTENTION** and **THINKING** can be omitted if known from the context. It is often a sort of anaphoric zero, as in: *Do you remember that night when we all got drunk? Yes, I remember.*

(b) Verbs from the **AFFECT** type may omit statement of an O NP when describing an activity being performed continuously, over a period of time, to a typical patient of that verb. For example *He’s been sawing* [sc. wood] *all afternoon*, and *She’s knitting* [sc. a garment]. There is wide scope for what the unstated O might be. One may say *He’s polishing now* when the Patient could be any object which is habitually polished—a table, some silverware, or the floor. Turning now to the **CORPOREAL** semantic type, *John is eating* could be said irrespective of what in particular he may be putting into his mouth. The identity of the thing(s) consumed is here considered to be of little importance, the focus being on the fact that it is John who is indulging in the activity of eating.

(c) The verb *drink* has quite different properties from those of *eat*. If one hears someone say *I don’t drink*, the unstated O argument is taken to refer to alcoholic drink. *John’s drinking again* is likely to imply not only that he is partaking of alcohol, but doing so to excess. In our society, drinking alcohol is a social activity which can have bad results and about which, as a consequence, many people have strong feelings. Thus the linguistic convention has arisen that using the verb *drink* without a specified O NP implies that it is alcoholic drink which is being referred to by the implicit O argument. (This is typically referred to just by the noun *drink*.)

13.6 Non-canonical marking of core arguments

Languages which use marking of NPs to identify core arguments fall into two sets. In some, the A for every instance of every transitive verb is marked in the same way, and similarly for O (and for S of intransitive verbs). In others there is a canonical marking pattern, which applies most of the time, and in

Scenario (a) can be illustrated for Chechen-Ingush, from the North-East Caucasian family. The canonical marking for transitive clauses in this language is ergative case for A argument and absolutive for O. However, verbs of the ATTENTION and LIKING types have their A argument marked with dative case, and verbs referring to possession have their A shown as genitive. Nichols (1980, 1983) demonstrates that there is a syntactic unity to ergative-marked A, dative-marked A, and genitive-marked A. The A argument of all verbs behaves in a consistent way, with respect to reflexivization, constraints on clause linking, and so on. Another example of scenario (a) comes from Icelandic, which has nominative/accusative canonical marking. However, for verbs like ‘see’ and ‘think’, the NP in A function is marked by dative in place of the canonical nominative. Zaenen, Maling, and Thráinsson (1985) show that nominative-marked A and dative-marked A have identical syntactic properties; for example, in control of reflexivization.

(42) ayer vi [a tu hermana]_O
 yesterday see:Past:lsgA PREPOSITION your sister
 Yesterday I saw your sister

In Finnish, the NP realizing an O argument is generally in accusative case, as in (44), but may alternatively be marked by partitive case, as in (45):

(44) liikemies_A kirjoitti kirj_{en}_O
 businessman:NOMINATIVE wrote letter:ACCUSATIVE
 valiokunnalle
 committee:ALLATIVE
 The businessman wrote a letter to the committee

- (45) liikemies_A kirjoitti kirjettä_O
 businessman:NOMINATIVE wrote letter:PARTITIVE
 valiokunnalle
 committee:ALLATIVE
 The businessman was writing a letter to the committee

Hopper and Thompson (1980: 262) explain that accusative marking on the O, in (44), indicates that the activity is completed (the letter is finished) whereas partitive, in (45), shows that it is not yet completed (the letter has not been fully written).

Hopper and Thompson (1980: 264–5) also mention that, in Estonian, an NP is marked by partitive case when it refers to ‘some of X’, but takes the canonical accusative (which doubles as genitive marker) when it refers to ‘all of X’. Partitive has further uses—including marking the O NP of a negative verb, and the O NP for verbs of ATTENTION and LIKING (discussed further in §13.6.1).

The different markings for ‘some’ and ‘all’ reference of the NP, and for completion, introduce parameters which appear to apply only for non-canonical marking. The ‘specific human’ reference condition on O NPs in Spanish is reminiscent of parameters discussed in relation to split marking systems, under (i) of §13.5.4.

Under (iv) in §13.5.5, ‘follow’ was quoted as an example of a verb which does not involve any participant being saliently affected. We noted that, while *follow* is a transitive verb in English, Jarawara only has intransitive *joto -na-* ‘follow’. Correlating with this, in Icelandic *fylgja* ‘follow’ is a transitive verb with non-canonical marking, its O NP taking dative inflection rather than the canonical accusative. In Table 13.2, ‘praise’ was shown as different from ‘like’ in that it is volitional (and ‘praising’ might be said to constitute an action whereas ‘liking’ could not be so described). We can predict that ‘praise’ is more likely than ‘like’ to be coded as a transitive verb.

Only a small sample of the different kinds of non-canonical marking has been mentioned here. Further examples—including non-canonical marking on S arguments—are provided in Onishi (2001a).

13.6.1 ATTENTION and LIKING verbs

There have been a number of mentions in this chapter of the special treatment accorded by some languages to verbs from the ATTENTION and LIKING semantic types. Recurrent use of non-canonical marking broadens the picture. It will be convenient here to summarize the varying ways in which grammars deal with these semantic types. They do have similar characteristics—bearing two semantic roles, not necessarily involving volition, not really describing

anything which could be called an action, and with no participant being saliently affected.

The various ways in which verbs from the ATTENTION and LIKING types are dealt with cross-linguistically can be summarized.

A. Nothing special. Shown by transitive verbs, with canonical marking.

Ai. In an accusative language, Perceiver role for ATTENTION and Experiencer role for LIKING are both in A syntactic function, marked by nominative case; Impression role for ATTENTION and Stimulus role for LIKING are both in O syntactic function, marked by accusative case. This applies to Latin, among very many other languages.

Aii. In an ergative language, Perceiver and Experiencer roles are again in A syntactic function, marked by ergative case; Impression and Stimulus roles are in O function, marked by absolutive case. This applies to Dyirbal, among very many other languages.

B. Shown by extended intransitive verbs.

Perceiver and Experiencer roles are in S syntactic function (marked by nominative or absolutive case, depending on the language); Impression and Stimulus roles are in E function (often marked by dative case). This was illustrated for Tongan by examples (1–4) of §13.1.

C. Shown by transitive verbs, but with non-canonical marking for A and O.

— Ci. Where standard marking is nominative for A and accusative for O (non-canonical markings are shown in bold type):

Perceiver (A) for ATTENTION	Experiencer (A) for LIKING	Impression (O) for ATTENTION	Stimulus (O) for LIKING	
nominative		partitive		Estonian
dative		accusative	nominative	Icelandic
genitive		accusative		Bengali

The Perceiver and Experiencer roles (each in A function) are coded in the same way in all three languages. Similarly, Impression and Stimulus are coded in the same way in Estonian and Bengali. However, Icelandic retains accusative case for the Impression role (O argument for ATTENTION) while using nominative for Stimulus (O for LIKING). Note that in Estonian it is the O argument which is marked non-canonically, whereas in Bengali (and in Icelandic for ATTENTION verbs), the A argument receives non-canonical marking. LIKING verbs in Icelandic have both core arguments marked in a way different from the norm—dative

for A (rather than nominative) and nominative for O (rather than accusative).

- Cii. Where standard marking is ergative for A and absolutive for O (non-canonical markings are again shown in bold type):

Perceiver (A) for ATTENTION	Experiencer (A) for LIKING	Impression (O) for ATTENTION	Stimulus (O) for LIKING
locative	dative	absolutive	Avar
dative		absolutive	Chechen-Ingush
'affective'		absolutive	Archi
ergative		dative	Yawuru

Yawuru, an Australian language, has non-canonical dative marking (in place of absolutive) for the O argument of verbs from the two semantic types. The other three languages, all from the North-East Caucasian family, pursue a different strategy in using non-canonical marking for the A argument (in place of the expected ergative case). Avar has locative case on the A argument of ATTENTION verbs and dative for LIKING, while Chechen-Ingush has dative for both. Archi has a special case inflection, termed 'affective', used only for the A argument of ATTENTION and LIKING verbs.

The contrast between B and C is an important one. Under B, extended intransitive verbs are used, with Perceiver and Experiencer being in a syntactic function which has all the grammatical properties—in terms of the grammar of that language—of S, intransitive subject. Under C, Perceiver and Experiencer show the grammatical properties of an A argument, albeit that they may be accorded different case marking from a prototypical A.

The difference between 'listen' and 'hear' was mentioned under (ii) in §13.5.5. (These remarks are basically extendible to 'look' and 'see'.) Many languages have a single verb corresponding to both *hear* and *listen to* in English, and similarly for *see* and *look at*. A number of such languages employ two case frames—one with canonical and the other with non-canonical marking—to distinguish the senses. For example, Lezgian (a further language from the North-East Caucasian family) has verbs *akun* 'see, look at' and *van akun* 'hear, listen to'. Each occurs in two case frames:

	Perceiver (A)	Impression (O)	SENSE
CANONICAL	ergative	absolutive	'look at', 'listen to'
NON-CANONICAL	dative	absolutive	'see', 'hear'

Punjabi is similar in that the A argument is followed by dative marker *nüüü* for the sense ‘hear’ but not for ‘listen to’. Both languages draw attention to the non-volitional nature of the A for ‘hear’ and ‘see’ by employing dative marking in place of the canonical ergative.

13.7 Summary

Many factors weave together to create the transitivity profile for a language.

Each clause has a transitivity value—intransitive, transitive, plus extended transitive in many languages, plus extended intransitive in a few. For each of these, there are a number of core arguments, whose functions must be marked. It is rare to encounter S (intransitive subject), A (transitive subject), and O (transitive object) all marked differently; where this does occur, it is only in one part of the grammar. Most often, S is marked in the same way as A, a (nominative–)accusative system; alternatively, S may be marked like O, an (absolutive–)ergative system. In some languages, S is marked like A (Sa) for some verbs and like O (So) for others, a split-S system. A variant on this is where S may be marked like A or like O on the same verb, with a difference of meaning; this is a fluid-S system.

There is a semantic explanation for which semantic roles, for verbs of a given semantic type, are related to which of the syntactic arguments A, O, S, E (and also Sa and So). That argument whose referent could initiate or control an activity or state (if anything could) is placed in A function. For a simple transitive verb, the other core argument is in O function. If there are further candidates, then that whose referent is likely to be saliently affected by the activity will be in O function.

Each verb has one or more transitivity values, depending on what clause types it may occur in. This transitivity value relates to (i) the number of semantic roles associated with the verb; (ii) whether there is volitional control; (iii) whether or not the verb describes an action; (iv) whether one role is saliently affected.

Many languages combine accusative and ergative marking systems. Such splits are semantically motivated, depending on the referents of core arguments (according to a ‘nominal hierarchy’), to the tense and/or aspect of the clause, or to the syntactic status of clauses.

Syntactic function A or O may have canonical marking, which applies in most circumstances, or one or more kinds of non-canonical marking which is used (a) for the A or O of verbs from certain semantic types, or (b) when the referent of this argument has certain characteristics. For example, the A argument of a verb from the ATTENTION or LIKING types may be marked as dative, instead of the canonical nominative or ergative case.

13.8 What to investigate

When commencing work on a language, one must perforce look first at surface structure.

1. Examine the ways of indicating which core argument is in which syntactic function. This generally involves one or more of the following: (a) marking on an NP that provides whole or partial realization of a core argument (that is, a case system, which may be realized by inflections, clitics, or adpositions; see §5.4); (b) constituent order; (c) bound pronouns, generally attached within the predicate.

2. Work out the underlying grammatical relations in the language. Suppose the Perceiver role for an ATTENTION verb is marked by dative case but has the same grammatical properties as the Agent role of an AFFECT verb (which is marked by nominative case)—in terms of constraints on the formation of complex sentence types, of the operation of valency-changing derivations, and in exercising control within reflexive and/or reciprocal constructions. Then the Perceiver role can be taken to be in A function, like the Agent role, within a transitive clause type.

But if the Perceiver has similar grammatical properties to the Moving role of a MOTION verb (such as ‘go’, ‘return’, or ‘jump’), then it should be identified as in S function, presumably within an extended intransitive clause type.

3. Working steadily in this way, as every aspect of the grammar takes shape—each part interacting with each other part, and each part having existence and meaning with respect to the whole of which it is a segment—the various layers of transitivity (and their semantic motivations) will gradually become clear. That is, what the transitivity types of clauses are (§13.1), how core arguments are marked (§13.2 and §13.6), and what the transitivity classes of verbs are (§13.3). And, of course, the semantic bases for all grammatical distinctions, as outlined in §13.5. Throughout, the universal notion of semantic types, and the semantic roles associated with each, will provide a foundation for investigation of syntactic relations and of their contrastive meanings.

The process of analysis necessarily begins with examination of surface structure, and moves on to hypothesizing the underlying grammatical structures and relations in terms of which the language works. The final descriptive statement should be organized according to the principles set out in this chapter.

Appendix 1 Beyond ‘accusative’ and ‘ergative’

Terms ‘nominative’, ‘accusative’, ‘ergative’, and ‘absolutive’ are properly used to describe the marking of core arguments, as explained in §13.2. The terms are appropriately

extended to constraints on inter-clausal linking, described in §3.21. If a certain complex sentence construction requires that two clauses share an argument which is in S or A function in one (or both) of them, this is said to be an instance of ‘accusative syntax’; and if a constraint relates to S and O, it is said to constitute an instance of ‘ergative syntax’.

There are a number of totally different ways in which S and A may be associated within a grammar, and also many further kinds of association between S and O. For none of these is it appropriate to append labels ‘(nominative–)accusative’ or ‘(absolutive–)ergative’.

Further associations between S and A

- (a) In an imperative construction the most common—often, the only—referent for S (in an intransitive) or A (in a transitive imperative) is second person. Moreover, many languages allow the S or A argument of an imperative not to be explicitly stated when it is second person (or, when it is second person singular).
- (b) When a concept such as ‘can’, ‘try’, or ‘begin’ is realized by a lexical verb, it is likely to have the same subject (S or A) as the verb to which it is linked. (For example, in English, *John tried to run*, and *Mary began writing*.)
- (c) If a language has a reflexive construction which involves a reflexive pronoun or some other reflexive marker being placed in one argument slot, it will always go into O slot with a transitive verb and into a peripheral slot with an intransitive verb. (For example *John_A cut himself_O* and *Mary_S looked at herself in the mirror* in English.) That is, it is always the A or S argument which retains its normal form (rather than being replaced by a reflexive pronoun or marker).
- (d) As described in §13.3, many languages have ambitransitive verbs of type S = A. An example from English is *eat*, as in *John_S has eaten* and *John_A has eaten lunch_O*.

These associations between S and A apply equally to languages with accusative or ergative marking of core syntactic functions. (They relate to the fact that the topic around which a discourse is organized is in the great majority of instances human, and generally the controller of an activity.) It is not appropriate to describe them as ‘accusative’ features.

Further associations between S and O

- (a) Many languages have suppletive forms for a small number of the most common verbs, the choice of form depending on whether the S argument for an intransitive or the O argument for a transitive clause has singular or plural reference. For example, in Jarawara we find *-sona-* (singular S) and *-foro-* (plural S) for ‘fall’, and *-ibana-* (singular O), *-joka-* (plural O) for ‘roast’.
- (b) When there is an affix to the verb which specifies the number reference of a core argument, it typically relates to an S argument in an intransitive and an O argument in a transitive clause. For example, *-tu-* ‘all’ in Manambu.

- (c) Many languages derive nouns by compounding a verb with a core argument—this is always in underlying S or O function. Examples from English are *punch-ball* (from X_A *punches* [*the ball*]_O), and *hovercraft* (from [*The craft*]_S *hovers*).
- (d) If nominal incorporation relates to a core function, this is almost always S or O (and not A).
- (e) Verbal classifiers are morphemes which occur on the verb and characterize a core argument in terms of its shape, form, consistency, and other semantic properties (often to the exclusion of animacy and humanness). They typically categorize S and O (hardly ever A).
- (f) In most languages, a demonstrative with deictic (or pointing) effect can occur as any core argument. But there are languages with constraints on the functions in which a demonstrative may occur. In all the instances that have been reported, demonstratives are restricted to S and O functions; see (6) in §15.2.3.
- (g) As described in §13.5.5, many languages have ambitransitive verbs of type $S = O$. An example from English is *spill*, as in [*The water*]_S *spilled* and *John*_A *spilled* [*the water*]_O.

Associations (a)–(g) apply equally to languages with accusative or ergative marking of core syntactic functions; there are various explanations for them (see Aikhenvald and Dixon forthcoming). They should not be accorded the label ‘ergative’. It was mentioned in §13.2 that around one-quarter of the world’s languages show an ergative pattern for the marking of core arguments (with just a few of these also showing syntactic ergativity in terms of constraints on clause linking). But in fact just about every language shows some or all of (a)–(g). To accord such disparate associations the label ‘ergative’ would obscure the nature and function of ‘ergativity’ in its established sense, as described in §13.2.

Appendix 2 Confusing uses of terms ‘unaccusative’ and ‘unergative’

Understanding the multi-layered phenomenon of transitivity has in recent years been obscured by use of a particular type of terminology.

The terms ‘unaccusative’ and ‘unergative’ were introduced in Perlmutter (1978) and then spread through several brands of formal linguistics. These labels tend to be applied in a wide and vague way, to different sorts of phenomena in different languages. Their uses include:

- (a) If a language has consistent marking of S, and ambitransitive subclasses of verbs, then ambitransitives of type $S = A$ may be called ‘unergative’ and those of type $S = O$ ‘unaccusative’.
- (b) If a language has strict transitivity plus split-S marking, intransitive verbs of type S_a may be styled ‘unergative’ and those of type S_o ‘unaccusative’.
- (c) If a language has strict transitivity plus consistent marking of S, the label ‘unergative’ may be used of transitive verbs that are particularly open to

antipassivization, or of intransitive verbs that are particularly open to undergo an application derivation, and the label ‘unaccusative’ may be used of transitive verbs that are particularly open to passivization, or of intransitive verbs that are particularly likely to be used in a causative derivation. Or these labels may be used to relate to some quite different grammatical parameter. Or they may be used to classify verbs on an entirely semantic basis—those that are intuitively thought to be basically non-volitional are labelled ‘unaccusative’, etc.

‘Unaccusative’ and ‘unergative’ are vague labels used to describe different phenomena in different languages, without any explicit acknowledgement that this is being done. Linguists employing the labels tend to imply that certain verbs are inherently unaccusative and others inherently unergative. In fact a verbal meaning that belongs to the Sa class in one language may be in the S = O class in another; and so on.

We also find that some languages show more than one of the phenomena for which the labels ‘unaccusative’ and ‘unergative’ are employed. As mentioned in §13.4, Warekena combines split-S marking with ambitransitives of types S = A and S = O (in fact, many of type Sa = A and a few of types So = A and Sa = O). If one came across the label ‘unaccusative’ in a grammar of Warekena it would be hard to know whether it was referring to So intransitives or to S = O ambitransitives. Again, the properties that were listed under (c) may also apply for a language with an Sa/So split or for one with two varieties of ambitransitives. That is, the labels ‘unaccusative’ and ‘unergative’ are used for such a wide variety of phenomena as to be essentially imprecise and unclear.

The use of these labels, far from explaining anything, obscures certain key differences between languages. Their employment provides the false sense of a universal semantic basis for varied grammatical properties. They are best avoided.

Sources and notes

The discussion throughout this chapter builds on my previous publications on the topic (very occasionally, there is verbatim repetition from them), particularly Dixon (1989, 1994, 1999b).

Hopper and Thompson (1980) include much useful material; some of which I have used. Note, though, that their examples relate to several different levels of transitivity, and to non-canonical marking. And they do not describe the transitivity profile of the languages from which illustrations are taken.

13.1. Note that in recent years some formal theorists have used the label ‘adjuncts’ for peripheral arguments. Further discussion of extended intransitive clause types, and references to languages in which they are well attested, will be found in Dixon (1994: 122–4).

13.2. One language which shows tripartite marking in main clauses is Dhalanji, from Western Australia. At an earlier stage: (i) nouns showed an

absolutive–ergative pattern, with A indicated by ergative suffix *-ŋgu ~ -lu*, and S and O arguments being formally unmarked; (ii) pronouns showed a nominative–accusative system with O marked by accusative suffix *-nha*, while A and S were unmarked. Ergative suffix *-ŋgu ~ -lu* (for A function) and accusative suffix *-nha* (for O function) have now been generalized to apply to both nouns and pronouns, with just S function being left unmarked. There is one exception—the 1sg pronoun retains a single form for both A and S functions. (See Austin 1981b.) Further discussion of tripartite marking is in §13.5.4 and Dixon (1994: 34–41, 44–5).

Information on Mali from Stebbins (forthcoming; personal communication). Note that in Mali the O/So argument after the predicate is shown either by an NP or by a pronoun (not by both), whereas the A/Sa argument before the predicate can be shown by both an NP and a pronoun. Information on Bats is repeated from Dixon (1994: 79–80), being entirely based on Holisky (1987).

13.3. In some languages there is a further class of verbs, with just a handful of members, which take no core arguments at all (and could be said to have ‘zero valency’); they typically include weather verbs such as ‘rain’.

Data on Manambu from Aikhenvald (2008a), on Tariana from Aikhenvald (2003). The Jarawara data is from Dixon (2004a: 549, 82–3); Warekena from Aikhenvald (1998 and personal communication).

13.5.1. Fuller information on verbs in Yidiñ which may only have inanimate A arguments (plus example sentences) is in Dixon (1991b: 273–4). The concept of ‘liking’ is shown through a nominal root in two Australian languages: Martuthunira (Dench 1995: 208) and Burarra (Glasgow 1994: 358).

13.5.2. Fuller discussion and exemplification of the points made here for Fijian is in Dixon (1988a: 215–21). Note that Fijian double transitive suffixes sometimes have a simply semantic role such as ‘do lots of times’ or ‘do intensively’. Information on Jarawara is from Dixon (2004a: 550–7).

13.5.3. For discussion of the syntactic possibilities for GIVING verbs in English, see Dixon (2005: 119–24). The syntax of *wuga-l* ‘give’ in Dyirbal is described in Dixon (1972: 300). Information on Bantu languages is in Hyman and Duranti (1982), Gary and Keenan (1977), Duranti and Byarushengo (1977), and Dryer (1983). Comrie (2003) describes how some languages have two suppletive forms for the verb ‘give’, either (i) with one form for when the recipient is 1st or 2nd person and the other for when it is 3rd person, or (ii) with one form for 1st and another for 2nd or 3rd person recipient. Further discussion on ‘give’, ‘tell’, and ‘show’ cross-linguistically—plus source information for the languages mentioned here—is in Dixon (1989).

13.5.5. *Diingal* ‘laugh at’ in Guugu Yimidhirr is from Haviland (1979: 168, 178). A fuller discussion of the semantic basis for assignment to $S = A$ and $S = O$ subclasses in Fijian is in Dixon (1988a: 204–14), from which the account here is abstracted. A fuller discussion on the conditions for omitting an O NP in English will be found in Dixon (2005: 305–9).

13.6. Most of the information in this section comes from Onishi (2001a) and other chapters in Aikhenvald, Dixon, and Onishi (2001). Information concerning ‘follow’ in Icelandic is from John Maling (personal communication). Fuller information on the partitive case in Finnish is in Kiparsky (1998).

13.6.1. Information on Lezgian was supplied by David Kilby from Mejlanova (1960). That on other North-East Caucasian languages is based on, among other sources, Černý (1971), Paris (1985), Comrie (1981), Nichols (1982), and Simon Crisp (personal communication). Punjabi from Bhatia (1993: 170–1).

Appendix 1. This is a condensation of Aikhenvald and Dixon (forthcoming).

Appendix 2. In a 1995 book entitled *Unaccusativity: At the syntax–lexical semantics interface*, Levin and Rappaport Hovav state ‘the hypothesis [is] that the syntactic properties of verbs are determined by their meanings’. It is amazing that at this date such a fallacy should be revitalized. It is *never* possible to predict, with certainty, the syntactic status of a word in a given language from its meaning (if it were, all languages would have correspondingly iconic grammars). See the discussion in §1.8, illustrated for ‘hungry’, ‘mother’, ‘father’, and ‘two’.

A further, more esoteric, illustration, comes from Baniwa of Içana (an Arawak language spoken in Brazil). This is a split-S-type language; the verb *-hmanika* ‘play’ is of type Sa in the Hohôdene dialect but of type So in the Siuci dialect (Alexandra Aikhenvald, personal communication)—same meaning, different grammatical status.

Copula Clauses and Verbless Clauses

14.1 Introduction

Each language has intransitive and transitive clause types. There is frequently a further minor—but still important—type: copula clause. This has as predicate a copula verb, taking two core arguments, Copula Subject (CS) and Copula Complement (CC). The predicate in an intransitive or transitive clause has reference. A copula verb as predicate is different in that it has relational rather than referential meaning.

In each specific language, a copula construction marks a range of relations between CS and CC, depending in large part on the nature of the CC. These are illustrated for English in Table 14.1.

A copula will always cover relations A1, Identity, and/or A2, Attribution; often also A3, Possession, and A4, Benefaction. (In a number of languages these last two merge as a single relation.) In some languages the copula construction also covers A5, Location, but in others a stance verb must be employed (literally ‘The apple tree stands in the garden/outside/over there’).

TABLE 14.1. Outline of the semantic relations shown in copula constructions for English

NATURE OF CC		RELATION	CS	COPULA	CC
A1	NP or complement clause	Identity	<i>This man</i> <i>The basic idea</i>	<i>is</i> <i>was</i>	<i>a doctor</i> <i>that John</i> <i>should lead</i>
A2	Adjective	Attribution	<i>This man</i>	<i>is</i>	<i>clever</i>
A3	Possessive phrase	Possession	<i>This book</i>	<i>is</i>	<i>John’s</i>
A4	NP marked by appropriate adposition or affix	Benefaction	<i>This present</i>	<i>is</i>	<i>for John’s birthday</i>
A5	NP marked by appropriate adposition or affix, or locational adverb	Location	<i>The apple tree</i>	<i>is</i>	<i>in the garden/</i> <i>outside/over</i> <i>there</i>

A defining feature for a copula verb is that it *must* be able to occur in a construction with two core arguments, CS and CC. In some languages—not including English—there is an alternative construction for a Copula verb, with a single argument, CS. This was illustrated from Latin under (c) in §3.2, and can be repeated here:

NATURE OF CC		RELATION	CS	COPULA
B	none	Existence	<i>Deus est</i>	‘There is a god (lit: God is)’

Note that if a putative copula *always* occurs with just one core argument, CS—and not also with a CC—then it is not a copula verb at all, but a straight-forward intransitive verb. If a putative copula verb occurs just in relation A5, with an NP marked by a local case (and assuming that CS is marked in the same way as S) then it should be regarded as an intransitive verb with an oblique, local NP. Similarly for a putative copula which occurs only in relation A3 and A4, with an NP marked as genitive, etc.

That is, for a verb to be identified as a copula, it must occur with two core arguments, CS and CC, covering at least A1, Identity relation, and/or A2, Attribution relation.

A number of languages have a special existential marker ‘there is’; this should not be considered a type of copula. English uses *there is/are/was/were*; for example, *There is a solution*. As Matthews (1997: 77) puts it ‘this is an existential use of *be*’, to be distinguished from the Copula use of *be*.

Some languages lack a copula verb, but show some of the semantic relations illustrated in Table 14.1 simply by apposition. That is, the copula slot is left blank and we have ‘verbless clauses’. Like copula clauses, these have two core arguments, a Verbless Clause Subject (VCS) and a Verbless Clause Complement (VCC). The Australian language Yidiñ is of this type. Examples of verbless clauses include:

	VCS		VCC	
(1) A1, Identity	jugu	yiju	jundu	‘This tree is (just) a stump’
	tree	THIS	stump	
(2) A2, Attribution	mayi		mamba	‘The fruit is sour’
	fruit		sour	
(3) A3, Possession	yiju	guda:ga	waga:l-ni	‘This dog is (my) wife’s’
	THIS	dog	wife-GENITIVE	
(4) A4, Benefaction	mayi	miwur	ɲajin	‘The gathered fruit is for me’
	fruit	gathered	1SG:GENITIVE	

Note that both Possession and Benefaction involve genitive inflection of the VCC; they are distinguished on pragmatic grounds. In (4) adjective *miwur* ‘gathered’ modifies noun *mayi* ‘fruit’ within the CS. In context, the sentence means ‘the fruit has been gathered for me’. (Unlike copula clauses, verbless clauses do not—save in exceptional cases—mark tense.)

A verbless clause construction is unlikely to be used for relation A5, Location. A stance verb, such as ‘sit’, ‘stand’, or ‘lie’ will be employed instead, as it is in Yidiñ:

- (5) dundun_S jana-ŋ ŋun_{gu}
Java.cedar stand-PRESENT THERE
A Java cedar tree is there (lit. a Java cedar tree stands there)

There is one semantic relation, shown by copula clauses in some languages, which is seldom found in verbless clauses. In most languages, a clause cannot consist just of a VCS, indicating B, Existence. In place of this, an intransitive clause is likely to be used, involving a stance verb, or else a verb ‘exist’.

In some languages with no copula, verbless clauses show very limited possibilities. For example, in Fijian they are only used for A1, Identity. Attribution is shown by a verb functioning as the head of an intransitive predicate, while Possession, Benefaction, and Location all employ intransitive verbs of stance (Dixon 1988a: 128).

We can now, in Table 14.2, complete the chart of clause types (adding to that presented in §3.2). As mentioned before, predicates in intransitive and transitive clause types may always be filled by a verb of the appropriate transitivity class. In some languages, an intransitive predicate slot may alternatively be filled by an adjective, or a noun, or even an NP (of any kind).

Rows (b), (c), (e), and (f) each involve two core arguments. As described in §3.2 and §13.1, the core arguments in an extended intransitive are distinct from those in a plain transitive. In Tongan, for example, S is indicated by absolutive preposition ‘a (in both plain intransitive and extended intransitive),

TABLE 14.2. Clause types

CLAUSE TYPE		REFERENTIAL PREDICATE	CORE ARGUMENTS		
a	intransitive	intransitive	S		
b	extended intransitive	extended intransitive	S	E	
c	transitive	transitive	A	O	
d	extended transitive	extended transitive	A	O	E
RELATIONAL PREDICATE					
e	copula	copula verb	CS	CC	
f	verbless	<zero>	VCS	VCC	

A by ergative '*e*' (in both plain transitive and extended transitive), O again by absolutive '*a*' (also in both plain transitive and extended transitive), and E by dative '*ki*' (in both extended intransitive and extended transitive).

Although it has two core arguments, a copula clause construction is quite distinct from a transitive (and from an extended intransitive). Hindi and Sumerian use ergative for one core argument in a transitive clause (that in A function) but no argument in a copula clause receives ergative marking. Similarly, Russian uses accusative for one core argument (that in O function) in a transitive clause, but in this language no argument in a copula construction is marked as accusative.

In the great majority of languages, CS is marked like S, but not in all (see §14.3), showing that copula clauses should not be regarded as a special subtype of intransitives. And CC sometimes has different syntactic properties and/or different form from all other core syntactic functions.

For many languages which show a copula construction, the copula verb may be omitted in specifiable circumstances, producing what is effectively a verbless clause. In fact, copula clauses (in languages which have them) and verbless clauses (in languages which have these) are often remarkably similar, so that there may be a temptation to describe them as variants of the same construction type. But some languages have both, with distinct and established semantic effects (see §14.6). It is most prudent to treat copula clauses and verbless clauses as distinct construction types.

As mentioned in §3.2, all kinds of intransitive and transitive clauses may include optional peripheral arguments. Generally, these are not found in copula or verbless clauses.

§14.3 discusses the syntax of copula and verbless clauses, in particular the nature and marking of their core arguments. There is then, in §14.4, detailed consideration of types and subtypes of relational meanings that may be expressed. §14.5 looks at the form of copula verbs, and grammatical categories which they may carry. Their modes of occurrence, and when a copula can be omitted, are dealt with in §14.6. Then §14.7 considers where copulas may have evolved from and what they may develop into. But before all this, it will be useful to compare and contrast the use of an adjective or a noun (i) as head of an intransitive predicate, and (ii) as copula complement.

14.2 Contrasting functions of adjectives and nouns

As pointed out under (f) in §2.5, the term 'predicate' is applied in a variety of ways. Most linguists nowadays shun the sense 'everything in a clause but the subject'. However, there is still a tendency to say that in a copula

construction the predicate consists of copula verb plus what we call the Copula Complement. And there is a stronger tendency to say that in a verbless clause construction the Verbless Clause Complement *is* the predicate, a ‘nominal predicate’. (For us, ‘nominal predicate’ is used for a nominal functioning as head of the predicate in an intransitive construction, such as *kuphe* in (8a) below.)

It is important to make a clear distinction between the two elements of clause structure—predicate (which can have zero surface realization) and argument (which may also have zero surface realization). *A doctor* is just as much an argument in $[My\ son]_{CS}\ is_{\text{PREDICATE}}\ [a\ doctor]_{CC}$ as it is in $[A\ doctor]_S\ lives_{\text{PREDICATE}}\ next\ door$ or $We_A\ need_{\text{PREDICATE}}\ [a\ doctor]_O$. To say that *is a doctor* is the predicate of *My son is a doctor* can only lead to confusion.

Some languages allow a noun or adjective to function as head of an intransitive predicate, as in Fijian (repeating (2) from §12.3):

- (6) [E balavu]_{INTRANSITIVE.PREDICATE} [a tuuraga]_S
 3sgS tall ARTICLE chief
 The chief is tall (lit. The chief tall-s)

Other languages, including English, do not permit this, but have noun or adjective as CC in a copula construction. Thus, (6) would be translated into English as:

- (7) [The chief]_{COPULA.SUBJECT} [is]_{COPULA.PREDICATE}
[tall]_{COPULA.COMPLEMENT}

There are languages which combine both possibilities, as shown in the following sentences from Tariana:

- (8) (a) *čiāri(-ne)_S kuphe-pidana_{INTRANSITIVE.PREDICATE}*
 man(-FOCUS) fish-REMOTE.PAST.REPORTED
 A man was a fish
- (b) *čiāri(-ne)_{CS} kuphe_{CC}*
 man(-FOCUS) fish
di-dia-pidana_{COPULA.VERB}
 3sg.NON.FEM.CS-become-REMOTE.PAST.REPORTED
 A man became a fish

Both clauses refer to the kind of transmutation which can occur in a myth. In (8a) the noun *kuphe* ‘fish’ is head of the intransitive predicate, and takes a tense-evidentiality suffix (just as a verb would do in this slot). In (8b) *kuphe* is the copula complement, an argument outside the predicate of the clause. The predicate is here copula verb *-dia-* ‘become’, and it is this which carries the tense-evidentiality suffix.

The possibilities for case marking on NP arguments in Tariana are:

A, S, CS	focus subject marker - <i>ne</i> (optional)	—
O, non-core arguments	—	topical non-subject marker - <i>nuku</i>
CC	—	—

That is, both S in the intransitive clause (8a) and CS in the copula clause (8b) may take suffix *-ne*, if that NP is in focus. The NP *kuphe* in (8b) is in CC function and may take neither suffix *-ne* nor suffix *-nuku*. Note that it is not possible to treat (8b) as a type of extended intransitive clause, with *čĩārĩ* ‘man’ as S argument and *kuphe* as an oblique argument; if this were a valid analysis then *kuphe* should be able to take topical non-subject marker *-nuku*, which in fact it cannot do.

In Fijian, an NP functioning as head of an intransitive predicate can take all the tense, aspect, and other modifiers available for a verb in this slot. In Tariana, a nominal as head of an intransitive predicate takes tense-evidentiality, mood, aspect, and most other suffixes that would be available for a verb in the slot. Different types of clause nuclei have varying properties with respect to prefixes in Tariana. In brief, pronominal prefixes are used with transitive and with active intransitive (Sa) verbs and with the copula verb *-dia* ‘become’, but not with stative intransitive (So) verbs, nor with the copula verb *alia* ‘be’, nor with a nominal as head of an intransitive predicate.

Suppose that there was a language like Tariana but with the additional property that a copula verb may optionally be omitted. There would still be a clear distinction between a clause with a noun as head of the intransitive predicate, as in (8a), and a copula clause with the copula omitted, such as (8b) without *di-dia-pidana*. In the first clause the noun *kuphe* ‘fish’ takes a fair selection of the affixes available to a verb as predicate head; in the second example *kuphe* takes none of these (in fact, as a CC in Tariana, it cannot take any affixes).

14.3 Syntax

When there is a fixed or preferred order for the constituents within a transitive or intransitive clause, a requirement for fixed order generally carries over into copula and verbless clauses. For example, NPs in A, S, and CS function precede the verb in English, while O and CC follow it.

Northern Subanen (Austronesian, Philippines) has canonical constituent orders VS and VAO. In verbless clauses, VCC generally precedes VCS. That is, CS is like S and O in occurring as the last core argument. CS is also like S and O in being marked by the absolutive preposition *su* (an argument in A function is marked by ergative preposition *nə*). The VCC receives no marking. For example (Daguman 2004: 137):

- (9) $\eta miinit_{VCC}$ [*su* $kpədəs$] $_{VCS}$
 hot ABSOLUTIVE sun
 The sun is hot

In languages where the order of constituents in a transitive or intransitive clause has a degree of freedom, it is not unusual for there to be fixed ordering within a copula construction. In Jarawara, for instance, the predicate comes last and NPs in A and O function may occur in either order before it. In a copula clause the CC constituent must occur immediately before the verb (save in the Naming subtype of Identity, when CC follows the predicate—see §14.4). In each instance, CS is clause-initial.

In Dolakha Newar (Tibeto-Burman, Nepal; Genetti 2007: 275) there is considerable freedom in the order in which A and O NPs may occur before a transitive predicate, but in a copula construction the CC NP directly precedes the copula verb. Cavineña (Tacanan, Bolivia; Guillaume 2008: 91–7) has a fairly free constituent order in transitive and intransitive clauses, but a CC almost always precedes the copula predicate.

In many languages, valency-changing derivations apply to non-copula verbs. Causative and/or applicative processes may derive a transitive from an intransitive, while passive and/or antipassive may derive an intransitive from a transitive. In the great majority of cases, copula verbs do *not* enter into such derivational processes. (If this does occur, it is likely to concern a copula with the meaning ‘become’, rather than just ‘be’.) The Tibeto-Burman language Qiang is unusual in that it can form a causative on its copula verb $\eta uə$ ‘be’. Example (10) shows $\eta uə$ in a straightforward copula clause, and (11) has it with causative suffix *-z* (LaPolla and Huang 2002: 2; LaPolla 2003: 127):

- (10) $the:_{CS}$ mi_{CC} $\eta uə$
 3sg man COPULA
 He is a man
- (11) $khuə-le:$ $puñu$ $hə-\eta uə-z$
 dog-DEFINITE cat DIRECTIONAL-COPULA-CAUSATIVE
 Treat the dog as a cat (lit. Make dog be cat)

In some languages, copula clauses may trigger a special allomorph for a grammatical category. Valenzuela (2003: 376, and personal communication)

reports that in Shipibo-Konibo (Panoan family, Peru) a polar question involves the clitic *-ki* added to the first major constituent of the clause, except in a copula clause where clitic *-ri(n)* is used (this is its only function in the grammar).

We can now look in turn at the properties of the two core arguments in copula and verbless clauses.

(a) Copula subject (CS) and verbless clause subject (VCS)

The CS and VCS slots generally have exactly the same structural possibilities as S, A, and O—they can be filled by an NP, or by a complement clause (in languages which have these), and so on. When a language has bound pronouns (generally attached to the predicate) as partial realization of S and O, or of S and A (or of all three), there is invariably a bound pronoun for CS.

If one member of a case system is functionally unmarked, it is generally this that is assigned to S, and also to CS and VCS. In all absolutive–ergative systems, absolutive is functionally unmarked (being used for citation, etc.) and is used with CS or VCS as well as with S and O. (It was exemplified in (9) for Northern Subanen.). In the great majority of nominative–accusative systems, nominative case is functionally unmarked, being used for S and A, and also for CS and VCS. In a language with split-S, CS and VCS are generally marked like So; this is as would be expected, since the referent of CS and VCS generally does not exhibit volition.

There are just a few languages which mark CS or VCS in a different way.

- (i) Ainu (isolate, Japan) has a system of bound pronouns in which S and A functions sometimes have the same form (for example, for 1sg), and sometimes different forms (as for 1pl). Where they differ, CS is the same as A, different from S. Compare Tamura (2000: 50–1):

with intransitive subject (S)	ku-mina	‘I laugh’	mína-as(-pa)	‘We laugh’
with copula subject (CS)	ku-ne	‘I am’	ci-ne(-pa)	‘We are’
with transitive subject (A) (O is here 3sg)	ku-nukar	‘I see him/ her/it’	ci-nukar(-pa)	‘We see him/ her/it’

It will be seen that while prefix *ku-* is used for 1sg in S, CS, and A functions, with 1pl we have prefix *ci-* for CS and for A, but suffix *-as* for S.

Note that while CS is marked like A in Ainu, CC is not treated like O. A transitive verb takes pronominal prefixes marking A and O

arguments (the O prefix is zero for 3rd person), but a copula verb simply has one pronominal prefix, for CS.

- (ii) There are different possibilities for VCS in the Australian language Djingulu (Pensalfini 1997: 185–7) depending on the nature of the VCC. If the VCC is an NP in an Identity relation, then—as in (12)—VCS is in ergative case, like A (reminiscent of CS marking in Ainu). And if the VCC is an adjective in Attribution relation then—as in (13)—VCS is in absolutive case (with zero marking), like S. In each instance, VCC is in absolutive case.

(12) njamina-ni_{VCS} wamalagardirni- \emptyset _{VCC}
 THAT:FEMININE-ERGATIVE virgin-ABSOLUTIVE
 She's a virgin

(13) [njima- \emptyset babirdimi- \emptyset]_{VCS}
 THAT:VEGETABLE-ABSOLUTIVE yam-ABSOLUTIVE
 kiyaljiyanu- \emptyset _{VCC}
 rotten-ABSOLUTIVE
 That yam is rotten

- (iii) In some languages the CS or VCS in a positive clause is marked like S but in a negative copula clause it takes some quite different marking. For instance, within a negative clause, the CS for some relations takes partitive case in Finnish and genitive case in Russian (these cases are also used to mark a type of object, in the two languages).
- (iv) There are a few languages in which nominative (for S and A functions) is the formally and functionally marked case, while accusative (for O function) is unmarked (see Dixon 1994: 63–7). There is fair variation concerning the ways in which CS/VCS and CC/VCC are marked in such 'marked nominative' languages. For example:

(a)	CS: accusative	CC: accusative	Kabyle (Berber branch of Afro-Asiatic; Vincennes and Dallet 1960: 99)
(b)	CS: accusative	CC: nominative	Mojave (Yuman language; Munro 1976: 269–70)
(c)	CS: nominative VCS: accusative	CC: accusative VCC: accusative	Tennet (Surmic family; Randal 1998: 233)
(d)	CS/VCS: nominative	CC/VCC: accusative	Oromo (Cushitic branch of Afro-Asiatic; Owens 1985)

For all these languages, accusative (used for O) is functionally unmarked, being used as the citation form, and also formally unmarked, having zero realization. It will be seen that in Kabyle both CS and CC are zero-marked (like O, and unlike S and A). In Mojave, CC bears the non-zero nominative affix (like S and A). In Tannet, there is a copula construction in which CS is marked by nominative when CC has non-specific reference, and a verbless clause construction in which VCS is left unmarked (accusative) when the referent of VCC is specific; see §14.4. Oromo employs a copula clause for past and future tenses but a verbless clause in present tense. CS/VCS is in nominative case, like A, while CC/VCC is accusative, like O. However, we can hardly say that this is an instance of a transitive construction since in present tense there is no verb.

Other languages mark CS in the same way as S, even when there are a number of alternatives for S. For example, in Mingrelian (from the Kartvelian family), both S and CS take nominative case in three of the tense series, and both take narrative case in the fourth series (Harris 1991: 375–6).

In languages where CS or VCS is marked like S, it generally has the same grammatical properties as S in terms of constraints on clause linking, etc. (Information is not available concerning the ‘subject properties’ of CS in languages which employ non-standard marking for CS, such as Ainu, Djingulu, and Mojave.)

(b) Copula complement (CC) and verbless clause complement (VCC)

The possibilities for what can be in CC or VCC slot depend on the range of relations expressed, as set out in Table 14.1. The defining criterion for a copula construction is that it should cover one or both of A₁, Identity—for which CC/VCC is an NP (or complement clause, etc.)—and A₂, Attribution—for which CC/VCC is an adjective. The CC slot differs from A, S, O, and CS in English (and in many other languages) in that it can consist just of an adjective (as in *John is tall*). For the other slots an adjective may function only as modifier to a head noun within an NP (for example, *the tall man*, *a tall one*).

In some languages, quite complex constructions may function as CC/VCC. An example from Spanish is (Travis 2002: 8):

- (14) Pero Jaime_{CS}, es [[el que]_{CS} está
 BUT NAME be:3sg:PRES ART:MASC WHO be:3sg:PRES
 [mal de carrito]_{CC}]_{CC}
 bad OF car.DIM
 But Jaime is the one who is badly off for a car

Spanish has two Copula verbs, *ser* and *estar* (their meanings are discussed in §14.4.1). The predicate of the main clause in (14) is a form of *ser*. Its CC involves a relative clause which is itself a copula construction using *estar* as predicate.

In most languages a copula or verbless clause is used for both Identity and Attribution. But in some it shows only one of these functions. §12.6 drew attention to languages in which an adjective may not function as CC. In Yoruba, one cannot say ‘Olu is good’, only ‘Olu is a good girl’. That is, the CC must be an NP, involving a head noun, which can be modified by an adjective; the copula construction indicates Identity, not Attribution. For some languages which allow an adjective to be head of an intransitive predicate, this is the only way of stating that something has a certain property (see §12.3). In such cases an adjective cannot be CC within a copula construction although an NP can be, for the Identity relation. (The Chadic language Mupun provides an example of this; see Frajzyngier 1993: 54, 252–5.)

When a copula or verbless construction covers some or all of A₃, Possession, A₄, Benefaction, and A₅, Location, then its CC/VCC will be marked as possessive, or by some appropriate adposition, affix, or etc.

One interesting property of copula and verbless clauses is that the CC/VCC is seldom (possibly never) marked by a bound pronominal attached to the verb. Even in a language such as Yimas (Papuan area; Foley 1991: 193–226), where A, S, O, and indirect object are marked on the verb, in a copula clause only CS—not CC—is included in the system of bound markers. The same applies for Manambu, another Papuan language (Aikhenvald 2008a).

A CC/VCC typically receives no explicit case or adpositional marking. When one term in the marking system has zero realization—absolute in an absolute–ergative system and nominative in a nominative–accusative one—then CC/VCC can be identified with this term, which is typically also the marking on CS/VCS. But sometimes CC/VCC does take no marking at all (rather than, say, a term from a case system which has zero realization). This is the situation in Japanese where a CC may not be followed by any postposition, unlike all other core and peripheral NPs. And, as mentioned in §14.2, in Tariana a CC is the only kind of NP which cannot take any case marker.

For the Australian language Diyari, the CC normally takes zero marking (absolute case). However, Austin (1981a: 104–5) states that, if the CC is one of a set of nominals referring to ‘more or less temporary mental or physiological states’, then it takes ergative-instrumental case marking, as in:

- (15) nganhi_{CS} mawa-li_{CC} ngana-yi
 1sg hunger-ERGATIVE.INSTRUMENTAL be-PRESENT
 I am hungry

The other forms selecting ergative-instrumental include ‘sleep’, ‘fear’, ‘danger’, ‘sadness’, ‘jealousy’, ‘strength’, and ‘cold’.

In Jarawara (Arawá family, Brazil; Dixon 2004a: 378) the pronominal paradigm for CC is different from that for any other core argument. For example:

	A, S, CS	O	CC
1sg	o-	owa	owa
2pl	tee	tera	tee

For singular pronouns CC has the same form as O, and for plural pronouns it has the same form as A/S/CS.

Another language in which CC has different form from other NP arguments is Zayse, from the Cushitic branch of Afro-Asiatic (Hayward 1990: 266). A sample paradigm of pronouns is:

	A, S, CS	O	CC	with postposition
1sg	tá[j]	tána	tánte	táa(-ro)
2sg	né[j]	néna	nénte	née(-ro)

We find further possibilities in individual languages and linguistic regions. For example, in a number of East European languages the CC in an Identity copula clause can be marked either by nominative case or by an oblique case, with a difference in meaning. Nominative indicates a permanent relation—for example, ‘he is a cleaner’ (as a profession)—whereas oblique marking refers to something that is temporary—‘he is a cleaner’ (as a fill-in job). The oblique case involved is *essive* in Estonian and *instrumental* in Russian. (Also see Timberlake 2004: 286–7; and Comrie 1997: 40 on Polish.)

14.4 Relational meanings

There are a number of subtypes of the Identity relation. These can be demonstrated by recognizing three referential possibilities for both CS/VCS and CC/VCC. They can be referred to as R, D and G:

- R—Specific referent, shown by a proper name, demonstrative, or pronoun.
- D—Specific description; that is, a description which specifies a particular person or thing. For example ‘the doctor at the hospital’ (when there is only one), ‘my father’.
- G—General description, which does not have a unique referent. For example ‘a doctor at the hospital’ (when there is more than one), ‘a member of the golf club’.

In English, each of R, D, and G can function as CS and as CC, as illustrated in Table 14.3. In relation to row (i), note that Michael Innes was the *nom de plume* assumed by J. I. M. Stewart when he wrote detective stories.

TABLE 14.3. Examples of types of Identity relation in English

	CS	CC	
(i)	R	R	Michael Innes is J. I. M. Stewart J. I. M. Stewart is Michael Innes
(ii)	D	D	My father is the doctor at the hospital The doctor at the hospital is my father
(iii)	G	G	A doctor at the hospital is a member of the golf club A member of the golf club is a doctor at the hospital
(iv)	R	D	John Smith is the doctor at the hospital
	D	R	The doctor at the hospital is John Smith
(v)	D	G	The doctor at the hospital is a member of the golf club
	G	D	A member of the golf club is the doctor at the hospital
(vi)	R	G	John Smith is a doctor at the hospital

It will be seen that in (i)–(v), CS and CC can be reversed, still producing an acceptable sentence (albeit with a shift of meaning in some cases). However, this does not apply for (vi), with R as CS and G as CC. The reverse of this, **A doctor at the hospital is John Smith*, is not acceptable. That is, a copula clause in English may not have a General description as CS and a Specific referent shown as CC.

There are further possibilities for the arguments of an Identity copula construction in English (as in many other languages) such as direct speech and a complement clause. For example (all arguments in these examples are D, specific descriptions):

- (16) (a) [What he said]_{CS} was ['Let's go']_{CC}
 (b) ['Let's go']_{CS} was [what he said]_{CC}
- (17) (a) [My idea]_{CS} was [that we should sell]_{CC}
 (b) [That we should sell]_{CS} was [my idea]_{CC}

Copula constructions showing Attribution may not—in most languages—have their core arguments interchanged, simply because an adjective cannot function as CS. Reversal is also not possible for Benefaction. Taking the illustrative sentences from Table 14.1, it is not permissible to say **Clever is this man* or **For John's birthday is this present*. Reversal is possible for clauses showing the Possession relation. For example, one could say either [*This Honda*]_{CS} is [*John's*]_{CC} or [*John's*]_{CS} is [*this Honda*]_{CC} (say, in answer to a question *Which is John's car?*). This suggests that Possession could perhaps be regarded as a special type of the Identity relation.

Reversal is also not possible for the Location relation. In place of *The apple tree is in the garden/outside/over there* one could scarcely say **In the garden/outside/over there is the apple tree*. However, it is possible to have a locational description as one argument in an Identity clause, and then reversal is quite acceptable. For example:

- (18) (a) [A good place to hide]_{CS} is [under the bed]_{CC}
 (b) [Under the bed]_{CS} is [a good place to hide]_{CC}

A good place to hide is here a General description (G), while *under the bed* is a Specific description (D).

Examples of what is here called the ‘Identity relation’ have received a variety of names in the literature, including ‘equative’, ‘equation(al)’, ‘descriptive’, ‘specification(al)’, ‘characterization(al)’, ‘identification(al)’, and ‘predication’. One could employ ‘equation’ for (i), where both arguments indicate a specific referent (R), or for (i), (ii), and (iv) where each of CS and CC is either R or D. But if multiple labels were to be used, there should surely be separate names for each of (i)–(vi). I prefer just to group them together under ‘Identity’.

One interesting question concerns what the basis is for placing one NP in CS and the other in CC function in (i)–(v) and in (16–18). The principle is basically pragmatic, relating to what is topic within a section of discourse, what has a known referent, and what is new, and so on.

Some languages have distinct construction types depending on the reference of CS or of CC. The Surmic language Tannet (Randal 1998: 233–4) is of the ‘marked nominative’ type, included in the tabulation in §14.3. If the CS is a general description (G), then a copula clause is likely to be used, as in:

- (19) *k-eéní anná_{CS} mót-tóh-t_{CC}*
 1sg-be 1sg:NOMINATIVE be.angry-AGENT.NOMINALIZER-SINGULAR
 I am a brave man (not bragging)

But if one wanted to say ‘I am the brave man’, with a specific description (D), then a verbless clause construction should be employed, as in:

- (20) *anet_{VCS} mót-tóh-t_{VCC}*
 1sg:ACCUSATIVE be.angry-AGENT.NOMINALIZER-SINGULAR
 I am the brave man (self-praise)

Note that in (19) CS is in the marked nominative case (also used for S and A), whereas VCS in (20), and CC/VCC in both sentences, are left unmarked (this is accusative case, used for O function and in citation).

The possibilities for CS may also vary depending on the kind of relation expressed. In English, for instance, a demonstrative (*this/these, that/those*)

which has animate reference can be the full CS NP only in an Identity copula clause, such as *This is my father*. If a demonstrative with animate reference is to be in CS function within an Attributive copula construction, it can only be modifier to the head noun of an NP, as in *This girl is beautiful*. At the least, there must be a dummy head, such as *one*. A speaker of English will say *This one is beautiful*, not **This* [animate reference] *is beautiful*.

In most languages with a copula construction, stating the name of a person or place is a straightforward instance of the Identity relation. The CS is ‘X’s name’ with the CC being the actual name. For example, in English:

(21a) [My father’s name]_{CS} is William_{CC}

In English—but not in every language—the CS and CC may be interchanged, in an appropriate pragmatic context:

(21b) William_{CS} is [my father’s name]_{CC}

However, in some languages naming is treated a little differently. Copula constructions in Jarawara generally have CS then CC then the copula verb. One day whilst in the field, I picked up a moth and put it outside the door. Someone commented:

(22) Jobeto_{CS} kiso_{CC} ama-ka
 name(m) capuchin.monkey(f) be-DECLARATIVE:m
 Jobeto is (sc. similar to) a capuchin (a species of monkey that
 eats moths)

But in a naming clause the CC (the proper name) must follow the copula, as in:

(23) [otaa taboro ino]_{CS} ama-ke Kasanofa_{CC}
 1exc:POSSESSOR village+m name+m be-DECLARATIVE:f name
 The name of our village is Casa Nova

Manambu does have a number of copula verbs (see §14.4.1), but prefers to employ a verbless clause for naming (Aikhenvald 2008a):

(24) [wun-a sə]_{VCS} Walinəm_{VCC}
 1sg-fem:sg name Walinum
 My name is Walinum

Some languages seldom use a copula clause for specifying a name. In Mandarin Chinese, for example, there is a copula construction but the preferred construction for naming involves the verb *jiào* ‘be called’, as in (Huang 2002: 5):

- (25) tā jiào Lǐsì
 3sg be.called Lisi
 Her name is Lisi (lit. she is called Lisi)

Many languages use a copula clause to express Possession. When there is a verb ‘have’, the two constructions will be used in complementary circumstances, as in English *I have a red Honda* and *That red Honda is mine*. However, many languages lack a verb ‘have’, using a copula or verbless clause construction of Possession in its stead. An example from Tamil (Asher 1985: 91) is:

- (26) avarukku_{CC} [neraya paṇam]_{CS} irukkutu
 he:DATIVE plentiful money be:PRESENT:3sg:NEUTER
 He has a lot of money (lit. plentiful money is to him)

(There is discussion of this construction type, across a variety of languages, in Benveniste 1971b. See also §16.9.2 below.)

In a language where a copula can occur with just a CS (in addition to the regular construction with both CS and CC), this may always indicate Existence, as exemplified in §14.1. In some languages it has wider uses. For example, one chilly morning I heard a Jarawara man say:

- (27) sire_{CS} ama-ke
 coldness(f) be-DECLARATIVE:f
 It is cold (lit. Coldness is)

A further alternative in Jarawara is for a nominalized clause to function as CS in monovalent use of copula *ama* ‘be’. This meaning is then ‘happens’, as in:

- (28) [ee to-ko-ma na-bi]
 1incS AWAY-IN.MOTION-BACK AUX-ALL.NIGHT:NOMINALIZER:fem
 -ro_{CS} ama-ke
 -RPef be-DECLARATIVE:fem
 We were travelling all night (lit. Our travelling all night happened)

The copula *ama* ‘be’ in Jarawara has restricted morphological possibilities; it may not take a tense/evidentiality suffix. What eventuates is that the appropriate suffix (here recent past eyewitness feminine, RPef) is added to the NP in CS function.

The copula *hona* ‘be’ in Hindi may also be used in a monovalent copula construction with the meaning ‘happen’; for example, ‘Nowadays many accidents happen’ is, literally, ‘Nowadays [many accidents]_{CS} are’ (Kachru 1968: 45–6).

14.4.1 Multiple copulas

Some languages have more than one Copula verb. Most commonly, one will just refer to a state and the other to coming into a state, similar to *be* and *become* in English. For example:

(29) [My son]_{CS} is [a doctor]_{CC}

(30) [My son]_{CS} became [a doctor]_{CC} (by passing his exams)

A copula ‘be’ is most often used for general statements (such as ‘John is fat’) whereas ‘become’—referring to change of state—is more likely to have temporal reference, as in ‘John has become fat’ or ‘John will become fat (if he doesn’t watch his diet)’. Dolakha Newar (Genetti 2007: 275–83) has two copulas—*khayŋ* ‘be’, which does not inflect for tense (or for person and number of CS), and *jur-* ‘be, become’, which does. We find that *khayŋ* can only be used for general statements such as:

(31) [u nis-mā]_{CS} [chana dāju]_{CC} khayŋ
 THIS two-CLASSIFIER 2sg:GENITIVE elder.brother be
 These two are your elder brothers

Copula *jar-* is typically used with the sense ‘becoming’, as in:

(32) āle kaimu_{CS} subbā_{CC} jur-a
 THEN husband official become-3sgCS:PAST
 Then my husband became an official

It must also be used—rather than *khayŋ*—for a statement of identity in the past. For example:

(33) optecā_{CC} ju [ān tākku]
 small be:3sgCS:PAST.ANTERIOR THAT time
 He was small (i.e. a child) at that time

As is well known, Spanish has two copulas. Basically, *ser* refers to a characteristic feature (a permanent state) whereas *estar* refers to a temporary state. Compare Travis (2002):

(34) (a) [La cáscara]_{CS} es verde_{CC}
 ARTICLE:sg:fem skin be:3sgCS:PRESENT green
 The skin is (*ser*) green
 (b) Está verde_{CC}
 be:3sgCS:PRESENT green
 It is (*estar*) unripe

Sentence (34a), with copula *ser*, indicates that the skin has a permanent green colour, but (34b), with *estar*, indicates that a fruit is at present ‘green’ on its way to becoming ripe.

Both copulas may be used for Attribution, with *ser* being preferred for Identity (including naming), Possession, and Benefaction. *Ser* and *estar* have different meanings in Location copula clauses, *ser* being used to identify the location of places and events while *estar* describes the location of physical objects and people. Compare (Butt and Benjamin 2004: 423):

- (35) (a) Dónde_{CC} es [la conferencia]_{CS}
 WHERE be:3sgCS:PRESENT ARTICLE:sg:fem lecture
 Where is (*ser*) the lecture being held?
- (b) Dónde_{CC} está [la conferencia]_{CS}
 WHERE be:3sgCS:PRESENT ARTICLE:sg:fem lecture
 Where are (*estar*) the lecture notes?

These two clauses have the same CC, *dónde* ‘where’, and the same CS, *la conferencia*. But the nature of the copula (*ser*) in (35a) leads to the reference of *la conferencia* taken as the lecture (an event) while the nature of the copula (*estar*) in (35b) leads to *la conferencia* being here interpreted as the lecture notes (physical objects).

Ser and *estar* behave similarly in Portuguese. We can give an example of how copulas may be switched for comical effect. If Paulo is a minister in the government one would usually say:

- (36a) Paulo_{CS} é ministro_{CC}
 Paul be:3sgCS:PRESENT minister
 Paul is (*ser*) a minister

During a period of political upheaval in Brazil, when ministers seemed to succeed each other at bewildering speed, people would jokingly say:

- (36b) Paulo_{CS} está ministro_{CC}
 Paul be:3sgCS:PRESENT minister
 Paul is (*estar*) a minister (that is: here today, gone tomorrow)

In Sranan (a creole from Surinam) there are two copulas, *de* and *na*; compare *Mi de botoman* ‘I am a boatmen (expressing current occupation)’ with *Mi na botoman* is ‘I am a boatman (expressing general qualifications, capability)’ (Faverey, Johns, and Wouk 1976). The Papuan language Manambu can use a noun or adjective as head of an intransitive predicate; it also has verbless clauses—exemplified at (24)—and it has a goodly number of copula verbs (Aikhenvald 2008a: chapter 4). Some are restricted to use with a particular

type of CC; for example, Copula verb *say-* requires as CC a noun referring to one of a number of mental and physical states, as in:

- (37) wap_{CC} say-na-wun
 shame be-ACTION.FOCUS-1sgCS
 I feel ashamed

Copula verb *yasa-* occurs with a CC which is a noun referring to hunger, thirst, sleep, etc.; copula *tay-* is used with the noun ‘coldness’ as CC; and so on. The default choice for a copula is *tə-*, which typically means ‘become’, as in:

- (38) kwakuli_{CC} tə-bər
 orphan become-3duCS
 They two became orphans

This contrasts semantically with copula *kur-* ‘become fully, reach, arrive at being something’, exemplified in:

- (39) du-a-ñan_{CC} ata kurə-d
 man-LINKER-child THEN become.fully-3sgCS
 He then reached the stage of being a boy

14.5 Forms

In isolating languages the copula is likely to have invariable form; this is as would be expected. But we also find that in some languages where transitive and intransitive verbs have a number of morphological processes available to them, the copula (or one copula) has invariable form. This was exemplified, in §14.4.1, for Dolakha Newar where copula *khayŋ* is invariable, while copula *jar-* inflects like other verbs. Describing the Nilotic language Nuer (Ethiopia and Sudan), Crazzolara (1933: 90) states that the copula à ‘be’ is ‘a defective verb, it is invariable for all persons and tenses; it has but this one form.’

In some languages a copula does have exactly the same morphological possibilities as other verbs. Tamil is an example of this. But in many languages there is a difference concerning the grammatical categories which may be associated with a copula and/or the way these are realized.

Some languages have a copula which shows more grammatical distinctions than other verbs. For instance, *be* in English has distinct forms depending on the person of CS—*I am, you/we/they are, he/she/it is*. This also applies in Hindi (Kachru 1968: 41). But in many languages the copula shows fewer forms than other verbs. In Modern Greek, for instance, it makes no aspectual distinctions

(Joseph and Philippaki-Warbuton 1987: 196). Describing Urarina (isolate, Peru), Olawsky (2006: 391) identifies more than a dozen types of verbal suffix. All are attested with the Copula verb *nia* ‘be’ save two—completive aspect and causative. For Motuna (Buin family, Papua New Guinea), the copula *tu(h)* ‘be’ occurs with most verbal categories, lacking just near past, perfect, and present progressive (Onishi 1994: 436–41).

The copula is frequently irregular in its forms. Indeed, Foley (1991: 226) states that in Yimas ‘the copula is the only truly irregular verb, and it is highly so’. Of the two copulas in Spanish, *estar* is mildly and *ser* wildly irregular. There are suppletive stems of the copula in a number of languages, including Kurukh, from the Dravidian family (Vesper 1968) and Mundari, from the Munda branch of Austro-Asiatic (Langendoen 1967).

Perhaps the most intriguing form for a copula is found in Kana (Benue-Congo family, Nigeria; Ikoro 1996a: 214–16). It basically consists of a low tone, which attaches to the CS. For example, the citation form of ‘child’ is *ɲwíí* (where ‘ indicates a high tone). When used as CS, this noun becomes *ɲwîí*, with low tone, ‘, replacing the high tone on the final vowel, as in:

- (40) *ɲwîí*_{CS+COP} *tám-bàrìcc*
 child work-god
 Children are works of God

If the CS ends in a vowel with low tone, the vowel is repeated in CS function; for instance *gbò* ‘friend’ becomes *gbòð* as CS. If the CS already ends in a VV sequence with low tones, the Copula low tone becomes invisible; ‘fear’, with citation form *bɔ̀ð*, remains unchanged when placed in CS function.

14.5.1 Negative copulas

In most languages, copula clauses are negated in the same way as transitives and intransitives. However, a number of languages have a distinct negative copula. In Koromfe (Gur family, Burkina Faso; Rennison 1997: 61), the positive copula has form *la* and the negative one has form *dɔ*. Awa Pit (Barbacoan, Colombia; Curnow 2002) has copula *i* ‘be’ used everywhere save in negative and polar interrogative clauses, where the form *ki* must be employed.

There are languages in which some of the distinctions shown through a copula in a positive clause are neutralized under negation. For Anywa (Nilotic, Sudan and Ethiopia), Reh (1996: 302–8) reports that there are two copulas, *ā* and *béɛ(y)*. She explains that *ā* is used to answer a question ‘what kind of thing is it’, as in:

- (41) āan=ā cōɔr_{CC}
 1sgCS=be blind.person
 I am blind

And *bée(y)* will be used to answer a question ‘which one’, as in:

- (42) pòl_{CS} bée [nɪlāl mən-ā-kwál gwél]_{CC}
 Paul be child RELATIVE-PAST-steal money
 Paul is the child who stole the money

The interesting feature is that, in a negative clause, *bée(y)* is replaced by *páa* and *ā* by either *páa* or *páthá*. That is, *páa* is used as negative correspondent for both copulas from positive constructions.

There are also examples in the opposite direction, when a negative copula exhibits more grammatical distinctions than its positive counterpart. For example, the copula *alia* ‘be’ is used in positive clauses in Tariana to cover all of semantic relations A1–A5 and B, described in §14.1. However, negative copula constructions are of two kinds (Aikhenvald 2003: 488–97):

- negative copula *sede* replaces *alia* for A3, Possession, A5, Location, and B, Existence.
- the regular negative suffix *-kade* is added to *alia* for A1, Identity, A2, Attribution, and A4, Benefaction.

In Mangghuer, a Mongolic language from north-west China, ‘finite verbs (with the exception of imperatives) must be marked for one of two pragmatically-determined categories’. Slater (2003: 116–17, 194–9) explains that what he calls the ‘subjective’ category is used when the A, S, or CS argument is 1st person and in control in a declarative clause, and 2nd person and in control in an interrogative. In all other circumstances the ‘objective’ category is marked. The interesting feature of copula constructions in this language is that separate forms are used for Identity and Attribution in negative copula constructions, but the same form is used for both in positive clauses:

		A1, Identity	A2, Attribution
POSITIVE	subjective	bi bang	
	objective		
NEGATIVE	subjective	puzhi	(u)gui
	objective	puzhang	(u)guang

14.6 Occurrence and omission

Since a copula lacks referential meaning, but instead indicates a semantic relation between CS and CC, it is at risk of being omitted whenever the nature of the semantic relation could be inferred from the referents of CS and CC. Some languages (including English, French, Basque, Finnish, and Jarawara) have a requirement that every clause must include a verb; as a consequence, the copula may never be omitted. However, in most languages with a copula construction the copula verb may be omitted in certain circumstances. Matthews and Yip (1994: 129) describe the copula verb *haih* ‘be’ (with invariable form) in Cantonese and say: ‘in most cases *haih* can be omitted without affecting the sense or structure of a sentence.’ They specify that *haih* is typically used to refute a claim. For example, if someone makes the claim in (43) this could be refuted with (44), which includes *haih*.

- (43) Kéuih_{CS} mhaih [hóu lengjái]_{CC} je
 HE NOT:COPULA very good-looking PARTICLE
 He isn’t very good-looking

- (44) Kéuih_{CS} haih [hóu lengjái]_{CC} léh
 HE COPULA very good-looking PARTICLE
 He is really good-looking

It appears that in the Dravidian language Malayalam a Copula verb is optionally omissible in every circumstance (Asher 1968: 97).

If a language has two copula verbs, ‘be’ and ‘become’, it is more likely that ‘be’ will be omissible than ‘become’. This correlates with the fact that ‘become’ often shows more verbal categories than does ‘be’. In fact, ‘become’ often has some of the characteristics of a regular verb while still functioning as a copula (that is, it takes copula arguments CS and CC, rather than S, or A, and O).

In keeping with this, a number of languages have a Copula verb for ‘become’ and use a verbless clause construction for ‘be’. This can be exemplified for the Australian language Yuwaalaraay (Williams 1980: 69). Compare the verbless clause in (45) with the copula clause in (46).

- (45) burul_{CC} [nhama dhayn]_{CS}
 big THAT man
 That man is big

- (46) burul_{CC} [nhama dhayn]_{CS} gi-nji
 big THAT man become-NON.FUTURE
 That man is getting big

A common explanation offered for the omissibility of a Copula verb is that it is, effectively, a ‘dummy’ element needed just to carry bound morphemes

providing information about TAM, person/number of CS, etc. In this view, the copula can be or must be omitted in the context of the unmarked choice from certain grammatical system(s). For example, if present is the functionally unmarked term in the tense system, then the copula may be omitted in present tense; its lack will signal that the clause is in present tense. And if 3sg is the unmarked term from the pronominal system, then a copula may be omitted when CS is 3sg.

For example, in Hungarian, the copula is omitted in present tense when the CS is 3rd person and the CC relates to Identity or Attribution (but is included when it relates to Location or Possession). In Russian the copula must be included in past and future tenses but is generally omitted in present tense; it is retained only in high-flown language and in mathematical formulae (but see Timberlake 2004: 292). And the present copula has a single form, based on the old 3sg, whereas in past and future tenses the copula agrees with the CS in number and person/gender. In Tarma Quechua, 'if the copula is in the (unmarked) present tense with a 3rd person subject and no other verbal morphology is added, it is not expressed' (Adelaar 2002: 3, see also Adelaar 1977: 178). And for Ojibwe (Algonquian), Valentine (2001a: 457) reports that the copula is typically omitted from a statement of (or polar question concerning) Identity.

For Cavineña, Guillaume (2008: 97) reports as follows. 'The main function of the copula predicate is to carry verbal affixes. Speakers very often leave out the copula predicate when they do not judge it necessary to express [tense, aspect, etc.] verbal categories coded by these affixes. This happens for example in generic statements...or when the verbal categories are understood from either the textual context...or the physical/visual context.'

For the Australian language Diyari, Austin (1978: 240) states that an imperative involving an adjective as CC *requires* a copula as a vehicle to host the imperative suffix. Thus:

- (47) wata malhanjtji_{CC} ngana-a-ni-mayi!
 NOT bad be-IMPERATIVE-NUMBER.MARKER-EMPHATIC
 Don't be bad!

Austin (1981a: 104) also states that, in non-imperative clauses, *ngana-* is 'only rarely used in the present tense' but 'when the attribute, equation or possessive equation is not located in the present, *ngana-* must be used to carry the tense or mood inflection'.

For some languages the conditions for inclusion or omission of a copula verb may be less specific. Sneddon (1996: 237–8) characterizes the copula in Indonesian as 'optional' and most common when either CS or CC is long, in which case the copula 'breaks up a string of nouns and adds a smoothness to

the construction'. The copula is also frequently used if the CC is a nominalized verb. For Urarina, Olawsky (2002: 12) noticed 'casual omission of the copula in colloquial speech'.

14.7 Historical development

It is instructive to look at what copulas may develop out of, and what they may develop into. A common origin is for a regular intransitive verb, with referential meaning, to develop into a Copula verb with relational meaning. Likely candidates are verbs of stance (typically 'sit', 'stand', and 'lie') or else 'go' or 'live' or just 'exist'. What happens is that such a verb may first of all be used for existence—for 'a man is in the garden' one would say, literally 'a man stands in the garden'. The verb then loses its referential meaning and comes just to mark a relation of Identity or Attribution. In the Australian language Arabana *thangka-* continues to be used as an intransitive verb, 'sit', but it has developed a second function as a copula verb which is 'neutral as to stance, and the notion "to sit" has totally faded' (Hercus 1994: 295). For example:

- (48) antha_{CS} minpaRu_{CC} thangka-ka
 1sg doctor be-PAST
 I was a (tribal) doctor (but I have now lost my special powers)

Hercus notes that the copula *thangka-* is only included 'when tense has to be expressed'. Thus a statement 'I am a doctor' will be just *antha minpaRu*, with no Copula verb stated.

Australian languages provide a revealing panorama of this kind of development. The verbs which have developed a copula function include (full details, and references to sources, are in Dixon 2002c):

- 'sit' in Yir-Yoront, Waga-Waga, and Pitta-Pitta (as well as in Arabana)
- 'stand' in Wirangu, and the Western Desert language
- 'lie' in Wardaman, Gooniyandi, and Guugu-Yimidhirr
- 'stay, exist' in Njangumarta and Walmatjarri
- 'live, dwell' in Lardil
- 'go' in Martuthunira, Gumbaynggirr, Warray, and Gaagudju

In some of these languages the original verb retains its referential sense, in addition to the new copula function. In others it is only a copula, but typically is cognate with a verb of stance or motion in a related language. A similar origin is reported for copulas in other parts of the world. For example, in Manambu, from Papua New Guinea, five of the copulas are homonymous with regular verbs 'stand', 'sit', 'stay', 'go', and 'do, take, get' respectively.

In the case of languages for which there are good historical records, the origins of copulas often seem to be complex. The etymology of Spanish copula *estar* is straightforward; it comes from Latin *stare* 'be standing'. But the other copula, *ser*, results from a fusion of two Latin verbs, *esse* 'be' and *sedere* 'be seated'; the future, conditional, present subjunctive, and imperative forms are from *sedere* with the remaining forms being from *esse* (Travis 2002; Corominas and Pascual 1980–91: vol. 19, 213; Moliner 1984).

There are other paths of development leading to a copula. One is from a demonstrative or a 3rd person pronoun. Suppose that, at one stage in a language, form X signified 'that' or 'he'. We would then have had:

- (49) 'John [pause] X a doctor' meaning 'John, he/that [one] is a doctor'

The pause may come to be omitted, with X being grammaticalized as a Copula verb, giving

- (50) 'John [no pause] X a doctor' meaning 'John is a doctor' (with X now being 'is')

This path of development is suggested for the present-day copula *shi* in Mandarin Chinese (C. N. Li and Thompson 1977; Feng Li 1993). In Egyptian, the form *pw* appears to have begun as a demonstrative, developed into a 3rd person pronoun, and then into a copula 'be'. Gardiner (1927: 105) quotes the following sentence:

- (51) Nwn pw it ntrw
Nun he/be father gods

He then states: 'nothing but the context can decide whether the intended meaning was 'it is Nun, the father of the gods' or 'the father of the gods is Nun'.

Once a form has developed into a copula, it may not stop there. In Spanish, *estar* is used as a marker of progressive aspect and *ser* as a marker of passive. *Be* in English extends to both these functions; besides the copula *be* we also have a *be* in imperfective aspect *be...-ing*, a *be* in passive *be...-en*, and (as mentioned in §14.1), a fourth *be* in existential *there is/are/was/were*. And there are complex lexemes beginning with *be*, including *be about to*, *be up to*, *be up for*, *be in on*, *be out of*, *be for*, *be over*, among many others.

The copula verb-form also functions as an auxiliary verb to show various TAM categories in Colloquial Welsh; in Malayalam it is used for progressive aspect (Variar 1979: 49). Sango is notable in that the copula, *yeke*, is the most frequent verb (with 23.6 per cent of occurrences in a corpus of 10,015 verbs); *yeke* also functions as an auxiliary element, marking imperfective and future (Thornell 1997: 134–41).

In Jarawara, copula *ama* ‘be’ is homonymous with the modal-type word *ama* ‘be extended in time’ (this is included within a transitive or intransitive predicate), and copula *-ha-* ‘become’ is homonymous with verbal auxiliary *-ha-*, which is required with a number of non-inflecting verbs. In Tarma Quechua, the copula *ka-* ‘plays a central role in the formation of compound tenses’ such as habitual past (Adelaar 2002: 6). And in Manambu, four of the Copula verbs are homonymous with auxiliaries, indicating anterior, durative, and prolonged durative aspects, plus imminent modality.

14.8 Summary

Besides intransitive and transitive clause types (whose predicates have referential meaning), many languages also have a copula clause construction where the predicate—here a copula verb—indicates a type of relation between its two arguments, copula subject (CS) and copula complement (CC). The relations shown by a copula always include identity and/or attribution, and often also possession, benefaction, and location. In some languages there is a further type of copula clause, with just a CS argument, indicating existence. Languages lacking a copula may utilize a verbless clause construction, where the predicate slot is unfilled but a copula-like relation is established between verbless clause subject (VCS) and verbless clause complement (VCC).

In many—but not all—instances, CS (and VCS) have similar grammatical properties to intransitive subject (S). However, CC (and VCC) typically have quite different properties from other core arguments. For example, no example is known of CC (or VCC) being realized through a bound pronoun.

There may be more than one copula verb, typically ‘be’, which just refers to a state, and ‘become’, which refers to getting into a state. Copula verbs often have rather different morphological properties from referential verbs. They may mark less, or more, grammatical categories and may show irregular (sometimes suppletive) forms. In some languages there is a special negative copula which may show a different grammatical profile from its positive counterpart. In some languages, a copula verb may be omitted in certain circumstances; this may be in present tense (where this is the unmarked term in the tense system) or when the CS is 3sg (this being the unmarked term from the pronoun system).

The historical origins of copula verbs include: intransitive stance verb, demonstrative, and 3rd person pronoun. They typically develop further functions, often as markers of aspect or tense or a modality.

14.9 What to investigate

The main questions to confront are as follows.

- A. Does the language have one or more copula verbs? The criteria for recognizing a verb as copula are:
 - It must be able to occur with two core arguments: copula subject (CS) and copula complement (CC).
 - It has a relational rather than a referential meaning.
 - It is used for, at least, relations of A1, Identity (NP as CC), and/or A2, Attribution (adjective as CC).
- B. Are there verbless clauses, with two arguments—verbless clause subject (VCS) and verbless clause complement (VCC)?

(Note that a verbless clause construction should be distinguished from a copula clause construction where the copula verb may be omitted in specified circumstances.)
- C. What are the relations covered by each type of copula clause (or by the verbless clause)?
 - (i) Those involving two arguments (CS and CC) may be Identity, Attribution, Possession, Benefaction, Location, as set out in Table 14.1 of §14.1: there could be further semantic relations in particular languages.
 - (ii) There may be secondary use of a copula with just one argument (CS): this generally indicates B, Existence (CS is then an NP) or may show ‘happening’ (with a complement clause as CS).

For each of these relations, if it is not realized through a verbless clause or a copula clause, the grammarian should say *how* it is realized in the language.

- D. What is the marking for CS(/VCS)?

Which of the subject properties (shown by A argument or by S argument or by both) does the CS argument also show?

Syntactic properties may include some or all of: subject possibilities in imperatives (and possibilities for omission); antecedent control over reflexive pronouns; role in coordination, with respect to S/O or S/A pivot, or with respect to ‘same subject’ and ‘different subject’ in switch-reference; constraints on coreferentiality and possibilities for omission in complement clause and relative clause constructions.

Coding properties may include: case marking on NP; place of NP in constituent order; marking by bound pronominal elements.

E. What is the marking for CC(/VCC)?

Does the CC argument show any properties in common with other arguments (A, S, O, CS, peripheral) or does it show significant differences?

Can the CC be complex, consisting of a noun or adjective plus a governed NP, or of a subordinate clause? For example, in English *She_{CS} is [clever with her hands]_{CC}*, *[My intention]_{CS} is [that my wife should never want for anything]_{CC}*.

- F. Do copula(/verbless) clauses behave like intransitive and transitive clauses with respect to such operations as: (a) negation; (b) imperative; (c) polar interrogative and content interrogative; (d) causative and applicative derivations (and other valency-changing derivations)?
- G. What are the forms of the Copula verb(s)? Do they show the same—or less, or more—TAM and other distinctions as non-Copula verbs? Do they have irregular or regular inflection?
- H. Do copula verb-forms have other functions in the grammar? (For example, *be* in English also features in progressive *be...ing* and in passive *be...-en*, as in *She is running*, *He was taken in (by the police)*.)
- I. Are there lexical homonyms of a Copula verb? For example, a single form may function as (i) a Copula verb, and (ii) an intransitive lexical verb of stance or motion ('sit', 'stand', 'lie', 'go', etc.)
- J. Can you say anything about the diachronic origin of a Copula verb? (Two common origins are: from a verb of stance or motion; and from a demonstrative or a 3rd person pronoun.)
- K. Can you say anything about copula and/or verbless constructions in (i) genetically closely related, and/or (ii) geographically contiguous languages? Do they have similar constructions? Do they have similar forms for copula verbs (perhaps with similar irregularities)? If there are similarities between languages in copulas and copula constructions, does it appear that these are due to shared genetic inheritance, or to areal diffusion?

If there are similarities in copula constructions but differences in copula forms between adjacent languages, does it appear that 'having the copula construction type' has diffused, with each language developing copula verbs from its internal resources in its own individual way?

Sources and notes

This chapter incorporates parts of Dixon (2002c). I have benefited a great deal from the presentations of participants in Local Workshops on this topic at the Research Centre for Linguistic Typology in 1994 and in 2002. There is a great

deal of good information on copula constructions in the five volumes edited by Verhaar (1967–72).

Pustet (2003) purports to be a typological study of copulas. However, she approaches the topic from a different angle from that espoused in this present volume. (a) The author does not distinguish between copula complements and an adjective or noun functioning as head of an intransitive predicate. (b) She states that copula-plus-CC *is* an intransitive predicate (there is no realization that CC is an argument, just as CS is). (c) She treats inchoative verbalizing morphological processes as copulas. (d) If a copula is homonymous with some other grammatical element (such as an auxiliary verb), then these other elements are taken as bona fide instances of the copula. (e) Some of the information is based upon the author's limited elicitation in particular languages (lacking a holistic view of the grammar), taking little or no notice of the established grammatical literature on those languages. (f) Some of the information presented is misquoted from sources.

14.1. The discussion here leaves aside minimal clauses such as a reply to a question, which can be just an NP; these occur in specific discourse environments. In Russian, a clause can consist of a single NP, for instance *dom* 'house', the meaning being 'There is a house (here)' (Shirjaev 1979).

14.3. A distinctive feature of CC in English is the continuing debate over the form of the 1sg pronoun in this function—whether one should say *It_{CS} is I_{CC}* or *It_{CS} is me_{CC}*.

14.4. Varied labels for subtypes of the Identity relation will be found in Stassen (1997: 100–6) and Mikkelsen (2005), among many other sources.

14.4.1 In English, the verbs *look*, *sound*, *smell*, *taste*, and *feel* have sometimes been designated as copula verbs when they occur in a sentence such as *This looks/sounds/smells/tastes/feels good*. These senses are more appropriately regarded as a type of intransitive verb; see Dixon (2005: 138–9, 203–5).

14.5. Veselinova (1999) provides some information concerning suppletive Copula verbs. Note, though, that a fair amount of what she says is misquoted or misunderstood; readers should always check back into the original sources.

14.6. A further example of the contrast between copula and verbless clause comes from Semelai (Aslian branch of Austro-Asiatic; Kruspe 2004a: 269–78). A verbless clause is used for 'be' and a clause with copula verb *sot* for 'become' in the sense of 'metamorphose into' (as in 'your wife metamorphosed into a birth vampire').

In Dumi (Tibeto-Burman, van Driem 1993: 168–76), a verbless clause is used for Identity (including naming); and a copula verb (there are two, depending

on whether the CS is animate or inanimate) for Attribution, Location, and also Existence.

14.7. Ng (2004) presents evidence for the development of copulas from demonstratives in Passamaquoddy (an Algonquian language). For further examples, see Hengeveld (1992: 237–56) and Devitt (1990). Some of the chapters in Newman (2002) discuss copulas and stance verbs. Further examples of the use of a copula form as aspectual marker across a variety of languages will be found in Heine and Kuteva (2002: 97–9). And see Benveniste (1971b) for insightful discussion on the origins and functions of copula verbs.

Pronouns and Demonstratives

This chapter deals with two of the varieties of shifters which were introduced in §3.7—pronouns, whose shifting reference relates to participants in the speech act, and demonstratives, whose shifting reference relates to spatial location. In each case, the reference is deictic, effectively ‘pointing’ at some person or thing. (The third type of shifter covers temporal terms such as ‘today’ and ‘now’; see §3.8.) In §15.3 there is discussion of anaphora and cataphora, which can involve 3rd person pronouns and demonstratives.

15.1 The category of pronoun

Leaving demonstratives for discussion in §15.2, we begin with what are often called ‘personal pronouns’. In this volume, ‘pronoun’ is defined as ‘a small closed class of grammatical words which vary for person’. Markers of reflexive constructions sometimes—but not always—reflect the person of the controlling argument (the subject); when they do, the term ‘reflexive pronoun’ is appropriate. Markers of relative clause construction may—but do not have to—code reference, number, and/or syntactic function of the argument which is common to main and relative clauses. They do not code person; nevertheless, the label ‘relative pronouns’ is well established (see (f) in §17.3.1). As the term ‘pronoun’ is used here, it does not apply to interrogative content words, to indefinite terms, to quantifiers, or to demonstratives.

There are two participants in a speech act: the person speaking—the ‘speaker’ or ‘1st person’—and the person spoken to—the ‘addressee’ or ‘2nd person’ (*not* ‘the hearer’, a term which includes people addressed and also anyone who happens to hear what was said, even if the message was not specifically intended for them). Many languages have just two persons in their pronoun system. In others there is a further item in the paradigm, with similar grammatical structure and function to 1st and 2nd person pronouns; this is called ‘3rd person’. It has been variously described as ‘that which is spoken of’ or (in the Semitic grammatical traditions) either ‘that which is hidden’ or ‘that which is absent’. But one can speak of ‘I’ or ‘you’, and the referent of a 3rd person may be present and perfectly visible to speech act participants. The

only appropriate definition of ‘3rd person’ is ‘some person or thing which is neither speaker nor addressee’. That is, it does not properly refer to a ‘person’ (in the sense of speech act participant), but is accorded the—one might say, honorary—title of ‘person’ when it occurs in the same grammatical paradigm as 1st and 2nd person pronouns.

In the Sanskrit tradition, which pre-dates that of Greek and Latin grammar, ‘he, she, it’ was referred to as *prathamah* ‘first’ (or ‘lowest’), ‘you’ as *madhymah* ‘middle’, and ‘I’ as *uttamah* ‘highest’ (or ‘most excellent’). The Greek and Latin tradition adopts a different convention. But why is the speaker called the 1st person and the addressee the 2nd person? Lyons (1977: 638) explains that ‘the Latin word *persona* (meaning “mask”) was used to translate the Greek word for “dramatic character” or “role”, and the use of this term by grammarians derives from their metaphorical conception of a language-event as a drama in which the principal role is played by the first person, the role subsidiary to his by the second person, and all other roles by the third person.’

‘Pro-noun’ suggests ‘substitute for a noun’ (although in fact it goes back to Latin *prō-nōmen* ‘substitute for a name’). It is sometimes maintained that only 3rd person pronouns—not 1st and 2nd persons—can properly substitute for a noun through anaphoric function. In *The fat man sat on the chair and he broke it*, pronoun *he* is used in place of a repetition of *the fat man*, and *it* for *the chair*. However, the original rationale for the term appears to have simply been that a pronoun could be used as an alternative to a noun in filling an argument slot in clause structure; in *X ran fast*, *X* could be a noun, such as *Spartans*, or else a pronoun, such as *we*.

Every language has a closed class of ‘free’ (or ‘independent’) pronouns which can be head of an NP in core argument function. Many languages also have a grammatical system of ‘bound pronouns’ which typically attach to a predicate and may be an obligatory component of it (irrespective of whether there is a free pronoun in the appropriate NP slot). For a language with obligatory bound pronouns, free pronouns are likely to be used sparingly—for contrast, or emphasis, or in copula complement or verbless clause complement function (functions which are generally not open to bound pronouns). There is discussion of bound pronouns in §15.1.9.

Most languages do have three persons in their pronouns—abbreviated as a ‘1/2/3’ system. Generally, pronouns of the three persons have similar grammatical properties, but there are languages in which 3rd person differs significantly from 1st and 2nd—see the nominal hierarchy in §13.5.4 (and illustrations in Dixon 1994: 86). A sizeable minority of languages lack a 3rd person pronoun per se, showing a ‘1/2’ system. The functional roles covered by 3rd person in other languages are likely to be taken over by demonstratives, or classifiers, or other grammatical forms.

TABLE 15.1. Forms of 2nd person pronouns in languages with different number systems

NUMBER OF INDIVIDUALS REFERRED TO	1	2	A FEW	MANY
(i) Akan (Kwa family, Ghana)	singular <i>wo</i>	plural <i>mo</i>		
(ii) Kayardild (Tangkic family, Australia)	singular <i>nyingka</i>	dual <i>kirra</i>	plural <i>kilda</i>	
(iii) Longgu (Austronesian family, Solomon Islands)	singular <i>oe</i>	dual <i>amurua</i>	paucal <i>amolu</i>	plural <i>amu</i>

15.1.1 Person and number

All languages have a distinction of number in either free or bound pronouns (or both). We can repeat here (as Table 15.1) Table 1.1 from §1.4, which illustrates 2nd person pronouns in number systems of different size. In each language, singular involves reference to just one person. We find, in the different-sized systems:

- (i) ‘sg/pl’—where plural refers to more than one.
- (ii) ‘sg/du/pl’—where dual refers to two and plural to more than two.
- (iii) ‘sg/du/paucal/pl’—where dual refers to two, paucal to a few, and plural refers to more than a few.

Four-term number systems are rather infrequent, and where they do occur by far the most common type is (iii) ‘sg/du/paucal/pl’. Much rarer is (iv) ‘sg/du/trial/pl’, where trial refers to three and plural to more than three (one language in which this is attested is Wunambal, from Australia; see Dixon 2002a: 246).

In languages with a system of type (iii) there is no fixed range of numbers associated with ‘paucal’ and with ‘plural’. What we find is that paucal and plural *contrast* within the context of speaking. In 1985, I encountered an instructive use of paucal and plural pronouns in an announcement about communal work in the Fijian village of Waitabu (Dixon 1988a: 52). The message was called out three times, each in a different part of the village; it should reach the ears of one-third of the villagers each time. The crier used the paucal pronoun in addressing his listeners—‘you (*dou*, 2.paucal) listen, our (*odatou*, 1.inclusive.paucal) people in this part of the village.’ Then he said, ‘I’m calling out the tasks of you (*onumuu*, 2.plural), the women, for today because this is our (*oda*, 1.inclusive.plural) day for village work.’ Here the pronouns were plural, referring to everyone in the village. Paucal was used when addressing

one-third of the adult villagers (twenty or so people), but plural when referring to the whole village (perhaps sixty adults). In other contexts, paucal could be used for reference to three or four people, and plural for a slightly larger number.

In the great majority of languages, pronouns are *not* analysable into person and number elements. This is illustrated in:

- (1) Kana (Benue-Congo family, Nigeria; Ikoro 1996a: 117)

Non-emphatic pronouns for A and S functions:

1sg	m	1pl	i
2sg	o	2pl	bì
3sg	a	3pl	bà

However, there are a fair few languages for which person and number morphemes can be clearly distinguished, as in:

- (2) Cantonese (Sinitic; Matthews and Yip 1994: 79)

1sg	ngóh	1pl	ngóh-deih
2sg	léih	2pl	léih-deih
3sg	kéuih	3pl	kéuih-deih

(The plural suffix on pronouns, *-deih*, is not used on nouns, save with *yành* ‘(other) person’.)

It has often been pointed out that what is called ‘plural’ of 2nd person, and most especially of 1st person, is not a plural in the same sense as the plural of a noun. *Dogs* refers to a number of animals each of which can be referred to as *dog*. But *we* does not refer to several people each of whom could be referred to, in this context, as *I*. A 1pl pronoun (such as *we* in English) generally relates to the speaker and one or more other people associated with the speaker (who are not in this instance speakers). A 2pl pronoun may refer to two or more people who are all being directly addressed, or else to one addressee and some other person(s) associated with the addressee but not themselves being currently addressed. However, for 3rd person, the relation of 3pl (*they*) to 3sg (*he, she, or it*) is parallel to that between *dogs* and *dog*. It is another instance of the difference between 1st and 2nd persons—the pronouns referring to speech act participants—and 3rd person.

This may help explain why in many languages, pl and sg pronouns of the same person have quite different forms. (Although there are a significant number of languages similar to Cantonese in which there *is* a segmentable plural element in pronouns.) A further factor is that pronouns are amongst the most frequently used words in any language, making them particularly susceptible to phonological reduction. What once were separate person and number components may, over time, become indissolubly fused, obscuring the original composition.

15.1.2 'Me and you'

We can use '1' to refer to a single speaker, '2' for a single addressee, and '3' for anyone other than a speech act participant. There are six possible two-person combinations:

- 11 two people simultaneously uttering something which they have agreed in advance pertains to each of them (such as 'We are Christians'); this is highly unusual
- 12 speaker and one addressee
- 13 speaker and a further person, not an addressee
- 22 two people, both addressees
- 23 one addressee and someone else who is neither speaker nor addressee
- 33 two people, neither a speech act participant

For very many languages, these meanings are coded in a canonical '1/2/3, sg/pl' pronoun system as:

- (3)
- | | | | |
|-----|---|-----|-----------------|
| 1sg | 1 | 1pl | 12, 13 (and 11) |
| 2sg | 2 | 2pl | 22, 23 |
| 3sg | 3 | 3pl | 33 |

(The system is easily extendible for reference to more than two people: 1pl will cover 122, 123, 133 (and 111), 2pl will cover 222, 223, and 233, 3pl will extend to 333, and so on.)

However, there are other schemes of organization. For example, Sanuma has a '1/2, sg/pl' system, but the assignment of person combinations to the plural pronouns differs from the canonical pattern in (3):

- (4) Sanuma (Yanomami, Brazil/Venezuela, Borgman 1990: 149)
Short forms of pronouns
- | | | | |
|----|---|------|-------------|
| sa | 1 | sama | 13 (and 11) |
| wa | 2 | ma | 12, 22, 23 |

We see that, rather than 12 being grouped with 13, as in the canonical pattern, (3), Sanuma groups 12 with 22 and 23. In this language *ma*, 2pl, is 'one addressee plus one (or more) other person(s), who can be another addressee, or the speaker, or neither of these'.

The 'me and you' pronoun, 12, has a special role for many languages. We can trace various sequences of development, leading off in different directions to a number of more complex pronoun systems.

The most common pronoun system—being found in languages from every part of the world—has just 'sg/pl' with either '1/2/3' or '1/2'. Just occasionally, we find a language of this type with an additional term referring to 'me and you':

- (5) Koiari (Papua New Guinea; Dutton 1996: 48–9)

1sg	da	1pl	no	
2sg	a	2pl	ya	plus 12 ‘me and you’ no ninavore
3sg	ahu	3pl	yabu	

Dutton states: ‘the special dual inclusive form *no ninavore* “we (2), you and me” is of restricted use; it cannot be used in possessive, reflexive or reciprocal constructions, for example.’

An extra-paradigmatic pronoun ‘me and you’, similar to that in Koiari, is likely to be incorporated into the pronoun system. There are several different ways in which this can happen.

(a) Development of a dual number in all persons

The original ‘me and you’ pronoun may extend its reference to cover ‘me and one other person, not you’; that is, 13 in addition to 12. It is now a regular 1du pronoun. A 2du and (in a ‘1/2/3’ system) a 3du form will be innovated, producing a regular ‘1/2(/3), sg/du/pl’ paradigm. In many Australian languages *ŋali* only refers to 12, ‘me and you’. But in some languages it now also refers to 13:

- (6) Baagandji (Australian area; Hercus 1982: 109–10)

1sg	ŋaba	1du	ŋali	1pl	ŋina
2sg	ŋimba	2du	ŋubu	2pl	ŋurda

Most pronoun systems of the type ‘1/2(/3), sg/du/pl’ are found in languages from South and North America, Australia, and New Guinea. Some have undoubtedly evolved in the way just described, while others may have followed a different path of development.

(b) Development of an inclusive/exclusive distinction

Non-singular 1st person pronouns may be of two types—inclusive (including the addressee) versus exclusive (not including the addressee). A system like that in (5) can be restructured as ‘1/2(/3), sg/pl, inc/exc’ through (i) the original ‘me and you’ being extended to refer to ‘me and you and optionally one or more others’, and (ii) what was originally 1st person plural (speaker and any one or more others) being reinterpreted as 1st person plural exclusive (speaker and one or more others, not including the addressee). That is:

- (7)
- | | | |
|-------|---|---------------------------------|
| 1sg | { | 1pl.inc (original ‘me and you’) |
| | | 1pl.exc (original 1pl) |
| 2sg | | 2pl |
| (3sg) | | (3pl) |

An example of such a system is:

(8) Motuna (Buin family, Papua New Guinea; Onishi 1994: 128)

1sg	ni	{	1pl.inc	nee
			1pl.exc	noni
2sg	ro		2pl	ree

(Some grammar-writers place 1sg in the same row as 1pl.inc while a rather larger number put it in the same row as 1pl.exc. Neither is appropriate. 1sg neither includes reference to the addressee nor to someone other than the addressee; that is, 1sg is neither inclusive nor exclusive. It is best located between rows.)

Pronoun systems including an inclusive/exclusive contrast—whether with a ‘sg/pl’ or a ‘sg/du/pl’ number system—are found in every continent save Europe.

(c) Development of an inclusive/exclusive distinction plus a dual number in all persons

In system (7), the original ‘me and you’ pronoun has had its reference extended to also cover ‘me and other(s), not you’. An alternative way of developing an inclusive/exclusive paradigm is to extend the number system from sg/pl to sg/du/pl. We can begin with a pronoun system like (5), that is:

- (9) 1sg 1nsg
 2sg 2nsg plus 12 ‘me and you’
 3sg 3nsg

(Non-singular (nsg) is used here to avoid confusion with pl, which is used in (10) to refer to ‘more than two’.)

We can now examine how (9) may be converted into a ‘1/2/3, sg/du/pl, inc/exc’ system:

- (10) 1sg { 1du.inc (original ‘me and you’) 1pl.inc
 1du.exc 1pl.exc (original 1nsg)
 2sg 2du 2pl (original 2nsg)
 3sg 3du 3pl (original 3nsg)

In this system the ‘me and you’ pronoun retains its original reference, as 1du.inc. The three original nsg pronouns (referring to more than one person) now become plural (referring to more than two). And four new terms—underlined in (10)—are required to complete the sg/du/pl. inc/exc system.

A number of Australian languages have pronoun systems of this type. For example:

- (11) Nhangu (Yolngu genetic group; Schebeck 2001: 22)

1sg	ɲarra	{	1du.inc	ɲali	1pl.inc	ɲalima
			1du.exc	ɲalinyu	1pl.exc	ɲanapu
2sg	nhuunu		2du	nhuma	2pl	nhurruli
3sg	ɲayi		3du	balay	3pl	walala

It will be seen that two of the new forms, 1pl.inc *ɲalima* and 1du.exc *ɲalinyu*, are plainly based on the familiar Australian ‘me and you’ form *ɲali*, now 1du.inc.

(d) Development of a minimal/augmented system

There is a further type of pronoun system, whereby ‘me and you’ may be placed in the same column as 1sg and 2sg (and 3sg, within a ‘1/2/3’ paradigm). This can be illustrated for the eight pronouns of Hanunóo, an Austronesian language from the Philippines. If we try to arrange these in a ‘1/2/3, sg/du/pl, inc/exc’ matrix, the result is:

(12)			{	1du.inc	tah	1pl. inc	tam
	1sg	kuh					
					1du/pl.exc	mih	
	2sg	muh			2du/pl	yuh	
	3sg	yah			3du/pl	dah	

Conklin (1962: 134–5) suggests that a better model of this pronoun system would be:

(13)	SPEAKER	HEARER		minimal	non-minimal
	✓	–	1	kuh	mih
	✓	✓	1+2	tah	tam
	–	✓	2	muh	yuh
	–	–	3	yah	dah

That is, the eight terms in this pronoun system are expressed in terms of three binary contrasts: \pm speaker, \pm hearer, and \pm minimal.

Many Australian languages have a pronoun paradigm similar to that described by Conklin but with a three-term number system. McKay (1975, 1978) suggested the terms ‘unit augmented (ua)’ for ‘one person in addition

to minimal' and 'augmented (aug)' for 'more than one person in addition to minimal'. For example:

- (14) Gurindji (Australian region; McConvell 1990)
- | | minimal | unit augmented | augmented |
|-----|---------|----------------|-----------|
| 1 | ɲayu | ɲayirra | ɲantipa |
| 1+2 | ɲali | ɲali-wula | ɲaliwa |
| 2 | njun-tu | njun-pula | nju-rrulu |
| 3 | njan-tu | njan-pula | nja-rrulu |

As far as 1st, 2nd, and 3rd person are concerned, unit augmented and augmented are equivalent to dual and plural (in a 'sg/pl' system). The minimal term 1+2 refers to two people, its unit augmented term to three, and its augmented correspondent to four or more.

Some recurrent morphemes are segmented out in (14): *-pula* for three ua forms (the initial *p* lenites to *w* after a vowel in *ɲali-wula*), *-tu* for two min, and *-rrulu* for two aug forms. In addition, 1+2.aug *ɲaliwa* appears to be based on 1+2.min *ɲali*, but *-wa* occurs only in this one pronoun. The recurrent Australian 1+2.min form *ɲali* is familiar from the pronoun systems in (6) and (11).

Minimal/augmented systems are fairly widespread. They are found in further Philippines languages besides Hanunóo (including Ilocano) and in a fair number of South American languages, including Aymara, Quechua, and members of the Carib family. For example:

- (15) Trio (Carib, Surinam; Carlin 2004: 144)
- | | minimal | augmented |
|-----|---------|-------------|
| 1 | wĩ | ainja |
| 1+2 | kĩmẽ | kĩmẽ-injamo |
| 2 | ẽmẽ | ẽmẽ-injamo |

These systems are also found in North American languages, including Southern Paiute (Sapir 1930–1) and others from the Uto-Aztecan family, and some from the Siouan family (see (18) below). And they are found in various languages of Africa, including those from the Chadic family. The system in Hdi was given in §3.7. To this can be added:

- (16) Margi (Chadic branch of Afro-Asiatic, Nigeria; Hoffmann 1963: 73)
- | | minimal | augmented |
|-----|---------|-----------|
| 1 | nàyù | nà'yà |
| 1+2 | nàmà | nàmòr |
| 2 | nàgù | nànyì |
| 3 | nàjà | nàndà |

All minimal/augmented systems from languages in the Philippines, Africa, and the Americas have just two numbers. In Australia there are three or so languages with two-number systems and around a dozen with three numbers, as illustrated in (14) and (17): see Dixon (2002a: 249–52, 266). Only limited information is available on the hundreds of non-Austronesian ('Papuan') languages in New Guinea, but they do include some 'min/ua/aug' paradigms; see Boxwell (1967) on Weri.

In the same way that 'plural' has a different significance depending on the size of the number system it occurs in—illustrated in Table 15.1—so the reference of 'augmented' depends on the size of its number system:

(i)	3-term number system	MINIMAL 1, 1+2, 2 (3)	UNIT AUGMENTED (one person in addition to minimal)	AUGMENTED (more than one person in addition to minimal)
(ii)	2-term number system	MINIMAL 1, 1+2, 2 (3)	AUGMENTED (one or more persons in addition to minimal)	

To avoid having the same kind of ambiguity attached to 'augmented' as to 'plural', one could follow Conklin and just use the labels 'minimal' and 'non-minimal' for two-term systems.

As with 'sg/pl' and 'sg/du/pl', both 'min/aug' and 'min/ua/aug' systems may be used with just two persons, '1/2', or with three, '1/2/3'. And, as with other systems, the segmentation of pronouns into person and number elements varies from language to language. The partial analysis possible for (14), Gurindji, has already been discussed. In (15), Trio, there is an aug increment to min forms for 1+2 and for 2 but not for 1. A further pattern is shown in:

(17) Mangarrayi (Australian area; Merlan 1982: 102, 160)

Free pronouns for S and A functions

	minimal	unit augmented	augmented
1	ŋaya	ŋi-rr	ŋi-rla
1+2	ŋi	ŋa-rr	ŋa-rla
2	njaŋgi	rnu-rr	rnu-rla

Here ua involves suffix *-rr* and aug suffix *-rla*, both to a non-minimal root: 1 *ŋi-*, 1+2 *ŋa-*, and 2 *rnu-*. But note that these are quite different from the minimal roots.

There are, as with other pronominal systems, various kinds of variation on the basic pattern, such as the following fascinating paradigm:

- (18) Lakota (Siouan; Boas and Deloria 1941: 76; Mithun 1999: 70; Pustet 2001: 69)

	minimal	augmented
1	miyé	} ūkiye – pi
1+2	ūkiye	
2	niyé	niyé – pi
3	iyé	iyé – pi

The 1/1+2 distinction is made in minimal but neutralized in augmented number. What is interesting is that the combined 1/1+2 aug form is based on 1+2.min.

15.1.3 Neutralization

In every language, there are distinct forms for 1sg and 2sg—or just for 1st and 2nd person—free pronouns. Beyond this, every variety of pronoun system may allow limited neutralization of person or of number. This was exemplified in the minimal/augmented paradigm for Lakota in (18), where 1aug and 1+2.aug fall together. The same neutralization occurs in the Australian language Tiwi (Lee 1987: 105).

Contrasts within a person system may be neutralized in a non-singular number, although this is rather rare. Examples include:

- In a ‘sg/pl’ paradigm, there may be no distinction between 2nd and 3rd persons for plural. This was exemplified for the Algonquian language Chipewyan in §3.7. It is also found in the Athapaskan language Slave (Rice 1989: 253).
- In the ‘sg/du/pl’ system for Manambu, a Papuan language from New Guinea, 2du and 3du have the same canonical form, used in subject function—see Aikhenvald (2008a). (However, 2du and 3du differ in the way they take case markers for oblique functions.)
- In the ‘sg/du/paucal/pl’ paradigm for the Papuan language Yimas, 1st and 2nd persons are neutralized in paucal number (Foley 1986: 74, 1991: 111).

The interesting point about the person neutralizations just listed is that each occurs in that term from the number system which is functionally most marked—plural in a ‘sg/pl’, dual in a ‘sg/du/pl’, and paucal in a ‘sg/du/paucal/pl’ system.

In a number of Papuan languages, 2nd and 3rd person fall together in both dual and plural, from a ‘sg/du/pl’ system; they include Amele (Roberts 1987: 208) and Hua (Haiman 1980: 215).

Amongst the neutralizations within a sg/pl number system we find:

- 2sg = 2pl in Standard English (see §3.7).
- 3sg = 3pl in a number of languages, including Koasati (Muskogean family, Louisiana; Kimball 1991: 417).
- 2sg = 2pl and 3sg = 3pl in Kuman, a Papuan language from New Guinea (Piau 1985).

In the '1/2/3, sg/du/pl, inc/exc' paradigm for Yagua, spoken in Peru, we find 1du.inc = 1pl.inc (Payne and Payne 1990: 370). And in the identical system for the Austronesian language Ponapean, 1du.exc = 1pl.exc (Rehg 1981: 158).

This is only a small sample of the kinds of neutralization which occur in systems of free pronouns. In languages which also show bound pronouns, these often (although not always) make fewer distinctions than their free congeners and, as a consequence, show a wider range of neutralizations.

15.1.4 Gender

Another category which can interrelate with person and number in a pronoun system is gender. In most instances, there is simply a contrast between masculine and feminine, occasionally between human and non-human, or animate and non-animate (the latter two only apply to 3rd person). There can, in 3rd person, be a three-way division between masculine, feminine, and neuter. The Dravidian language Kannada has such a distinction in 3sg, but in 3pl the m/f contrast is neutralized (Bhat 2004: 109):

(19)		singular		plural
	1	na:nu		na:vu
	2	ni:nu		ni:ru
	3.masc	avanu	}	avaru
	3.fem	avalu		
	3.neuter	adu		avu

It is almost unknown for a gender contrast to apply for all person/number combinations. In many instances, it is confined to just one pronominal value. For example:

- Just 3sg, in very many languages, including English and Somali (Cushitic; Saeed 1993: 173).
- Just 2sg in Iraqw (Cushitic, Tanzania; Mous 1993: 112); for full details see (46) below.

Or it may be found just in two or more persons, only in singular number:

- 2sg and 3sg in Miya (Chadic, Nigeria; Schuh 1998: 187) and Manambu (Ndu family, New Guinea; Aikhenvald 2008a).
- 1sg, 2sg and 3sg in Gala (Ndu family, New Guinea; Aikhenvald 2008a).

Or a masculine/feminine contrast may occur in all numbers for certain person(s):

- All 3rd (in sg and pl numbers) for Rumanian (Mallinson 1986: 257) and Modern Greek (Joseph and Philippaki-Warbuton 1987: 156).
- All 1st and 3rd (in sg and pl numbers) in Karaja (Macro-Jê grouping, Brazil; Wiesemann 1986: 361).
- All 2nd and 3rd (in sg and pl numbers) for Tunica (isolate, Louisiana; Haas 1941: 37) and Semitic languages (Gray 1934: 62).

It is interesting to find that some languages make a gender distinction in dual or unit-augmented number, plus 3sg or 3min:

- Dual in 1/2/3 and also 3sg in Bora (Peru; Thiesen 1996: 33).
- Unit augmented in 1/2/3 and also 3 min in Nunggubuyu (Australia; Heath 1984: 243).

There are many other placements of a gender distinction at limited locations within a pronoun paradigm, beyond the representative sample presented here.

15.1.5 *Social niceties*

Different societies have varying conventions for personal interaction, which may sometimes be coded in grammar. It is all a matter of expressing respect—respect for someone in a certain kinship relation to you, or respect for someone occupying a superior position in a social hierarchy. A plain pronoun may be used between social equals, or between people whose kinship link allows for ‘familiar’ interaction. In other social contexts, use of a plain pronoun may be considered socially offensive.

Japanese is spoken in a highly structured society. There are words which can be recognized as personal pronouns, but their use is rather restricted. We find instead that kin terms may be employed, or else the name for a profession—‘mister doctor’ or ‘mister teacher’ or ‘mister shoe-shop-salesperson’—or a proper name. If the identity of a putative pronoun could be inferred from the grammatical context, it is likely to be omitted. In fact, the use of honorific

forms of verbs and of nouns may provide clues as to who is being spoken of (as indeed may bodily gestures).

Many of the major languages from South, South-East, and East Asia have special 'honorific' pronouns. In Bengali, for instance, there are three forms of the 2nd person pronoun (in both sg and pl): 'intimate', 'ordinary', and 'honorific'. The 3rd person pronoun has two forms, 'ordinary' and 'honorific'; and there is a single 1st person pronoun. Languages such as Thai and Khmer show a multiplicity of pronouns for all three persons, and in addition special forms depending on the identities of speaker and addressee: a layman speaking to a monk, a monk speaking to a layman, a monk speaking to the king, a layman speaking to the king, and so on. In these languages, as in Japanese, many of the words used for referring to speech act participants are basically nouns (so that it may sometimes not be an easy task to distinguish a class of pronouns).

An alternative technique is to use a term from a regular pronoun paradigm in an unusual way in order to indicate kinship relationship or social rank. An example may be taken from Fijian, which has a regular '1/2/3, sg/du/paucal/pl, inc/exc' pronoun paradigm. Everyone in a Fijian village is related to everyone else through a classificatory kinship system, and each type of relationship carries certain social obligations and requires particular patterns of behaviour. While living in the Boumaa village of Waitabu, in 1985, I noted the following ways of employing 2nd person pronouns (Dixon 1988a: 53):

- For most types of kin linkages (including parent/child, two brothers and two sisters) there are no special conditions. A single addressee will then be referred to with the 2sg pronoun, *i'o*.
- There are two types of 'avoidance' relations, with whom social contact must be formal (no joking!) and often indirect. If such a relative should be directly addressed, a 2du or a 2paucal pronoun is employed for referring to a single person:
 - 2du *mudrau* for an actual or classificatory mother-in-law, father-in-law, son-in-law, or daughter-in-law.
 - 2paucal *mudou* for an actual or classificatory brother or sister of the opposite sex.

There is also a well-defined social hierarchy, with the village chief being accorded prime status:

- The 2pl pronoun, *munuu*, must be used when addressing the chief.

Similar conventions apply for other languages in Oceania, in some from Australia, and in many from Europe. In French the sg/pl distinction in 2nd person (*tu/vous*) has been reinterpreted as a marker of social distance. As

a schoolboy learner I was told that one should only use *tu* for addressing servants, children, animals, relatives, and friends (acronym SCARF). But how to measure the intimacy of a friendship, in order to decide whether to employ *vous* or *tu*?

In English, the original 2pl pronoun, *you*, was used for formal reference to a single person and eventually ousted the original 2sg form, *thou*. We now have the highly unusual situation of a single form being used for 2nd person in all numbers. Different varieties of English have innovated a new 2pl (*youse* or *yous*, *you-all* or *y'all*, *you-uns*, *you guys*, or just *guys*, among a number of others) but this is resisted in formal varieties. A similar shift occurred in Basque—the old 2pl form is now 2sg formal but here a new 2pl form has been created, using suffix *-ek* which is similar to plural marking on nouns:

(20)	EARLIER SYSTEM				PRESENT-DAY SYSTEM			
	1sg	ni	1pl	gu	1sg	ni	1pl	gu
	2sg	hi	2pl	zu	2sg informal	hi	} 2pl	zu-ek
					2sg formal	zu		

Saltarelli (1988: 208) comments that the distinction between the two 2sg pronouns is ‘apparently falling into disuse, *zu* being used more or less exclusively in urban areas, *hi* being used more or less exclusively in rural areas’.

Some languages combine the two techniques we have discussed: (i) having special ‘honorific’ pronouns, and (ii) using a non-singular pronoun to refer to a single person. For example, Purki, a Tibeto-Burman language spoken in Jammu and Kashmir (Rangan 1979: 66), has a ‘1/2/3, sg/pl, inc/exc’ paradigm with two 2sg and two 2pl pronouns—honorific for addressing someone of higher status and non-honorific for someone of lower status. There is also the convention of using the 3pl form for referring to a single person if they have a higher social status or are older than the speaker.

15.1.6 A further ‘person’, and impersonal/indefinite

As pointed out in §15.1, in the sense of ‘speech act participant’, there are only two ‘persons’, 1st (speaker) and 2nd (addressee). As Benveniste (1971a: 198, 217) insists, ‘the “third person” is not a “person”, it is really the ... form whose function is to express the *non-person*’. (That is, it is not a speech act participant.) However, when there is a term referring to ‘non-person’ in the same paradigm as 1st and 2nd person pronouns, then the custom is to refer to it as ‘3rd person’.

Some languages have a further row in the paradigm, with varied meanings and functions. As mentioned in §15.1.4, there may be a gender distinction just for ‘3rd person’ or else two 3rd person forms, referring to human/non-human or animate/inanimate. Two ‘3rd persons’? Why not call one of them

'4th person'? Unfortunately, some have followed this path, using the label '4th person' in ways totally incompatible with each other. They include the following.

- (i) Quechua has a minimal/augmented system. There has arisen the habit of referring to the 'me and you' minimal pronoun as '4th person'—see Adelaar (2004: xx, 211). The 1+2 augmented form is now called '4th person plural' (referring to three or more people, while plurals of 1st, 2nd, and 3rd persons refer to two or more people).
- (ii) In Algonquian languages, there are two 3rd person forms, 'proximate', referring to an established topic, in the centre of attention, and 'obviative' for someone or something less central in that part of the discourse. The label '4th person' has been used for obviative, reserving '3rd person' for proximate (Anderson and Keenan 1985: 262).
- (iii) In some languages from the unrelated Eskimo and Pomoan families, there may be 'two different third person categories...one basic and one for arguments coreferent with the subject of the matrix clause'. These are sometimes called '3rd person' and '4th person' respectively (Mithun 1999: 73–4).
- (iv) For Matsigenka (Panoan family, Brazil), Fleck (2008: 282) uses '4th person' to 'refer to both the pronominal forms which fill this fourth position in the pronominal paradigm and to the coreferential function...this function is to stand in for a nondiscourse participant that is coreferential with a core argument in its matrix or higher clause...particularly as opposed to the simple third person, which, where it occurs in paradigmatic contrast with the fourth person, entails a lack of such coreference.'
- (v) Some Athapaskan languages have an 'inverse' system. If the A and O arguments for a transitive clause are both non-speech-act-participants, there are two possible prefixes to the verb, depending on whether or not the A argument is the expected controller of the activity. These prefixes have been referred to as '3rd person' and '4th person' (see Akmajian and Anderson 1970, and further references therein).

It can be seen that if a grammar refers to a '4th person', the reader will not know which of the many senses of the term is intended. It is surely better to avoid employing this label.

The list does continue. In some Arawak languages, there is a further term in the pronoun system, with 'impersonal' import. This roughly corresponds to *on* in French and to *one* and the impersonal sense of *you* in English, as in a sentence like 'While one is travelling in a canoe, one doesn't sleep'. The unfortunate label '4th person' has been used for such impersonal terms; for

instance, Ramirez (1992: 41) writing on Bahuana, and Rowan and Burgess (1979: 18) on Parecis.

'Impersonal' and 'indefinite'—the labels are often used interchangeably—do show one recurrent and fascinating trait: they may relate to a 1st person inclusive referent. Ainu (isolate, northern Japan; Tamura 2000: 49–73) has a 'sg/pl' system for 1st, 2nd, and 3rd persons and also 'indefinite', glossed as 'someone'. The interesting feature is that the plural indefinite can be used for 'some people' and also for 1pl inclusive, the regular 1pl then being restricted to exclusive reference. The Caddo language (Oklahoma; Chafe 1990; Mithun 1999: 71) behaves in a strikingly similar way. Bound pronoun *di-* means 'someone' when used in the singular, but when the regular dual and plural pronominal prefixes are added, it indicates 1du.inc and 1pl.inc respectively.

In some languages, things go in the opposite direction. Jarawara (Arawá family, Brazil; Dixon 2004a: 77) has a '1/2/3, sg/pl, inc/exc' system. Here the 1pl.inc pronoun, *ee*, has secondary function as impersonal 'one'. For instance, if a Jarawara villager is shown a picture of a strange animal, they invariably enquire 'Is it edible?', which is rendered as:

- (21) *ee*_A *kaba-tee* *awa?*
 1pl.inc eat-HABITUAL QUESTION
 Does one eat [it]?

A similar phenomenon is found in Acehnese (Austronesian; Indonesia; Durie 1985: 120–1), where the 1pl.inc pronoun has a secondary function as an indefinite which is particularly used for 'stating general truths about human activity or behaviour'. For instance, the way to say 'Coconut shells are made into ladles' is:

- (22) *bruek*_O *ta*_A =*peugöt* *keu* =*aweuek*
 coconut.shell 1pl.inc make to ladle
 One makes coconut shells into ladles

Lichtenberk (2005: 277–80) documents similar instances in languages belonging to the Oceanic branch of Austronesian.

Tariana (Aikhenvald 2003: 122) has '1/2/3m/3f' in 'sg/pl' and also an impersonal term which is not marked for number. As befitting an Arawak language, originally it did not show an 'inc/exc' distinction. However, Tariana has close association with East Tucanoan languages which do have an 'inc/exc' contrast. The language contact situation in the multilingual Vaupes River region may be responsible for the fact that Tariana appears to be developing an 'inc/exc' parameter (Aikhenvald 2002a: 62; 1999b: 88)—the impersonal pronoun may nowadays be used with the meaning 1pl.inc (and the original 1pl is then restricted to 1pl.exc reference).

15.1.7 Grammatical properties

Each language has its own individual character, which demands a particular strategy of presentation. That is, there is no universal scheme for how the components of a grammatical description should be ordered for maximal clarity and explanatory potential. In some languages, pronouns show a multi-faceted paradigm which should be presented very early in the grammar. But in many instances it is easiest to first describe the properties of nouns, and follow on to pronouns.

Nouns typically undergo a number of derivational processes, and may be required to select one from a system of inflections, in a fairly regular paradigm. Pronouns typically share only some of these morphological properties, and tend to have more irregularities. Of the derivational processes which may apply to nouns (perhaps involving ‘another’ or ‘really’ or ‘be like a –’) few or none may be applicable to pronouns.

If a noun has a possessive form—which functions (rather like an adjective) as modifier within an NP—there are likely to be possessive forms of pronouns, although these are typically irregular. For example, in Dyirbal suffix *-ŋu* derives the possessive form of a noun (however many syllables it contains). For non-singular pronouns, suffix *-ŋu* is added to a disyllabic and *-nu* to a trisyllabic subject form, and for singular pronouns the possessive form is not related to the subject form:

(23)		FORM IN S FUNCTION	POSSESSIVE FORM
nouns	‘man’	yara	yara-ŋu
	‘rainbow’	yamani	yamani-ŋu
pronouns	1du	ŋali	ŋali-ŋu
	2du	ñubala	ñubala-nu
	1sg	ŋaja	ŋaygu
	2sg	ŋinda	ŋinu

As mentioned in §3.9 and in §13.5.4, many languages have a ‘split-ergative’ profile, whereby pronouns show nominative (S and A functions) versus accusative (O) case marking, while nouns show absolutive (S and O) versus ergative (A) marking. Sometimes, 3rd person pronouns pattern like nouns and unlike their 1st and 2nd person confrères.

For case marking other than that relating to the core functions, pronouns typically pattern like nouns—for dative, purpose, causal, and so on. There is likely to be a set of local cases on nouns—typically locative (‘at’), allative (‘to’), and ablative (‘from’)—and these are generally *not* found on pronouns; this applies for Dyirbal. In the Papuan language Manambu, the 3rd person pronoun takes the same set of nine case inflections as nouns, while 1st and 2nd persons take only seven (omitting the two ‘transportative’ cases, meaning ‘move by means of –’).

While pronouns are likely to lack some of the morphological possibilities available to nouns, they may exhibit processes of their own. In Manambu, for instance, a feature of pronouns which assists in defining them as a distinct word class is that only they may undergo full reduplication, indicating emphasis; for example *wun* ‘I’, *wun-a-wun* ‘I myself, really me, as for me’.

Pronouns also tend to have more restricted possibilities than nouns at the syntactic level. It is true that there are some languages in which a pronoun as NP head takes the same range of modifiers as a common noun in NP head function. In Dyirbal, for example, both types of head may occur with a demonstrative, with one or more adjectives, and with a relative clause (Dixon 1972: 60–3, 100). But in many languages, a free pronoun as NP head may take only a limited set of modifiers—or else none at all, as is reported for Ainu (Tamura 2000: 47).

Pronouns in English allow few possibilities for modification when head of an NP. One can follow a noun with a relative clause, as in:

- (24) [That man [who was drunk]] crashed the car

But a singular pronoun cannot be modified with a relative clause; that is, one cannot say *[I [who was drunk]] *crashed the car*. A temporal subordinate clause may be used instead:

- (25) I crashed the car when I was drunk

However, a relative clause may be used with a plural pronoun:

- (26) [We [who believe in free speech]] abhor dictators

An interesting property of English is that the head of an NP may be made up of a plural pronoun and a noun, as in:

- (27) [We teachers] have a responsibility to look after [you students]

15.1.8 Pronoun elaboration

It was pointed out in §15.1.1 that what are called ‘plurals’ of 1st and 2nd person pronouns are not really plural in the way that *dogs* is the plural of *dog*. 1pl does not mean ‘lots of me’s’, but instead ‘me and some other person(s)’ as in the Yidiñ sentence:

- (28) ɲaɲjɪs gali-ɲ
 1pl go-PRESENT
 We are going (lit: I and one or more others are going)

Hearing (28), one might wonder who the ‘one or more others’ might be. Many languages have a way of elaborating on the reference of the pronoun and providing this information. The NP in S function in (28) can be elaborated (Dixon 1977a: 177):

- Suffix *-ba* ‘member(s) of a group’ can be added to any noun; in fact, this is the way nouns are coordinated in Yidiñ. For instance, an NP could be *Jani-ba Mari-ba* ‘John and Mary (either alone or with other(s))’; this is, literally, ‘John, being a member of a group, and Mary, being another member of the group’. The NP in S function for (29) is, literally, ‘I being a member of a group, and the woman/women, being other member(s) of the group’.

How can we tell that *ḡaṇji* and *buña-ba* make up one NP in (29)? Yidiñ has a typical split-ergative system with pronouns having nominative/accusative and nouns absolutive/ergative marking. Thus:

- When *ñañji* is in transitive subject function (A) it is elaborated by *buña-ba* plus ergative inflection *-ngu*:

- (31) [ɲaɲji buña-ba-ŋgu]_A wagu:ja_O gali:ɲa-l
 ɪpl woman-member(s).of.a.group-ERG man take-PRESENT
 The woman/women and I are taking the man

This illustrates one variety of pronoun elaboration—adding to a non-singular pronoun a noun which is marked by a suffix such as ‘member(s) of a group’ or ‘one of a pair’ or ‘with’ or something similar. An alternative is to just juxtapose a noun (without any marking) to the pronoun as in the Australian language Kayardild (Evans 1995: 299):

- (32) [ngarra kajakaja]_S warra-ja thaa-th
 idu daddy go-ACTUAL return-ACTUAL
 Daddy and I will go (lit. we two (including) daddy will go)

The examples from Yidiñ and Kayardild have involved free pronouns, and the elaboration occurs in the same NP as the pronoun. In other languages there are bound pronouns as part of the predicate, and elaboration must be through an extra-predicate NP. Fijian has a '1/2/3, sg/du/paucal/pl, inc/exc' system. An S argument can be shown just by the 1du.exc subject bound pronoun *'eirau* at the beginning of the predicate (Dixon 1988a: 157), as in:

- (33) ['eirus aa sota vata]_{PREDICATE} mai Viidawa
 1du.exc PAST meet TOGETHER AT Place
 We two met at Viidawa

This is literally ‘I and someone else (not you) met at Viidawa’. But who was the other person? They can be specified through a post-predicate NP:

- (34) [‘eirau_S aa sota vata]_{PREDICATE} [‘ei Jone]_S mai
 1du.exc PAST meet TOGETHER WITH John AT
 Viidawa
 Place

John and I met at Viidawa (lit: We two met at Viidawa with John)

The S argument is here shown discontinuously, by pronoun ‘eirau ‘we two (not including you)’ within the predicate and by the extra-predicate NP ‘ei Jone ‘with John’.

In fact, it is possible to specify ‘I’ within the extra-predicate NP:

- (35) [‘eirau_S aa sota vata]_{PREDICATE} [o yau ‘ei Jone]_S
 1du.exc PAST meet TOGETHER ART 1sg WITH John
 mai Viidawa
 AT Place

John and I met at Viidawa (lit: We two met at Viidawa, me with John)

That *o yau ‘ei Jone* constitutes one NP is shown by the fact that its component words must occur together, in this order, and that it has the structure of an NP.

Sentence (35) is judged as totally grammatical and acceptable. However, in normal discourse, the article-plus-free-pronoun *o yau* will generally be omitted, giving (34).

Just as with free pronouns, an elaboration to a bound pronoun may consist simply of a noun without any marker such as ‘with’ or ‘one of a group’. This can be illustrated from another Australian language, Nunggubuyu (Heath 1984: 542):

- (36) nu:-yanggi waramijbura:yuŋs
 I.augmented-go:PAST children
 We went with the children

Here the argument in S function is a combination of 1st person augmented verbal prefix *nu:-* plus the post-predicate NP *waramijbura:yuŋ* ‘children’

If a pronoun system does not involve an inclusive/exclusive distinction, then all non-singular 1st person pronouns may be elaborated. If there is an inclusive/exclusive contrast, the elaboration is possible for exclusive 1st person. 1du.inc is fully specified as ‘me and you’; however, 1pl.inc, ‘me and you and other(s)’, could be elaborated to supply the identity of the ‘other(s)’. In a minimal/augmented system, non-minimal 1 or 1+2 may be elaborated, as in (36).

Third person non-singular pronouns can also be elaborated. The established topic of the discourse will clearly be one of the referents and the other

can then be specified. For example: ‘John came and 3dual (they two) talked with/and Mary’. Here the 3dual pronoun naturally refers to John with the elaboration adding ‘with/and Mary’. And there is no reason why non-singular 2nd person pronouns could not also be elaborated.

15.1.9 Bound pronouns

An intransitive verb must have its core argument—in S (intransitive subject) function—stated. For a language which lacks bound pronouns, this is achieved through an NP. If the S argument refers to speaker or addressee, then a pronoun is used. If not, a full NP will be appropriate. Thus, in English:

- (37) (a) [I] looked
 (b) [You] looked
 (c) [The big woman] looked

The NP must be included in (37a–c); one cannot just say **looked*.

In a language with obligatory bound pronouns, a sentence can consist just of verb plus bound pronominal affix. For example, in Tariana (Aikhenvald 2003):

- (38) (a) nu-ka-ka ‘I looked’
 (b) pi-ka-ka ‘You looked’
 (c) du-ka-ka ‘She looked’

The initial element in (38a–c) is a bound pronominal prefix, the first *-ka-* is verb root ‘look’, and the final *-ka* is a suffix marking recent past tense plus visual evidential.

Verbs like *-ka-* ‘look’ in Tariana must carry a person prefix, chosen from the system in:

- (39) Subject prefixes to verbs (which also function as possessive markers on nouns):

1sg	nu-	1pl	wa-
2sg	pi-	2pl	i-
3sg feminine	du-	3pl	na-
3sg non-feminine	di-		

The bound pronoun provides a full specification for speaker or for addressee, in (38a–b). But in (38c) *du-* only indicates that the S argument is neither speaker nor addressee (that is, 3rd person) and feminine. To convey the same information as (37c) an NP must be added:

- (40) [inaru hanu-ma]_S du-ka-ka
 woman big-FEMININE 3sg.fem.S-look-RECENT.PAST.VISUAL
 The big women looked

The S argument is here jointly specified by the pre-predicate NP *inaru hanuma* and by the 3sg.fem prefix to the verb, *du-*.

There are in Tariana free forms corresponding to each bound pronoun, which simply involve the addition of *-ha*; thus, 1sg *nuha*, 2sg *piha*, etc. Any of these could make up an NP preceding the verb; for example, (38a) could be extended to *nuha nukaka* ‘I looked’. This just adds a degree of emphasis to the sentence.

It can be seen that 3rd person has a totally different role in free and in bound pronoun systems. An obligatory system of bound pronouns deals with *all* possibilities for a core argument; there *must* be a 3rd person term covering non-speech-act-participants. However, many systems of free pronouns only include 1st and 2nd persons. As is evident in (37a–c), 1st and 2nd person pronouns are needed for specification of speaker and addressee, whereas other types of reference will be through an NP (here, *the big woman*). If there is a 3rd person pronoun it is likely to be used either (i) deictically, as when someone says *She looked*, pointing at a woman nearby; (ii) anaphorically, to refer back to a stated NP, as in *The big woman looked and she saw everything*. Note that in the latter instance, the pronoun *she* could be omitted, giving *The big woman looked and saw everything*. From the grammatical conventions of English, one knows that if a subject slot is left blank it must be identical to the subject of the previous, coordinated, clause. In contrast, a bound pronoun such as *du-* in Tariana—in (38c) and (40)—can *never* be omitted.

In most languages in which they occur, bound pronouns are—in some or all functions—obligatory. Their reference can always be extended by a full NP. In other languages, bound pronouns are optional. For instance, in the Luritja dialect of the Western Desert language in Australia, core arguments may be shown by free pronouns (which are separate words preceding the verb), as in (41a), or by bound pronouns (which are enclitics to the first word in the clause), as in (41b), or by both (Heffernan 1984: 25).

- (41) (a) η ayulu_A nyuntu-nya_O pu- η ju
 1sg:NOMINATIVE 2sg-ACCUSATIVE hit-PAST
 I hit you
- (b) pu- η ju=rna=nta
 hit-PAST=1sg.A=2sg.O
 I hit you

The two sentences have essentially the same meaning, the choice of which to use being pretty much a matter of style.

Sometimes an argument may be shown by a bound pronoun, or by an NP, but *not* by both. Escalante (1990: 17) describes Yaqui (Uto-Aztecan) as being of this type and remarks that ‘the Yaqui speaker may choose between free and

clitic pronouns according to discourse factors, in placing more or less focus on an argument'. A free pronoun is used for 'contrastive emphasis' and a bound one for 'topic continuity in discourse'. And there can be more complex patterns. For example, in Mende (Sierra Leone; Creissels 2005: 47), a 3sg subject argument can be expressed by a bound pronominal prefix *or* by an NP, but *not* by both, whereas for a 3pl subject argument a bound pronominal prefix must always be stated, and this can optionally have its reference extended by an NP.

How can one tell what is a bound pronoun? How does one distinguish between free and bound pronoun forms? There is no universal set of criteria for deciding this question. However, Table 15.2 contrasts the typical properties

TABLE 15.2. Contrastive properties of free and bound pronouns

	FREE PRONOUN	BOUND PRONOUN
FUNCTION	Functions as head of an NP, and may be substitutable by a noun. In many languages, may take some modifiers within its NP.	Is not head of an NP. Is not substitutable by a noun. Takes no modifiers.
POSITION	Has the freedom of positioning of words within an NP, and of NPs within a clause (and thus of words within a clause).	If affix to verb, has fixed position within the verb. If clitic, may only occur at one (or perhaps one or two) positions in the clause.
FORM	Is one grammatical word and, in most cases, is also one phonological word. If so, it must satisfy the requirements on a phonological word (e.g. in many languages, must be of at least two moras).	Almost always has status of clitic (separate grammatical word but not a distinct phonological word) or of affix. May be less than a syllable in extent (e.g. just a syllable-closing or syllable-opening segment).
CAN THERE BE A ZERO FORM?	Can never have zero form, or a zero allomorph.	One term in the system may have zero realization, or a zero allomorph (this is most frequently 3sg).
DISCOURSE PROPERTIES	Typically used for first mention of a participant, for special emphasis, and/or as copula or verbless clause complement.	Obligatory in some languages. Optional in others; then, typically used to mark inter-clausal cohesion.
PERSONS	Sometimes free pronouns for all three persons; other times only for 1st and 2nd persons.	There are generally bound pronouns for all three persons.

of free and bound pronouns. When one takes account of all of these properties, it is almost always possible to come to a decision.

We can now briefly discuss properties and parameters of variation for bound pronouns.

(a) Number and function

Some languages allow just one bound pronoun in each clause. Most typically, it relates to A in a transitive and S in an intransitive construction, as in Latin, Spanish, and many further Indo-European (and some other) languages. In a few languages, the single bound pronoun relates to S and O function; this applies to Avar, from the North-East Caucasian family (Anderson 1976: 4 and Simon Crisp, personal communication).

The commonest situation is for there to be two bound pronouns, covering all core arguments. The most frequent type here is for one bound pronoun to mark S and A and the other O function. Sometimes the second bound pronoun can also be used for a particular oblique argument in an intransitive clause. There are languages where one bound pronoun always relates to A and S while the other codes a benefactive argument (shown by dative case on an NP) if there is one, and codes O only when there is no benefactive. Other languages have one bound pronoun for S and O, a second one for A. Kamaiurá (Tupí-Guaraní branch of Tupí family, Brazil; Seki 1990: 369; 2000: 65–9) has split-S marking (§13.5.4): one set of bound pronouns marks A and Sa arguments and another set marks O and So (plus, possessor on a noun).

It will be seen that some of these systems are ‘accusative’ in that A and S are treated in the same way and O differently, while others are ‘ergative’ in treating S and O in the same way and A differently. When there is an accusative system for bound pronouns, this generally correlates with an accusative system of case marking on free pronouns and on nouns (as in Amharic, for example). Sometimes bound pronouns, free pronouns, and nouns all show an ergative system (as in the Australian language Adjnyamathantha). A number of types of combinations of systems are found. Dixon (2002a: 349–50) provides details of the following in Australian languages:

NOUNS	FREE PRONOUNS	BOUND PRONOUNS	EXAMPLE LANGUAGE
ergative	accusative	accusative	Watjarri
ergative	ergative	accusative	Warlpiri

Some languages may have three bound pronouns in a transitive clause, relating to A, O, and an oblique argument (typically marked by dative case in its NP realization). This applies to Basque (Saltarelli 1988: 238–45) and to half a dozen Australian languages, including Ngiyambaa (Donaldson 1980: 131) and Kugu-Muminh (Smith and Johnson 2000: 402–3).

In Abaza, from the North-West Caucasian family, there are two paradigms for bound pronouns: prefixes in column P1 come first in the verb and those in P2 come second (Allen 1956: 153).

(42)	P1	P2
1sg	s-	s-
2sg.masc	w-	w-
2sg.fem	b-	b-
3sg.human.masc	d-	y-
3sg.human.fem		l-
3sg.non-human	y-	a-
3pl		r-
2pl	ʃ ^w -	ʃ ^w -
1pl	h-	h-

An intransitive clause has just one prefix, from the P1 column, marking the S argument. For a simple transitive clause there will be prefixes from the P1 and P2 columns, marking A and O respectively. A clause with three arguments (such as ‘I gave it to her’ or ‘I made her kill it’) includes an initial prefix from the P1 columns and then two from P2.

Allen (1956: 139) provides the only example I am aware of—from any language—of a verb with four bound pronouns. (It is, admittedly, elicited.) The sentence ‘The old man couldn’t make the boy give the girl her dog back’ has NPs ‘old man’, ‘boy’, ‘dog’, and ‘girl’ followed by verb form:

(43)	y-	g ^y -	y-	z-	d-	n-
	3sg.non.human-	NEG-	3sg.masc-	COULD-	3sg.masc-	NEG-
	l-	r-	ótxd			
	3sg.FEM-	MAKE-	give			

The first *y-* prefix (from column P1) refers to the dog, the second *y-* (from P2) to the old man, the *d-* prefix (from P1) to the boy, and the *l-* prefix (from P2) to the girl.

The paradigm shown in (42) also includes bound relative pronouns, which are discussed under (f) in §17.3.1.

(b) Nature and position

A bound pronoun can be realized as an affix, a clitic, or a separate grammatical word. If it is an affix, the bound pronoun will attach to the verb as prefix—as in (36), (38), and (40)—suffix, or infix. Some languages combine prefixes and suffixes. For example, Hoijer (1933: 67) reports that in Tonkawa, an isolate from Texas, object pronouns are prefixes (except 2nd person which involves suffixes) while subject bound pronouns are suffixes. The bound pronouns occur on the rim of the word, as first prefix or as last suffix.

In many languages each bound pronoun is a clitic—a separate grammatical word which has to attach to something else to form a phonological word. They can attach to the verb as in (22) or to some other item, often the first constituent or first word of the clause, as in (41b). Alternatively, bound pronouns may have the status of separate words, such as 1du.exc 'eirau from Fijian in (33–5). These are clearly recognizable as bound pronouns according to the criteria in Table 15.2—they cannot be head of an NP, are not replaceable by nouns, and may not take modifiers. In Fijian, a subject bound pronoun has fixed position as first element in the predicate, while an object bound pronoun must immediately follow the verb. In addition, bound pronouns are obligatory. (Fijian has a quite separate set of free pronouns.)

There can be combinations of types. For example, in Jarawara, 1sg and 2sg subject bound pronouns are prefixes to the verb which follows, while the plural bound pronouns are separate grammatical words placed immediately before the verb (3sg is zero). Compare (noting also (21) in §15.1.6):

- (44) (a) o-kaba 'I am eating'
 (b) otaa kaba 'We (exclusive) are eating'

When bound pronouns are affixes, they may fuse with other affixes. These fusions are of two types:

- A and O are shown by a 'portmanteau' bound pronoun affix. That is, there is a separate form for each combination of A and O values; they are not segmentable and do not relate in a straightforward way to bound pronoun affixes for S function. This can be exemplified with the forms used for 1sg and 2sg in non-future clauses for the Australian language Patitjamalh (Ford 1990: 121–31):

- (45) Bound pronoun prefixes (in non-future tense) for S function:

1sg *ŋa-*

2sg *kañV-* (where *V* copies the following vowel)

Fused bound pronoun prefixes (in non-future tense) for:

1sg A and 2sg O *ŋañ-*

2sg A and 1sg O *ñen-*

- One affix fuses a bound pronoun with information on some of tense, aspect, modality, mood, voice, etc. A prime example is Latin, where the verb form *ama-t* ‘he/she loves’ involves verb root *ama-* and suffix *-t* which codes: 3rd person singular subject argument, plus present tense, indicative mood, and active voice. In Spanish the verb root is again *ama-*, taking zero suffix to show 3sg present indicative: *ama* ‘he/she loves’ (voice is no longer shown through an inflection).

The two types of fusion can, as would be expected, be combined. In Patj-tjamalh the verbal prefixes combine information on A and O arguments and also on tense (future versus non-future).

(c) Complexity of systems

Most frequently, a system of bound pronouns is less complex than the corresponding free pronouns. This can be illustrated for Iraqw (Mous 1993: 112–24, 126, 155 ff.):

(46)

	free pronouns	bound pronouns for O function	subject marking on verb ‘leave’ in present indicative
1sg	aníng	i	a máw
2sg.masc	kúung	u	a méér
2sg.fem	kíng	i	
3sg.masc	inós	u	i máy
3sg.fem		i	i méér
1pl	atén	ti	a mawáan
2pl	kuungáʔ	nu	a meeráʔ
3pl	inoʔín	i	i mayáʔ

There are seven distinct free pronouns. Bound pronouns in O function cannot be modified and immediately precede the verb; we find only four forms here:

- *i* for 1sg, 2sg.fem, 3sg.fem, and 3pl
- *u* for 2sg.masc and 3sg.masc
- *ti* for 1pl
- *nu* for 2pl

The last column of (46) shows that the subject argument is shown by (i) a bound pronominal prefix before the verb, and (ii) a fusion of subject marking, tense, and mood (here present indicative) on the verb root. There are only two markers at (i)—*a* for 1st and 2nd, and *i* for 3rd person. The same form is used under (ii) for 2sg and 3sg.fem, these being distinguished, by combination with *a* and *i* respectively, from (i). Note that a gender distinction is made just in 2sg for free pronouns, in both 2sg and 3sg for bound pronouns in O function, and just in 3sg for the final column of (46).

Most systems of bound pronouns are less complex than that for Iraqw, and they generally show neutralizations not found in the paradigm of free pronouns. The Australian language Djamindjung has a '1/2/3, sg/du/pl, inc/exc' system for free pronouns and for bound pronouns in S function but has a single form covering 2du and 3du and another for 2pl and 3pl for bound pronouns in O function, a neutralization of person (Schultze-Berndt 2000: 64, 85–91; Dixon 2002a: 366). Other languages show neutralization of number. Wambaya, also spoken in Australia, has a '1/2/3, sg/du/pl, inc/exc' system for free pronouns and for bound forms in S function, but just two bound pronouns for O function -*ŋ*- for 1st person and -*ñ*- for 2nd person, irrespective of number (Nordlinger 1998: 126, 139). In Coast Tsimshian (British Columbia; Stebbins 2001) free pronouns show a '1/2/3, sg/pl' system. There are two sets of bound pronouns, one for S and O and the other for A function; each set has a single 3rd person form, the number distinction being neutralized for 3rd person in bound forms. The Austronesian language Tabambo (Vanuatu) has a '1/2/3, sg/pl' system, plus an inc/exc contrast for free pronouns and bound pronouns in object function but not for bound forms in subject function (Jauncey 1997: 102).

In languages with obligatory bound pronouns, these play the major role in identifying core arguments. Free pronouns may be used sparingly—sometimes, like a special kind of proper name—to introduce a new participant, for contrast or emphasis, or in a function for which there is no bound pronoun (typically, copula complement and verbless clause complement). No language is known which has dispensed with free pronouns altogether, but there are some which maintain a minimal set of two forms. Kiowa (Watkins 1984: 115, 100) has a '1/2/3, sg/du/pl' system (with no dual/plural distinction for 1st person) in bound pronouns, but just two free pronouns (making no distinction for number): *nó*: 1st person and *ám* 2nd person. Cayuga (Iroquoian; Sasse 1999: 42, 66) shows a '1/2/3, sg/du/pl, inc/exc' system for bound pronouns but again has just two free forms, 1st person *í*? and 2nd person *í:s* (with no distinction for number or for inc/exc). A further language with just two free pronouns is Acoma Keresan (Maring 1967: 43, 113).

(d) Form and development

At one time, in the history of every language, there were just free pronouns. Bound pronouns developed out of free ones. First of all, pronouns would have to begin to appear in a fixed position, normally within the predicate. They could continue as free words (but as terms in an obligatory system, which cannot be modified, etc.) as in Fijian. Or they could become attached to some specific element, generally the verb—first as a clitic, later taking on the characteristics of an affix.

In the beginning, bound pronouns would have the same form as free pronouns from which they developed. But then—as typically happens with frequently used forms—shortening is likely to occur. For example, the Flinders Island language (spoken off the north-east coast of Australia) has two paradigms of pronouns which include (Sutton n.d.):

(47)	FREE PRONOUNS		ENCLITIC BOUND PRONOUNS	
	S/A FUNCTIONS	O FUNCTION	S/A FUNCTIONS	O FUNCTION
1sg	ŋayu	ŋanini	=yu	=nini
2sg	yundu	yudun	=(n)du	=dun
3sg	ŋulu	ŋuŋun	=lu	=ŋun

It will be seen that each bound pronoun simply drops the initial CV(C) from the corresponding free form (the same rule applies to duals and plurals). The clitic pronouns are attached to the verb after tense–mood suffixes (with A preceding O).

As further phonological reductions apply, bound pronouns are likely to diverge in form and in function from their free pronoun progenitors, sometimes to the extent that a connection is unrecognizable. Sapir (1922: 251) wrote: ‘The independent personal pronouns of Takelma, differing in this respect from what is found to be true of most American languages, show not the slightest etymological relationship to any of the various pronominal series incorporated in noun and verb, except in so far as the second person plural is formed from the second person singular by the addition of element *-p'* that we have found to be characteristic of every second person plural in the language.’ In this language, bound subject pronouns are fused with terms from the tense–mode system (aorist, future, potential, inferential, present imperative, and future imperative), each in several intransitive and several transitive verbal conjugations (Sapir 1922: 157–83).

Bound pronouns may also come to diverge from free ones in their syntactic orientation. The Australian language Warlpiri originally had ergative case marking on nouns and accusative marking on free pronouns. Bound pronouns then evolved, as reduction from the free forms, and retained the accusative profile. As bound pronouns became obligatory, free pronouns were

used sparingly, and they took on the ergative marking system of nouns. That is (Dixon 1994: 96):

(48)	NOUNS	FREE PRONOUNS	BOUND PRONOUNS
Stage 1	ergative	accusative	<none>
Stage 2	ergative	accusative	accusative
Present-day	ergative	ergative	accusative

When bound pronouns are obligatory they are invariably the key indicators of person and number. But free forms never quite fade away. There can be a minimal set of just two free pronouns—1st and 2nd person—as described above for Kiowa, Cayuga, and Acoma Keresan. An alternative is for the original set of free pronouns to be replaced by an entirely new paradigm which involves the addition of bound pronominal affixes to an invariable root. Wemba-Wemba is spoken over a large region in south-east Australia. It has a dozen or more dialects which can for the present purpose be arranged in three groups: (i) far northern, (ii) northern, and (iii) southern. It is instructive to examine the forms of a sample of free and bound pronouns in the northern and in two southern dialects. (Other pronouns behave in a similar way.)

(49)		1sg	2sg
	bound subject pronoun in northern and southern dialects	-an	-ar
	bound possessive pronoun in all dialects	-eg	-in
	free subject pronoun		
	in northern dialects	ŋatj	ŋin
	in southern dialect Tyeddyuwurru	waŋ-an	waŋ-ar
	in southern dialect Tjatjala	yurw-eg	yurw-in

Free subject pronouns in northern dialects are cognate with forms in many other Australian languages. However, in southern dialects these have been replaced by new forms created by adding bound pronouns to an invariable root. The fascinating point is that bound subject suffixes are used in some dialects, such as Tyeddyuwurru, but possessive bound pronouns in others, such as Tjatjala. The invariable root varies from dialect to dialect; in one it is *beŋ-*, a lexeme ‘human being, body’, and in others *win-*, *waŋ-*, *waluŋ-*, *nhuŋ-*, *djurm-*, and *yurw-*, whose lexical meanings have not been identified. (Further examples of pronouns in Wemba-Wemba, and of other newly created paradigms of free pronouns, are in Dixon 2002a: 360–3; 2006e: 88.)

The full array of types of pronouns across the three dialect areas of Wemba-Wemba is:

(50)		free subject pronouns	bound pronouns in core function	possessive bound pronouns
(i)	far northern dialect	original set (regular Australian forms)	<none>	yes
(ii)	northern dialects	original set (regular Australian forms)	optional	yes
(iii)	southern dialects	newly created forms from either subject or possessive bound pronouns	obligatory	yes

It is likely that at one time the system in row (ii) applied across the whole language. In the southern dialects, row (iii), bound pronouns have become obligatory, with the original pronouns being replaced by new forms which involve either subject or possessive bound pronouns. The far northern dialect, Madi-Madi, retains possessive bound pronouns but has lost those bound pronouns which related to core functions in the clause. Why should this have happened? The answer is—probably as a result of areal diffusion. Madi-Madi is surrounded on three sides by languages with no bound pronouns of any sort; it has lost bound pronouns marking core arguments simply to become more like them.

There are two main reasons for the development and for the loss of bound pronouns—one external and the other internal to the language. First, a language may adjust its grammatical structure in order to accommodate to that of languages with which it is in contact. If X is surrounded by languages which use bound pronouns, it is likely to develop its own bound pronouns, not by borrowing forms but from its own resources. Similarly, if X has bound pronouns and moves into a new location so that all its neighbours lack them, it is likely to eliminate bound pronouns from its grammatical inventory. (See the case studies in Dixon 2006e.)

Internal factors may trigger the development of bound pronouns—fixed positioning of pronouns, obligatory inclusion in every clause, phonological reduction of frequently occurring forms, fusion of affixes, and the like. Phonological reduction may also lead to the loss of bound pronouns. For example, the verb in Old English had a portmanteau inflection showing person and number of the subject argument combined with information on tense/aspect and on mood. Phonological attrition at the end of a word has led to all trace of person and number specification being lost, save only for *-s* marking 3sg subject in present tense (and this replaced the original *-eth* only in the sixteenth century).

(e) Possessive bound pronouns

It often seems that languages like to make the fullest use of whatever is available. A category whose major role is with one word class may develop a secondary function with another class. A case system always relates to nouns (and often also pronouns) but in many languages cases can be extended for use with verbs, where they may mark varieties of clause linkage, or add aspectual or modal meanings to the clause as a whole. Tense and aspect markers are typically attached to a verb but in some languages they may also be used with nouns. (See §§11.6–7, Aikhenvald 2008b; and Nordlinger and Sadler 2004.)

If a language has a system of bound pronouns, attached to a verb as realization of a core argument, it may also have a set of bound pronouns which may attach to a noun, marking its possessor. Not infrequently, the same set of affixes or clitics may combine the two functions. Subject pronominal prefixes in Tariana, illustrated in (38–40), also mark the obligatory possessor on a noun referring to a body part or a kin term. Thus:

(38a) nu-ka-ka I looked (51) nu-kawa my leg

Jarawara is like Tariana in that, for a subset of ‘inalienably possessed’ nouns, a possessor must be stated, and this is marked on the noun by the set of forms used to indicate the subject on a verb:

(44a) o-kaba I am eating (52) o-fanako my thigh

There are languages which have bound pronouns just for possession, not for marking core arguments, like the Madi-Madi dialect of Wemba-Wemba. And—rather more often—we find the opposite situation. Patitjamalh has bound pronouns indicating subject and object, illustrated at (45), but no possessive bound pronouns. And where a language does show both varieties of bound pronouns, their forms may be quite different. Iraqw is of this type—compare the forms presented in (46) with possessive suffixes 1sg -ʔéeʔ, 2sg -ók, 3sg -ós, and so on (Mous 1993: 92).

As illustrated for Tariana and Jarawara, if possessive pronouns have similar form to one set marking core syntactic relations, it is likely to be those marking subject (S and A) functions. But not always. In keeping with its split-S profile, Kamaiurá has one set of bound pronouns for A and Sa and another set for O and So; it is the latter forms which are also used to mark a possessor to a noun.

Tunica has two sets of bound pronouns, used for marking different types of possession. A sample of forms, together with the corresponding free pronouns, is (Haas 1941: 37–8):

(53)	BOUND POSSESSIVE PREFIXES		FREE PRONOUNS
	inalienable	alienable	
1sg	ʔi-	ʔi-hk-	ʔi-'ma
1pl	ʔi-n-	ʔi-n-k-	ʔi-'n-ima
2sg.feminine	he-	he-hk-	he-'ma

It can be seen that the alienable possessive suffixes involve *-hk-* added to the inalienable forms (*-n-hk-* becomes *-n-k-* by a regular rule). The free pronouns are also based on inalienable forms, by adding *-(i)ma* in the examples given here.

An interesting feature of Tunica is that the alienable possessive prefixes are also used to mark the O argument with a transitive verb, while the inalienable possessive prefixes mark the S argument with what Haas calls ‘static verbs’ (including ‘be ashamed’, ‘be angry’, ‘want, wish, be willing’, ‘be happy, glad, pleased’, and so on). The remaining verbs (called ‘active’) have their S or A argument shown by a portmanteau suffix also indicating mood, modality, etc.

15.1.10 ‘Conjunct/disjunct’ marking

An unusual phenomenon is found in two groups of languages from opposite sides of the world—a number in the Tibeto-Burman genetic grouping and also languages from the Barbacoan family of Colombia and Ecuador. Basically, there are two verbal affixes (or clitics), a kind of bound pronoun. In a statement, what is called ‘conjunct’ is used when the subject is 1st person, and ‘disjunct’ otherwise. But in a question, conjunct is used for 2nd person and disjunct otherwise. That is:

(54)	CONJUNCT	DISJUNCT
	STATEMENT	1st person 2nd and 3rd persons
	QUESTION	2nd person 1st and 3rd person

Woodbury (1986: 192) suggests an explanation: ‘2nd person forms in questions anticipate the use of 1st person in the answer.’

In the Tibeto-Burman language Kaiké (Watters 2006), conjunct/disjunct marking is fused with tense. Suffix *-pa* marks past-plus-conjunct while *-bo* is past-plus-disjunct. Statements are illustrated in (55) and questions in (56):

- (55) (a) ɲə-i yim doŋ-pa
 1sg-ERGATIVE house make-PAST.CONJUNCT
 I built a house
- (b) na-i yim doŋ-bo
 2sg-ERGATIVE house make-PAST.DISJUNCT
 You built a house

- (c) nu-i yim doŋ-bo
 3sg-ERGATIVE house make-PAST.DISJUNCT
 He built a house
- (56) (a) ŋə-i yim doŋ-bo-yo
 1sg-ERGATIVE house make-PAST.DISJUNCT-QUESTION
 Did I build a house? (I can't remember)
- (b) na-i yim doŋ-pə-yo
 2sg-ERGATIVE house make-PAST.CONJUNCT-QUESTION
 Did you build a house?
- (c) nu-i yim doŋ-bo-yo
 3sg-ERGATIVE house make-PAST.DISJUNCT-QUESTION
 Did he build a house?

In some languages, conjunct marking is only used if the subject exercises active control; a result achieved accidentally by 1st person is described with disjunct marking. We can illustrate this again from Kaike:

- (57) (a) ŋa hoŋ-bo
 1sg fall-PAST.DISJUNCT
 I fell (without meaning to)
- (b) ŋa hoŋ-pa
 1sg fall-PAST.CONJUNCT
 I (made myself) fall

One grammatical role for conjunct/disjunct marking is to provide disambiguation for same versus different subjects in clause combining. This is not unlike the function of logophoric pronouns, in §15.3.4.

There are example sentences including conjunct/disjunct marking at (92) in §16.9.1 and at (62) in §18.5.2.

15.2 Demonstratives

As mentioned at the beginning of this chapter, 1st and 2nd person pronouns are inherently deictic, effectively ‘pointing’ at speaker or addressee (or both). Where they exist, what are called ‘3rd person pronouns’—described as ‘honorary pronouns’ in §15.1—are used primarily for anaphora, referring back to a fully stated 3rd person argument (anaphora and cataphora are the topic of §15.3). A 3rd person pronoun may sometimes—in some languages—be used deictically, but only if accompanied by a pointing gesture. In English, one could say *He is the culprit!*, indicating some person in the vicinity through

pointing with hand or lip and chin. This is, however, a minor function for any 3rd person pronoun.

The class of shifters with deictic reference to some person (or some thing) other than speaker or addressee is ‘demonstratives’. These fall into three well-defined types:

- (a) **Nominal demonstratives**—can occur in an NP with a noun or pronoun (e.g. ‘[this stone] is hot’) or, in most languages, can make up a complete NP (e.g. ‘[this] is hot’).
- (b) **Local adverbial demonstratives**—occur either alone (e.g. ‘put it here’) or with a noun taking local marking (e.g. ‘put it (on the table) there’).
- (c) **Verbal demonstratives** ‘do it like this’, with an accompanying mimicking action—can occur as the only verb in a predicate, or together with a lexical verb.

Types (a) and (b) are found in every language whereas type (c) is relatively rare; see §15.2.1.

In some languages there is a series of manner adverbial demonstratives, such as ‘(do) in this way/manner’. These are almost always morphologically derived from nominal demonstratives, and generally function as non-inflecting modifiers to verbs. Interestingly, in most of the languages in which this occurs, local adverbial demonstratives are also morphologically derived from nominal demonstratives; see §15.2.2. Individual languages can have further specifications. For instance, in Telugu (Dravidian family, South India; Krishnamurti 2001), there is a series of plain manner adverbial demonstratives (‘in this/that manner’) and also a series of causal manner adverbial demonstratives (‘for this/that reason’); these involve different suffixes to the basic nominal demonstrative roots.

In some languages, nominal and/or local adverbial demonstratives have a secondary temporal sense; for example, ‘this’ or ‘here’ may also relate to ‘now’, and ‘that’ or ‘there’ to ‘then’ (which may refer to past or to future, depending on the language). In just a few languages there are separate temporal forms. See §15.2.4.

There may well be further types of demonstratives in some languages. Note that, in English, an angler may boast of the size of a catch either by holding his hands wide apart and saying *It was this big* or *It was so big*, or else by holding up a number of fingers and saying *It was this many* or *It was so many*. It appears that *so* has deictic reference as an alternative to *this* in contexts like this.

Before entering on discussion of the three types of demonstratives, a note on terminology is in order. This concerns terms ‘demonstrative pronoun’ and ‘demonstrative adjective’. In most languages nominal demonstratives can (1) make up a complete NP (as in ‘This is hot’); and (2) occur in an NP with a

noun or a personal pronoun (as in ‘This stone is hot’). Sometimes different—but related—forms are used for (1) and (2) but generally the same forms are employed. There is a tradition of using the term ‘demonstrative pronoun’ for (1); in fact, a nominal demonstrative is in most languages nothing like a personal pronoun in either form or function. And there is a tradition of using the term ‘demonstrative adjective’ for (2); in fact, demonstrative nominals are almost always totally different from adjectives, both in formal categories and in function. A more satisfactory label sometimes used for (2) is ‘demonstrative determiner’. As the discussion below shows, it is most appropriate to sidestep the tradition and just use the label ‘nominal demonstrative’.

Much modern work in linguistics is burdened by the yoke of traditional grammar and Eurocentricism; this applies particularly to work on demonstratives and related items. I have heard it asserted that ‘all languages have demonstratives “this” and “that”’, and that ‘in all languages demonstratives have anaphoric as well as deictic function’. It will be shown below that neither of these *a priori* assertions is correct.

The first and most important step when describing a new language is to find criteria for recognizing demonstratives, and for distinguishing them from other items with similar properties. There is a tendency to use the label ‘demonstrative’ for anything which could be translated by *this* and *that* in English. For an analysis to be valid it must be based on internal criteria in the language under study, rather than on the analysis of translation equivalents in another language.

In §15.2.1 there is discussion of the attested types of demonstratives, and then §15.2.2 deals with formal relationships between different types of demonstratives, and between demonstratives and related items. The functions of demonstratives are discussed in §15.2.3, and their referential possibilities in §15.2.4.

15.2.1 Types

As mentioned above, demonstratives can be divided into three main types, depending on whether they occur in an NP (nominal), as a local adverb to a clause (adverbial), or in a predicate (verbal). We can discuss the types in turn.

(a) Nominal demonstratives

It appears that every language has at least one nominal demonstrative. A nominal demonstrative may always appear in an NP with a noun, and in some languages with a pronoun. In most languages it may also make up a complete NP (the ‘demonstrative pronoun’ function) but this is sometimes not permitted. In Ainu, for instance, nominal demonstratives may not be used

without an accompanying noun (Tamura 2000: 261–2; Refsing 1986: 97, 154–5); see also (78) in §15.2.4.

In many languages where a nominal demonstrative may be a full NP, this usage is relatively infrequent. In the Australian language Yidiñ, for example, less than 10 per cent of the textual occurrences of the demonstratives *yijū* ‘this’ and *guju* ‘that’ are as a complete NP; about 80 per cent of the remainder involve the demonstrative occurring with a noun, as in (58), and about 20 per cent with a pronoun, as in (59) (examples are from texts given in Dixon 1977a: 537).

- (58) ɲaɲji_A waɲi:ɲja-l [yijū dundu]_O
 1pl do.what-NON.PAST THIS head
 What shall we do with this head? (which we have cut off from the
 dead man, after burying the corpse, and which his returning spirit
 has smelled the stench emanating from)

- (59) ɲayu_A [ɲundu:baɲ yijū]_O badja-r=ala
 1sg 2pl THIS leave-NON.PAST=NOW
 I’m now leaving these-you

Note that it is impossible properly to render (59) by English translation; it is literally ‘I’m now leaving this you-all’ (note that number is not marked on demonstratives in Yidiñ).

Other languages, such as English, do not allow a demonstrative to co-occur with a pronoun, just with a noun; and this is the most frequent context of use for nominal demonstratives. In English *this* and *that* can only be used as a full NP in limited circumstances. An NP *this* or *that* which has animate reference is restricted to copula subject in an identity clause, e.g. *That’s my wife*. In other contexts, a demonstrative with animate reference *requires* a following noun, such as the dummy item *one*, e.g. *That one* [animate] *is beautiful*, *I like that one* [animate]. Indeed, although an NP *this* or *that* with inanimate reference may occur in any syntactic function, a following *one* generally makes the sentence sound more felicitous. (*I’ll have that one* is likely to be preferred to *I’ll have that* when, say, pointing to a chocolate bar.)

There are a number of links between nominal demonstratives and 3rd person pronouns, and between nominal demonstratives and articles. As stated before, all languages have 1st and 2nd person pronouns but some lack a 3rd person in the system; nominal demonstratives may fill part of this functional role. It appears that nominal demonstratives may be more likely to have anaphoric function in a language with no 3rd person pronouns than in a language which does have these items. Within the Australian linguistic area, there are a number of examples of a certain form functioning as a

demonstrative in one language and as 3rd person pronoun in another (Dixon 2002a: 306, 335–6).

There can also be obligatory (or almost obligatory) co-occurrence of nominal demonstratives with 3rd person pronouns. Gragg (1976: 178–9) describes how in the Wellegga dialect of Oromo (a Cushitic language spoken in Ethiopia), when a nominal demonstrative does not have an accompanying noun it is generally used with a 3rd person pronoun (perhaps functioning something like *one* in English).

Turning now to articles, there can be a synchronic or diachronic connection with nominal demonstratives. In standard German, the forms *die* (f and pl), *der* (m), and *das* (n) have definite article function when unstressed and demonstrative function when stressed. Old English had two nominal demonstratives, which showed two numbers, three genders, and five cases; from these have developed the modern definite article *the*, and the nominal demonstratives *this/these* and *that/those* (Sweet 1898: 112–5). The definite article in modern French has developed out of the nominal demonstrative *ille* ‘that’ in Latin; and the nominal demonstratives in modern French come from the Latin nominal demonstrative *hic* ‘this’ with strengthening from deictic particle *ecce* (Pope 1934: 322–7).

We also find languages where a nominal demonstrative can co-occur with the definite article. Newman (2000: 143–9) recognizes a definite article in Hausa; this is ‘an enclitic that indicates that the NP to which it is attached is a definite item previously referred to in the discourse or contextually inferable therefrom’. It can co-occur in an NP with a nominal demonstrative, as in:

- (60) wannàn yār’ò-n
 THIS.m.sg boy-DEFINITE.ARTICLE
 this boy (that we were referring to) (lit. this boy-the)

Defining properties of nominal demonstratives are:

- (i) Having deictic function (the defining criterion for an item to be a demonstrative);
- (ii) Occurring in an NP with a noun (the criterion for distinguishing this type from other types of demonstratives).

Other recurrent properties are not criterial. That is, they apply in most—but not all—languages:

- (iii) In many languages, a nominal demonstrative can make up an NP by itself.
- (iv) In many languages, a nominal demonstrative may have anaphoric function (see §15.3).

- (v) In many languages, there are at least two terms in the nominal demonstrative system, with spatial contrast; this is exemplified by English *this* and *that*.

(b) Local adverbial demonstratives

Whereas nominal demonstratives point to an object, local adverbial demonstratives point to a place. There are languages with a single nominal demonstrative (see §15.2.3), but no language is known which lacks two adverbial demonstratives, contrasting in terms of spatial function (like English *here* and *there*) or in terms of visibility.

In some languages, local adverbial demonstratives must occur with a local adposition or case affix, whereas in others they require no marking. There is a tendency for there to be less adpositional or case marking on adverbial demonstratives than on other kinds of locational specification. Thus, in English we get:

(61) LOCAL NP WITH NOUN HEAD	ADVERBIAL DEMONSTRATIVE
He lives at the coast	} He lives there
He lives in the mountains	
He went to the mountains	He went there
He went from the mountains	He went from there

That is, a locative preposition (*at* or *in* or *on*, etc.) and the allative preposition (*to*) are not normally used before *here* and *there*. However, the ablative preposition (*from*) must be retained, to distinguish between allative and ablative specifications, since these may occur with the same verbs (verbs of motion, etc.). (The interrogative *where* behaves like *there* and *here*; see §15.2.2.)

We saw that a nominal demonstrative generally occurs with a noun or pronoun in an NP, although in most languages it can also make up a full NP. A local adverbial demonstrative is most often the sole locational specification in its clause but it can, in most (or all) languages, co-occur with an NP bearing local marking, as in *John lives here in the mountains*. It is probably best to consider *here* and *in the mountains* as distinct, apposed constituents (which must occur next to each other, but in either order), rather than as making up a single constituent. For example, their order can be reversed, giving *John lives in the mountains here* (with a slight difference of meaning). These remarks apply to English; other languages may differ.

In English there is an obvious referential connection between *this* and *here* and between *that* and *there*. This can lead to a temptation to define the adverbial demonstratives in terms of the nominal ones, or vice versa. The

temptation should be resisted, since the two types of demonstrative, while similar in reference, are not equivalent.

Lyons (1977: 646) says: ‘roughly speaking ... “this book” means “the book (which is) here”.’ However, when the noun is changed the equation becomes less viable. For example, *this afternoon* can scarcely be rephrased as *the afternoon (which is) here*. And difficulties arise even when we confine ourselves to nouns referring to concrete things. Compare:

- (62) (a) This hospital has a very fine reputation
 (b) The hospital here has a very fine reputation

These two statements certainly do not have the same meaning. Example (62a) is focusing attention on the hospital in or near which the sentence is uttered, whereas (62b) draws attention to the location of the hospital, that it is ‘here’. There could be an implication in (62b) that the fine reputation is due, at least in part, to the location of the hospital—that it is, say, part of a university system, or situated in a rich suburb.

(c) Verbal demonstratives

A small number of languages lack manner adverbial demonstratives but have a subclass of verbs with demonstrative meaning, involving deictic reference to an action. Verbal demonstratives can be illustrated for the Boumaa dialect of Fijian (Dixon 1988a: 61, examples from text 4, lines 206, 86, and 24, pp. 308–27). They are also reported for the Australian language Dyirbal (Dixon 1972: 56; 2003: 101–3), and for Mapuche, from Chile (Smeets 1989: 424–6).

The Boumaa Fijian verb *'ene(ii)* can be glossed ‘do like this’; it has three basic functions. First, it may have deictic reference to an activity, either actual or mimicked. One story tells how, in a battle, the Prince of Boumaa speared to death many soldiers of the opposing army:

- (63) [o 'ea]_S ['eneii tuu gaa 'eneii]_{PREDICATE}
 ARTICLE 3sg DO.LIKE.THIS ASPECT JUST DO.LIKE.THIS
 He did just like this [narrator mimes a spearing action]

The second function is anaphoric. Earlier in the same story a bridegroom is slain on his marriage bed and the bride cries uncontrollably. The narrator describes the intensity of her wailing and then tells how a relative was told to go and investigate what had happened, since it was the first time that, on the first night of a marriage:

- (64) [saa tagi ti'o 'eneii-maa]PREDICATE [a
 ASPECT CRY ASPECT DO.LIKE.THIS-THAT/THERE ART
 wati-na]_s
 spouse-3sgPOSS
 The wife had continually cried like that

Here the verbal demonstrative *'eneii* accompanies the lexical verb *tagi* 'cry', the whole predicate referring back to the earlier description of uncontrolled wailing.

The third function is to introduce direct speech, as in (from the same text):

- (65) [aa 'eeneii sara 'eneii-qee]PREDICATE
 PAST DO.LIKE.THIS IMMEDIATELY DO.LIKE.THIS-THIS/HERE
 [o Tabu]_s <direct speech follows>
 ARTICLE attendant
 The (Prince's) attendant then spoke like this <direct speech follows>

Looking now at the syntactic possibilities for the verbal demonstrative in Boumaa Fijian, there are again three. First, note that a predicate in Fijian involves a verb which may be preceded and followed by a string of elements, including bound pronouns, tense and aspect markers, and quantifiers.

A verbal demonstrative may be the head of a predicate. One story describing traditional customs concludes with:

- (66) [e 'eneii-mayaa]PREDICATE [a 'e-na iva'arau]_s
 3sgS DO.LIKE.THIS-THAT/THERE ART CLASS-3sg custom
 [i Taveuni]
 ON place

That's the way the custom (of greeting a visiting high chief) is carried out on Taveuni (island) (lit. the custom is done like this on Taveuni).

Alternatively, the verbal demonstrative can be a modifier to the head and is then placed last in the predicate, as in (64), where the predicate head is *tagi* 'cry', preceded by aspectual marker *saa* 'happening now but not in the past' and followed by aspectual element *ti'o* 'continuous' and then the verbal demonstrative *'eneii*.

The third alternative, which is illustrated by (63) and (65), is for the predicate to include two tokens of the verbal demonstrative, one as head and the other as modifier.

Boumaa Fijian has a set of forms which function (without inflection) both as nominal and as local adverbial demonstrative. There are three terms: 'this/here (near speaker)', 'that/there (near addressee or mid-distance)', and 'that/there (far)'. One of these may be added to a verbal demonstrative,

whether functioning as predicate head or as predicate modifier. Thus, we get *-mayaa* ‘that/there (far)’, in (66); *-maa*, a variant of ‘that/there (far)’ from another dialect, in (64); and *-qee* a variant of ‘this/here (near speaker)’ from another dialect, in (65). (Note that speakers of Boumaa Fijian show more dialect mixing in the use of demonstratives than in any other part of the grammar; see Dixon 1988a: 58–9.)

15.2.2 Forms

We can now discuss the ways in which languages have the same or different forms for the various types and subtypes of demonstratives. Table 15.3 presents a sample of the kinds of similarities and differences encountered between (a) nominal demonstratives functioning as a full NP; (b) nominal demonstratives functioning in an NP with a noun or pronoun; and (c) local adverbial

TABLE 15.3. Forms of nominal and adverbial demonstratives, illustrated for ‘this’ and ‘here’

	nominal demonstratives		local adverbial demonstrative	
	as full NP	in NP with noun/pronoun		
I	<u>ko</u> -re celui/celle + -ci	<u>ko</u> -no ce(t)/cette ... + -ci	<u>ko</u> -ko ici	Japanese—Coulmas (1982) French
II	in̩gi	ti		Mangap-Mbula (Austronesian)—Bugenhagen (1994: 94)
III	a-n-a	a-n	a-ŋ	Awa Pit (Barbacoan, Colombia)—Curnow (1997a: 87, 94, 108)
IV	me-t̩ mâ-n	-e-t̩ -ni	ie-t̩ èn	Ponapean (Austronesian)— Rehg (1981: 143–54) Lango (Nilo-Saharan)— Noonan (1992: 86, 108, 334)
V	this		here	English
VI	i-n̩i n̩ih		s-i-n̩i ʔi-n̩ih	Indonesian (Austronesian)— Sneddon (1996: 160, 189) Khmer (Austroasiatic)—Jacob (1968: 142)
VII	d̩ə + s̩ə		s̩ə	Mupun (Chadic)—Frajzyngier (1993: 84–9)
VIII	yai			Boumaa Fijian (Austronesian)— Dixon (1988a: 58)

demonstratives. Each language has two or more items in each set; just the ‘this’ and ‘here’ forms are given in the table.

When a spatial element can be segmented out, it is underlined. For example, in Japanese the syllable *ko-* indicates ‘this/here’ (near speaker); it can be replaced by *so-* ‘that/there’ (near addressee), by *a-* ‘that/there’ (not in conversational space), or by *do-* for interrogatives. In Ponapean the final *-t* indicates ‘this/here’ and is replaced by *-n* for ‘that/there’. In Khmer *nih* is ‘here’ and *nùh* is ‘there’. Mupun marks the spatial dimension by tone—low tone on *sə* ‘here’ and high tone on *só* ‘there’.

Looking now at the sets of rows in Table 15.3:

- I. We here get a different form in each of the three columns. Note that in French the forms are diachronically related but synchronically distinct. This language has an unusual way of marking ‘this’ and ‘that’ with nominal demonstratives. An element *-ci* ‘this’ (related to *ici* ‘here’) or *-la* ‘that’ (identical to *là* ‘there’) attaches to the demonstrative in the first column and to the accompanying noun in the second column; for example *celui-ci* ‘this (m)’, as a full NP, and *cette plume-ci* ‘this pen (f)’. (French also has special forms *voici* and *voilà* which can be used in place of *ceci est* ‘this is’ and *cela est* ‘that is’, respectively (*ceci* and *cela* are alternatives to *celui-ci* and *celui-là*); they have a slightly different grammar and a distinctive stylistic effect; see Ferrar 1967: 214.)
- II. In Mangap-Mbula, one form functions as an adverbial demonstrative and as a nominal demonstrative used with a noun, with a separate form for a nominal demonstrative used alone.
- III. In Awa Pit the nominal demonstratives have basic forms *a-n* ‘this’ and *su-n* ‘that’. When functioning as a full NP the focus marker *na* is added, with reductions *an-na* > *ana* and *sun-na* > *suná*. The adverbial demonstratives are *ay* ‘here’ and *uy* (where *suuy* would be expected) ‘there’.
- IV. In Ponapean and Lango, the nominal demonstrative has its full form when making up a complete NP but when used with a noun it reduces to be a suffix to the noun. In Ponapean the initial consonant is omitted, whereas in Lango the initial syllable *má-* is generally omitted, except that ‘this’ (the form in the table) is irregular, with *mân* reducing to *-ni*. The adverbial demonstrative appears to be paradigmatically related to the nominal in Ponapean, but not in Lango.
- V. Many languages are like English in having one form for the nominal demonstrative (in all contexts) and another for the adverbial.

- VI. Here the adverbial demonstrative is derived from the nominal form. In Indonesian an initial *s-* is added: *i-ni* ‘this’, *s-i-ni* ‘here’; *i-tu* ‘that’, *s-i-tu* ‘there’ (there is a third adverbial demonstrative—see (79) in §15.2.4). Khmer simply forms a compound with the preposition *ʔae* ‘at’, with reductions *ʔae-nih* > *ʔi-nih* and *ʔae-nùh* > *ʔi-nùh*.
- VII. There is here a derivation in the opposite direction. In Mupun nominal demonstratives are formed from adverbial demonstratives by adding an initial *dʔ-*.
- VIII. In the Boumaa dialect of Fijian a single form functions as nominal and as adverbial demonstrative; the full set of forms is at (77) in §15.2.4.

For Koasati (Muskogean family), Kimball (1991: 486) identifies ten nominal demonstratives; five of these are used unchanged as local adverbial demonstratives while two add *-á:li* (the remaining three appear not to have an adverbial correspondent).

There is one other kind of conditioning for demonstrative forms. Whereas the Boumaa dialect of Fijian has a single form for nominal and local adverbial demonstratives, the Bau (or standard) dialect has two sets of forms—one is used for nominal demonstratives and for local adverbials when no preposition is employed, and the other is used after a local preposition ‘at’, ‘to’, or ‘from’ (Churchward 1941: 28).

Where there is a set of manner adverb demonstratives, these are typically derived from the nominal forms. For example, in Mandarin Chinese corresponding to nominal demonstratives *zhè-* ‘this’ and *nà-* ‘that’ there are manner forms *zhè-yàng* ‘this way, like this’ and *nà-yàng(zi)* ‘that way, like that’ (Chappell 2001). In Japanese, the manner adverbial demonstratives *koo*, *soo*, and *aa* correspond to the nominal demonstratives *ko-re* ‘this, near speaker’, *so-re* ‘that, near addressee’, and *a-re* ‘that, distant from speaker and addressee’ (Coulmas 1982). It appears that languages with manner adverbial demonstratives derived from nominal demonstratives typically have local adverbial demonstratives also derived from nominals, e.g. *zhèr* ~ *zhèli* ‘here’, *nàr* ~ *nàli* ‘there’ in Mandarin Chinese; and *koko* ‘here, near speaker’, *soko* ‘there, near addressee’, *asoko* ‘there, distant from speaker and addressee’ in Japanese. (See also Diessel 1999: 74–5.)

We can now look at formal similarities between demonstratives and other items. In a number of languages, interrogatives fall into the same paradigm as demonstratives. In Japanese there is a series of interrogative forms parallel to the demonstratives: nominal *dore* ‘which’, local adverbial *doko* ‘where’ and manner adverbial *doo* ‘in what way, how’. In Tamil the paradigm includes (Asher 1985: 150):

(67)

	DEMONSTRATIVES					
	PROXIMATE		REMOTE		INTERROGATIVES	
nominal	i-nta	‘this’	a-nta	‘that’	e-nta	‘which’
adverbial	i-ŋke	‘here’	a-ŋke	‘there’	e-ŋke	‘where’
temporal	i-ppa	‘now’	a-ppa	‘then’	e-ppa	‘when’
quantity	i-ttane	‘this many’	a-ttane	‘that many’	e-ttane	‘how many’
manner	i-ppati	‘in this way’	a-ppati	‘in that way’	e-ppati	‘how’

Similar paradigms are found in many languages from South Asia, both Indo-Aryan and Dravidian.

Newman (2000: 147) provides a paradigm for Hausa which links together nominal demonstratives (with four degrees of distance), three varieties of interrogative (‘who’, ‘which’, and ‘which one’) plus two indefinites (‘some/other’ and ‘so and so’), each in three genders. Other languages which have interrogatives in the same paradigm as demonstratives include Bengali and Vietnamese.

In English the nominal interrogatives (*who*, *what*, *which*) are quite different in form from the nominal demonstratives (*this*, *that*) but adverbials do show similarity in form: *where*, *here*, *there*. There is also a syntactic similarity—*where* is like *here* and *there*, shown in (61), in that it can omit a preceding allative or locative preposition.

In some languages demonstratives have a quite different pattern of inflection from personal pronouns. In Yidiñ, for instance, 1st and 2nd person pronouns follow a nominative(SA)/accusative(O) system (there is no 3rd person pronoun); inanimate demonstratives have an absolutive(SO)/ergative(A) system (like nouns), while animate demonstratives have a distinct form for each of the three core syntactic functions, S, A, and O (Dixon 1977a: 187).

15.2.3 Functions

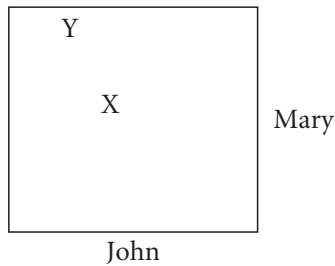
We first discuss the defining properties for demonstratives: (a) deictic reference and (b) syntactic function. And then further possible properties: (c) identification, (d) new information, and (c) discourse organization. Anaphora and cataphora are the subject of §15.3.

(a) Deictic function

We can illustrate this for English, a language which has two nominal demonstratives, *this* and *that*, contrasting in terms of the relative spatial location of their referents.

Consider the following situation. John and Mary are sitting at a table on which are placed two bowls of strawberries, X and Y; bowl X is nearer to both

John and Mary with Y being further away. This can be shown diagrammatically:



Mary offers John a bowl of strawberries. She could offer X or Y, saying in either case *Would you like this one?* Suppose that John prefers the other bowl; his reply would vary, depending on whether he had been offered X or Y. Thus:

- (68) (a) Mary: Would you like this one? [pointing at X]
 John: No, I'd rather have that one [pointing at Y]
 (b) Mary: Would you like this one? [pointing at Y]
 John: No, I'd rather have this one [pointing at X]

In (68a) Mary offers John the nearer bowl and refers to it by *this*; he prefers the further one, and refers to it by *that*. In (68b) she offers him the further bowl and again refers to it by *this*. He again prefers the other one; however, he cannot refer to it by *that*, since it is the nearest to him of the two bowls, and must use *this*. Note that the response in (68b) could be expanded to: *No, I'd rather have this one* [pointing at X] *than that one* [pointing at Y].

Mary uses *this* in each of (68a) and (68b) since in English *this* is typically employed to introduce new information. The spatial sense of *this* only comes into play when there is an explicit spatial contrast between two objects, at different distances from the speaker. In the second utterances of (68a) and (68b), John is comparing X and Y and so uses *this* for the bowl that is nearest to him (X) and *that* for the one which is further away (Y).

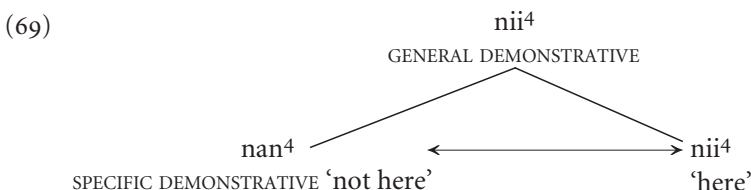
Suppose that you go to the dentist complaining of a sore tooth. The dentist taps a tooth and asks *Is it this one?* You point to a different tooth and say *No, it's this one*. In the case of a speaker's teeth, there is no contrast of 'relative distance from speaker'; thus, only *this* would be used (not *that*) in such circumstances.

In summary, it can be seen that—in deictic use—*this* is the primary nominal demonstrative in English. When only one object is being discussed, *this* is used. When there are two objects which cannot be distinguished in terms of relative distance from the speaker, *this* is used for each. When two objects vary in relative distance then *this* is used for the one nearer to the speaker and *that*

for the one further off. Interleaving with this, in the case of English, is the use of *this* to introduce new information (mentioned under (d) below). It is the ‘new information’ sense which motivates Mary to use *this* for referring to both X in (68a) and Y in (68b).

I would think it unlikely that this account of the use of *this* and *that* in English would apply, point by point, to other languages which have two nominal demonstratives distinguished in terms of distance. Each language needs to be investigated, in its own terms, with cross-linguistic generalizations then being put forward on an inductive basis.

As one further example, Enfield (2003) provides a detailed account of two nominal demonstratives in Lao, with photographic illustrations. ‘The term *nii*⁴ is a semantically general demonstrative, lacking specificity of *any* spatial property (such as location or distance). The term *nan*⁴ specifies that the reference is “not here” (encoding location but *not* distance).’ In effect we have:



That is, *nii*⁴ functions both as a superordinate term without regard to a specific location, and also as the complementary term to *nan*⁴ ‘not here’.

Some languages have a single nominal demonstrative. In Supyire (Gur family, Mali; Carlson 1994: 160–1) there is a single form (marked for gender and number) which is, presumably, used to translate all instances of *this* and of *that* in (68a/b). X and Y would be identified just by pointing. In the English versions of (68a/b) the two bowls are identified by pointing, with concomitant specification by the use of *this* and *that*.

Other languages with a single demonstrative include Dyirbal (see Dixon 2003: 94–9), colloquial Czech (Meyerstein 1972; Cummins 1998), and some dialects of German. It appears that all of the languages with just one nominal demonstrative do have two adverbial demonstratives, similar to *here* and *there* in English. (And see Diessel 1999: 50.)

The adverbial demonstratives in English, *here* and *there*, have parallel deictic use to *this* and *that*. Referring again to the table at which John and Mary are sitting, suppose now that X and Y are plates. Mary holds a cake in her hand and enquires as to which plate she should put it on. There are again two scenarios:

- (70) (a) Mary: Shall I put it here? [pointing at X]
 John: No, put it there. [pointing at Y]
 (b) Mary: Shall I put it here? [pointing at Y]
 John: No, put it here. [pointing at X]

Exactly the same discussion applies as for *this* and *that*. But it should not be assumed that adverbial demonstratives always have the same deictic functions as nominal demonstratives. This is something which must be investigated separately for each language.

(b) Syntactic function

In §15.2.1 we noted that a nominal demonstrative can always occur in an NP with a noun (and, in some languages, also in an NP with a pronoun). In the majority of languages it can make up an NP on its own, but this is not a universal property.

In most languages, an NP including a nominal demonstrative can appear in any core or peripheral function in a clause. But some languages do have restrictions. In Dyirbal, an NP including the nominal demonstrative may only occur in S or O function (the pivot functions in this syntactically ergative language). If a speaker needs to point out an object that is in underlying A function, then the antipassive derivation must be applied, bringing that argument into derived S function. It appears that a similar restriction applies in Warrgamay, Dyirbal’s southerly neighbour (Dixon 1981: 44–5) and also in the Philippines language Northern Subanen (Daguman 2004: 206–8). (See Aikhenvald and Dixon forthcoming.)

§15.2.2 discussed the different forms nominal demonstratives can take, depending on whether they make up a full NP or occur with a noun or pronoun. More rarely, nominal demonstratives can have different forms depending on the syntactic function of the NP in which they occur. For example, Hayward (1990: 273–4) lists the forms of the six nominal demonstratives in Zayse (a Cushitic language from Ethiopia) used for (a) functioning in an NP with a noun; (b) as a full NP in subject function; (c) ditto in object function; (d) ditto in copula complement function. These include (further terms are mentioned under (b) in §15.2.4):

(71)

		as full NP		
	in NP with noun	subject	object	copula complement
near speaker	ha	há(j)	hayá	háytte
near addressee	yi	yíi	yá	yítte
far away	so	só(j)	soyá	sóytte

The syntactic functions of local adverbial and of verbal demonstratives should also be examined, in terms of the clause types they may occur in. There appear to be no restrictions for most languages, but there may well be occasional instances where an adverbial or a verbal demonstrative is not permitted in certain types of subordinate clause. These will only be uncovered if a linguist working on the language is alert to the possibility.

(c) Identification

There are several ways in which nominal demonstratives can be used for identification or recognition (see Himmelmann 1996 and Lindström 2000). For example:

- (72) It was that sort of gluggy rice which the Japanese go in for.

Here *that* identifies what sort of rice it was; note that in this particular use the NP must include a relative clause providing a description of what sort of rice it was.

Some languages have a special demonstrative for identification. For example, *ce* in French can be used as subject of the copula *être* 'be' (or of *pouvoir être* 'can be' or *devoir être* 'must be'). In an example from Flaubert (Ferrar 1967: 215):

- (73) Un beuglement formidable s'éleva. C'était un taureau.
A frightful bellowing arose. It was a bull

Note that Diessel (1999: 78–88) recognizes a special category of 'demonstrative identifiers'—often translatable as 'here/there it is'—with examples from a range of languages.

(d) New information

The nominal demonstrative *this* in English is typically used as a mark of new information. Indeed, this use has in recent years expanded in colloquial speech. For instance, a child may come home from school and say to its parent:

- (74) There's this new girl at school today and she talks really funny

(e) Discourse organization

There are two local adverbial demonstratives in Jarawara, *ahi* 'here, visible' and *fahi* 'here/there, not visible'. These are used in deictic function, and also to mark the development of a discourse.

Fahi is employed to mark the climax of some particular segment of discourse. In one recorded story, the narrator spends sixteen clauses telling how his tape-recorder wouldn't work—he took the back off, fixed the insides, put it back together again, and then (including *fahi* on the final clause of this segment of discourse):

- (75) karafato_O jaro
 tape.recorder (f) start.up
 o-ka-na-ma-hara-ke
 1sgA-APPLIC-AUX-BACK-IMMEDIATE.PAST:EYEWITNESS:f-DEC:f
fahi
CLIMAX
 I switched it back on

Other instances of the ‘climax’ use of *fahi* include: (a) ‘he got hold of the snuff; he rocked the snuff back and forth in his hand; and then he sniffed the snuff *fahi*’; and (b) ‘he found the woman; he grabbed her; he put her down on the ground; and then he copulated with her *fahi*’.

We also find *ahi* as a discourse marker (although much less often than *fahi*). It can mark something that is a ‘lead up’ to a final climax. One or more *ahi* ‘lead up’ clauses are generally followed by a ‘climax’ clause marked by *fahi*.

Mithun (1987) provides an insightful discussion of the discourse function of demonstratives in the Iroquoian language Tuscarora.

15.2.4 Reference

The major parameters of reference for demonstratives are: (a) spatial—sometimes extended to temporal—location; (b) height and stance; and (c) visibility. Other, more abstract, senses are mentioned under (d). In (e) there is brief discussion of person, gender/noun class, classifiers, and number.

All of these kinds of referential information may also be coded at other places in a clause. There may be verbal affixes ‘up’ or ‘down’, or ‘coming’ or ‘going’. Gender, noun class, and number may be marked on other nominal and verbal elements. We are here concerned only with these categories as they are encoded in demonstratives—that is, in forms which can have deictic function.

(a) Spatial reference

The majority of languages have two demonstratives, relating to ‘near speaker’ and ‘not near speaker’. There are also a fair number of languages with a three-term system. This can be one of two types:

- (a) Near speaker; near addressee; not near speaker or addressee—found in Japanese, Basque, Quechua, Swahili, and Thai, among many other languages.
- (b) Near speaker; mid-distance to speaker; far from speaker—found in Georgian, Lango, Ponapean, and Hixkaryana, among many other languages.

In Abui, a Papuan language spoken on the Indonesian island of Alor, the two types of system are combined (Kratochvíl 2007: 162):

(76)

NEAR SPEAKER do	NEAR ADDRESSEE to
MID-DISTANCE FROM SPEAKER (l)o	MID-DISTANCE FROM ADDRESSEE yo
FAR FROM BOTH SPEAKER AND ADDRESSEE oro	

Some languages have a three-term system which is a blend of (a) and (b). For example, the nominal/adverbial demonstratives in Boumaa Fijian are (Dixon 1988a: 58–9):

- (77) yai ‘this/here’, near speaker
yaa ‘that/there’, mid-distance from speaker, often near addressee
mayaa ‘that/there’, far distance from speaker and addressee

If two people are talking together, each will use *yaa* for ‘near you’. But in a narrative *yaa* is employed simply for ‘mid-distance’. (Similar remarks appear to apply for Hopi—see Malotki 1982.)

An important point to note is that the distances involved for ‘mid’ and ‘far’ terms are relative. One story that I recorded in Boumaa Fijian took place entirely within the north-eastern part of the island of Taveuni. The ‘far’ term was used for a place about ten kilometres away from where the narrator was sitting. In another story, the action flowed across from Taveuni to the nearby island of Vanua Levu. The ‘mid-distance’ term was used of a place on the near side of Vanua Levu, about thirty kilometres from where the narrator was sitting, with the ‘far’ term used to refer to a place further away in Vanua Levu, about fifty kilometres from the place of narration.

Ainu has a demonstrative system of particular interest. Quoting from Tamura (2000: 261):

there are the following three kinds of noun-modifying spatial demonstratives.

- *tan*, emphatic form *tapan*. Expresses something that is present, visible, is the new topic of conversation, or that is located where the conversation is taking place: ‘this, here’.
- *taan*. Something that is in the immediate vicinity of the speaker: ‘this, here’.
- *toan*. Separated from oneself: ‘that, there’.

As illustrations of use, Tamura gives:

- (78) tan/tapan kampi 'this letter' (e.g. that one has in possession, or is writing now)
 taan kampi 'this letter' (e.g. on the desk, right next to the speaker)
 toan kampi 'that letter' (e.g. a letter in a letter rack over there)

Some languages have different deictic gestures for relating to varying distances and visibility. In the Tucanoan and Arawak languages of the Vaupés River basin (spanning the border between Brazil and Colombia), for instance, we find: (i) pointing with the lips for 'visible and near'; (ii) pointing with the lips plus a backwards tilt of the head for 'visible and not near'; (iii) pointing with the index finger for 'not visible' (if the direction in which the object lies is known).

In most languages, nominal and local adverbial demonstratives make the same spatial distinctions. However, there are rare exceptions. We saw that the few languages with a single nominal demonstrative appear all to have a spatial contrast for adverbial demonstratives. Other languages exhibiting a difference between the types include:

- (i) As illustrated in Table 15.3, Indonesian has just two nominal demonstratives. There is, however, a three-term system for the adverbials. That is (Sneddon 1996: 160, 189):

(79)	NOMINAL	LOCAL ADVERBIAL
	ini 'this', near speaker	sini 'here', near speaker
	itu 'that', not near speaker	situ 'there', mid-distance from speaker
		sana 'there', far from speaker

- (ii) Yagua (spoken in Peru; Payne and Payne 1990: 374–5, 257, 271) has two nominal demonstratives, *jíyi*- 'this' (near) and *ru*- 'that' (far); these are suffixed with an appropriate classifier. This language has adverbial demonstratives specifying four degrees of distance from the speaker: *jíyu* 'here', *díty* 'there (nearby)', *jásiy* 'there (mid-distant)', and *múúy* 'there (far distant)'.

In these and a variety of other instances, adverbial demonstratives make more distinctions than do the nominal variety. I know of two examples in the opposite direction. Tariana (Arawak family, Brazil; Aikhenvald 2003) has four spatially distinguished terms in the nominal system but only two for adverbials; and Lillooet (Salish family; van Eijk 1997: 168–77) has three nominal

demonstratives but only two local adverbials (each nominal and adverbial form being combined with a visible/invisible specification).

In some languages the spatial reference of nominal demonstratives can, by analogy, be extended to time. For example, Iraqw (a Cushitic language spoken in Tanzania) has four nominal demonstratives: 'near speaker', 'near addressee', 'mid-distant', and 'far'. The two latter are used anaphorically, 'mid-distant' in present tense and 'far' in past tense (Mous 1993: 90–1). For Sarcee (an Athapaskan language spoken in Alberta, Canada), Cook (1984: 73–5) describes a system of three nominal demonstratives; the 'near' term can refer to 'someone who is physically present near the speaker (and hearer) or to somebody being talked about at the narrative moment', and the 'far' term 'refers to someone who is away from the speaker (and hearer) or to someone mentioned in a story on an earlier occasion'. In some languages, local adverbial demonstratives may have a secondary sense referring to time. In others, there may be distinct forms for temporal adverbial demonstratives; this applies for Tamil, as shown in (67). (Anderson and Keenan 1985 provide further examples of spatial demonstratives also being used for temporal reference, in Australian and Austronesian languages.)

I know of only three languages with verbal demonstratives. In both Boumaa Fijian and Dyirbal there is just one verb 'do it like this' (with transitive and intransitive variants in Dyirbal). But Mapuche has two contrasting verbs: *fa*- 'become like this' and *fe*- 'become like that' (Smeets 1989: 424–6).

A topic which has not been systematically studied on a cross-linguistic basis is the iconicity or sound symbolism typically found in demonstratives which distinguish degrees of distance. As a first approximation, the 'near' term is more likely than 'non-near' terms to involve a front high vowel and/or laminal or dental consonant(s). Compare for instance, *this* and *that* in English, *-ci* and *-là* in French, *yigu* and *ɣuɣu* in Yidiñ, *nih* and *núh* in Khmer, *ii* and *aa* in Telugu. This is a topic which will require careful and systematic study. (Diessel 1999: 151–2 provides a useful list of references to the literature on sound symbolism in demonstratives.)

(b) Height and stance

Some languages spoken in hilly country code, onto their demonstratives, reference not only to distance but also to height. In the Arakul dialect of Lak (North-East Caucasian family) there are three nominal demonstratives (Khaidakov 1966: 12):

- (80) aha 'close to speaker'
 hava 'further from speaker, but on the same level'
 ho 'higher or lower than the speaker'

Lahu (Tibeto-Burman, spoken in northern Thailand) distinguishes between higher and lower in its system of five spatial demonstratives (Matisoff 1973: 51):

- (81) chó 'here'
 ô 'there'
 cô 'way over there'
 nô 'up there'
 mô 'down there'

'Uphill' and 'downhill' demonstratives are also found in Hua (Papuan type, spoken in the highlands of New Guinea; Haiman 1980: 258) and in South-East Tepehuan (Uto-Aztecan family, spoken around mountainous ridges in the state of Durango, Mexico; Willett 1991: 187–8). Example (71) in §15.2.3 listed three of the demonstratives for Zayse; the remaining terms are 'equally near speaker and addressee', 'at lower level than speaker', and 'at higher level than speaker' (Hayward 1990: 273).

Other languages indulge in further kinds of spatial specification. For instance, Boas (1911b: 41) mentions an Eskimo system with reference not only to 'near me', 'near thee', 'near him', 'above me', and 'below me', but also to 'behind me', 'in front of me', 'to the right of me', and 'to the left of me'.

In languages of the Waikurúan family (spoken in the Chaco area of Brazil and Argentina) demonstratives must specify the stance or motion of the referent—whether 'standing', 'sitting', 'lying', 'coming', or 'going' (Ceria and Sândalo 1995: 181). Similar specification is made in some languages of the Siouan family, in the set of classifiers which is used only with demonstratives (Rankin 1976, 2004); further examples are mentioned in Aikhenvald (2000: 176–83). Fortescue (1988: 25) describes a nine-term demonstrative system in one variety of Eskimo involving not only distance from speaker but also accessibility (whether or not across a barrier), inside/outside with respect to speaker, and up/down.

(c) Visibility

The criterion used here for a demonstrative is that it must be able to have deictic (or pointing) function. A number of languages have a grammatical system with one or more terms referring to visible objects and also one or more terms referring to something which is not visible. It is natural to ask how a term referring to something which is not visible can be used deictically, and qualify as a demonstrative. There are a number of possible responses. The 'non-visible' term could be used to refer to something which had been visible but has now moved out of sight (over a hill, say); one can point in the direction in which it has gone. Or the 'non-visible' term can refer to something which is audible—round a corner, perhaps, or behind a wall. If we can hear

something we can certainly point in the direction from which the noise is coming. In other instances, the pointing may be metaphorical, with a 'non-visible' term relating to something which is remembered from the past. In 1967, I accompanied some elderly speakers of Dyirbal on an expedition to a traditional site which had not been visited for several decades. They used the 'non-visible' form of the allative adverbial, *ŋalu*; literally 'to that which is remembered from the past'.

It is likely that the labels 'visible' and 'non-visible' have varying implications in different grammars. There is need for detailed study of the meaning of 'non-visible' terms, in languages in which they occur, and then comparison of them. (See Aikhenvald 2004b: 130–1.)

Boas (1911b: 41) drew attention to 'visible/invisible' as an obligatory distinction for demonstratives in Kwakiutl (a Wakashan language). Here it combines with three degrees of spatial distance, yielding a six-term system:

- (82) visible, near me visible, near thee visible, near him
 invisible, near me invisible, near thee invisible, near him

In the Salish language Lillooet there is a similar system with 'visible/invisible' applying for each of three distance terms: 'near', 'at some distance', and 'way over there' (van Eijk 1997: 168–9).

Other languages simply have one 'non-visible' term within the demonstrative system. For example, Shoshone (Uto-Aztecan; Miller 1996: 709) has a four-term system: 'near', 'not quite so near', 'far, but in sight', and 'not in sight, usually far'. Bengali has a three-term system consisting of 'near and visible', 'non-near and visible', and 'invisible' (Onishi 1997: 25–6). Palikur (Arawak family, Brazil; Aikhenvald and Green 1998: 437) has a five-term system: 'in speaker's hand', 'near to both speaker and addressee', 'near to only one of speaker and addressee', 'far from both but visible', and 'far from both and not visible'.

Some languages have complex demonstrative systems, involving a mixture of the parameters discussed here, and more besides (there are further examples in Anderson and Keenan 1985: 286–95). There is a seven-term system in Muna (Austronesian, Sulawesi; van den Berg 1997: 199–201):

- (83) aini 'near speaker'
 aitu 'near addressee'
 amaitu 'away from speaker and addressee, but nearby'
 awatu 'far away, lower than or level with point of speaking or
 orientation'
 atatu 'far away, higher than point of speaking or orientation'
 anagha 'not visible (may be audible), unspecified for time'
 awaghaitu 'not visible, was in view but no longer is'

(d) Other senses

In some languages, demonstratives can take on other kinds of meaning; for example, indicating an emotional attitude, or personal interest, or familiarity.

Quirk and Greenbaum (1973: 107) suggest that in English *this/these* may be used to ‘connote interest and familiarity’, whereas *that/those* may imply ‘a corresponding emotive rejection’. Zandvoort (1975: 148) is more perceptive in stating that: ‘the demonstrative pronouns, especially in their deictic function, are often used with emotional connotation. The kind of feeling implied (affection, vexation, contempt, disgust, etc.) depends on the situation.’ That is, both *this* and *that* may carry either a positive or a negative overtone. Examples include: *I can’t stand that/this mother-in-law of mine* (negative), *These/those modern poets publish a lot* (could be positive or negative), *This headache is killing me* (negative). Example (84), from a newspaper article, uses *that* with distinctly positive connotation:

- (84) Mums and Dads loved his on-screen persona: that nice boy who used to be in *The Andy Griffith Show* was now that nice young man with the sensible, short red hair and freckles

Chinese has two nominal demonstratives, *zhe* ‘this’ and *na* ‘that’. *Zhe* is used for ‘something close to the speaker, either literally in space or time or in their thoughts and present interests’. *Na* is used ‘when the speaker is referring to something in space or time which is at a distance relative to themselves, or when they are thinking of objects or situations removed from their present interests and of small importance to themselves’ (Li 1996: 22).

Mithun (1999: 132–6) provides a fine survey of complex demonstrative systems in some North American languages, involving such features as ‘familiarity to speaker’, and ‘whether stationary or moving and, if moving, whether towards the speaker’. There are also surveys of a number of complex systems in Anderson and Keenan (1985: 280–99) and Diessel (1999: 35–55).

(e) Person, gender/noun class, classifiers, and number

A number of grammatical categories may be realized at several places in the clause, including in demonstratives.

There are two ways in which demonstratives can relate to person. First, they can specify whether an object referred to is near the speaker or near the addressee. Secondly, a nominal demonstrative may actually refer to speech act participants, occurring in an NP with a 1st and 2nd person pronoun. This was illustrated in §15.1. Example (59) is from Yidiñ where there are three demonstratives, *yiju* ‘this’, *ɲuɲu* ‘that (far)’, and *yuru* ‘that’ (‘very far but visible’ in one dialect, ‘non-visible’ in another dialect); in (59), *yiju* ‘this’ co-occurs with 2nd person non-singular pronoun *ñundu:bañ* ‘you all’

Grammarians often do not specify whether or not a demonstrative may co-occur with a 1st or 2nd (or 3rd) person pronoun. It is important to investigate this matter.

In languages with a gender or noun class system, this is frequently marked on nominal demonstratives, as illustrated in (60) from Hausa. In Shilh (Berber branch of the Afro-Asiatic family; Stumme 1899: 92–3), a nominal demonstrative will be specified for gender when making up a full NP, but not when modifying a noun within an NP. And, just as a nominal demonstrative may agree with its (actual or ellipsed) head noun in gender or noun class, it will often also agree with it in number, as in Hausa and Supyire.

If a language has classifiers which are used in a multiplicity of contexts, occurrence with demonstratives is often one of them. There may be a different set of classifiers used with demonstratives than in other contexts, or they may have different forms, different orderings, or different degrees of obligatoriness—see Aikhenvald (2000: 206–41), and especially her table 9.3.

There can be a dependency between the spatial/visibility system and the number system. In Panare (a Carib language from Venezuela; Payne and Payne 1999: 97) there are parallel systems of demonstratives for animate and inanimate reference. The animate set is:

(85)

	SINGULAR	PLURAL
NEAR AND VISIBLE	měj	mějchanton
FAR AND VISIBLE	muku(j)	
INVISIBLE	kěň	kamonton/kaměnton

That is, a number distinction is made for the ‘near and visible’ and ‘invisible’ demonstratives, but is neutralized for ‘far and visible’.

One question that has not so far been mentioned is (formal and functional) markedness within demonstrative systems. For example, which term from a spatially determined system will be used in neutral circumstances, if spatial location is not relevant? Lyons (1977: 647) suggests that ‘generally speaking, in English “this” is marked and “that” is unmarked’. This may apply for anaphoric usage; however, our exploration in §15.2.3 of the deictic reference of *this* and *that*—in examples such as (68a/b)—would point to *this* being the unmarked term in deictic use.

The question of markedness is a difficult one, which may have different solutions depending on whether one is dealing with deictic or with anaphoric

functions. In some languages there may be no markedness within the demonstrative system(s), but in others there certainly is. There appears to be a tendency for the ‘that’ form from a two-term system to be functionally unmarked (it is in Mandarin Chinese and in Telugu, for instance); however, in Tariana the ‘this’ nominal demonstrative (from a four-term system) is functionally unmarked. The question of markedness requires careful examination, in each individual language.

Anaphora and cataphora are prime functions of 3rd person pronouns, and often also of demonstratives. If this applies to just one demonstrative, it is often that one which is functionally unmarked. We can now turn to discussion of anaphora and cataphora.

15.3 Anaphora and cataphora

In a survey of the two hundred or so grammatical accounts of English produced before 1800, Michael (1970: 320) found that more than 80 per cent of the definitions of pronoun ‘are based on the statement that a pronoun is used “instead of a noun”, and more than half of these definitions continue: “... in order to avoid its too frequent repetition”’. In fact this can only describe a 3rd person pronoun, not 1st or 2nd person forms.

As stated before, ‘3rd person pronouns’ may pattern with 1st and 2nd person pronouns in one morphological paradigm, but they do not refer to participants in a speech act and are semantically and syntactically quite different from 1st and 2nd persons. The main function of 3rd person pronouns is to substitute for a full NP in order to avoid repetition of it. There are two ways of achieving this:

- **anaphora** describes reference to something earlier in the text
- **cataphora** describes reference to something later in the text

These can be exemplified by English sentences (with underlining of both the anaphoric/cataphoric element and what it relates to):

- (86) John failed the exam because he [anaphora] hadn’t studied.
- (87) Because he [cataphora] hadn’t studied John failed the exam.
- (88) Mount Vesuvius destroyed Pompeii when it [anaphora] erupted
- (89) When it [cataphora] erupted, Mount Vesuvius destroyed Pompeii

Some linguists used the label ‘forwards anaphora’ for (87) and (89), but most prefer ‘cataphora’. (‘Anaphora’ is sometimes used as a generic term to cover both forwards and backwards reference, in all of (86–9).)

3rd person pronouns may always be used anaphorically, and often also cataphorically. And in most languages nominal demonstratives may have similar roles. The difference between these two grammatical categories can be summarized:

Third Person Pronouns

- (i) In all languages, may have anaphoric (and often also cataphoric) function.
- (ii) In every language, 3rd person pronouns may occur in any core syntactic function (and generally in peripheral functions as well).
- (iii) As pointed out in §15.2, it may be possible to use a 3rd person pronoun deictically (as in *He is the culprit!*) but only when accompanied by a pointing gesture.

Nominal Demonstratives

- (i) The major function is deictic. May be accompanied by a pointing gesture but this is not always necessary; for instance, one could say *That man who is standing behind me is the culprit*, with no pointing gesture at all.
- (ii) In many (but not all) languages, nominal demonstratives may have some anaphoric/cataphoric roles.
- (iii) In most languages, nominal demonstratives may occur in all core syntactic functions, but in other languages their functions are limited. For instance, in Dyirbal, Warrgamay, and Northern Subanen they are restricted to S and O functions.
- (iv) As mentioned under (c)–(e) in §15.2.3, nominal demonstratives are often employed for identification, for introducing new information, and for organization of discourse.
- (v) In some languages, nominal demonstratives may occur in the same NP as a 1st or 2nd person pronoun (this possibility is never available for 3rd person pronouns).

15.3.1 Substitution and textual types of anaphora and cataphora

Two distinct types of anaphora/cataphora can be distinguished:

- **Substitution anaphora/cataphora.** A pronoun or a nominal demonstrative is used as substitute for an NP, as is *he* in (86–7) and *it* in (88–9). In (90) the anaphoric element could involve *this* or *that*, either alone or as modifier within an NP, or pronoun *it*:

(90) He asked for tonic water, insisting that it/this/this drink/that/that drink [anaphora] was the best remedy against insomnia.

- **Textual anaphora/cataphora.** Here the anaphoric reference can be to a clause—such as the complement clause in (92)—or to any stretch of discourse (which may, potentially, be several pages long)—as in (91) and (93).

(91) John hadn't studied and failed the exam and Mary considered it/that/this [anaphora] a terrible shame.

(92) It [cataphora] annoyed Mary that John hadn't studied and failed the exam.

(93) These [cataphora] are the choices available: either study and pass the exam, or become a politician.

Note that *tonic water* could be used instead of *it/this/this drink/that/that drink* in (90), producing an acceptable—if clumsy—sentence. Similarly, *John* could be employed instead of *he* in (86–7), and *Mount Vesuvius* in place of *it* in (88–9). In contrast, the long underlined portions could not be substituted for *it* or *that* or *this* or *these* in (91–3). This is a critical difference between the substitution and textual types.

The possibilities in English for various kinds of 3rd person pronouns, and nominal demonstratives, in substitution and textual anaphora and cataphora, are summarized in Table 15.4, with example numbers illustrating the possibilities. It appears that nominal demonstratives may be used for textual anaphora and cataphora, and for substitution anaphora (but not for substitution cataphora). And that while all 3rd person pronouns may be used for substitution anaphora and cataphora, only *it* is available for textual anaphora and cataphora.

Note that—in all its possible anaphoric and cataphoric functions—a nominal demonstrative may either make up an NP on its own or modify a head noun within an NP. This is illustrated in (90). In (91), the

TABLE 15.4. Anaphora and cataphora possibilities for English

	NOMINAL DEMONSTRATIVES <i>this, these, that, those</i>	3rd PERSON PRONOUNS <i>he, she, they it</i>	
SUBSTITUTION ANAPHORA	(90)	(86)	(88), (90)
SUBSTITUTION CATAPHORA	—	(87)	(89)
TEXTUAL ANAPHORA	(91)	—	(91)
TEXTUAL CATAPHORA	(93)	—	(92)

anaphoric element could be *this/that happening*. And an alternative formulation of (93) is *These choices [cataphora] are available: either study and pass the exam, or become a politician*.

In some anaphoric/cataphoric contexts—such as (90) and (91)—either *this/these* or *that/those* may be used, in others only one of them. Consider:

- (94) *A fool and his money are soon parted, have you ever heard that/*this saying [anaphora]?*
- (95) *Have you heard this/*that saying [cataphora]: a fool and his money are soon parted?*

It appears that when referring to a preceding/following list or saying quoted verbatim, *this/these* should be used for cataphora, as in (93) and (95) and *that/those* for anaphora, as in (94). (See Jespersen 1933: 158.) However, this is only part of the story. Further, detailed work is needed on why both nominal demonstratives are acceptable in some contexts and only one in others (and which and why).

15.3.2 Anaphoric and cataphoric functions of nominal demonstratives

In some languages, nominal demonstratives do not have any anaphoric or cataphoric role. For example, Dyirbal from Australia, Jarawara from Amazonia, and Ainu from northern Japan and an adjacent region of Russia. In languages where they do, it is interesting to relate anaphoric and cataphoric functions to the distance specifications of the demonstratives.

Table 15.5 provides some examples from the literature. (The majority of sources do not specify what type of anaphora is involved; in most cases it is textual anaphora but may sometimes be substitution anaphora.) In both Quechua and Fox the ‘near’ demonstrative is used for cataphora. For anaphora, Fox uses the ‘non-near’ term from a two-term system, and Quechua the ‘mid’ term from a three-term system (the ‘far’ term apparently has neither anaphoric nor cataphoric use). In Cayuga only the ‘non-near’ term is used for textual reference, all of the available examples being anaphoric. Tamil again uses the ‘near’ term for (infrequent) cataphora, and both demonstratives for anaphora. Georgian appears to go against the pattern in using the ‘far’ demonstrative for cataphora (and all three for anaphora). In Basque, just the ‘near’ and ‘far’ terms have anaphoric reference, the ‘near’ demonstrative for something recently mentioned in the text and the ‘far’ term for something mentioned further in the past. (See also §15.2.4.)

As mentioned under (a) in §15.2.3, Supyire has just one nominal demonstrative. This can be used by itself for anaphora, and it is combined with a

TABLE 15.5. Anaphoric and cataphoric functions of nominal demonstratives

NOMINAL DEMONSTRATIVES	USED FOR ANAPHORA	USED FOR CATAPHORA	
near, mid, far	mid	near	Quechua—Weber (1989: 38)
near, non-near	non-near	near	Fox (Algonquian)—Dahlstrom (n.d.)
near, non-near	non-near	—	Cayuga (Iroquoian)—Sasse (1999)
near, non-near	both	near	Tamil—Asher (1985: 80–1)
near speaker, near addressee, far	all	far	Georgian (Kartvelian)—Imedadze and Tuite (1992: 105)
near, mid, far	near: recent anaphora far: distant anaphora	—	Basque—Saltarelli (1988: 213)
one only	demonstrative	demonstrative plus 3rd person pronoun	Supyire (Niger-Congo)—Carlson (1994: 190–4)

3rd person pronoun for cataphora. We saw, in (60), that in Hausa a nominal demonstrative used with the definite article can have anaphoric function.

Verbal and local adverbial demonstratives may also have anaphoric function. The anaphoric use of the verbal demonstrative *'eneii* 'do like this', in Boumaa Fijian, was illustrated in (64).

The anaphoric function of adverbial demonstratives in English is illustrated in:

- (96) John moved to Melbourne in 1959 and lived here [if the speaker is situated in Melbourne] / there [if the speaker is situated elsewhere] for the next ten years

In (96) the demonstrative has both anaphoric and deictic effect. That is, which of *here* and *there* is used in this instance of anaphora depends on the location of Melbourne with respect to the speaker.

Very little work has been done on the anaphoric (and cataphoric) functions of local adverbial demonstratives; this should be a priority for future research.

15.3.3 *Special anaphoric forms*

In most languages, some members of the system of demonstratives—all of which can have deictic reference—may also be used for anaphora or cataphora. In other languages, there is a separate set of forms used just for anaphora/cataphora. For example, in Aguaruna (Jivaroan family, Peru; Overall 2007) the paradigm for four cases is:

(97)	NOMINAL DEMONSTRATIVES			
	NEAR	MEDIAL	FAR	ANAPHORICS
NOMINATIVE	hú-u	an-ú	áu	nú-u
ACCUSATIVE	hú-na	aán-na	áu-na	nú-na
LOCATIVE	hú-ĩ	aan-ĩ	áw-ĩ	nú-ĩ
ALLATIVE	hú-ní	aán	a-ní	nu-ní

Only the near, medial, and far terms should properly be called ‘demonstratives’ (since only they have deictic function). But the anaphoric forms plainly belong in the same morphological paradigm. They can be regarded as a further variety of demonstrative in the same way that ‘3rd person’ forms are regarded as a kind of pronoun, although they do not refer to a ‘person’, in the sense of speech act participant (see §15.1).

The system of seven demonstratives in Muna was set out in (83). It can be seen that all commence with *a-*. This is in fact a prefix which derives a deictic form from an underlying root which, used with no prefix, has anaphoric function. For example (van den Berg 1997: 198–200):

(98)	ANAPHORIC	DEICTIC
‘nearby’	maitu	a-maitu
‘far (lower or level)’	watu	a-watu
‘far (higher)’	tatu	a-tatu

The anaphoric items are used to refer back to something already introduced in discourse and the deictic terms for something in the context of discourse (they are ‘usually accompanied by some kind of gesture such as pointing, looking or nodding’).

Another possibility is for a language to use a reduced form of a nominal demonstrative for anaphoric function. For example, in Longgu (Austronesian, spoken in the Solomon Islands; Hill 1992: 96–7), the deictics *nene* ‘this’ and *nina* ‘that’ are shortened to *-ne* and *-na* respectively when used anaphorically.

15.3.4 Logophoric pronouns

The reference of a 3rd person pronoun, used anaphorically, can only be inferred from study of its discourse context (and, perhaps, the situation of usage). Consider the English utterance consisting of:

- (99) Kofi hadn’t been at the six o’clock meeting.
 We wondered why.
 Then Kofi said that he had left at five o’clock.

The antecedent of *he* in the final sentence can only be Kofi, since that is the only 3rd person argument included in the discourse. And, with that

interpretation, the final sentence provides an explanation for what is reported in the first sentence.

Now consider an utterance consisting of:

- (100) John hadn't been at the six o'clock meeting.
 We wondered why.
 Then Kofi said that he had left at five o'clock.

On purely grammatical grounds, *he* in (100) could be taken as referring back to either *Kofi* or *John*. But in order to provide a coherent interpretation for the utterance as a whole, *he* should be taken as an anaphoric substitute for *John*. Under this reading, the final sentence in (100) again provides an explanation for what was stated in the first, as in (99).

There is a phenomenon called 'logophoric pronouns', found in about thirty languages (from a dozen or so distinct language families) spoken in a 'broad belt immediately south of the Sahara spanning the continent [of Africa] from the Ethiopian Plateau in the east to the Niger bend in the west' (Güldemann 2003: 384). Basically, there are two forms of some 3rd person pronouns—'neutral' and 'logophoric'. A logophoric pronoun, used in a complement clause, refers back to the subject of the matrix clause. For example, in Ewe (Kwa family, Ghana; Clements 1975: 142):

- (101) (a) Kofi be yè-dzo
 Kofi say 3sg.LOGOPHORIC-leave
 Kofi said that he (Kofi) left
- (b) Kofi be ɛ-dzo
 Kofi say 3sg.NEUTRAL-leave
 Kofi said that he (someone other than Kofi) left

There are a number of parameters of variation across the languages in which logophoric pronouns occur. These include the following.

1. Meaning of complement-taking verb (Culy 1994):
 - (a) Some languages only allow a logophoric pronoun in the complement clause of a verb of speaking, such as 'say' in (101a); this may extend to 'ask', 'tell'.
 - (b) Other languages also allow it with verbs of thinking, such as 'think' and sometimes also 'believe'.
 - (c) A further set of languages add verbs such as 'know' and 'see' (but not 'hear').

- (d) Just a few languages also have a logophoric pronoun in the complement clause of a verb of emotion, such as ‘be happy about’, ‘be afraid of’.
- 2. In some languages a logophoric pronoun must be in subject function within the complement clause. In others it may have wider possibilities, including object, indirect object, and even possessor.

Logophoric pronouns may be distinct grammatical words or clitics. And some languages within the ‘logophoric belt’ have an alternative marker—a suffix or clitic which attaches to the complement clause verb, indicating that the 3rd person pronoun in that clause is coreferential with the subject of the matrix clause. (See, for example, Hyman and Comrie 1981 on Gokana, and Ikoro 1996a: 283–8 on Kana.)

It is interesting to enquire why the languages in this geographical region should have the category of logophoric pronouns, when languages from other parts of the world do perfectly well without it. As shown for English in (99–100), discourse context will generally provide sufficient information to indicate the antecedent of a 3rd person pronoun. Ameka (2004) suggests an explanation. There is, in the logophoric belt, a tradition of indirect (or ‘triadic’) communication. If A wants to communicate with B, they are likely to use an intermediary—A tells C who in turn tells B, rather than A speaking directly to B. It is—in large part—to avoid the possibility of ambiguous reference of an anaphoric pronoun in such complex speech situations that logophoric pronouns are needed.

15.4 Summary

1st and 2nd person pronouns, 3rd person pronouns, and nominal demonstratives are interconnected:

(102)	IN ALL LANGUAGES	UNIQUE REFERENCE		ANAPHORA/ CATAPHORA
		TO SPEECH ACT PARTICIPANTS	TO PARTICIPANTS IN VICINITY OF SPEECH ACT	
1st and 2nd person pronouns	yes	yes	—	—
3rd person pronouns	no	no	sometimes	in all languages
nominal demonstratives	yes	no	in all languages	in most languages

'3rd person pronouns' are often in the same morphological paradigm as 1st and 2nd persons and for that reason are labelled as 'pronouns' although they do not represent a 'person' in the strict sense of referring to a speech act participant.

In some languages, all pronouns and nominal demonstratives have the same syntactic profile; in others they differ. Nominal demonstratives, or 3rd person pronouns (or both), may take absolutive/ergative inflection while 1st and 2nd person pronouns show a nominative/accusative system.

Pronouns involve a specification of person—1st and 2nd person in all languages, 3rd in some—generally in combination with a specification of number. Number oppositions can be singular/plural, or singular/dual/plural, or singular/dual/paucal/plural, or singular/dual/trial/plural (with 'plural' having different signification for each type of system). There may be a 'me and you' pronoun with special properties. This can lead to development of an inclusive/exclusive, or a minimal/augmented, opposition. Many languages have limited neutralization of some number contrasts in certain persons, or of some person contrasts in certain numbers. There is often a gender specification in some person/number choices (most often in 3sg). Many languages have special rules concerning use of a non-singular pronoun for singular reference, to mark respected social standing or a particular kin relationship. The label '4th person' has been used, ill-advisedly, for a variety of different features (see §15.1.6).

Pronouns typically occur in all core functions, in some peripheral functions, and as possessors. In some languages they may take modifiers, just like a common noun; in others they may not. There is often a technique for 'elaborating' the reference of a non-singular pronoun; for example 'we-two [and] John' for 'me and John'.

Many languages have, in addition to independent pronouns, a paradigm of 'bound pronouns', which have the form of affix or clitic (typically to a verb) and may fuse with each other, or with a tense, aspect, modality, mood, or voice element. Typical properties of bound pronouns—and criteria for distinguishing them from free pronouns—are summarized in Table 15.2 of §15.1.9. Bound pronouns typically (but not invariably) show fewer distinctions than their free counterparts. 'Conjunct/disjunct' marking involves a special bound pronoun indicating 1st person as subject of a statement, and 2nd person as subject of a question.

All languages have at least one nominal demonstrative ('this') and at least two adverbial demonstratives ('here' and 'there'). A few also have a verbal demonstrative ('do it like this'). All demonstratives may have deictic reference, pointing to some person or thing in the environment. In many languages nominal demonstratives may also be used for identification, for introducing

new information, and in the organization of discourse. Most nominal demonstratives can occur in all syntactic functions, but in a few languages they are restricted (sometimes just to S and O functions). There are languages in which demonstratives are in the same morphological paradigm as interrogatives. Nominal demonstratives typically indicate distance or location, sometimes also height and stance, visibility, etc.

3rd person pronouns are always used for anaphora (relating to something earlier in the text) and often also for cataphora (referring to something later in the text). In most—but not all—languages, demonstratives have similar functions. Anaphora and cataphora can be of two types: substitution, referring to an NP (which could be repeated instead of the anaphoric/cataphoric element), and textual, referring to a clause or a longer stretch of discourse (and this could *not* be used instead of the anaphoric/cataphoric element). Some languages have special anaphoric forms. Logophoric pronouns (found in a swathe of African languages) are a special variety of anaphoric pronoun which indicates that an argument in a complement clause (typically to a verb such as ‘say’) is coreferential with the subject of the matrix clause.

15.5 What to investigate

- A. Establish the paradigm for free pronouns, and also for bound pronouns if the language has these. Decide whether the number system is ‘sg/pl’, ‘sg/du/pl’, ‘sg/du/paucal/pl’, or ‘sg/du/trial/pl’. Investigate whether there is an ‘inc/exc’ distinction for 1st person non-singular pronouns, or whether there is a ‘minimal/augmented’ or ‘minimal/unit-augmented/augmented’ pattern.
- B. There may or may not be 3rd person forms in the free pronoun paradigm. If there are not, do not label nominal demonstratives as ‘3rd person pronouns’ (to fill the gap), as is sometimes done. There may be further terms in the pronoun paradigm, perhaps with indefinite or impersonal reference. It is not helpful to call these ‘4th person’.
- C. Examine whether there is neutralization of any person or number contrasts in the free and bound pronoun paradigms. There may be a gender contrast associated with some person/number combinations; check which they are.
- D. Study the syntactic possibilities of pronouns. Whether a free pronoun may, when head of an NP, take any or all of the modifiers open to a common noun as NP head. And what the functions of pronouns are in the organization of discourse.
- E. Look at the possibility of there being a mechanism for ‘elaborating’ the reference of a non-singular pronoun.

- F. Investigate whether, in addition to free pronouns, there is a grammatical system of bound pronouns. Describe the form of bound pronouns (whether affix, clitic, or separate grammatical word) and their positioning within the clause. Properties of bound pronouns, and criteria for distinguishing them from their free congeners, are set out in Table 15.2. It is important to work out the conditions for use of bound pronouns, in relation to those for free pronouns. Compare the two sets of pronouns in terms of person, number, gender, and other parameters, and types of neutralization.
- G. Describe nominal and adverbial demonstratives; also verbal demonstratives, if the language has these. What are the similarities in form between the types of demonstrative? Are there formal and functional similarities between 3rd person pronouns and nominal demonstratives? Is it possible to recognize one term in a demonstrative system as (formally and/or functionally) 'unmarked' with respect to the others?
- H. What are the deictic and syntactic functions of each type of demonstrative? Can a nominal demonstrative occur (i) in an NP with a noun; (ii) in an NP with a 1st or 2nd person pronoun; (iii) in an NP with a 3rd person pronoun; (iv) making up a full NP?
- J. Describe the parameters of reference for each type of demonstrative, in terms of spatial reference, height and stance, visibility, etc. Are there further senses to demonstratives, such as temporal reference, emotional attitude, familiarity, personal interest?
- K. Investigate the possibilities of substitution and textual anaphora/cataphora for 3rd person pronouns and, if applicable, for demonstratives. Are there any special anaphoric-only forms, or any logophoric pronouns?
- L. Check whether different kinds of pointing or similar gestures relate to the linguistic system in systematic ways.

Sources and notes

15.1. Within the scope of this section, it is only possible to cover the main parameters underlying pronominal systems; there are many variations on the patterns presented here. Bhat (2004) provides an excellent and, on the whole, reliable account. Mithun (1999: 68–79) has good information on North American languages. Dixon (2002a: 243–401) presents a comprehensive account of the nature and development of pronoun systems in Australian languages.

A number of other surveys cannot be wholeheartedly recommended, since they contain numerous errors of quotation and interpretation; these include Forchheimer (1953), Ingram (1978), Cysouw (2003), and Siewierska (2004). As on many other topics, the information provided in Haspelmath et al. (2005) is of mixed quality, being quite often erroneous and inconsistent. For instance, the map on p. 149 shows Acoma as having ‘no independent subject pronouns’. This may be based on looking at Miller’s (1965) grammar, where no independent pronouns are mentioned. However—as stated in §15.1.9—they are clearly identified in Maring’s (1967: 43, 113) grammar of this language.

It is not impossible to devise a scheme whereby 1st and 2nd person pronouns can be considered substitutes for NPs. For example, *The person who is speaking the current utterance is hungry* could be taken as underlying *I am hungry*, and *The person to whom the current utterance is being addressed is tall* as underlying *You are tall*. Such games matter little for an understanding of the nature and function of pronoun systems.

15.1.1. In some languages for which pronouns can be segmented into person and number elements, their number components may also be used with nouns. Dixon (2006: 87–90) describes how—in some dialects of the Western Desert language, from Australia—the original unanalyzable dual and plural pronouns were replaced by new forms created through adding nominal number suffixes to singular pronoun forms.

15.1.2. Haas (1969) is an admirable account of the history of terms ‘inclusive’ and ‘exclusive’. Jacobsen (1980) surveys the inclusive/exclusive contrast in languages of western North America, suggesting that it is readily diffusible. LaPolle (2005) provides a useful survey with respect to Tibeto-Burman languages. He notes that in Dolakha Newar the 1pl.inc ‘me and you’ pronoun (in a ‘1/2, sg/pl, inc/excl’ system), *chaji*, is made up of 1sg *ji* and 2sg non-honorific *chi* (see also Genetti 2007: 130). The volume edited by Filimonova (2005), which deals with the inclusive/exclusive distinction—under the label ‘clusivity’—includes chapters of mixed quality.

Thomas (1955) was a precursor to Conklin (1962), suggesting a minimal/non-minimal analysis for pronouns in Ilocano, although without using these labels. Greenberg (1988) provides a useful list of languages with a ‘minimal/augmented’ pronoun system without recognizing it as such. He refers to Conklin’s (1962) classic analysis but misquotes this, using ‘restricted’ in place of Conklin’s ‘minimal’.

The pronoun systems of a number of Australian languages relate to the relative generation levels of the speech act participants, and in some cases reflect the actual kin relationship between speaker and addressee; see Dixon (1980: 80, 490; 2002a: 70) and further references therein.

15.1.4. A rare example of a gender contrast being reported for every person/number combination is Korana, also called !Ora (South Africa, Khoisan grouping; Meinhof 1930: 43). See Aikhenvald (2000: 252–5) for discussion of gender neutralizations with respect to person.

15.1.5. Information on Japanese is largely taken from Hinds (1986: 238–61). Enfield (2007: 77–84) provides a perceptive account of the use of pronouns according to social level in Lao.

15.1.8. Relating to examples (28–9), the 1sg pronoun in Yidiñ is *ɲayu*, and 1pl *ɲaɲji* is equivalent to *ɲayu* plus *-ba* (there is in fact no form **ɲayu-ba*). However, the 2sg pronoun is *ñundu* and 2pl is *ñundu-ba*, precisely ‘you (singular) plus one or more others’.

Lichtenberk (2000) provides a seminal study of pronoun elaboration in Oceanic and other languages. However, he uses the labels ‘inclusory pronominal’ and ‘inclusory construction’. Elaboration is more common with exclusive than with inclusive pronouns, so that ‘inclusory’ can be a misleading description. And it does not represent a separate construction, rather the addition of an elaborating element to a standard construction.

15.1.9. Some of the discussion here is based on that in Dixon (2002a: 337–401; 2006e). Sentence (43) is glossed by Allen (1956: 139) as ‘The old man couldn’t make the boys give the girl her dog back’. However, from examination of Allen’s paradigms and discussion, it seems that it should be ‘boy’ rather than ‘boys’.

An infixed bound pronoun is rather rare; it is attested for Sorowahá, from the Arawá family; see Dixon (1999a: 304).

Detailed discussion of how Australian languages gain and lose bound pronouns due to areal pressure from neighbours will be found in Dixon (2002a: 379–93; 2006e). Creissels (2005) is an invaluable survey of bound pronouns in African languages, and their stages of development.

15.1.10. ‘Conjunct/disjunct’ has also been called ‘locutor/non-locutor’ and ‘congruent/non-congruent’. There is a succinct survey in Aikhenvald (2004b: 123–30); and see further references therein.

15.2. The discussion of demonstratives here is basically a condensation of Dixon (2003), with some additional material included. There is some valuable material in Diessel (1999). A set of practical questions concerning demonstratives is in Comrie and Smith (1977: 44–6). Wilkins (1999) and Levinson (1999) are useful documents, which were intended as a guide for the fieldworker’s observation of natural usage. But, unfortunately, they have sometimes been used as the basis for elicitation. The study of demonstratives, as of any other aspect of a language, can only properly be conducted through analysis of texts

and by participant observation, with at most minor augmentation through direct elicitation.

Some varieties of English employ a demonstrative-type word *yay*, as in *It was yay big/long*. This is always accompanied by a gesture indicating the (often exaggerated) size.

15.2.1. There is in the literature the report of one language which appears not to have nominal (or local adverbial) demonstratives. According to Dickens (n.d.: 30) the Khoisan language Ju|'hoan simply has two verbal demonstratives, *hè* 'be here, be this one' and *to ·à* 'be there, be that one'. 'When qualifying a noun they must, like any other verb, be preceded by a noun with a relative suffix.' For example, *Jùà hè lūá mí* ('person.who is.here told me'), that is, 'this person (or, the person who is here) told me'. This should be further checked.

Emeneau (1961) describes how in the Dravidian language Brahui, the original 'near' demonstrative base has developed into the base of an enclitic pronoun. Diessel (1999: 115–55) provides a discussion of the diachronic development of demonstratives.

15.2.4. In some languages, demonstratives may also include information on evidentiality; see Aikhenvald (2004b: 130–1).

It does seem that there is some sort of inverse correlation between size of demonstrative system and size of language community. The most complex systems are found in languages with a relatively small number of speakers, while languages spoken by a very large number of people tend to have just a two-term system. For example, within the Dravidian language family, demonstrative systems with three or four members are found in the small tribal languages whereas each of the four languages with tens of millions of speakers (Telugu, Tamil, Malayalam, and Kannada) shows just two terms. At an earlier stage, English had a three-term system (*this/here*, *that/there*, and *yon/yonder*) but the third term was lost as English developed into a world language. Spanish and Portuguese are said to have three or four demonstratives but in daily usage there is just a two-term contrast (see, for example, Jungbluth 2001). (I am grateful to D. N. S. Bhat for discussion of this matter, and for drawing attention to the Dravidian data.) See also Diessel (1999: 160–1) and references provided there to the relevant literature.

The volume edited by Kita (2003) includes a number of insightful studies of deictic gestures, particularly the chapters by Haviland and Wilkins.

15.3. Halliday and Hasan (1976) provide a classic study of anaphora and cataphora in English.

15.3.1. Lyons (1977: 667) restricts the label ‘anaphora’ to what I call ‘substitution anaphora’ and employs ‘textual deixis’ for ‘textual anaphora’. Levinson (1983: 83) adjusts the latter label to ‘discourse deixis’. (Use of the term ‘deixis’ is so confused that it seems wisest not to use it here.)

15.3.3. Terrill (2003: 172–90, 205–6) provides interesting information on demonstratives and anaphora in Lavukaleve (Papuan area; Solomon Islands).

Possession

16.1 Introduction

The term ‘possession’ is used to cover a wide range of relationships. Every language has—in its grammar—a ‘possessive construction’ within an NP; for example *the surgeon’s knife* or *the trunk of the tree* in English. There is cross-linguistic variation concerning what kind of person, animal, or thing may be the possessor, what may be the possessed, and what kind of possessive relationship is involved. Also, as to whether there is some formal marking on the possessor, or on the possessed, or on both, or on neither (possessor and possessed then being simply apposed).

Since ‘possessor’ and ‘possessed’ begin with the same letter, it is convenient to abbreviate reference to them in terms of their final letter:

R—possessor

D—possessed

In English, the possessive construction within an NP can be expressed as *R’s* *D* or as *(the) D of R* (see §16.7.1 for discussion of when to use these alternatives). It spans a wide range of relationships, not all of which could be said to involve actual possession. The range includes:

A. Ownership (and also temporary possession):

- (1) [John’s car] runs smoothly
- (2) We have to vacate [our house] since the lease has expired

B. Whole–part relationship, whether a body part of a human or animal, or a part of an object:

- (3) [Mary’s teeth] hurt
- (4) [The door of the cabin] won’t close

C. Kinship relationship, whether blood relations or relations through marriage, as in:

(5) [My mother] is dead

(6) [Mary's husband] works in a bank

D. An attribute of a person, animal, or thing:

(7) [John's temper] is terrible to behold

(8) [The age of that fossil] is indeterminate

E. A statement of orientation or location:

(9) [The front of the van] is dented

(10) [The inside of the Easter egg] is full of toffee

F. Association:

(11) [Paul's dentist] lives in Perth

(12) [Carol's ancestral village] was destroyed by the volcano

Nominalizations employ a construction which is superficially very similar to that for possession:

G. A nominalization, which can relate to an object, location, or activity, etc.:

(13) [John's discovery] won him the prize

(14) [The refugees' settlement] is on high ground

(15) [The kidnapping of their prince] stunned the nation

In some other languages, the NP-internal possessive construction has a similarly wide range of meanings. For instance, Nikiforidou (1991: 153) lists twelve uses for the genitive construction in Classical Greek. These include A–D and G of the senses just illustrated for English, plus more besides. But, although it appears that every language does have an NP-internal possessive construction, it seldom covers such an extensive semantic range as in English and Classical Greek.

Can a semantic definition of 'possession'—as this is coded through a grammatical construction—be provided cross-linguistically? The answer is clearly in the negative. Indeed, a definition (which does not include any *or*'s) of what 'possession' is in a single language seems scarcely feasible.

There are three central semantic relationships which are typically covered by an NP-internal grammatical construction of possession. These are A, ownership, B, whole–part relationship, and C, kinship relationship; illustrated for English by (1–6). They are quite dissimilar in nature.

- A. **Ownership.** The description *John's car* states that, in the social context within which he lives, John (the R) is acknowledged to have control over a car (the D). It may be that he has full ownership—John can drive the car, and he may sell it. Or he may have only temporary possession, as when the car is leased to him—John can drive it, but he may not legally sell it.
- B. **Whole–part relationship.** Here, the ‘whole’ is always coded as R and the ‘part’ as D. In some languages the possessive construction is only used for body parts of humans (such as *Mary's teeth*). In other languages, it may be extended to apply to parts of animals (*the monkey's tail*, *a bird's feathers*), and of things (*the door of the cabin*).
- C. **Kinship relationship.** There are two subdivisions:
- (a) **Blood** (or consanguineal) **relationship**—a kinship link involving childbirth. For example, *John's mother* or *Jane's grandfather*. There is always a reciprocal relationship. If Mary is John's mother then John is Mary's son; if Tom is Jane's grandfather, then Jane is Tom's granddaughter; and so on. Of the two people in such a reciprocal relationship, either can be coded as R with the other then being D.
 - (b) **Affinal relationship**, involving marriage; such as *Mary's husband* or *John's mother-in-law*. Reciprocity applies once more—if Fred is Mary's husband then Mary is Fred's wife, and if Kate is John's mother-in-law then John is Kate's son-in-law. As with consanguineal links, either of the two people in an affinal relationship can be coded as R within a possessive construction, with the other being D.

It can be seen that there is little in common to the three types of relationship that are typically treated within grammar through an NP-internal possessive construction. Yet there is constancy across languages in that it is the owner, under A, and the whole, from B, which are coded as R. That which is owned could not be R, nor could a part. Then we have kinship relationships, which can be approached from either direction and for which either of the two people involved in a relationship may be coded as R.

Many grammars do code ownership, whole–part relationship, and kinship relationship in the same way. However, as shown in §1.3, a number of languages have several possessive constructions. There can be one covering A, ownership, and B, whole–part relationships, with a different construction for C, kinship. Or one for A and C and a second construction for B. Or one for B and C and another for A. Or three distinct possessive construction types—one for each of A, B, and C. All this is discussed further in §16.5.

And an NP-internal possessive construction may be extended to mark further kinds of relationship, such as attribute, orientation/location, and

association, illustrated for English in (7–12) above. These could be referred to as further varieties of possession. In many languages, the mark on R in a possessive construction is also used in a nominalization, as shown for English in (13–15). It is *not* appropriate to describe nominalization as a type of possession, but rather to note that this shows a further function of a grammatical marker used in a possessive construction.

In every language, most grammatical elements bear a variety of functions. As will be shown at many places in the chapter, this is particularly true of possessive markers. In English, for instance, preposition *of* is also used for NP-internal specifications such as quantity/collectivity (as in *two cups of tea*, and *a bunch of bananas*) and material (*a house of straw*, and *the crown of gold*). These are quite distinct from possession. *Of* is also used to introduce an argument following various adjectives (*aware of*, *afraid of*) and verbs (*consist of*, *approve of*). In Classical Greek, genitive is also used to mark the Standard in a comparative construction—‘better Plato-GENITIVE’ for ‘better than Plato’. (Nikiforidou 1991: 153 mentions that this would be coded by ablative case in Latin.)

Besides an NP-internal possessive construction, each language also has some means for stating a relationship of possession, through a ‘predicative’ possessive construction. In English this is achieved through the verb *have*, which takes the possessor (R) as its argument in A syntactic function and the possessed (D) as its O argument. Corresponding to (1) and (2), one can say:

(16) John has a car

(17) We have a house

In each NP-internal possessive construction, illustrated by (1–12), the possession is assumed as known information—in (4) that the cabin has a door, in (6) that Mary has a husband, and so on. Rather than *presupposing* a relationship of possession, a ‘have’ construction *establishes* one.

Less than half the world’s languages include a verb like *have* for asserting that a certain relationship of possession holds. Whereas the NP-internal construction typically covers A, ownership, B, whole–part relationships, and C, kinship relationships, the only sense common to ‘have’ constructions in all languages is just A, ownership. Individual languages show varying extensions to this. Languages lacking a verb ‘have’, but with a copula construction, often utilize this for establishing possession. (16) might be rendered as, literally, ‘A car is to John’ or ‘John is with a car’. Another technique is to employ an intransitive verb ‘exist’—‘John’s car exists’. Some languages resort to other means. A full survey of grammatical techniques for establishing a possessive relationship is in §16.9.

Returning now to English, a *have* construction is likely to be employed only for statement of some significant fact. Not everyone has a daughter. As a consequence, one can felicitously say:

- (18) Mary has a daughter

Speakers would be unlikely to employ a 'have' construction for describing some possession or attribute which every relevant possessor would be expected to have. In normal circumstances people do not say *John has a father* or *Mary has teeth* or *Paul has a dentist*. However, if some extra information were added, then a 'have' construction becomes acceptable. For example

- (19) John has a rich father
 (20) John has a father in jail
 (21) Mary has no teeth
 (22) Mary still has her teeth (whereas everyone else in the family has lost theirs)
 (23) Paul has a good dentist
 (24) Paul does now have a dentist (it took him a long time to find one)

In similar fashion, one would not expect to hear *The Easter egg has an inside* or *That fossil has an age*. But when further information is included, a 'have' construction becomes felicitous:

- (25) The Easter egg has a hollow inside
 (26) That fossil has an indeterminate age

Corresponding to (7), one can say *John has a temper* since not every person has one. An alternative is *John has a bad temper*.

In English, a 'have' construction can be used for any of meanings A–F associated with the NP-internal possessive construction, if there is need to state some significant fact concerning possession.

Nominalizations are rather different in that they do not occur in a 'have' construction. Instead, the underlying verb is used to establish a relationship:

- (27) John discovered an inland sea
 (28) The refugees settled in a place which is on high ground
 (29) Their prince has been kidnapped

This confirms that nominalizations, such as those illustrated in (13–15), should not be regarded as a kind of possession. Similarly, expressions of material and of quantity/collectivity—although also shown by an NP-internal of construction—cannot be rephrased with ‘have’. Corresponding to *a house of straw* one cannot say **the straw has a house*, and corresponding to *two cups of tea* and *a bunch of bananas* one cannot say **the tea has two cups* or **the bananas have a bunch*. This confirms that these fall outside the gamut of ‘possession’, even in its widest interpretation.

The bulk of this chapter is concerned with a cross-linguistic study of NP-internal possessive constructions. How these are marked is discussed in §16.2. The parameters of variation relating to R, to the nature of the possessive relationship, and to D are dealt with in §16.3–5. Following these, §16.6 discusses how markers of NP-internal possession often have further function in the grammar, typically marking the functions of arguments at clause level. Illustration of complex (and irregular) marking for possession is in §16.7, together with discussion of *’s* and *of* in English. The topic of §16.8 is the internal syntax of NPs which include possession—what is head of the NP? Then §16.9 deals with ‘have’ and other techniques for stating the establishment of a possessive relationship. As with other chapters in the volume, there is a summary in §16.9 and then a final section, §16.10, which offers suggestions concerning ‘what to investigate’ when working on possession, within the writing of a comprehensive grammar of a language. The appendix provides a brief overview of possessive constructions in Fijian, which take account of the nature of R, of the possessive relationship, and of D.

16.2 Marking an NP-internal possessive construction

How can one tell that an NP involves a possessive construction—that it includes two components: a possessor (R), which can be a noun or pronoun or a full NP, plus the possessed (D), which is most often a noun and can be modified by adjective(s), relative clause, etc.? Sometimes R and D are simply apposed within the NP. More frequently, a morphological process applies to R or to D or to both. These possibilities will be discussed in turn.

The first alternative involves no marking at all, save for the ordering of elements. We find R followed by D in Angami (Tibeto-Burman, Nagaland; Giridhar 1980: 48):

- | | | | |
|------|------------|------|------------------|
| (30) | â zēù | (31) | mīzā phi |
| | 1sg friend | | table leg |
| | my friend | | leg of the table |

The opposite ordering, D followed by R, is found in Indonesian (Sneddon 1996: 144):

- | | | | |
|------|--------------|------|------------|
| (32) | rumah Tomo | (33) | mobil saya |
| | house Tomo | | car 1sg |
| | Tomo's house | | my car |

A similar 'D R' construction is found in Acehnese, another Austronesian language from Indonesia (Durie 1985: 109–10). In all three of these languages the R may be a noun or a cardinal pronoun (there are no special possessive pronouns).

The most common pattern is to have a morphological marker on R or on D or on both. The following labels will be used in this chapter:

marker on R—genitive (the familiar term)

marker on D—pertensive (based on the Latin verb *pertinēre* 'to belong')

'Genitive' has sometimes been used for a marker on D, as well as for that on R, which can lead to confusion. It is because of the absence of an established label that 'pertensive' is introduced. As pointed out under (f) in §1.10, the designation 'case' is best reserved for marking function within a clause; it is not helpful to extend it to mark function within an NP, by genitive, pertensive, comitative, etc.

In Mandarin Chinese, genitive enclitic =*de* is added to the end of an NP expressing the R (Li and Thompson 1981: 113–14):

- | | | | |
|------|--------------|----------|--|
| (34) | wǒ=de | chènshān | |
| | 1sg=GENITIVE | shirt | |
| | | my shirt | |
-
- | | | | | |
|------|-----------------|-----------------------------|-----------|------|
| (35) | [nèi=ge | fāndiàn] | =de | cài |
| | THAT=CLASSIFIER | restaurant | =GENITIVE | food |
| | | the food of that restaurant | | |

Note that in Mandarin, a noun modified by a demonstrative must involve a classifier, here the classifier =*ge*.

Whereas the genitive enclitic is just added to the last word of the NP in R function for Chinese, in Dyirbal the genitive suffix -*ŋu* is added to every word in an NP in R function. For example:

- | | | |
|------|----------------------|-------------------|
| (36) | ŋali-ŋu _R | yabu _D |
| | 1du-GENITIVE | mother |
| | | our (dual) mother |

- (37) [ba-ŋu-l yara-ŋu midi-ŋu]_R guda_D
 DETERMINER-GEN-MASCULINE man-GEN small-GEN dog
 the small man's dog

Note that determiner *ba-ŋu-l* adds masculine marker *-l* after genitive *-ŋu*. In Dyirbal, every word in an NP must take the case suffix appropriate to the function of that NP in its clause. For example, when NP (37) is placed in A function within a sentence—as at (38)—ergative suffix *-du* ~ *-ŋgu* is added to each word of the NP. (A linking suffix *-(n)jin* is inserted between genitive and a case inflection.)

- (38) η ayguna [ba- η u-l- η in-du yara- η u-n η in-du
 1sgO DET-GEN-MASC-LINK-ERG man-GEN-LINK-ERG
 midi- η u-n η in-du]_{POSS} guda- η gu]_A baja-n
 small-GEN-LINK-ERG dog-ERG bite-PAST
 The small man's dog bit me

The words in (38) may be permuted in any order, since Dyrbal not only has free constituent order (order of phrases in a clause) but also free word order (order of words in a phrase and also order of words in a clause). The function of each word in (38) is clear from the suffixes it takes. In contrast, for a language like Mandarin which just adds the genitive clitic at the end of the NP in R function, words in this NP must be kept together.

In some languages there is a perturbative marker on D, to indicate that what we have here is a possessive construction. Karbi (Tibeto-Burman, Assam; Jeyapul 1987: 78) uses prefix *a-*:

- | | | | | | | | |
|------|------------------|----------------|--|------|---------|------------------|--|
| (39) | tebul | a-keŋ | | (40) | la | a-hem | |
| | table | PERTENSIVE-leg | | | 3sg | PERTENSIVE-house | |
| | leg of the table | | | | his/her | house | |

Genitive and perturbive marking need not, of course, be just by affixes or clitics; other morphological processes may be available. In Igbo, R and D are apposed in that order but the tones on each may be adjusted to mark that they make up a possessive construction (Emenajio 1978: 75; Ikoro 1996c: 30–1). M̃yky, a language isolate from Brazil, employs initial mutation as perturbive marking on the D (Montserrat 2000: 162–3). If the first segment of the noun in D function begins in *t*, *p*, or *k*, it is palatalized, as in (41). If the first segment is *m* or *m'*, this is replaced by *k* or *k'*, as in (42). If the noun in D function begins with a vowel, then either *k* or *y* is added before it, as in (43).

- (41) Tapura PLUS poku 'bow' GIVES Tapura p^yoku 'Tapura's bow'

(42) *sēi* ‘2sg’ PLUS *mʷuku* ‘tooth’ GIVES *sēi kʷuku* ‘your teeth’

(43) *are* ‘1sg’ PLUS *atohu* ‘sieve’ GIVES *are yatohu* ‘my sieve’

In Angami and M̐ky (and also in Igbo), there is no paradigm of possessive pronouns, the cardinal form of a pronoun being used in possessive constructions.

The genitive and pertensive markings illustrated in (34–43) simply indicate that a certain constituent is in R or D function within an NP-internal possessive construction. Quite a number of languages have a more sophisticated kind of pertensive marker—an affix which (i) shows that the noun to which it is attached is in D function, and (ii) specifies the person and number of the R. That is, it is a pronominal pertensive affix. In Fijian, if R is a pronoun or common noun and D is a ‘bound noun’ (inalienably possessed), then a pronominal pertensive suffix is added to D:

D-pertensive.suffix.showing.person.and.number.of.R NP_R

For example:

(44) [a ulu-na]_D [a cauravou]_R
 ARTICLE head-3sgR.PERTENSIVE ARTICLE youth
 the youth’s head (lit. head-his, the youth/youths)

(45) [a ulu-dra]_D [a cauravou]_R
 ARTICLE head-3plR.PERTENSIVE ARTICLE youth
 the youths’ heads (lit. head-their, the youth/youths)

And, exemplifying when R is 1st or 2nd person:¹

(46) [a ulu-mu]_D ([o i’o]_R)
 ARTICLE head-2sgR.PERTENSIVE (ARTICLE 2sg)
 your (sg) head

Here, *a ulu-mu* provides a full specification. The 2sg free pronoun (plus article) *o i’o* can optionally be included in (46), in the same slot as *a cauravou* in (44–5), as a kind of emphasis.

Other languages have the components in an opposite order, literally ‘the youth head-his’; this occurs in Buru (Austronesian, Indonesia; Grimes 1991: 190–2). And, within either order, the pronominal pertensive marker can be prefix rather than suffix.

We have examined two kinds of pertensive affix. There is the plain variety which just marks a noun as D—this will be shown in glosses just as ‘pertensive’. And the kind which includes information about person and/or number of the R—this person and number information is in the gloss. In Hixkaryana

(Carib, Brazil; Derbyshire 1979: 68–70; 1985: 199–200) both types of pertensive attach to a noun in D function, the plain variety as suffix and the pronominal pertensive as prefix:

- (47) Waraka y-owa-ni
 Waraka 3R-chest-PERTENSIVE
 Waraka's chest (lit. Waraka his-chest-POSSESSED)

With a 1st or 2nd person as R we can get, for example, *r-owa-ni* 'my chest', where *r-* is the 1sg pertensive prefix.

A different kind of double marking is found when there is a genitive affix on R and also a pronominal pertensive affix on D, as in Bolivian Quechua (Crapo and Aitken 1986: vol. 2, 13):

- (48) runa-q alqu-n
 man-GENITIVE dog-3sgR
 the man's dog (lit. man's dog-his)

'My dog' is *alqu-y* ('dog-1sgR'). Number specification on D involves suffix *-kuna* added after the pertensive suffix; for example *alqu-y-kuna* 'my dogs'. The same kind of genitive plus pronominal pertensive marking is found in Turkish (Kornfilt 1997: 185).

It was pointed out in the last chapter (at (b) in §15.1.9) that in some languages a pronominal affix to a transitive verb may fuse information concerning both A and O arguments. In similar fashion, one affix to D may combine pronominal information concerning both R and D. In Aleut, the same fused suffixes mark A and O arguments for a transitive verb, and R and D elements in an NP-internal possessive construction (Leer 1987: 11). For example suffix *-ning* may be added to a verb root, such as *aguqa-* 'make', indicating 1sg A and 3pl O, as in (49). Or the same suffix can be added to a noun root, such as *ukina* 'knife', and indicates 1sg R and 3pl D, as in (50).

- | | | | |
|------|---|------|--|
| (49) | aguqa-ning
make-1sgA.3plO
I made them | (50) | ukina-ning
knife-1sgR.3plD
my knives |
|------|---|------|--|

We next turn to parameters of variation for NP-internal possessive constructions relating to the possessor (R), the nature of the possessive relation, and the possessed (D), in §16.3–5.

16.3 The nature of the possessor (R)

In some languages, an NP-internal possessive construction is marked in the same way whatever the nature of R. This applies to some of the languages

TABLE 16.1. Different marking of possessive constructions according to the nature of the possessor

	I Yagua	II Tialo	III Hua	IV Jersey Norman French	V Awa Pit	VI Fijian
pronoun	A	A		A		D
proper noun		B	A		A	B (or A)
kin term				B		
common noun human	B					D
animate		C	B	C	B	E (or D)
inanimate						E

discussed in the last section—Angami, Acehnese, Mandarin Chinese, Karbi, and M̐ky. In Dyirbal, genitive suffix *-ŋu* (~ *-nu*, ~ *-u*) marks alienable possession for all nouns and non-singular pronouns; as described in §15.1.7, there are special possessive forms just for singular pronouns.

But in many languages, there are two or more alternative marking schemes depending on the nature of the possessor—whether it be a pronoun, a proper name, a kin term, or some other common noun which is human, non-human animate, or inanimate. Table 16.1 illustrates this for six languages.

I. In Yagua (Peru; Payne and Payne 1990: 348) we find:

- A. If R is a pronoun then a pronominal possessive prefix (identical to the A/Sa pronominal prefix on verbs) attaches to D; for example, *sa-rooriy* (3sgR-house) ‘his/her house’.
- B. If R is not a pronoun, then R and D are simply apposed; for example *Tomáása rooriy* ‘Tomáása’s house’.

A fair number of languages are like this, with genitive marking just on pronouns. Note that we never find the reverse—genitive marking on everything but pronouns. That is, if there is genitive marking on all types of nouns, there will also be genitive marking on pronouns.

II. In Tialo (Austronesian, Indonesia; Yoshimura 1993):

- A. If R is a pronoun then there is a pronominal pertensive suffix to D; for example *soobuan-oqu* (friend-1sgR) ‘my friend’.

- B. If R is a proper noun or kin term, then marker *ni* comes between D and R:

(51) [si ama]_D ni Hairune_R
 NOMINATIVE father POSSESSIVE Hairun
 Hairun's father

- C. If R is any other common noun, the possessive construction is marked by pertensive suffix *-nu* (after a vowel) ~ *-u* (after a consonant) attached to D:

(52) labong-u_D soobuan-oqu_R
 house-PERTENSIVE friend-1sgR
 My friend's house

Note that in (52) the R is itself a possessive construction.

- III. In Hua (Gorokan family, Papua New Guinea; Haiman 1980: 238), for alienable possession:

- A. If R is a pronoun, proper name or unaffixed kin term, then R and D are simply apposed; for example, *Buro' fu* 'Buro's pig'.
 B. For all other R, genitive suffix *-ma'* is added to R; for example, *de-ma' fu* '(the) man's pig'.

- IV. In the Saint-Ouen variety of Jersey Norman French (Liddicoat 1993):

- A. If R is a pronoun, then its possessive form is placed before D; for example, *mā frɛð* 'my brother'.
 B. If R is a noun with human reference (including proper nouns and kin terms), possessive marker *a* comes between D and R, as in:

(53) [lɛ:ð ɛ:fã:z]_D a
 ART.DEF.PL children POSSESSIVE
 [ma fil]_R
 1sg.POSSESSOR.FEMININE daughter
 My daughter's children

- C. If R is a noun with non-human reference, the possessive marker is *d*. For example:

(54) [l ɛ:gji:ð]_D d
 ART.DEF.SING.FEM church POSSESSIVE
 [la pɑ:rɛ:s]_R
 ART.DEF.SING.FEM parish
 The church of the parish

V. Awa Pit (Barbacoan family, Colombia/Ecuador; Curnow 1997a: 123):

- A. If R is a pronoun, or a noun with human reference, it bears genitive enclitic *=pa*; for example, *Santoas=pa pimpul* ‘Santos’s leg’. Singular pronouns fuse with *=pa*; 1sg *na* plus *=pa* gives *ap* (for example, *ap pimpul* ‘my leg’) while 2sg *nu* plus *=pa* gives *up*.
- B. If R is a noun with non-human reference, then R and D are simply apposed, as in *kwizha pimpul* ‘dog’s leg’.

VI. Like Yagua, Fijian has two classes of D, roughly ‘alienable’ (free nouns) and ‘inalienable’ (bound nouns). There are, in all, five types of possessive construction; these are described in the appendix to this chapter. Column VI summarizes the four different constructions for types of R when D is a free noun.

Column VI is included in Table 16.1 to show that there is no universal hierarchy relating to R. In columns I, II, and IV, pronouns behave in one way and proper names and kin in another. But in column VI, pronouns are grouped with kin terms and with common nouns with human reference, being treated differently from proper nouns.

16.4 The nature of the possessive relationship

A possessive construction may encompass a range of meanings. *John’s picture*, in English, can refer to a picture which John owns, or one which he painted, or one which someone painted of him—or any two of these, or all three; see also the end of §16.7.1. The relation between R and D may be in the present or only in the past; it may be temporary or permanent, close or distant; or it may relate to the reason for the possessive relation holding. Some of the parameters relating to a possessive relationship will be briefly surveyed.

(i) **Temporal.** In Apalai (Carib, Brazil; Koehn and Koehn 1986: 85–6), R can be realized either by a pronominal pertensive prefix to D, or by an NP preceding D (not by both). For example:

- (55) *i-kyry-ry*
 3sgR-thing-PERTENSIVE
 his/her possession (lit. thing)
- (56) *nohpo kyry-ry*
 woman thing-PERTENSIVE
 the woman’s possession (lit. thing)

It will be seen that in addition to a pronominal prefix, the noun in D function takes a plain pertensive suffix. There are in fact two choices here—suffix *-ry*

~ -*ny* refers to present possession, and -*ny* followed by - \bar{V} *pyry* to something possessed in the past:

- | | |
|---|--|
| (57) <i>y-tapyi-ny</i>
1sgR-house-PERTENSIVE
my house | (58) <i>y-tapyi-ny-\bar{V}pyry</i>
1sgR-house-PERTENSIVE-PAST
my former house |
|---|--|

Further examples are included within a general discussion of tense marking on NPs, in Nordlinger and Sadler (2004: 780–3, 788–9).

(ii) **Temporary/permanent.** If John owns something (say a boomerang) and lends it out to Tom, then Tom has it in his temporary possession. Dyirbal has two genitive suffixes which are added to R and distinguish these situations:

- (59) *Tami-ŋu wajal* Tom's boomerang (which he has in his present possession)
- (60) *Jani-mi wajal* John's boomerang (which he owns, but does not have in his present possession, and which he can reclaim)

What may be called the 'simple genitive' suffix -*ŋu* occurs in all dialects and can be regarded as the default marker of possession. 'General genitive' suffix -*mi* is found only in central and southern (not in northern) dialects and is used to indicate special kinds of possession. These include: (a) for something which the acknowledged owner does not have in their current possession, as in (60); (b) to describe something (such as a hut) temporarily abandoned by its owner; (c) for something which has been lost by the owner and found by someone else; (d) to describe something which belonged to someone who is now deceased (*ŋuma-mi wajal* 'the boomerang which belonged to father, who has died'). Full details and exemplification are in Dixon (1972: 105–10).

In West Greenlandic (Eskimo; Fortescue 1984: 172) 'the nominal affix *ut(i)* expresses personal possession that is alienable or temporary. It usually refers to something acquired and/or disposable, most commonly ... a catch or supply of game, skins, etc, relating to the hunting way of life; in such cases it is obligatory ... The affix appears to have been extended in use in recent years to include the notion of legal ownership (or other forms of formally or tacitly recognised possession) e.g. for houses and girlfriends. Thus ... *saviutaa* "the knife he has (now)/owns" may be opposed to *savia* "his knife", regarded as an inalienable part of a hunter's equipment.' And see example (101) below.

(iii) **Closeness of relationship.** In Warekena (Arawak, Brazil; Aikhenvald 1998: 293–7), alienably possessed nouns take a pronominal pensive prefix coding the R, and also a plain pensive suffix. Two of the latter refer to

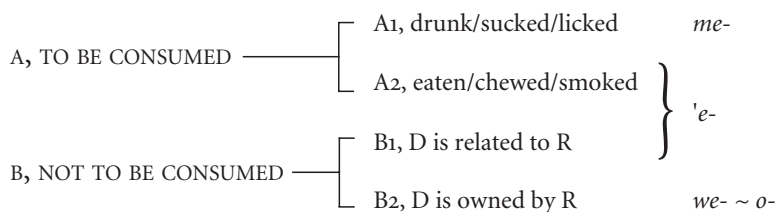
distance—‘suffix *-ne* is used when the object is closer to the speaker, and the suffix *-te* is used when it is at a certain distance from the speaker and is closer to the hearer’. Thus:

- (61) nu-waru-ne
1sgR-parrot-PERTENSIVE.CLOSE
my parrot (right here)
- (62) pi-waru-te
2sgR-parrot-PERTENSIVE.NOT.CLOSE
your parrot (over there)

The neighbouring and related language Baniwa has cognate pertensive suffixes which relate not to spatial but to emotional distance. For example (Aikhenvald 2000: 143):

- (63) nu-tʃinu-ni
1sgR-dog-PERTENSIVE.CLOSE
my dog (this one I brought up)
- (64) nu-tʃinu-te
1sgR-dog-PERTENSIVE.NOT.CLOSE
my dog (the one I found)

(iv) **General type of possession.** Many languages belonging to the Oceanic subgroup of Austronesian have a set of classifiers used in certain kinds of possessive construction. Fijian is typical in that two quite different principles are involved. First, if something is to be consumed, the mode of doing so is indicated—roughly, whether drunk or eaten. Then, for things not to be consumed the classifier shows whether the D is owned by the R or relates to the R in some other way. In summary:



Interestingly, A2 and B1 are shown by the same form, 'e-. The contrasts can be illustrated:

- (65) a *me-na* *waqona*_D
ARTICLE A1-3sgR kava
his kava (to be drunk)

- (66) *a* *o-na* *wagona_D*
 ARTICLE B2-3sgR *kava*
 his *kava* (to be sold)
- (67) *a* *o-na* *italanoa_D*
 ARTICLE B2-3sgR *story*
 his *story* (which he tells)
- (68) *a* *'e-na* *italanoa_D*
 ARTICLE B1-3sgR *story*
 his *story* (which is told about him)
- (69) *a* *o-na* *da'ai_D*
 ARTICLE B2-3sgR *gun*
 his *gun* (which belongs to him)
- (70) *a* *'e-na* *da'ai_D*
 ARTICLE B1-3sgR *gun*
 his *gun* (which will be used to shoot him)

The appendix to this chapter provides an integrated account of possessive constructions in Fijian; see also Dixon (1988a: 135–40, 119–29).

There are useful surveys of the relational classifiers in Oceanic and other languages in Pawley (1973), Lynch (1973), Lichtenberk (1985), and Aikhenvald (2000: 133–43). The last source also draws attention to the set of twelve relational classifiers in Kiriri (Brazil; Mamiani 1877: 59–62; Rodrigues 1997) which relate to obtaining food (by gathering, cultivation, etc.), to preparing food (by boiling, roasting, etc.), and to types of acquisition of goods (finding, sharing, receiving as gift, or taking as booty).

16.5 The nature of the possessed (D)

It is well known that in many languages nouns can be divided into two sets, called ‘alienably possessed’ and ‘inalienably possessed’, according to how they function as D within an NP-internal possessive construction. However, as pointed out in §1.3, these labels are used with a range of denotations. The semantics of the two sets (or three or even four sets, in a few languages) and the different markings associated with each are addressed in this section.

An ‘alienable/inalienable’ distinction is lacking from the familiar languages of Europe—those from the Indo-European and Finno-Ugric families, as well as Basque—from many major languages of Asia and Africa—including Japanese, Korean, Khmer, Vietnamese, Swahili, Akan, and Igbo—and from a

number of minor languages around the world. There is a definite tendency for the distinction to occur most often in languages spoken by relatively small speech communities. (There are, however, notable exceptions—a distinction between two sets of nouns on the basis of their occurrence as D is found in Mandarin Chinese and in Ewe.)

In a number of languages, body-part terms may behave in a different way from other nouns. For example, in French one would say *Il m'a cassé le bras* for 'He broke my arm', but for 'He broke my stick' it would have to be *Il a cassé mon bâton* not **Il m'a cassé le bâton* (Hyman, Alford, and Akpati 1970). That is, a body-part noun such as *bras* 'arm' takes the definite article *le* and requires accusative pronoun *me* ~ *m'* within the predicate, whereas a non-body-part noun such as *bâton* 'stick' is preceded within its NP by *mon* 'my'. (Hyman, Alford, and Akpati report a similar phenomenon in Igbo.) Such special grammatical behaviour of body-part terms is not strictly a function of NP-internal possessive constructions and lies outside the scope of discussion here.

16.5.1 Semantic basis

Table 16.2 summarizes—for eight representative languages—the two sets of nouns in relation to how they function as D within an NP-internal possessive construction. Set I covers nouns which, as D, are perceived to be closely related to R; set II includes all others. The table is arranged according to the six types of possessive relationship, A–F, exemplified for English by (1–12) at the very beginning of this chapter. In some cases the source materials provide a full list of set I nouns (108 for Amele and 175 for Jarawara). For other languages all the information given in the sources is summarized in the table. Boas (1911a) referred to set 1 in Tsimshian as 'inseparable' and set 2 as 'separable'. Other sources use newer terms—for each of Amele, Slave, Panare, Creek, Lango, Ewe, and Yidiñ, set 1 is referred to as 'inalienable' and set 2 as 'alienable'.

However, the semantic content of set 1 varies widely. In some of the languages in Table 16.2, set 1 includes all body-part terms, in some only a portion of them. In some languages all kin terms are in set 1, in others only a selection are. We find that D, Attributes, is represented in set 1 for three languages and E, Orientation/Location, in five. Set 1 in Yidiñ includes a full set of part terms but no kin, while set 1 in Ewe has all kin but no body parts.

The label 'inalienable' may be appropriate to describe set 1 in each individual language but, cross-linguistically, this label for a grammatically defined class has no invariable semantic content. There are also languages for which nouns divide into three sets with respect to their functioning as D; three of these are illustrated in Table 16.3. And Table 16.4 shows Dakota, which has four

TABLE 16.2. Languages with two sets of nouns for D function

	set 1, D closely associated with R	set 2, others
Amele (Gum family, Papua New Guinea; Roberts 1987: 139, 171–5)	B. Part of human or animal—‘ear’, ‘heart’, ‘penis’, ‘tears’, ‘tail’, ‘wing’ C. Kin—‘daughter’, ‘husband’, ‘brother-in-law’ D. Attributes—‘shame’, ‘pride’, ‘bravery’, ‘humility’, ‘fame’, ‘wealth’ E. Location—‘space behind a person’ (only one in list) F. Association—‘name’, ‘friend’, ‘shadow’, ‘dirt’	A. Ownership
Slave (Athapaskan, Canada; Rice 1989: 226–34)	B. Part of human or animal—‘eye’, ‘intestines’, ‘excrement’, ‘blood’, ‘tears’, ‘scab’, ‘leaf’ C. Kin—‘elder sister’, ‘daughter’, ‘mother’, ‘uncle’ F. Association—‘name’, ‘nest’, ‘goods/possessions’	B. Some ‘body parts’—‘excrement’, ‘scab’ A. Ownership
Panare (Carib, Venezuela; Payne and Payne 1999: 90–3)	B. Part of human—‘shoulder’, ‘tongue’, ‘heart’, ‘blood’ C. Kin—‘father’, ‘wife’ F. Association—‘name’, ‘friend’, ‘shaman’, ‘clothing’, ‘pestle’	A. Ownership
Tsimshian (British Columbia; Boas 1911c: 392–3)	B. Part of human—‘nose’, ‘belly’ C. Kin—‘father’, ‘sister’ E. Orientation/Location—‘the place behind’, ‘the place near’ F. Association—‘master’, ‘friend’	A. Ownership

(Cont.)

TABLE 16.2. (*Cont.*)

	set 1, D closely associated with R	set 2, others
Creek (Muskogean; Martin 1993)	<p>B. Most parts of humans and animals—‘hand’, ‘liver’, ‘bone’, ‘tail’, ‘wart’</p> <p>C. Most kin—‘father’, ‘mother’, ‘sister’, ‘aunt’</p> <p>E. Orientation/Location—‘in front of’, ‘behind’, ‘inside of’, ‘next to’</p> <p>F. Association—‘thing’, ‘lord/master’</p>	<p>B. Some body parts—‘rib’, ‘penis/vagina’, ‘sweat’, ‘urine’</p> <p>C. Some kin—‘child’, ‘grandchild’, ‘mother-in-law’ (and special terms for ‘father’ and ‘mother’ used just by children)</p> <p>A. Ownership</p>
Lango (Nilotic, Uganda; Noonan 1992: 77–83; 156–9)	<p>B. Part of human—‘head’, ‘tongue’, ‘liver’, ‘blood’</p> <p>C. Blood relations—‘sister’, ‘father’, ‘mother’s brother’</p> <p>D. Attributes—‘character’, ‘bad luck’</p> <p>E. Orientation/Location—‘beside’, ‘on’</p> <p>F. Association—‘name’, ‘picture’, ‘owner’, ‘house’</p>	<p>C. Affinal relations—‘wife’</p> <p>A. Ownership</p>
Ewe (Kwa, Ghana and Togo; Ameka 1996)	<p>C. Kin—‘grandfather’, ‘wife’</p> <p>E. Orientation/Location—‘middle of’, ‘top of’, ‘inside surface of’</p> <p>F. Association—‘name’, ‘friend’, ‘suitor’, ‘homeland’</p>	<p>D. Attributes—‘character’</p> <p>B. Part of human and animal—‘face’, ‘brain’, ‘urine’; and part of object—‘branch (of tree)’, ‘handle (of hoe)’</p> <p>A. Ownership</p>
Yidiñ (Australia; Dixon 1977a: 357–64)	<p>B. Part of human, animal, plant, or thing—‘eye’, ‘heart’, ‘penis’, ‘tears’, ‘tail’, ‘leaf’, ‘handle (of shield)’</p> <p>D. Attributes—‘fever’, ‘cramp’, ‘toothache’</p> <p>F. Association—‘name’</p>	<p>C. Kin</p> <p>A. Ownership</p>

TABLE 16.3. Languages with three sets of nouns for D function

	set 1	set 2	set 3
Koyukon (Athapaskan; Thompson 1996)	B. Parts of human, animal, plant—‘head’, ‘dandruff’, ‘antler’, ‘(fish) scales’, ‘branch’ F. Association—‘track’, ‘den’, ‘clothes’	C. Kin—‘father’, ‘husband’	B. Bodily fluids and solids:—‘blood’, ‘tears’, ‘urine’, ‘faeces’ A. Ownership
Jarawara (Arawá family, Brazil; Dixon 2004a: 295–301, 310–35, 343–60)	B. Part of human, animal, thing—‘ear’, ‘liver’, ‘penis’, ‘sweat’, ‘tail’, ‘egg’, ‘leaf’; plus general terms such as ‘piece of’, ‘hole in’ D. Attributes—‘itch’, ‘smell’, ‘being angry’ E. Orientation/Location—‘top surface of’, ‘inside of’ F. Association—‘name’, ‘companion of’, ‘ornament’, ‘food’, ‘path’, ‘grave’	C. Kin—‘mother’, ‘younger sister’, ‘grandmother’, ‘male cross-cousin’, ‘husband’	B. Body part—‘vagina’ (a compound, literally ‘thing eye’) A. Ownership
Gapapaiwa (Austronesian, Papua New Guinea; McGuckin 2002)	B. Part of human—‘hand’ C. Kin who are not in authority over ego: spouse, younger same-sex sibling, child, grandchild, in-law D. Attributes: feelings E. Orientation/Location—‘on top of’, ‘inside’, ‘beside’	D. Attributes: sickness, physical characteristics and character traits F. Association: clothing to wear, food to eat	C. Kin who are in authority over ego: parent, grandparent, elder same-sex sibling A. Ownership (including clothing to sell, food to sell)

TABLE 16.4. Languages with four sets of nouns for D function

	set 1	set 2	set 3	set 4
Dakota (Siouan; Boas and Deloria 1941: 127–33)	B. ‘Body parts conceived of as particularly subject to will-power’—‘mouth’, ‘lips’, ‘eye’, ‘arm’, ‘hand’, ‘foot’, ‘penis’, ‘ear (as sense organ)’, ‘spirit’	B. Other body parts—‘nose’, ‘shoulder’, ‘knee’, ‘hair’, ‘ribs’, ‘liver’, ‘lungs’, ‘blood’, ‘bone’	C. Kin terms	A. Ownership

sets. For Dakota, Boas and Deloria (1941: 127–33) refer to sets 1 and 2 as variants of ‘inseparable’, set 3 simply as ‘terms of relationship’, and set 4 as ‘separable’. For Koyukon (in Table 16.3), Thompson (1996) describes set 2—which is just kinship nouns—as a type of ‘inalienable’, linked to set 1. Dixon (2004a: 295) regards set 2 in Jarawara—which again consists just of kin terms—as ‘essentially a subtype of alienable possession’, akin to set 3. In the case of Gapapaiwa, McGuckin (2002) uses the label ‘inalienable’ for set 1, ‘semi-alienable’ for set 2, and ‘alienable’ for set 3.

In every language there will be some nouns for which it is unlikely (perhaps impossible) that they should function as D within a possessive construction. This is basically a matter of common sense (in terms of the speaker’s world view) but it can be hard to circumscribe exactly which nouns are ‘unpossessable’. Someone who has a house which nestles beneath a mountain may well refer, in English, to *my mountain* (this falls under F, Association). An Aboriginal Australian may use ‘my mountain’ to refer to their conception site, from where their spirit is believed to have come at birth and where it will return when they die. Nevertheless, some grammars do make useful comments on this. Concerning Dakota, Boas and Deloria (1911: 128) state: ‘natural objects like land, water, animals including the dog but excepting the horse cannot take the possessive pronoun, because under aboriginal conditions they could not be exclusive property of anyone.’ And for Koyukon, Thompson (1996: 651) says: ‘Unpossessable nouns cannot have an overt possessor without some morphological modifications. Terms for animals, people, and loan words are typical of this class.’

The set including A, Ownership, is always an open class consisting of free nouns (which can be used without any marking of a possessor)—this is set 2 in Table 16.2, set 3 in Table 16.3, and set 4 in Table 16.4. The remaining sets are

all closed classes, with limited membership. In some—but not all—languages their members are bound forms which cannot occur without some marking of a possessor. Typically, an ownership (alienable) possessor may function as the common argument of a relative clause construction, whereas an inalienable possessor may not; see the end of §17.2.2.

We can now survey, in turn, the five types of possessive relationship, B–E, which may fall within set 1.

B. Whole–part relationship. Seven subvarieties can be distinguished:

- (i) external body parts such as ‘eye’, ‘head’, ‘arm’, ‘foot’, ‘belly’
- (ii) internal body parts such as ‘heart’, ‘liver’, ‘brain’, ‘bone’, ‘rib’
- (iii) genitalia, chiefly ‘penis’ and ‘vagina’
- (iv) bodily fluids—‘blood’, ‘tears’, ‘sweat’, etc., and also ‘faeces, excrement’
- (v) parts of animals, such as ‘tail’, ‘feather’, ‘scale (on fish)’
- (vi) parts of plants, such as ‘branch’, ‘leaf’, ‘root’, ‘flower’
- (vii) parts of artefacts and other objects, such as ‘handle’, ‘point (of arrow or spear)’, ‘wheel’

Full information is not available for every language in Tables 16.2–16.4 but we do know that in some languages—including Yidiñ—all of (i)–(vii) belong to set 1. This set in Amele appears to cover only (i)–(v). There are languages where genitalia are placed in a different class from other body parts; this applies in Creek (and in Jarawara just ‘vagina’ is in set 3). Bodily fluids (and faeces) are excluded from set 1 for Creek and Koyukon.

For Dakota, Boas and Deloria (1941: 128) suggest that those body-part terms which belong to set 1 are ‘conceived of as particularly subject to will-power’. However, the semantics appears not to quite fit the grammar; one wonders why ‘knee’ is in set 2. In Slave there are just a few terms, such as ‘excrement’ and ‘scab’, which can be used in either a set 1 or a set 2 possessive construction. In a number of languages, including Lango, a body-part term is in set 1 while it is attached to the owner, but is treated as a set 2 term once it becomes detached—for example, a leg taken off or blood taken out of an animal.

Kiowa is particularly interesting. Its possessed nouns fall into an ‘inalienable’ set, covering kin terms, and an ‘alienable’ set, dealing with ownership. Watkins (1984: 1–2) reports that ‘possessive affixes do not occur with body parts. The possessor is indicated instead by a patient prefix’. For example, ‘My head hurts’ is rendered by, literally ‘Head it:me-hurts’.

Many Australian languages are like Yidiñ in having genitive suffix added to R when D is a set 2 noun, but simply apposing R and D when D is set 1. Set 1 covers just body parts, ‘name’, and illnesses. Thus:

- | | |
|---|--|
| (71) <i>wagu:ja-ni_R</i> <i>guda:ga_D</i>
man-GENITIVE dog
man's dog | (72) <i>wagu:ja_R</i> <i>jina_D</i>
man foot
man's foot |
|---|--|

The head of the NP in (71) is the D component, *guda:ga* 'dog', while in (72) the head is the R component *wagu:ja* 'man'. But it is really not helpful to refer to (72) as an instance of 'possession'. Rather than being 'possessed', the body-part noun *jina* 'foot' in (72) simply provides further specification of the NP head, *wagu:ja* 'man'. 'Man' and 'man foot' count as coreferential for purposes of clause linking within discourse organization; see §16.8.

C. Kinship relationship. For five of the languages in Table 16.2, all kin terms are in set 1. Koyukon and Jarawara, in Table 16.3, have all kin in set 2, while the sole language in Table 16.4, Dakota, places all kin in set 3. In Creek, kin terms are distributed between sets 1 and 2 without any principle of allocation being apparent. But it is interesting that whereas the regular terms for 'mother' and 'father' (*-cki-* and *-łki*) are in set 1, there are special children's terms (*-má:ma* and *-tá:ta*) which belong to set 2.

Like many Oceanic languages, Fijian has two sets of kin terms, one used for reference and the other primarily to address the kinsperson in question. These include (Dixon 1988a: 127):

REFERENCE	ADDRESS	
<i>tama-</i>	<i>tata</i>	'father'
<i>tina-</i>	<i>nana</i>	'mother'

The reference terms fall into set 1 (together with primary body-part nouns) and take a pronominal pensive suffix indicating the R—*tama-qu* 'my father', where *-qu* indicates 1sg R. *Tata* and *nana* are used for addressing one's parents but nowadays they are also employed, in informal style, for reference. The critical point is that *tata* and *nana* are then treated as set 2 nouns—*qou tata* 'my daddy', parallel to *qou 'olii* 'my dog' (here *qou* is a fusion of 1sg and the 'ownership' classifier).

In two of the languages from our tables, kin terms are divided between classes in a principled way. In Lango, terms for blood relations fall into set 1 while those for affines are in set 2. A blood relationship is unalterable, whereas one through marriage depends on the persistence of the marriage (which may be dissolved). Gapapaiwa operates on a different basis—terms for kin who, within the social sphere, are in authority over ego are placed in set 3, while those not in authority are in set 1—see Table 16.3. For Mohawk (Iroquoian; Mithun 1996), set 1 includes body-part terms and elder kin while younger kin are in set 2.

D. Attributes. In some languages, a handful of nouns describing physical or other attributes are included in set 1. Tables 16.2–16.4 illustrate the possibilities, including ‘pride’, ‘bad luck’, ‘being angry’, ‘fever’, ‘itch’, and ‘smell’. Otherwise, attributes are described through adjectives, or verbs, or set 2 nouns.

E. Orientation/Location. This may be shown by a variety of means, including grammatical elements such as adpositions, and lexemes such as adjectives and various sorts of nouns. For Tsimshian, Creek, Lango, Jarawara, and Gapapaiwa, there are a fair number of orientation/location terms in set 1 (just a single term is reported for Amele).

Ewe is particularly interesting in that set 1 includes many orientation/location terms but no words for body parts, all of which are in set 2. In fact, the same lexeme may function in both sets (Amele 1996: 811):

noun	meaning when used in a set 1 construction (R D)	meaning when used in a set 2 construction (R <i>phi</i> D)
ta	fore part, top, above	head
ŋkúme	front part	face
ŋúti	outer surface, near	skin
nũ	front, edge	mouth

Amele explains that B, Part of human, is grouped with A, Ownership, since one can do things with parts of the body in the same way that one can do things with material possessions.

F. Association. Many languages include in set 1 a limited number of nouns referring to social associations, habitats, valued goods, and the like. A representative sample is listed in Tables 16.2–16.4. Surveying a wider range of languages, the following terms recur:

- ‘friend’, ‘companion’, ‘master’, ‘shaman’
- ‘house’, ‘nest’, ‘den’, ‘homeland’, ‘path’, ‘grave’
- ‘clothing’, ‘ornament’, ‘goods’

In Nanai (Tungusic; Russia and China; Nichols 1988: 573, 591, quoting Avrorin 1959: 122), terms for domesticated animals are in set 1.

‘Name’ is a further recurrent member of set 1; it occurs in seven of the languages in Tables 16.2–16.3. Set 1 in Hua involves body parts, kin, and also *hamu* ‘namesake’ (Haiman 1980: 217). Interestingly, in Fijian noun *yaca* can occur in a set 1 construction, then meaning ‘name’ (for example *yaca-qu* ‘my name’), and also in a set 2 construction, then meaning ‘namesake’ (*qou yaca* ‘my namesake’).

Quite a few languages include ‘friend’ in set 1; these include Amele, Panare, Tsimshian, and Ewe. ‘Enemy’ occurs in set 1 for Hua (Haiman 1985: 130) and

also for Wiyot (Algic, California; Teeter 1964: 81). Fijian differs, placing *meca* 'enemy' in set 2, contrasting with *we'a-* 'friend', which is in set 1. A speaker explained the reason—one can choose who one has as a friend, so this noun is in set 1, whereas there may be no such control over who one has as an enemy.

Association may involve an inanimate R. In Lango, for example, we find *gúlú dēk* ('pot stew') for 'a pot of stew' (Noonan 1992: 158).

16.5.2 Marking

As part of a general study of iconicity, Haiman (1983: 793–5; 1985: 130–6) observed that a relationship which is semantically close—such as inalienable possession (sets 1 in Table 16.2)—is likely to employ a simpler or tighter grammatical marking than one which is semantically neutral—such as alienable possession (sets 2 in Table 16.2). This observation is borne out by examination of the grammars of a considerable number of languages.

We find the following: (i) alienable possession is marked in a similar way to inalienable possession but with an additional grammatical element; (ii) the grammatical marking for alienable is longer than that for inalienable possession; (iii) in some languages a classifier is required for all intra-NP possession, in others only for one type and this is then always alienable possession; (iv) alienable possession is shown by a marker, inalienable just by apposition. What is noteworthy is that none of the contrasts listed under (i)–(iv) are reversed; for example, with the marking for inalienable possession being longer than that for alienable, or involving an extra grammatical element.

We can now briefly illustrate each of (i)–(iv).

(i) An alienable possession construction is similar to that for inalienable possession, with an added grammatical element

- In the Nass dialect of Tsimshian (Boas 1911c: 392), inalienable possession involves a pronominal pertensive suffix (referring to R) added to D, i.e. D-possR, as in *huxdāg-întk^u-t* ('grandchildren-3sgR') 'his grandchildren'. Alienable possession is shown in the same way plus the inclusion of 'passive' suffix *-tk-* between D and possessive suffix, i.e. D-*tk*-possR, as in *g-ibō'-tk-t* ('wolf-*tk*-3sgR') 'his wolf'.
- Guajiro (Arawak, Venezuela and Colombia; Alvarez 1994: 87, 67; Adelaar 2004: 120) marks inalienable possession by a pronominal pertensive prefix to D, i.e. possR-D, as in *ta-ye:* ('1sg-tongue') 'my tongue'. Alienable possession requires, in addition, an invariable plain pertensive suffix *-se* to D, i.e. possR-D-*se*, as in *ta-kulu:t-se* 'my cloth'.
- Maká (Mataguayan, Argentina; Gerzenstein 1994: 147–8) is like Guajiro in that inalienable possession involves a possessive prefix added to D, as in

yo-qofol ('1sgR-fingernail') 'my fingernail'. For alienable possession this prefix must be followed by plain pertensive prefix *-q(V)-* as in *yo-qo-koyoyoy* ('1sgR-qo-car') 'my car'.

(ii) The grammatical marking for alienable possession is longer than that for inalienable possession

- Slave (Rice 1989: 207–34) uses the same paradigm of pronominal pertensive prefixes for both types of possession. There is also a plain pertensive suffix, *-é*, for alienable possession. Inalienable possession is marked not by a suffix but by the addition of high tone, *´*, to the last vowel of D. Thus 'my tears' is *senatú*, made up of 1sg prefix *se-*, compound noun *natu* 'tears' (comprising *na* 'eye' plus *tu* 'water'), and pertensive marker consisting of high tone on the final *tu* of *natu*. A further difference is that an inalienably possessed noun must take a pronominal prefix (indicating R) whether or not R is shown by an NP, whereas for alienable possession one would not usually include a pronominal prefix if R is shown by an NP.
- Creek (Martin 1993) has two sets of pronominal prefixes, added to D, which show the person and number of R. The forms are:

	SET 1 (INALIENABLE)	SET 2 (ALIENABLE)
1sg	ca-	am-
1pl	po-	pom-
2	ci-	cim-
3	i-	im-

For example *ca-pósi* 'my grandmother', *am-á:to* 'my car'. It will be seen that the alienable possession prefixes basically involve the addition of *-m* to the inalienable set, save that 1sg alienable is *am-* instead of the expected **cam-*. (Both prefix paradigms also mark function of a phrase in a clause, see §16.6.)

- Wiyot (Teeter 1964: 41, 79) has three distinct paradigms of possessive pronouns. 1st and 2nd person forms are:

	set 1 (body parts, 'wood', 'enemy' etc.)	set 2 (kin terms)	set 3 (alienable)
1st person	d-	ø-	duh-
2nd person	kh-	Ch ^a -	khuh-

It will be seen that the prefix forms for sets 1 and 2 (varieties of inalienably possessed nouns) are shorter than those for set 3.

- Under (e) in §15.1.9, we mentioned Tunica, where the alienable possessive prefix involves an increment to the inalienable forms.

(iii) The alienable possessive construction requires a classifier, the inalienable construction does not

- For Panare (Payne and Payne 1999: 90–3) the inalienable possessive construction uses a pertensive pronominal prefix to D, plus a plain pertensive suffix *-n*, i.e. possR-D-*n*, as in *yí-mata-n* ('1sg-shoulder-PERTENSIVE') 'my shoulder'. Alienable possession requires a classifier to which the possessive prefix and plain pertensive suffix are attached, this is followed by D which does not bear affixes, i.e. possR-CLASSIFIER-*n* D, as in:

- (73) *y-u'ku-n* *wanẽ_D*
 1sgR-CLASSIFIER(LIQUID)-PERTENSIVE honey
 my honey

Here the classifier relates to the nature of the D. Oceanic languages such as Fijian behave in a similar way, although here the classifier generally relates to the nature of the possessive relationship; see the appendix to this chapter.

(iv) Overt marking only in an alienable possessive construction

- In Lango (Noonan 1992: 156–7) inalienable possession just involves D followed by R, as in *léb lócə̃* ('tongue man') 'the man's tongue'. For alienable possession, D and R occur in the same order with an 'attributive particle' *à* between them, as in *gwòkk à lócə̃* ('dog PARTICLE man') 'the man's dog'.
- Ewe (Ameke 1996) follows a similar pattern with R preceding D. Thus, for set 1, *Kofi srɔ̃* 'Kofi's wife'. Set 2 possession requires possessive marker *ɸeé* between R and D, as in *Kofi ɸé awu* 'Kofe's garment'.
- Yidiñ again has R and D in apposition (generally, but not invariably, in this order) for set 1, as in (72), and for set 2 has R marked by genitive suffix *-ni*, as in (71).
- In Amele (Roberts 1987: 139, 171–5) inalienable possession involves a pronominal suffix to D, as in *cot-i* ('brother-1sg') 'my brother'; if R is 3rd person, an NP referring to R may be included before D-possR. For alienable possession, marker *na* is placed between R and D, as in *Banag na jo* 'Banag's house' and *ija na jo* ('1sg(free pronoun) *na* house') 'my house'.

We can now examine the markings in languages with three or four sets of nouns for D function, from Tables 16.3–16.4. It will be seen that they combine various of the distinctions (i)–(iv).

Dakota (Boas and Deloria 1941: 127–33; Pustet 2001) uses pronominal pertensive prefixes to D for sets 1–3. The 2nd person prefix has constant form,

ni-, but there are variant forms for 1st person. This can be illustrated for 1sg where the prefixes are *ma-* and *mi-*.

- For set 1 ('body parts conceived of as particularly subject to will-power') we get *mi-D*; for example, *mi-i'ha* 'my lips'.
- For set 2 (other body parts) 1sg prefix *ma-* is used in place of *mi-*, i.e. *ma-D*; for example, *ma-we'* 'my blood'.
- Set 3 (kin terms) are like set 1 in taking prefix *mi-* for a 1sg R. But they differ in that set 1 has no marker for 3sg R while kin terms take a 3sg suffix, as in *c'iyē-ku* ('elder.brother-3sgR') 'his elder brother'.
- Set 4 (ownership) takes longer possessive prefixes—1sg *mi-tha-*, 2sg *ni-tha-*, 3sg *tha-*. These are made up of those from set 1 with the addition of *-tha-*. (Alternatively, independent possessive pronouns can be used.)

Koyukon (Thompson 1996)

- Set 1 (parts, etc.) and set 2 (kin) take a pronominal prefix to D; for example *se-tlee'* ('1sg-head') , *be-to'* ('3sg-father') 'his father'. However, sets 1 and 2 differ in that only kinship nouns 'require a possessive prefix whether or not there is a full noun as possessor'.
- Set 3 (mostly ownership) takes the same pronominal prefixes and also requires a plain pertainive suffix *-e'* to D, as in *se-tel-e'* ('1sg-socks-PERTENSIVE') 'my socks'.

Jarawara (Dixon 2004a: 79–80, 99, 295–301)

- Set 1 (parts, etc.). R (which is head of the NP) has inherent gender and is followed by D; about half of the inalienably possessed nouns have distinct masculine and feminine forms and agree with R in gender. See example (99) and:

(74)	mat _R	nok _D	(75)	bat _R	nok _D
	mother(f)	eye:feminine		father(m)	eye:masculine
	mother's eye			father's eye	

- Set 3 (ownership) requires a quite different structure, R *kaa* D; for example *Okomobi kaa taokana* 'Okomobi's gun'.
- Set 2 (kin) differs in subtle ways from set 3. Only some kin terms require *kaa*, and then only with certain possessors; for example, there is no *kaa* in *Okomobi mati* 'Okomobi's mother'. In addition, the verb *-kiha-* 'have' may be used only with set 3 nouns, not with those from sets 2 or 1.

Gapapaiwa (McGuckin 2002)

- Set 1 (parts, kin who are not in authority, etc.). R may be shown by an NP before D and must be indicated by a pronominal suffix to D, as in *ima-na* ('hand-3sg') 'his/her hand'.
- Set 2 (attributes, association) have the pronominal suffix attached to possessive marker *ka-*, making up a word that precedes D, as in (76).
- Set 3 (kin who are in authority, ownership, etc.) have the same structure as set 2 save that the possessive marker is *i-*, as in (77).

(76) *ka-na* *gana*
 SET.2.MARKER-3sg clothing
 his/her clothing (to wear)

(77) *i-na* *gana*
 SET.3.MARKER-3sg clothing
 his/her clothing (to sell at the market)

It is clear that *ka-* and *i-* developed out of a set of classifiers (typically found in languages from the Oceanic subgroup of Austronesian, to which Gapapaiwa belongs) employed to indicate the general nature of the possessive relationship—food to be eaten/sold, clothes to be worn/sold, etc. These were discussed under (iv) in §16.4. Their function has been extended and reinterpreted so that they are now most appropriately regarded as markers for sets of nouns functioning as D.

This chapter began with the statement that the grammar of every language includes an NP-internal possessive construction. However, this does not always cover *all* types of possession. In the Hokkaido dialects of Ainu (isolate, Japan; Tamura 2000: 55, 86–7) inalienable possession is shown as (R) possR-D, as in *ku-sikihi* ('1sg-eye') 'my eye(s)'. But alienable possession may only be shown through a relative clause whose verb is *kor* 'have':

(78) *acapo* [*kor sake*]_{RELATIVE.CLAUSE}
 uncle have liquor
 uncle's liquor (lit. the liquor which uncle has)

16.6 Further functions of possessive markers

Many elements in a grammar have multiple functions. Markers of a possessive construction within an NP may also be used to mark the functions of an NP within a clause. A fair number of languages employ the same form for dative case (a clausal function) and genitive (an NP-internal function). The semantic

link between dative and genitive is that something which is ‘for X’ (dative) is likely soon to be ‘X’s’ (genitive). For example, in the Australian language Mangarayi (Merlan 1982: 66–7), dative and genitive are shown by the same prefixes to a noun, masculine *rna-* and genitive *ɲaya-*. Consider:

- (79) *nara-bayi*_{VCS} *rna-bugbungu*_{VCC}
 THAT-FOCUS DATIVE/GENITIVE:m-old.man

If prefix *rna-* has its dative sense, then (79) will mean ‘That is for the old man’. If it is accorded a genitive interpretation, the sentence means ‘That is the old man’s’. Here a verbless clause complement can be just a dative NP, or else just a genitive noun from an NP-internal possessive construction, with the D (the head) ellipsed.

Within a longer sentence, genitive and dative senses are clearly distinguishable. A full NP-internal possessive construction includes statement of R by an NP with genitive prefix and also by a pronominal pertensive suffix to D, as in:

- (80) [*rna-bugbungu* *barnam-rnawu*]_O
 GENITIVE:m-old.man camp-3sgMASC.POSS
 rna-wa-b
 1sgA-visit-PAST.PUNCTUAL
 I visited the old man’s camp (lit: the old man’s camp-his I-visited)

In contrast, when the prefix has dative sense, there is no ‘cross-referencing’ possessive suffix, such as *-rnawu* in (80). For example:

- (81) *rna-bugbungu* *barnam-garlama* *ga-ɲa-yag*
 DATIVE:m-old.man camp-ALLATIVE PRESENT.REALIS-1sgS-go
 I am going to the camp for the old man (i.e. to do something for his sake)

Similar to many other languages, in Mangarayi dative and genitive fall together on nouns, but have different form with pronouns.

In some languages genitive (marking R within an NP) has the same form as ergative (marking A function within a clause); these include Eskimo (de Reuse 1994: 30) and Ladakhi (Tibeto-Burman, Jammu and Kashmir; Koshal 1979: 65–75). Australian languages show a wide range of further functions for an affix which marks R in a possessive construction (usually, an alienable one). They include ablative (‘from’) in Anindilyakwa (Leeding 1989: 288, 298), causal (‘because of’) in Gog-Nar, both ablative and causal in Gugadj (Breen 1976: 155), and purposive (‘in order to’) in Watjarri (Douglas 1981: 218–20). Further examples are listed in Dixon (2002a: 167–75).

When a language uses non-canonical marking for subject function—in specific grammatical and semantic circumstances—genitive may be employed;

see, for example, Onishi (2001b) on Bengali. In some languages, each adposition requires the noun it marks to be in a particular inflected form, generally a case or genitive; this applies in, for example, German and Russian.

A genitive marker is often used to mark one argument in a nominalization. This was illustrated for *'s* and *of* in English in §16.1—in (13) genitive marks the underlying A argument, in (14) the location, and in (15) the underlying O. Cross-linguistically, genitive is most often used for the underlying subject of a nominalization. And it may also mark the subject for one variety of complement clause (illustrated by *I_A dislike [Mary's_A drinking whisky_O]_{COCL:O}* in English).

I do not know of any plain pertensive marking having an additional function within its grammar (further work may well reveal examples of this). However, pronominal pertensive affixes—added to D and marking person/number of R—frequently also mark functions of arguments within a clause. A fine example of this, in Aleut, was illustrated at (49) and (50) in §16.2, where fused pronominal suffix *-ning* indicates 1sg A plus 3pl O when added to a verb, and 1sg R plus 3pl D when added to a noun (the D) within an NP-internal possessive construction.

In other languages a possessive marker may also be used to mark one core constituent at clause level. There is considerable variation concerning which core constituent is involved.

(i) In languages lacking an inalienable/alienable distinction, the affix which is added to D and indicates R is identical with marker of the following clausal functions:

- A—Jacaltec (Mayan, Guatemala; Craig 1977: 108–11).
- A and Sa—Yagua (Payne and Payne 1990: 348, 361–4).
- O—Indonesian (Sneddon 1996: 166–7). In Cairene Egyptian Colloquial Arabic (Gary and Gamal-Eldin 1982: 84, 101) all bound pronouns are identical save that 1sg is *-ni* for O function and *-i* for possessor marking on D.

(ii) In languages with an inalienable/alienable distinction, where a single set of pronominal affixes is used for both with an additional element added for alienable. The affix which is added to D and refers to R is identical with marker of the following clausal functions:

- O and S—Tsimshian.
- O and So—Dakota (set 2 prefixes).
- O and indirect object—Koyukon.
- O and So (and VCS)—Panare.
- A and Sa—Warekena (Arawak; Aikhenvald 1998: 293–4).

(iii) Pronominal affixes marking inalienable possession are also used for:

- A—Ainu.
- S and A—Jarawara and Tariana; see (e) in §15.1.9.
- O and So—Kamaiurá—see (a) in §15.1.9.
- O—Hua (Haiman 1980: 199–200).

(iv) Inalienable and alienable possessive affixes, each of which is also used to mark a clausal function:

Inalienable possession affix, also used to mark clausal function(s):	Alienable possession affix, also used to mark clausal function(s)	
So	O	Tunica—see (e) in §15.1.9
So, O	dative	Creek, and other Muskogean languages
S, A	O	Tuscarora (Iroquoian, Mithun 1999: 253)

Surveying all of (i)–(iv), it will be seen that no sure overall generalizations are possible concerning a correlation between type of possessive marking and type of clausal marking.

Amele presents a further scenario. Alienable possession marker *na* is also used, at the clause level, to mark time, location, and instrument (Roberts 1987: 80–1).

16.7 Complex mechanisms

Most parts of a grammar are organized on regular principles (otherwise we would not have a grammar) but there may be some degree of irregularity. In English, for instance, morphological irregularity involves a few nouns (*woman/women*, *child/children*) and a fair number of verbs (*sing/sang*, *take/took*, *buy/bought*, and so on).

The types of possessive constructions described so far have, by and large, been pretty regular. However, a minority of languages employ a number of possessive constructions which do not seem to be semantically or grammatically motivated. This can be exemplified from Ndjébbana, from northern Australia. McKay (1996) distinguishes four sets of nouns, as D, each occurring in a distinctive possessive construction.

- **Set 1.** R is shown by a pronominal pertensive suffix to D (there may also be an NP indicating R), as in:

- (82) marnákarna-njabba
rib.bone-1minR
my rib bone

These possessive pronominal suffixes have no further function in the grammar. (Note that the pronoun paradigm is organized in a minimal/unit-augmented/augmented basis; see (d) in §15.1.2.)

This is a closed set consisting, in McKay's corpus, of twenty-three body-part nouns plus 'strength', 'speed', 'country (of origin)', and 'one, alone'.

- **Set 2.** R is shown by a pronominal pertensive prefix to D, as in:

- (83) nga-ngardabbámba
1minR-liver
my liver

The prefix paradigm here is identical to that which attaches to a verb, showing S function, save that the 3rd person minimal masculine form is *ka-* for S but is *n-*, *nga-*, *ka-*, *ø-*, *ma-*, or *na-* for set 2 possession. Indeed, set 2 nouns fall into six closed subsets—each with between one and thirty-one members—according to the form of the 3min.masc prefix they take. Set 2 includes seventeen body parts, many abstract nouns, plus 'adolescent', 'friend', 'shellfish', 'spouse', etc.

- **Set 3.** Shows R by the S pronominal prefix to the present/recent past form of verb *-réndeyi* 'stand, be', with D shown as a separate word:

- (84) kála_D ka-réndjeya
ear 3min.mascR-stand/be:PRESENT/RECENT.PAST
his ear

This is a closed set with twenty members, all body parts.

- **Set 4.** A pronominal prefix (identical to the S prefix to a verb), indicating D, is added to a free-form pronoun, indicating R, as in:

- (85) nja-ngáyabba kíkka
3min.femD-1minR mother
my mother

Set 4 is an open class, including a dozen body-part terms ('backbone', 'heart', 'blood', etc.), kin terms, and 'owned' things such as 'dog', 'bottle', and 'place, home'.

Just as one has to learn which conjugation a verb belongs to in Latin, in order to know how to use it, so it appears that a speaker would have to learn—for

each body-part noun in Ndjébbana—which set it belongs to, in order to know which possessive construction it may be employed in.

16.7.1 's and of in English

English has two ways of marking an NP-internal possessive construction—by suffix 's on R (which precedes D), R 's D, or by the preposition *of* before R (which follows D), D *of* = R. In each case the marker is attached to R, by a suffix in the first instance and in the second by preposition *of*, which is generally proclitic to the word which follows. For example, *the door of the cabin* is /ðə=dɔː əv=ðə=kæbɪn/.

With some choices of R and D, either 's or *of* is acceptable:

- | | | |
|------|--|---|
| (86) | EITHER | OR |
| | (a) <i>the president's private plane</i> | <i>the private plane of the president</i> |
| | (b) <i>my friend's sister</i> | <i>the sister of my friend</i> |
| | (c) <i>the table's leg</i> | <i>the leg of the table</i> |
| | (d) <i>the jumbo jet's length</i> | <i>the length of the jumbo jet</i> |
| | (e) <i>the Tsotsi tribe's homeland</i> | <i>the homeland of the Tsotsi tribe</i> |

However, there is only sometimes a choice between 's and *of*. For instance, it is in most circumstances infelicitous to say *the car of John*, *the husband of Mary*, *the foot of Bill*, *the anger of Jane*, *the dentist of Fred*. When R is a pronoun, only a possessive pronoun (the equivalent of 's on an NP) is allowed, never *of*. One must say *his eye* rather than **the eye of him*.

In essence, when R is an NP (which can be just a noun), the 's alternative is preferred (and the *of* alternative dispreferred) according as:

- (i) R is human (or at least animate), specific, and singular. A proper name always takes 's. And whereas *the boy's leg* is preferred over *the leg of the boy* (singular human R), *the legs of the boys* (plural human R) is more acceptable, with *the legs of the antique tables* (plural inanimate R) sounding better still.
- (ii) D is specific and singular. For example, *my friend's sister* is preferred over *the sister of my friend* (singular D) but *the sisters of my friend* (plural D) sounds considerably better.
- (iii) R has few words. The 's alternative is not liked on a long possessor, and here *of* may be preferred. For example, *the gun of that evil character who lives in the tumbledown shack down the road*, rather than *that evil character who lives in the tumbledown shack down the road's gun*.
- (iv) R is familiar information. For instance, in a discussion about my wife I might say *my wife's jewels*, since my wife is familiar information and this is the first mention of the jewels. But if in a discussion about jewels

I suddenly mention those belonging to my wife, I would be more likely to say *the jewels of my wife*, since this is the first mention of my wife (it is not familiar information).

Thus, for some instances of possession only 's is considered felicitous. For others—such as those in (86)—either 's or *of* is acceptable. And for others only *of* is likely to be used in normal circumstances; for example, one hears *the names of mountains, the virulence of the mosquitoes, the haunts of evil spirits*. (There are just a few idiomatic phrases which transgress principles (i) and (ii); for example, one generally says *a summer's day* rather than *a day of summer*.)

A possessive modifier (NP plus 's, or a possessive pronoun) is mutually exclusive with the article *a* and demonstratives, *this, that, these, those*. But one might want to include both *a* or a demonstrative, and a possessive modifier, in the same noun phrase. This is achieved by placing the possessive modifier after the head of the noun phrase, linked to it by *of*. Thus, *Mary's husband* but *that husband of Mary's, my daughters* but *these daughters of mine*. Here the possessive relation of *of* is shown by 's or *mine*, with the *of* simply a linker.

At the beginning of §16.4, we mentioned that *John's picture* is multiply ambiguous. When the *of* construction is used, ambiguity is partly resolved. *That picture of John's* would be used for a picture which John owns or one which he painted, while for a picture which someone else painted of John one would say *that picture of John*.

16.8 Internal structure of an NP which includes possession

As outlined in §5.6, one item in each phrase will be its 'head'. The head dictates agreement on other items in the phrase, and determines the properties of the NP as a whole. We need to enquire what is the head for an NP which includes a possessive construction.

In alienable possession, it is always the possessed (D) which is head (this is underlined). In Jarawara, *Bati kaa bati* ('father possessor axe') 'father's axe' refers to an implement, not to a person. If a language has only one possessive construction—with no inalienable/alienable distinction—the D will always be head: *John's mother, Mary's foot*.

In some languages 'inalienable possession' is shown just by apposition. Discussing this for Yidiñ, in §16.5.1, it was pointed out that in (72) *wagu:ja jina* ('man foot') it is the R, *wagu:ja* 'man', which is head. The NP refers to a man, with respect to his foot, not to a disembodied foot. The topic through a section of discourse could be 'man', then 'man foot' (i.e. man with respect to his foot) then 'man' again, as in 'the man fell down, the man's foot was hurt, the man limped home'.

Gender provides a clear criterion for deciding what is the head of an NP. It is the head which provides the gender of the whole NP, and which determines gender of modifying words within the NP. Jarawara has two genders—feminine (f), which is the functionally unmarked term in the system, and masculine (m). All pronouns are cross-referenced as feminine (irrespective of the sex of the person(s) they refer to).

In §5.6, we presented an NP in Jarawara which involves alienable possession. This can be placed as S argument in a sentence:

- (87) [mati_R kaa jomee_D]_S tafa-ka
 mother(f) POSSESSIVE dog(m) eat-DECLARATIVE:MASCULINE
 Mother's dog is eating

Here the whole NP is m, taking its gender from that of the D, *jomee* 'dog', which engenders the m form of the declarative suffix on the verb, -ka.

Now consider an NP involving a whole-part relationship ('inalienable possession'):

- (88) [mati_R tame_D]_S koma-ke
 mother(f) foot(f) be.sore-DECLARATIVE:FEMININE
 Mother's foot is sore

Here the 'whole' noun, *mati* 'mother', is head of the NP. It dictates the gender of the 'part' noun, so that the f form of 'foot', *tame*, is used (the m form is *teme*, as in *jomee teme* 'the dog's foot'). And the whole NP counts as f, requiring the f form, -ke, of the declarative suffix on the verb.

In fact, Jarawara has rather complex rules for the gender on a 'part' noun, depending on whether the 'whole' it modifies is shown as a free noun, as the 3rd person non-singular pronoun, or as a 1st or 2nd person pronoun. This is shown in Table 16.5.

A 'part' noun modifying a free noun as 'whole' agrees with it in gender, as in rows (a) and (b) of Table 16.5. The 3n-sg pronoun *mee*, in row (c), requires its part modifier to be f, probably because this pronoun evolved from free noun *madi* 'person', which was f. But all 1st and 2nd person pronouns require a modifying part noun to be in m form, as in rows (d) and (e). (I do not know the reason for this. It is just the way the language is.)

The important point is that the complete NP for rows (c)–(e) has f gender, requiring the f form of declarative suffix on the verb: [mee tame]_S *koma-ke* 'their feet hurt', [ee teme]_S *koma-ke* 'our (inc) feet hurt', and [o-teme]_S *koma-ke* 'my feet hurt'. Non-singular pronouns are expressed as separate words, but sg pronouns (as possessor, and in S and A functions) are prefixes, 1sg *o-* and 2sg *ti-*. Thus in *o-teme* the head of the NP is the pronominal possessor (R) *o-*,

TABLE 16.5. Gender in and on NPs involving whole–part possession in Jarawara

	NP HEAD (‘whole’)	GENDER OF MODIFIER (‘Part’)	EXAMPLE	GLOSS	GENDER OF COMPLETE NP
(a)	m free noun, e.g. <i>bat</i> i ‘father’	m	<i>bat</i> i <i>teme</i>	‘father’s foot’	m
(b)	f free noun, e.g. <i>mat</i> i ‘mother’	f	<i>mat</i> i <i>tame</i>	‘mother’s foot’	f
(c)	3n-sg pronoun, <i>mee</i> ‘they’	f	<i>mee</i> <i>tame</i>	‘their feet’	f
(d)	1st or 2nd n-sg pronoun, e.g. 1-nsg (inclusive) <i>ee</i>	m	<i>ee</i> <i>teme</i>	‘our (inclusive) feet’	f
(e)	1st or 2nd sg pronoun, e.g. 1sg <i>o-</i>	m	<i>o-</i> <i>teme</i>	‘my foot’	f

since the whole NP engenders f agreement on the verb. The ‘part’ (D) *teme* ‘foot’ is in m form and so could not be the head.

Set 1 in Jarawara includes at least 175 inalienably possessed nouns—parts of people, animals, things; attributes; orientation; and various types of association. In each of *o-nowati* ‘my behind (i.e. the space behind me)’, *o-tesene* ‘my companion’, and *o-tefe* ‘my food’ it is the 1sg possessor, *o-*, which is head.

Interestingly, all of those languages which mark inalienable possession (set 1) by apposition, with R as head of the NP, place kin terms outside this set. That is, for kinship possession it is always the D which is head—*ɲajin wagal* ‘my wife’ in Yidiñ and *oko jibotee* ‘my spouse’ in Jarawara, alongside *my wife* in English. It is hard to imagine it being otherwise.

16.9 Constructions which establish a possessive relationship

As we have shown, every language has an NP-internal possessive construction with roughly similar structure. There is more variation concerning what is often called a ‘predicative possessive construction’. Many languages do not have a verb similar to English *have*, French *avoir*, and German *haben*. Finnish, Latvian, and Japanese do not. Of the languages mentioned in Tables 16.1–16.4, at least the following lack a verb ‘have’—Yagua, Tialo, Fijian, Amele, Slave, Lango, and Yidiñ. Languages with no ‘have’ employ other strategies for establishing a possessive relationship—with copula ‘be’, or with intransitive

verb ‘exist’, or some other means. These are dealt with in turn in the following sections.

16.9.1 Using ‘have’

As pointed out in Chapter 14, copula verbs have relational rather than referential meaning. One could argue whether ‘have’ is like ‘be’ in being basically relational, or whether it is referential. In the first instance it should occur in a copula-type construction, in the second it would be a transitive verb.

In most languages in which it occurs, ‘have’ is basically a transitive verb, although generally with syntactic and morphological deficiencies and/or irregularities. In English, for instance, the construction is ‘R_A have D_O’, with D being marked for O function. In (89)—reporting a remark at a slave auction—‘object’ pronoun forms *him* and *them* realize the D’s:

- (89) I’ll let you_{R:A} have him_{D:O} if I_{R:A} can have them_{D:O} (pointing at the slaves in question)

In Jarawara, *-kiha-* ‘have’ patterns like a normal transitive verb, both syntactically and morphologically. But *have* in English does not occur in a passive construction (save with a quite different sense), and has an irregular morphological paradigm; the same applies for *avoir* in French and *haben* in German. In Oromo (Cushitic, Ethiopia and Kenya; Owens 1985: 74) there is no past tense form for *k’ab* ‘have’ (as there is not for *jir-* ‘exist’). For Tamambo (Oceanic subgroup of Austronesian, Vanuatu; Jauncey 1997: 297) *-noha* ‘have’ is the only transitive verb which cannot take tense or aspect marking. In Kana (Benue-Congo, Nigeria; Ikoro 1996a, 1996b: 30) the verb *ērē* ‘have’ shows an irregular tone pattern; other low–mid-tone verbs retain their inherent tone pattern in past tense whereas ‘have’ changes to mid–mid, *ērē*. ‘Have’ in Dagbani (Gur family, Ghana; Olawsky 1999: 50) is a regular transitive verb *mali*; but its negation is realized by suppletive form *ka* (which is also the negative form of locative-existential copula *be*). And in Dhaasanac (Cushitic, Ethiopia and Kenya; Tosco 1999a: 50, 2001) *ǵáb* ‘have’ may not be used in the negative; the invariable form *mán* ‘there is not’ is employed. In some languages a verb ‘have’ may not be used in the imperative. Many other examples of the irregular nature of ‘have’ verbs could be added.

In quite a number of languages, ‘have’ is a grammaticalization from a regular transitive verb with referential meaning; this may be at least part of the reason why in these languages ‘have’ is a transitive verb. Isačenko (1974: 42) recounts that ‘verbs meaning “have” are secondary acquisitions in all IE languages and . . . such verbs stem from transitive verbs with the general meaning “to hold, to grasp”’. He exemplifies from Greek and Latin and then mentions that ‘the semantic change “take” → “have” occurred in historical times in

Spanish'. *K'áb* in Oromo functions as possessive verb 'have' and as referential verb 'grasp, hold'. Dhaasanac verb *ǵáb* 'have' is cognate with verbs in other East Cushitic languages with primary meaning 'get, hold'. In Dakota (Pustet 2001: 62), *yuhá* 'have' also means 'take along'. Mithun (1999: 249) reports that in Tuscarora (Iroquoian) 'having' may be expressed through D being incorporated into the verb 'lie'; for example, 'him-pet-lie-STATIVE' (literally, 'a pet lies on him') for 'He has a pet'.

In Manambu (Ndu family, Papua New Guinea; Aikhenvald 2008a: 62, 176), *tə-* functions both as a copula 'become, be', illustrated in (90), and as transitive verb 'have', in (91):

- (90) [a-bər ñədi]_{CS} kwakuli_{CC} tə-bər
 THOSE-DUAL children:DUAL orphan become/be-3duCS
 Those children became/were orphans

- (91) [nəma kabak]_O tə-na-di
 lots.of money have-ACTOR.FOCUS-3plA
 They have lots of money

In Thai, and a number of other languages from South-East Asia, one verb functions both as 'have' and as intransitive verb 'exist'—a special kind of S = O ambitransitive. In Motuna (Buin family, Papua New Guinea; Onishi 1994: 407–8, 427–8), *tuh-ee* 'have' is derived from intransitive *tu(h)-* 'exist' by addition of applicative suffix *-(j)ee*.

As mentioned above, it is a moot point whether 'have' should be regarded as a referential verb or as simply relational, in the way that copulas are. There are languages in which 'have' is used in a copula construction, with R in copula subject (CS) and D in copula complement (CC) function. This applies for both of the verbs 'have' in Awa Pit (Barbacoan, Ecuador and Colombia; Curnow 1997a: 77). For example:

- (92) [ap tií]_{CS} =na [paas paynkul]_{CC} mij
 1sg:POSS uncle =TOPIC two son have:CONJUNCT
 My uncle has two sons

An NP in O function, if its referent is definite and specific, requires accusative enclitic =*ta*. The fact that there is no =*ta* on *paas paynkul* in (92) indicates that this NP is not in O function but rather in CC function within a copula construction.

In some languages, 'have' has restricted scope, being used only for certain kinds of D. As mentioned above, the O NP for *-kiha-* 'have' in Jarawara (Table 16.3) may only be a set 3 noun (ownership), not one from set 2

(kin)—even though set 2 is very similar to set 3 in its NP-internal possessive construction—nor one from set 1 (body parts, attributes, etc.). The verb *-ga-* in Ute (Uto-Aztecan; Givón 1980: 271) is quite different, being used only for kin and body parts (inalienable possession). In Koyukon (Thompson 1996: 668–9)—also from Table 16.3—the verb *-t'aanh* ‘have’ is confined to nouns from set 3 (ownership, plus bodily fluids and solids), as in (93). For set 1 (parts, etc.) and set 2 (kin), a copula construction with *laanh* ‘be’ must be employed, as in (94).

- (93) *leek es-ø-t'aanh*
 dog 1sgSUBJECT-CLASSIFIER-have
 I have (own) a dog
- (94) *s-ode hoo-ø-laanh*
 1sg-elder.sister area-CLASSIFIER-be
 I have an elder sister

Some languages have two verbs ‘have’ with differing scopes. In Somali (Cushitic; Tosco 1999b: 45), *léeyahay* ‘have’ is used for all kinds of possession while *qáb* ‘get’ is confined to alienable possession. For Awa Pit (Curnow 1997b: 22) the two verbs ‘have’ show complementary scopes, *waj* being restricted to body parts while *mij* is used for all else (kin, ownership), as illustrated in (92).

Generally, more varieties of possession are covered by an NP-internal possessive construction than by a ‘have’ construction. As mentioned at the end of §16.5.2, Ainu is unusual in that the verb *kor* ‘have’ can be used for all types of possession while the NP-internal construction is confined to the inalienable variety. To code alienable possession within an NP, one must use a relative clause with *kor* ‘have’, as illustrated in (78).

It has been shown that in many languages ‘have’ evolved from a transitive referential verb with a meaning like ‘hold, grasp’ or ‘take’. In a number of languages, ‘have’ has developed a further function, as a tense and/or aspect marker. (The two sets do overlap.) A recurrent scenario is for phonological erosion to reduce the inflectional complexity of a verb, new categories then being created periphrastically, typically involving ‘have’ and/or ‘be’ in new functions. For example, *John has gone* and *Mary is coming* in English. ‘Have’ may develop into a marker of perfect aspect, or of past tense, or of obligation modality (‘have to’) or of future tense. Illustrations of these changes—for Indo-European, African, and Asian languages—are in Heine (1997a: 187–208), Kuteva (2001: 37–43), Heine and Kuteva (2002: 242–5), Allen (1964), Benveniste (1971a: 178–9), Isačenko (1974: 73–4), and Bybee, Perkins, and Pagliuca (1994: 68–9, 260–4).

There may be further verbs with a meaning similar to 'have'. In English these include *possess*, *own*, *lack* ('not have'), and also verbs relating to 'coming to have' such as *get*, *obtain*, and *come by*. Each has its own semantic character. For instance, *possess* indicates that there is a strong emotional or mental connection between R and D, as in *She possesses a fine brain*. Unlike *have* and *possess*, *own* does occur in the passive; for example, *This car has only been owned by two old ladies*.

And there is *belong to*, for which the syntactic functions relating to R and D are effectively reversed. [*The boss*]_R *has* [*a red car*]_D makes a statement about the boss, whereas [*That red car*]_D *belongs to* [*one of the bosses*]_R makes a statement about the red car. Of the languages which include a verb 'have' only a small minority also have 'belong to'. Just occasionally, one finds a language which lacks 'have' but does include a verb 'belong to'; Colloquial Welsh is one such (Jones and Thomas 1977: 199–200). See Heine (1997a: 29–33) for discussion of 'belong'.

It is interesting to examine the frequency of occurrence of the different types of possessive construction. Velazquez-Castillo (1996: 69–77) undertook a study of texts in Guaraní and found that of NP-internal constructions 80.5 per cent involved inalienable possession (body parts, kin, etc.) whereas for constructions with 'have', 92.3 per cent involved alienable possession.

16.9.2 Using a copula construction

Some languages—including the familiar tongues of Western Europe—include both possessive verb 'have' and Copula verb 'be'. Some—including Yidiñ—lack both. A few show 'have' but not 'be' (although they may include an intransitive verb 'exist'). In Chrau (Mon-Khmer branch of Austro-Asiatic, Vietnam; Thomas 1971) a single verb functions both as intransitive verb 'exist' and as transitive verb 'have', but there is no copula 'be'. And there are languages—including almost all of those spoken in the South Asian subcontinent—which have 'be' but not 'have'. As mentioned in §14.4 (with exemplification from Tamil), many languages of this type establish a possessive relationship through a copula construction.

Each of R and D may be mapped onto CS, with the other being mapped onto CC. That is:

either (i) D_{CS} R_{CC} copula or (ii) R_{CS} D_{CC} copula

That which is placed in CS function is identified as topic. (The ordering of the three constituents will depend on the grammar of the language in question.)

The copula verb may take a bound pronoun which provides further realization of the CS. The CC—whether it be R or D—is likely to be marked by an oblique case or by an adposition.

Both of these possibilities are used in Lango, from Table 16.2 (Noonan 1992: 148, 170):

- (95) gwòk_{D:CS} tíê [bòt òkélò]_{R:CC}
 dog 3sgCS:be:PRESENT.HABITUAL TO Okelo
 The dog is Okelo's (lit. The dog is to Okelo)
- (96) òkélò_{R:CS} tíê [ì gwòk]_{D:CC}
 Okelo 3sgCS:be:PRESENT.HABITUAL ON dog
 Okelo has a dog (lit. Okelo is on the dog)

In (95) D is the topic, and in CS function, while the R, Okelo, is marked by preposition *bòt* 'to'. In (96) R is the topic, coded as CS, and the D, *gwòk* 'dog', is marked by the wide-ranging preposition *ì*, variously glossed as 'on, at, in, about, to, from'.

Evenki (Tungusic, Siberia; Nedjalkov 1997: 126) has construction type (i), in which the R, as CC, is in dative case:

- (97) min-du_{R:CC} purta_{D:CS} be-si-n
 1sg-DATIVE knife be-PRESENT-3sgCS
 I have a knife (lit. a knife is to me)

The details of the copula construction used for establishing a possessive relationship may depend on the nature of the D, and also on the nature of the R. Punjabi employs copula *hoNaa* 'be' in a construction of type (i). The marking on the R is determined as follows (Bhatia 1993: 146–8):

R	D	POSTPOSITION ON R
animate	alienable (concrete objects)	<i>de kol</i> 'near/possession'
animate	inalienable (relationships, body parts)	possessive <i>daa/de/dii/diāā</i>
inanimate	alienable (concrete objects)	<i>vicc</i> 'in'
animate	permanent (qualities, emotions)	<i>vicc</i> 'in'
animate	temporary (feelings, sensations, sicknesses)	<i>nüü</i> 'to'

Other languages show further variants on these patterns, sometimes involving a verbless clause rather than a copula clause.

16.9.3 Using intransitive verb 'exist'

When there is no copula verb—and sometimes when there is—an intransitive verb 'exist' may be employed to establish a possessive relationship. For

example, in Japanese there is verb *aru* ‘exist’ which may be used with D in S function and R as a peripheral NP marked by dative postposition, to indicate ‘R has a D’ (literally, ‘D exists to R’).

Jacaltec has no copula verb but there is an intransitive verb *ay* ‘exist’. This can take, as S argument, ‘R:POSSESSOR D’, meaning ‘R has a D’, as in (Craig 1977: 21):

- (98) *ay* [no’ *hin_R* *txitam_D*]_S
 exist CLASSIFIER:ANIMAL 1SG:POSSESSOR pig
 I have a pig (lit. My pig exists)

Here *hin* is the 1sg pronominal form used for possession and also to mark A function within a clause (see (i) in §16.6).

We noted (in §16.5.2) that in Jarawara verb *-kiha-* ‘have’ is only used with set 3 nouns (ownership). For set 1 (body parts, etc.) and set 2 (kin), the intransitive verb *wata-* ‘exist’ may be employed, as in:

- (99) [*Wero_R* *tese-ne_D*]_S
Wero(m) companion-m
wata-ra-re-ka
 exist-NEG-IMM.PAST.EYEWITNESS:m-DEC:m
 Wero had no companion (lit. Wero’s companion did not exist)

This language has copula verb *ama-* ‘be’ which generally occurs with two arguments, CS and CC. However, it can be used with just CS, then meaning ‘exist’, and used to establish possession (of a set 2 or set 3 noun as D). For example:

- (100) *o-koto_{CS}* *ama-ke*
 1sg(R)-daughter(D) be-DECLARATIVE:f
 I have a daughter (lit. My daughter is)

Wata- and *ama-* (when used with just a CS argument) may be substituted one for the other with little or no difference in meaning. But it appears that *wata-* is used more often in negative statements, such as (99), and *ama-* more often in positive ones, such as (100).

16.9.4 Other techniques

Individual languages have further techniques for establishing a possessive relationship. We can exemplify two (see also the appendix). West Greenlandic Eskimo (Fortescue 1984: 171) is a highly synthetic language with a large number of derivational suffixes. One such is *-gar-* ‘have’, which can be added to D and followed by a bound pronoun indicating R, to give a sentence ‘R has a D’. For instance:

- (101) aningaasa-ati-qar-punga
 money(D)-ALIENABLE-HAVE-1sg(R):INDICATIVE
 I have (some) money

The ‘alienable’ suffix *-ut(i)-* was discussed under (ii) in §16.4. In (101) it has the form *-ati-*, the initial *-u-* assimilating to the final *a* of *aningaasa* ‘money’.

At the end of §16.2 we showed that in Aleut a fused pronominal prefix can indicate R-plus-D when attached to a noun in an NP-internal construction and A-plus-O when attached to a verb at clause level; see (49) and (50). In other languages, A-plus-O pronouns may be used in a special kind of ‘having’ construction. Allen (1964: 340) states for Abaza (North-West Caucasian), ‘the system of verbal infixes is identical with the system of possessive prefixes; a form *d-l-š’ə-d* (“him-she-kill-ed”) is structurally comparable with *d-l-pa-b* (“he-her-son-is”). The similarity of expression for transitivity and possession could hardly be closer.’ (Pronominal prefixes are from the paradigm given at (42) in §15.1.9.).

16.10 Summary

Every language has some means for showing possession within an NP (for example, *John’s foot* or *the barrel of the gun* in English). This can be shown by simply apposing possessor (R) and possessed (D), by genitive marking on R, by pertensive marking on D, or by both of these at once. Not uncommonly, pertensive includes information concerning the person and number of R (literally ‘John gun-his’). Like many other grammatical elements, possessive markers may have a further function in the grammar, typically for indicating a core relation at clause level.

There may be a number of possessive constructions depending on the following. First, the nature and reference of R—whether a pronoun, proper noun, kin noun, or common noun which is human, animate, or inanimate. Secondly, the nature of the possessive relation, relating to time, permanence, or closeness of relationship, etc. Thirdly, according to whether D is a part term, a kin term, or something else (with many further subtypes interrelated with these). There can be subtle patternings; for instance, blood relations may be treated as ‘inalienable’, like body parts, and relations through marriage as ‘alienable’. Marking for inalienable possession tends to be grammatically tighter (and, often, shorter) than that for alienable possession.

If a language has only one type of intra-NP possession, it is always D which is head of the NP. But if there is a distinction, D will be head of an alienable possession construction and R may be head of an inalienable construction; this is especially likely when R and inalienable D are simply apposed.

There are also constructions which establish a possessive relationship. Some—but far from all—languages include a verb ‘have’ which often has irregularity morphology and syntax. Other languages may use a copula expression, such as ‘A dog is to John’ or ‘John is with a dog’ for ‘John has a dog’. Or an intransitive verb ‘exist’ may be employed; for example, ‘John’s dog exists’.

16.11 What to investigate

As far as we know, every language has an NP-internal possessive construction (which presupposes a possessive relationship). The points to investigate are:

- I. How is this shown? (§16.2.)
 - Simply by apposition of possessor (R) and possessed (D). If so, what is the ordering between them?
 - By an invariable affix on R (genitive) or on D (plain pertensive).
 - By a bound pronoun referring to R and attached to D (pronominal pertensive).
 - By a combination of these techniques.
- II. Does the construction used depend on:
 - (a) The nature of the R—whether pronoun, proper noun, kin term, or other common noun with human, animate, or inanimate reference (§16.3).
 - (b) The nature of the possessive relationship—temporal, temporary/permanent, closeness of relationship, general type of possession (§16.4).
 - (c) The nature of D. There may be two or more sets of nouns which may function as D, divided in terms of semantic sets which include: whole–part relationship, kinship, attribution, orientation/location, association, and ownership (§16.5).
- III. Do the possessive markers have further functions in the grammar? Typically, they may also mark core (and sometimes peripheral) arguments in a clause. If so, specify which arguments (§16.6).
- IV. What is the internal structure of an NP which includes possession? Specifically, is the head of the NP the D (as it always is for alienable and often also for inalienable possession) or is it the R (as it sometimes is for inalienable possession, especially when this is expressed just by apposition)? (§16.8.)
- V. Can a noun which is inherently possessed also be used as a non-possessed item? For example, when ‘head’ refers to a part of a person an obligatory pronominal affix may be required, indicating the person

and number of the possessor. But there may be an affix which can be added to an inherently possessed noun such as 'head' which enables it to be used without a specified possessor; for example, a head which had been severed from its body.

Every language also has some means for establishing a possessive construction (§16.9). Check whether this involves:

- (i) One or more verbs 'have'. Do these occur in transitive or in copula-type constructions? Do they show morphological irregularities and/or syntactic restrictions? Check if there are restrictions on what can be R and what can be D. Does a 'have' verb show any further function in the language—say as an auxiliary element marking tense and/or aspect?
- (ii) Through a copula construction. What is in copula subject (CS) function—D or R? How is the copula complement (CC) marked?
- (iii) Through an intransitive verb 'exist'.
- (iv) By some other means.

If circumstances permit, investigate diachronic and diffusional matters. What did a verb 'have' develop from? What did a genitive or plain pertensive marker develop from (or what did it develop into)? Does your language of study have similar construction(s) to a neighbour, which may indicate areal borrowing of grammatical structures?

Appendix Possessive constructions in Fijian

Fijian has a particularly rich set of NP-internal possessive constructions, relating to the nature of the possessive relationship, the nature of R, and the nature of D. These are here illustrated for the Boumaa dialect, which differs only in minor ways from the Standard Fijian dialect.

- (i) Repeating the information presented under (iv) in §16.4 (and exemplified there), Fijian has three classifier prefixes relating to four semantic choices—whether D is to be eaten or drunk by R, and whether D is owned by R or relates to R in some other way:

A, TO BE CONSUMED	—	[A1, drunk/sucked/licked	}	<i>me-</i>
			A2, eaten/chewed/smoked		'e-
B, NOT TO BE CONSUMED	—	[B1, D is related to R		
			B2, D is owned by R		<i>we- ~ o-</i>

- (ii) The possessive construction to be used depends on whether R is a personal or place name, a pronoun, a human noun, an animate noun, or an inanimate noun.
- (iii) It also depends on whether the D is a bound noun (inalienable possession) or a free noun (alienable possession). Bound nouns comprise referential kin terms, primary body-part terms, and also *yaca*- ‘name’, *ti’i*- ‘piece, portion of’, *vu’u*- ‘cause of’, *bete*- ‘use, purpose of’, and *vatu’a*- ‘form, shape of’.

Table 16.6 summarizes the possessive constructions, [A]–[E], employed for these various choices of R and D. Each of these can now be briefly characterized, for its major function.

- [A] When R is a name, and D is a bound noun, then we get D-*i* R, as in (*a* is the article used when the head of its NP is a common noun, here *liga*- ‘hand’):

(102) [a *liga*]_D-*i* Jone_R ‘John’s hand’

- [B] If R is a name and D is a free noun, we get D CLASSIFIER-*i* R:

(103) [a *madrai*]_D ‘e-*i* Jone_R ‘John’s bread (to be eaten)’

In this example, classifier ‘e- indicates that the D (*madrai* ‘bread’) is to be eaten.

TABLE 16.6. Possessive constructions in Boumaa Fijian

POSSESSOR (R)	POSSESSED (D)	
	BOUND NOUN	FREE NOUN
personal or place name	[A] suffix - <i>i</i>	[B] classifier with suffix - <i>i</i> OR [A] suffix - <i>i</i>
pronoun	[C] pronominal suffix OR [A] suffix - <i>i</i>	[D] classifier plus possessor pronoun
human noun	[C] pronominal suffix, expanded by post-head R NP OR [A] suffix - <i>i</i> OR [E] NP _D <i>ni</i> NP _R	[D] classifier plus possessor pronoun, expanded by post-head R NP
animate noun	[E] NP _D <i>ni</i> NP _R OR [C] pronominal suffix, expanded by post-head R NP	[E] NP _D <i>ni</i> NP _R OR [D] classifier plus possessor pronoun, expanded by post-head R NP
inanimate noun	[E] NP _D <i>ni</i> NP _R	

- [C] If D is a bound noun and R a pronoun, it is shown by a pertensive pronominal suffix to D:

(104) [a liqa]_D-qu_R 'my hand'

If D is a bound noun and R a common noun with human reference, then the appropriate pronominal pertensive suffix is used, and is expanded by a separate NP which follows it. For example (44–6) and:

(105) [a liqa-na levu]_D [a gone yai]_R
 ARTICLE hand-3sgR.PERTENSIVE big ARTICLE child THIS
 the big hand of this child (lit. his/her big hand, this child)

- [D] If D is a free noun and R a pronoun, then there will be a separate possessor word preceding D. It has two components—an initial classifier element (indicating the nature of the possessive relation) and a final pronominal suffix, as in (65–70) and:

(106) [a 'e-na madrai]_D
 ARTICLE CLASSIFIER(EDIBLE)-3sgR bread
 his/her bread (to be eaten)

When D is a free noun and R a common noun with human reference, the same construction is used, with the pronominal suffix expanded by a following NP, as in:

(107) [a 'e-na madrai]_D [a gone]_R
 ARTICLE CLASSIFIER(EDIBLE)-3sgR bread ARTICLE child
 the child's bread (to be eaten) (lit. his/her bread (to be eaten) the child)

- [E] If R is a noun with non-human reference, then a quite different construction is used, D *ni* R (literally, 'D associated with R'). For example *yaca ni manuaa* 'name of the battleship'. (Fijian *manuaa* is based on English *man-of-war*.)

At first glance it might appear that—as just described—there is just one possible construction type for each combination of R and D. However, during a six-month period of immersion fieldwork in the Vanua of Boumaa, I noted a number of alternative constructions in daily use. These are preceded by 'or' in Table 16.6. For example:

- When R is a pronoun and D a bound noun, we normally get construction [C], *a liqa-na* 'his/her hand'. But construction [A] may be used instead—*a liqa-i 'ea* 'his/her hand' (where *'ea* is the free form 3sg pronoun).
- When R is a name and D a free noun, construction [B] is generally employed, *a madrai 'e-i Jone* 'John's bread (to be eaten)'. But construction A is sometimes used—*a madra-i Jone* 'John's bread'.

Full details and exemplification for all the alternative constructions are in Dixon (1988a: 119–28, 134–40).

To establish a possessive relationship, Fijian uses verbs *ti'o* 'reside, remain, sit' and *tuu* 'stand'. The D is in S function and the R (which must be human) is marked by the general preposition *vei* 'to'. There is a difference in meaning—*tuu* refers to a state of permanent ownership while *ti'o* can just indicate that the R has possession of the D at that moment. For exemplification see Dixon (1988a: 128–9).

Sources and notes

There are a number of excellent general publications on possession. Seiler (1983) is a trifle old-fashioned but full of interesting facts and ideas. Heine (1997a) discusses possession in terms of a number of 'event schemes', paying particular attention to processes of grammaticalization and other diachronic considerations; it includes rich exemplification. Chapter 5 of Heine (1997b) is essentially a condensation of this material. Chappell and McGregor (1996) is a mine of information, containing twenty lengthy studies of 'inalienability' in languages of Australia, the Pacific, Asia, North America, Europe, and Africa. Nichols (1988) is a useful general survey, focusing on North American languages. Sapir (1917) is a classic early study of the range of alienable/inalienable contrasts.

16.1. NP-internal possessives have sometimes been referred to as 'attributive' possessive constructions. The name is scarcely appropriate since in none of 'John's car', 'Mary's foot', or 'Tom's mother' could either of the constituents be said to be an attribute of the other.

There have been attempts to provide an inclusive definition of 'possession' as it is found in NP-internal constructions, none of which have provided really satisfactory results. I suggest that there is little or nothing in common between, say, relationships A–F as illustrated by sentences (1–12), and that any attempt to discover some shared factors is likely to prove futile.

Two of the more enlightened attempts may be mentioned. Thompson (1996: 674) suggests, for Koyukon, three factors relating to whether an item is considered inalienable: '(1) the degree of its connectedness to the possessor (or the degree of its separability), (2) the degree of its cultural association with the possessor, and, (3) the degree of its salience apart from the possessor.' For Guaraní (Tupi family, Paraguay), Velazquez-Castillo (1996: 39) suggests (i) conceptual dependence of D on R; (ii) inseparability between R and D; (iii) permanency of the relation; (iv) inherency of the relation.

16.2. What is here termed 'pertensive' corresponds in part to 'construct state' in the Semitic grammatical tradition.

In languages where a limited set of nouns may be incorporated into a verb, a body-part term from a possessive construction (as O NP) may be so

incorporated. Rather than saying 'I slapped John's arm', one may say, literally, 'I arm-slapped John'. Useful discussions of this include Mithun (1984: 858–9) on Blackfoot (Algonquian), Velazquez-Castillo (1996) on Guaraní, and Evans (1996) on the Australian language Mayali (Gunwinygu).

It might be suggested that there is a further kind of marking within a possessive construction, where a possessive marker simply intervenes between R and D. An example of this would be *of* in English. But in fact *of* is a proclitic to the first word of the NP which follows, indicating R; *a glass of beer* is /ə=glɑ:s əv=biə/. Possession for sets 2 and 3 in Jarawara is 'R *kaa* D'. But the association between R and *kaa* is clearly shown when R is a singular pronoun. 1sg *o-* plus *kaa* gives *oko* (with assimilation), while 2sg *ti-* plus *kaa* gives *tika*. It remains now to investigate other languages from this point of view,

16.4. A contrast between present and past possession similar to that in Apalai is reported for another Carib language, variously called Tiriyó (Meira 1999: 211–20) and Trio (Carlin 2004: 103–13). Meira quotes *ji-pawana-ri-ja* 'to my friend' and *ji-pawana-hpë-ri-ja* 'to my former friend'. (This includes 1sg prefix *ji-* and dative suffix *-ja*.) Note that here the plain pertensive suffix *-ri* follows past suffix *-hpë*, whereas in the related Apalai—shown in (55–8)—past *-Vpyry* follows plain pertensive *-ny* ~ *-ry*.

16.5.1. The inalienable/alienable distinction has been accorded various names throughout the literature. Alternative labels for inalienable include inseparable, inherent, partitive, intimate, and relational, while those for alienable include separable, acquired, accidental, and transferable—see Nichols (1988: 561), where references are given, and Mithun (1999: 251).

Svorou (1993) includes a wide-ranging survey of how the same set of forms can refer to body parts and also location/orientation.

In some languages an R within a possessive NP (for which the D is head) in O function may become the whole O NP with the D placed in a peripheral NP. For instance, instead of *I slapped [John's hand]*, one can say *I slapped John_O on the hand*. This has been termed 'possessor raising' or 'possessor ascension'. See, among many other sources, Hyman (1977) on Haya (Bantu, Tanzania), Seiler (1983: 45–8), and the papers in Payne and Barshi (1999).

16.5.2. In some languages there is a special affix which can be added to what are basically inalienably possessed nouns, giving them 'unpossessed' forms. One language family in which this is found is Arawak; see Aikhenvald (1999b: 82) and further references therein.

16.7.1. *Of* is normally a proclitic /əv=/ to the word that follows, but in the clipped expression *cuppa* (short for *cup of tea*) it has fused with *cup*, giving /kʌpə/.

This subsection is closely based on Dixon (2005: 318–19). The many useful discussions of factors motivating the choice between *'s* and *of* in English include Deane (1987), Anschutz (1997), and Johansson and Oksefjell (1998); and see further references therein.

16.8. This discussion of possession in Jarawara is slightly simplified (without obscuring any of the critical points). Fuller treatment is in Dixon (2000) and especially in Dixon (2004a: 295–301, 310–35, 343–60).

16.9. Good general discussions of 'have' constructions include Benveniste (1971: 163–79), Isačenko (1974), and Heine (1997a).

Relative Clause Constructions

There are three recurrent patterns for the ways in which clauses may be combined. These can be illustrated from English:

- (1) [The boy]_A ate [the mango]_O, after I_S left
 (2) I_A know [that [the boy]_A ate [the mango]_O]_{COMPLEMENT.CLAUSE:O}
 (3) I_A know [the boy [who_A ate [the mango]_O]_{RELATIVE.CLAUSE}]_O

Sentence (1) illustrates the linking of two clauses, neither of which is contained within the other:

- (1')

MAIN CLAUSE	SUBORDINATE CLAUSE
The boy ate the mango	after I left

The second clause is marked by linker *after*, indicating that this is a temporal linkage type; see §3.11.

Sentences (2) and (3) both involve ‘embedding’, where one clause is included within another; but these are embeddings of quite different kinds. The O argument of *know* can be an NP—as in *I_A know French_O* or *I_A know [that man]_O*—or a complement clause, as in:

- (2')

MAIN CLAUSE		
A ARGUMENT	PREDICATE	O ARGUMENT = COMPLEMENT CLAUSE
I	know	that the boy ate the mango

Complement clauses are discussed in §3.10 and Chapter 18.

Sentence (3) illustrates a relative clause construction, where the relative clause is a modifier to the head noun within an NP which is, here, O argument in the main clause:

- (3')

MAIN CLAUSE				
A ARGUMENT	PREDICATE	O ARGUMENT = NP		
I	know	ARTICLE	HEAD	RELATIVE CLAUSE
		the	boy	who ate the mango

17.1 The canonical relative clause construction

There are in the literature a variety of definitions for ‘relative clause’ and a range of typological studies. These predominantly relate to surface structure. The approach in the present volume seeks to investigate the underlying structure of a language, in order to study the fundamental ways in which its grammar operates. The surface realization of underlying elements is looked upon as a secondary matter.

From this standpoint, the following characterization of a canonical relative clause construction is adopted:

A Canonical Relative Clause Construction

- (a) The construction involves two clauses—a main clause (MC) and a relative clause (RC)—making up one sentence which consists of a single unit of intonation.
- (b) The underlying structures of these two clauses must share an argument. This can be called the common argument (CA). The CA is understood to function as an argument in the MC, and as an argument in the RC. It may be stated in both clauses, or in just one, or in neither.

In (3), the CA is *(the) boy* which is in O function in the MC and in A function in the RC. It is stated in the MC, and indicated by *who* in the RC.

- (c) The RC functions as a syntactic modifier of the CA in the MC. At the semantic level, it will normally provide information about the CA which assists in focusing—or restricting—the reference of the CA. This is a ‘restrictive RC’, as in (3). Alternatively, the RC may provide further, background, information about a CA which is already uniquely identified (say, if it is a pronoun or a proper name). This is a ‘non-restrictive RC’ (discussed further in §17.3.4).
- (d) The RC must have the basic structure of a clause—involving a predicate and the core arguments required by that predicate. In some languages, an RC may also include peripheral arguments, although in others it cannot do so (see §17.3.2). And it may not be marked for all the grammatical categories which apply for an MC—tense, aspect, modality, etc. In some cases, an RC is specified for categories not available to MCs.

To exemplify the contrast under (c), suppose that I have two daughters and one son. The relative clause in (4) is restrictive since it specifies which of the daughters is being referred to:

- (4) [My daughter [who lives in Paris]_{RC}]_A drives [a Citroen]_O

Here, and throughout this chapter, realizations of the CA are underlined.

However, the RC in (5) is non-restrictive. *My son* already has unique reference (since I have only one son), with the relative clause simply supplying extra background information:

(5) [My son [, who lives in London,]_{RC}]_A drives [a Ford]_O

In English, a non-restrictive RC is set off by commas in writing and by appositional intonation in speech.

There are a number of parameters of variation for relative clause constructions:

Concerning the Common Argument

- Its nature—discussed in §17.2.1
- Its allowed functions in MC and in RC—§17.2.2
- Its possibilities for realization—§17.2.3

Concerning the Relative Clause

- How it is marked; for example, by a relative marker, by a relative pronoun, by a clitic or affix on the RC predicate, by positioning, by intonation (or by a combination of these)—§17.3.1
- Its internal structure—§17.3.2
- Its syntactic functions—§17.3.3
- Its meaning possibilities—§17.3.4

There is then discussion in §17.4 of the clues for recognizing what is a relative clause construction. §17.5 deals with a number of non-canonical types of relative clause constructions. These include: ‘co-relatives’ (or ‘correlatives’) where there is no embedding; and, following on from this, adjoined clauses which can have either a temporal or a relative clause interpretation. We then discuss what Henry Sweet (1891) dubbed ‘condensed relatives’, such as *What you say is true* in English (which is roughly an alternative expression for *That which you say is true*). And then the type of quasi-relative clause introduced by *to* in English, such as *The man to see is Fred* (this is roughly an alternative realization of *The man who you should see is Fred*).

In §17.6 there is brief consideration of some of the diachronic paths for development of relative clause constructions and of markers of relative clauses. A summary of the chapter is provided in §17.7, and in §17.8 there are hints on how to go about finding and distinguishing a relative clause construction.

Many grammatical forms have multiple function. As is mentioned in the next chapter, the great majority of markers of a complement clause—across every kind of language—have some other role in the grammar. The same applies for markers of relative clauses. A form which is used as a relative pronoun, or as

an invariant relative marker, or as a marker on the verb of an RC—all of these are likely to have further function(s) in the language, which must be carefully distinguished.

In a number of languages, some relative pronouns are homophonous with interrogative pronouns—illustrated by *The man [who did it] is dead* and *Who did it?* in English. Of course, this does not mean that a relative clause construction is a type of interrogative. There are languages in which one of the markers of a relative clause is homonymous with a demonstrative (for instance, *that* in English). This does not imply that a relative clause has deictic effect. In some languages, a verbal affix marking a relative clause is homophonous with a nominalizer. This should not be taken to mean that a relative clause is a type of nominalization.

It is important to distinguish between an RC—which has the structure of a clause, and can function as modifier within an NP—and a participle—which is an adjective derived from a verb, and can also function as modifier within an NP. Compare, say, the RC in *The priests [who had missionized the naked natives] went to heaven* and the participle in *The missionizing priests went to heaven*. In English, an RC includes a full set of core arguments, which is not possible for the participle (one can scarcely say **The native-missionizing priests* let alone **The naked-native-missionizing priests*).

In Russian and German, participles can take objects and all other clausal constituents. But an RC shows subject agreement on the same principles as an MC, whereas a participle agrees in number and gender with the head noun (as would an adjective). And while RCs allow the same tense choices as MCs, participles have only past and present forms (no future).

In some languages RCs have a similar structure to other types of subordinate clause, but there are always some distinguishing factors. For example, in Manambu it is possible to question a constituent from a temporal subordinate clause, but not one from a relative clause (Aikhenvald 2008a: 231).

Relative clauses have sometimes been called ‘adjectival clauses’, since they modify the head of an NP in a similar way to an adjective. However, in every language relative clauses do have properties significantly different from those of adjectives, so that this label is unhelpful.

Detailed studies of the grammars of a wide range of languages do suggest that each one has something which can be recognized as a relative clause construction, in terms of the characterization presented above. Some grammarians have suggested that their language of study lacks a relative clause construction—for example Kimball (1991: 525–7) on Koasati and Derbyshire (1979: 26) on Hixkaryana. But neither of these scholars posits a definition for ‘relative clause’, or mentions the criteria which they consider not to be satisfied.

17.2 The common argument (CA)

The CA is the kingpin—or binding element—of a relative clause construction. It functions as an argument within the underlying structure of the MC and of the RC. At the level of surface realization, the CA may be stated in its fullest form within the MC, or within the RC, or in both, or in neither. Irrespective of which of these applies in a particular language, within a canonical relative clause construction the RC is always a modifier to the CA in the MC (even if this occurrence of the CA has zero surface realization).

In recent years there has arisen the unfortunate habit of referring to the fullest statement of the CA as the ‘head’ of the RC. If the fullest statement of the CA is in the MC, the RC is said to be ‘externally headed’; if in the RC, it is said to be ‘internally headed’; and if the CA is not stated in either clause, then the term ‘headless’ is used.

The term ‘head’ is used in linguistics in a variety of different senses which, taken together, can be confusing and contradictory. As summarized in §3.4 and §5.6, we find the following.

- (a) In well-established usage, one item in each phrase will be called its ‘head’. The head (which can be stated or understood) is the only obligatory component, and may make up a complete phrase on its own. It is the head which dictates agreement on other items in the phrase, and it is the head which determines the properties of the phrase as a whole (such as gender in the case of an NP). For instance, in *Mary_A hates [dirty pigs [who smell badly]_{RC}]_O*, it is *pigs* which is head of the NP in O function; one could just say *Mary_A hates pigs_O*.

This is the only sense in which the term ‘head’ is used in the present volume.

- (b) Recently, new usages have been adopted by some—but by no means by all—linguists. Most notably, the predicate is dubbed the ‘head’ of a clause, although none of the criteria given for being head of a phrase applies here. Other instances of use are even more outré. To mention just one instance, when an NP is marked by an adposition, the adposition is said to be ‘head’ of that phrase.
- (c) And, as just noted, a further use of ‘head’ is to describe the fullest statement of the CA within a relative clause construction.

Consider two alternative ways of saying ‘The girls who are sitting on the ground are my friends’ in the Tibeto-Burman language Hmar (spoken in Assam; Kumar and Subbarao 2005: 130–1). The CA *nuhmei* ‘girl’ is in S function in the RC and it is the head of the NP in CS function in the MC. Only the realization differs—*nuhmei* is stated in the MC for (6) and in the RC for (7).

- (6) [[hlo cunga Ø_S in TSung]_{RC} nuhmei hai cu]_{CS} [ka
 grass LOC girl PART sit girl PL DEM 1sg
 rol hai]_{CC} ani
 friend PL be

The girls who are sitting on the grass are my friends

- (7) [[hlo cunga nuhmei_S in TSung]_{RC} Ø hai cu]_{CS} [ka
 grass LOC girl PART sit girl PL DEM 1sg
 ral hai]_{CC} ani
 friend PL be

The girls who are sitting on the grass are my friends

Use (c) of ‘head’ would label *nuhmei* in (6) as ‘head of the RC’ (called an ‘externally headed RC’). *Nuhmei* is in fact the head of that NP within the MC which includes the RC as modifier. For (7), usage (c) would again have *nuhmei* as ‘head of the RC’ (this is called an ‘internally headed RC’). *Nuhmei* is in fact the S argument within the RC.

Under use (b) of ‘head’, the head of every kind of clause is the predicate—this is *ani* for the MC and *TSung* for the RC in both (6) and (7). But under use (c), it is *nuhmei*—appearing within the MC in (6) and within the RC in (7)—which is said to be the head of the RC. So the RC has two ‘heads’!

This is plainly not a helpful use of terminology. It is most satisfactory to restrict ‘head’ to its established signification, (a), as the nuclear item within a phrase. In addition, usage (c) is based on surface structure (the place where the CA has its fullest realization), whereas an explanatory account of relative clause constructions needs to focus on underlying structure.

Note that the convention followed throughout this chapter, of marking the null realization of a CA by the zero sign, \emptyset , is in keeping with the established practice of thus marking coreferential omission within a coordinate construction; for instance *John_S came in and Ø_S sat down*.

17.2.1 The nature of the CA

There are a number of possibilities for the head of an NP. These typically include:

- common noun
- proper noun (name of a person or place)
- demonstrative
- generic term (such as *one* in English)
- pronoun

An important question to ask, in relation to every language, concerns what type of NP head may function as CA in a relative clause construction. This is

a matter which is seldom mentioned in grammars. However, Watters (2002: 201) states that for the Tibeto-Burman language Kham, the CA may only be 3rd person. The same restriction applies in Jarawara (Dixon 2004a: 525). And Genetti (2007: 129) mentions that in another Tibeto-Burman language, Dolakha Newar, the CA cannot be a pronoun. In contrast, Dyirbal permits any kind of NP head to be CA.

Languages where all RCs are of the restrictive type plainly cannot have a proper name or singular pronoun as CA, since all these have fully specified reference. In fact, the suitability of a pronoun to be CA varies widely between languages. In English, the non-subject form of the 1sg pronoun, *me*, may readily function as CA; for example:

- (8) Why do you_A denounce [me [, who have done so much for you]_{RC}]_O?

But there are more restricted possibilities for having an RC modify the subject form of 1sg, *I*. For example, the following sentence has a marginal degree of acceptability:

- (9) ???[I [, who have done so much for you,]_{RC}]_S do not deserve to be denounced

When I have heard spoken sentences like (9), in each case the relative clause has been set off by an exaggerated appositional intonation.

The generic sense of the 3sg pronoun, *he*, may be CA, as may generic term *one*, and also the generic sense of 2nd person pronoun, *you*:

- (10) [He/one/you [who venture(s) nothing]_{RC}]_S stand(s) to gain nothing

In English, the plural reading of *you* may be CA for a restrictive relative clause construction, as in (11), or for a non-restrictive one, as in (12). A candidate at a pre-election rally could say:

- (11) [You [who vote for me]_{RC}]_S will be rewarded, and [you [who vote for John Doe]_{RC}]_A will regret that act

Here *you* is 'those of you who . . .'. But at a victory party the elected person could use *you* in the sense 'all of you':

- (12) [You [, who voted for me,]_{RC}]_S will be rewarded

17.2.2 Allowed functions for CA in MC and in RC

The integrating factor in a relative clause construction is the CA. This has syntactic function in the MC and also in the RC. Sometimes the CA may be in any function in each clause. But in many languages there are restrictions, in

one or both clauses. What the allowed functions of the CA are—in MC and in RC—constitutes an important question concerning how the grammar of a language handles its relative clause construction.

In one of the most influential papers published during the second half of the twentieth century, Keenan and Comrie (1977, 1979) presented an Accessibility Hierarchy (AH) relating to the possible functions of the CA in an RC. Languages vary as to whether the CA in an RC may be:

just in subject (A, S) function	as in Malagasy
also in object (O) function	as in Welsh
also in 'indirect object' function	as in Tamil
also in 'oblique' function	as in Korean
also in possessor function	as in French
also as the Standard of Comparison ('object of comparison')	
in a comparative construction	as in English

(Note that a fair number of languages do not have an explicit comparative construction with Standard of Comparison marked; see §3.23.)

The hypotheses are (1) A language must be able to relativize subjects (that is, the subject of the RC may function as CA); (2) Any RC-forming strategy must apply to a continuous segment of the AH; and (3) Strategies that apply at one point of the AH may in principle cease to apply at any lower point. Fuller details, and information on the primary sources used, are in Keenan and Comrie (1977, 1979).

Later work has thrown up a number of exceptions to the hierarchy. Most notably, a fair number of languages with an ergative orientation allow the CA to be in just S or O (not A) function within the RC. As a consequence, it has been suggested that the two top entries in the hierarchy should be restated:

just in S function, and in whichever transitive core function is linked with
S in that language; that is, A in a language of accusative and O in one of
ergative profile
the other transitive core function—O or A respectively.

That is, some languages work in terms of S, A/O and others with S, O/A.

Following Keenan and Comrie's lead, many writers of grammars have investigated the possible functions of the CA in an RC. Fewer have attended to an equally important question—what are the possible functions for the CA in the MC? It is most helpful to link together information about possible functions in the two clauses. A sample of languages (chosen to be different from those considered by Keenan and Comrie) is given in Table 17.1.

Note that Warekena allows the CA in the MC to be in S or O function, or marked by instrumental or locative; it cannot be in A function. For Dyirbal

TABLE 17.1. Possible functions of the common argument in the two clauses of a relative clause construction

POSSIBLE FUNCTIONS OF CA		EXAMPLE LANGUAGES
in MC	in RC	
any core or peripheral function	any core or peripheral function	Fijian (Dixon 1988a: 251–5); Tariana (Aikhenvald 2003: 277–82); Nishnaabemwin dialect of Ojibwe (Valentine 2001b: 585)
any core or peripheral function	S, A, O	Jarawara (Dixon 2004a: 525–8); Longgu (Hill 1997b: 83; 1992: 275–80)
any core or peripheral function	S, O	Ilocano (Rubino 1998: 89–90)
any core, plus instrumental, dative, locative	S, O	Dyirbal (Dixon 1972: 99–103)
S, O, instrumental, locative	S, O	Warekena (Aikhenvald 1998: 273–8)
S, O	S, O	Yidiñ (Dixon 1977a: 323, 334)

the CA in the MC can be in any core function or marked by instrumental, dative, or locative; but it cannot be marked by allative or ablative.

A transitive clause in Jarawara must be in one of two forms—an A-construction, where the mood suffix on the verb agrees in gender with the A argument, or an O-construction, where mood agrees with the O argument. An A-construction is used when the A argument is ‘pivot’ (grammaticalized topic) of the stretch of discourse in which it occurs, and the O-construction when the O argument is pivot. (The two construction types play a similar functional role to active/passive or active/antipassive in other languages, but they are both fully transitive.) There is a constraint that the CA *must* be in pivot function in an RC—that is, it can be A within a transitive A-construction, O within a transitive O-construction, or S within an intransitive construction. There is no such requirement on the CA in the MC.

This ties in with a recurrent characteristic of relative clause constructions, that the CA in the RC is frequently ‘topicalized’ or ‘focused’ (see Schachter 1973). In keeping with this, when an RC is marked by a relative pronoun or by an invariable relative marker this tends to come first in the clause (which is likely to be the ‘focus’ or ‘topicalized’ position).

It is true that many languages do allow the CA to be in any function in the MC, certainly more than allow the CA to be in any function in the RC. But, as illustrated in Table 17.1, there are plenty of languages which restrict the functions of the CA in the MC. One generalization can be suggested:

The number of functions possible for the CA in the MC will always be the same as or more (never less) than those available for the CA in the RC.

A further generalization might be: each of the functions allowed for the CA in the RC will also be allowed for the CA in the MC. However, there is need for careful examination of relative clause constructions in many more languages before this could be confirmed.

There are a number of other grammatical distinctions which should be taken account of. Copula subject (CS) is often subsumed under intransitive subject (S) but, as shown in Chapter 14, S and CS are not always marked in the same way morphologically, and always have different syntactic properties. Besnier (2000: 74) states that in the Polynesian language Tuvaluan, the CA can be in S but not in CS function within the RC. And what about Copula Complement (CC) function? This typically has fewer grammatical properties than A, O, S, or CS (for example, it cannot be realized by a bound pronoun) and one might expect it to have more limited possibilities for being CA in a relative clause construction. Further work is needed to see whether this expectation is borne out.

A couple of studies have counted frequency of different argument types as CA within an RC. Based on a text count in the Panoan language Shipibo-Konibo, Valenzuela (2002: 49) gives:

O—32% S—30% A—25% Oblique—13%

Fox (1987: 858–9) undertook a text count on English relative clause constructions:

O—50% S—23% CS—16% A—11%

Note that, for each count, the order of frequency is the same: O more than S, which is in turn more than A, although the difference between S and A is smaller in Shipibo-Konibo than in English. (The data on Shipibo-Konibo includes no examples of the CA in an RC being in CS function.)

It is natural to enquire whether languages with a limited set of possible functions for the CA within a relative clause construction may perhaps be grammatically impoverished. The answer is: not at all. There are generally syntactic derivations which ‘feed’ a constraint on CA functions. This may be illustrated for Dyirbal, where the CA must be in S or O function in an RC. Consider the two simple sentences:

- (13) Jani- \emptyset _S banaga-n̄u
 John-ABSOLUTIVE return-PAST
 John returned

- (14) mani- \emptyset_O Jani- ηgu_A budi-n
 money-ABSOLUTIVE John-ERGATIVE get-PAST
 John got the money

Suppose that we want to combine these with (14) as RC within (13). The two clauses do have an argument in common, *Jani* 'John', which is in S function in (13), the MC, but in A function in (14), the RC. What we must do is derive the antipassive version of (14). The underlying A goes into S function (which, like O, is marked by absolutive case, with zero realization), the underlying O is placed in dative case, *-gu*, and the verb is marked by antipassive suffix *- ηa -* between root and tense inflection:

- (14') Jani- \emptyset_S mani-gu budil- ηa - $\tilde{n}u$
 John-ABSOLUTIVE money-DATIVE get-ANTIPASSIVE-PAST
 John got the money

The CA, *Jani*, is now in S function in (14'), and a relative clause construction can be created from (13) and (14'):

- (15) [Jani- \emptyset [\emptyset_S mani-gu
 John-ABSOLUTIVE John money-DATIVE
 budil- ηa - ηu - \emptyset]_{RC}]_S banaga- $\tilde{n}u$
 get-ANTIPASSIVE-RELATIVE-ABSOLUTIVE return-PAST
 John, who had got the money, returned

An RC in Dyirbal is marked by relative suffix *- ηu* replacing the tense suffix on the RC verb; thus *- ηu* in (15) replaces past tense *- $\tilde{n}u$* from (14'). The case appropriate to the CA within the MC is copied onto the verb of the RC, following the relative ending *- ηu* ; this is absolutive, with zero realization, in (15). In (16) the CA is in dative case in the MC, and the dative suffix *-gu* is also added to the RC verb:

- (16) jaban- \emptyset_O yibi- ηgu_A wuga-n [yara-gu [\emptyset_S
 eel-ABSOLUTIVE woman-ERGATIVE give-PAST man-DATIVE man
 mani-gu budil- ηa - ηu -gu]_{RC}]
 money-DATIVE get-ANTIPASSIVE-RELATIVE-DATIVE
 The woman gave the eel to the man who had got the money

There is a further syntactic derivation which places an instrumental NP into O function, so that it may be the CA within an RC; see example (282) in Dixon (1972: 100).

Examination of 176 relative clause constructions in Dyirbal texts reveals that the CA is in O function in 38 per cent, in underived S function in 51 per cent, and in derived S function—corresponding to an underlying A argument, as in

(15–16)—in 11 per cent. (Note that there are no copula clauses in Dyirbal.) The function of the CA in the MC is also of interest: 83 per cent are in S or O function—marked by absolutive case, as in (15)—and 16 per cent are marked by dative case *-gu*, as in (16), with an occasional example in A function (ergative case) or instrumental or locative.

It will be seen, from Table 17.1, that Yidiñ requires the CA to be in S or O function in both MC and RC. Just as in Dyirbal, there are syntactic derivations placing underlying A into surface S function, and underlying instrumental into surface O; see Dixon (1977a: 323, 334). Generally speaking, languages which allow a limited set of functions for a CA in RC or in MC (or in both) are likely to employ syntactic derivations which will put an argument originally in a non-CA-allowed function into one that is permitted. Keenan (1972) discusses a number of languages including Malagasy, where the CA can only be in S or A function in an RC; he describes passive, putting underlying O into derived S function, and circumstantial (or applicative) derivations which can put any peripheral argument into subject function.

We should also enquire whether the CA can be a constituent *within* a core or peripheral argument in the MC and/or in the RC. It appears that in English the CA can be any NP within an MC. For example, it can be a coordinand in a coordinate NP, as in:

- (17) I_A saw [[the doctor] and [the nurse [who_S lives next door]_{RC}]]_O

Or the CA can be an NP within an RC, or an NP within a complement clause, as in:

- (18) John_A decided [that [the boy [who_O solved [the puzzle]_O]_{RC}]_A should receive [the prize [which was donated by our rich patron]_{RC}]_O]_{CoCl:O}

It can also be an NP marked by a preposition, either directly within the MC or within an NP of the MC, as in:

- (19) [The dog of [the doctor [who_S lives here]_{RC}]]_A bit me_O

There is one exception—it is scarcely felicitous to have as CA in the MC a possessive phrase marked by 's, as in:

- (20) *[[The doctor [who_S lives here]_{RC}]]'s dog]_A bit me_O

Little loss is attached to the lack of this possibility. Since one can generally say either *X's Y* or the *Y of X*, the construction in (19) is available to fill the gap.

In English, the possible functions of the CA in an RC are much more restricted. Essentially, it must be a full core or peripheral argument of the RC. The CA cannot be an element within an NP. For instance, it cannot be an element within a coordinated NP (where \emptyset indicates the putative position of the CA in the RC):

- (21) *I_A saw [the doctor [who_O you like [Ø and the nurse]_O]_{RC}]_O

Nor can it be an NP within a complement clause which functions as core argument within an RC:

- (22) *I_A know [the woman [who_O the doctor told [that Ø was dying]_{CoCl:O}]_{RC}]_O

Nor may it be an NP marked by a preposition:

- (23) *I_A saw [the doctor [[the dog of whom]_A bit me_O]_{RC}]_O

Possessor NPs have special properties as CAs of RCs as they do as CAs of MCs. If an NP includes a possessor element, then it is this which may be CA, as in:

- (24) I_A saw [the doctor [[whose dog]_A bit me_O]_{RC}]_O

This is because English has a special possessor relative pronoun, *whose*. Note that the 'possessed' head of an NP cannot be CA. One cannot say either of:

- (25) *I_A saw [the dog [[which the doctor's Ø]_A bit me_O]_{RC}]_O

- (26) *I_A saw [the dog [[the doctor's which]_A bit me_O]_{RC}]_O

Some other languages do allow the CA to be an embedded element within a core or peripheral NP of an RC. That is, they include constructions similar to those shown in (21–3) and (25–6), which are not possible in English. In most such cases, the CA must be realized by a pronoun within the RC. (This is sometimes called a 'resumptive pronoun'.) For example, Modern Hebrew permits the CA to be a coordinand within an NP, similar to (21), or an argument of a complement clause within an RC, similar to (25–6), or an NP within an RC to an argument of the RC which is itself modifier to the CA in the MC, as in Borer (1984: 221):

- (27) raʔiti [ʔet ha-yeled [she=dalya_A makira
 see.PAST.1sgA ACC the-boy RELATIVE=Dalya know.PRES.sg.fem
 [ʔet ha-ʔisha [she=xashva
 ACC the-woman RELATIVE=think.PAST.3sg.fem
 ʔalay]_{RC}]_O]_{RC}]_O
 about.him

I saw the boy that Dalya knows the woman who thought about him

In this and similar relative clause constructions in Modern Hebrew, the CA must be stated as a pronoun within the RC (here 'about him'). Without this, the sentence would not be acceptable. (Note that in this language proclitic *she=* is attached to the beginning of a relative clause.)

Similar constructions, where the CA is embedded within an RC, are quoted by Keenan (1975: 407; 1985: 156) for Welsh and Egyptian Arabic. In Japanese, relative clause constructions of this type do not require the CA to be realized as a pronoun—see Kuno (1973: 238–42). Keenan (1975: 406–10) includes an instructive table (covering eighteen languages) of how the inclusion of a resumptive pronoun relates to various possibilities for the CA being embedded within the RC.

It is striking that Keenan and Comrie's hierarchy includes 'possessor', a term referring to an element within an NP, rather than directly to a core or peripheral argument. Some languages are like English in having a single type of possessor, but many languages distinguish two or more types. Typically, an 'alienable possessor' may function as CA whereas an 'inalienable possessor' (covering body parts, and perhaps also kin relations—see §1.3 and §16.5) may not; this applies in the case of Sinhala (Henadeerage 1997), among other languages.

17.2.3 Possibilities for realization of CA

In some languages, the arguments of a predicate can only be realized as NPs; the CA must then be stated as an NP. But, as described in §15.1.9, in a language with bound pronouns, an argument may be realized at two places in surface structure—by an NP and by a bound pronoun. When bound pronouns are obligatory, an NP may be employed as well, extending reference (if 3rd person) or intensifying it (if 1st or 2nd person).

For example, in Abkhaz (North-West Caucasian; Hewitt 1979: 156), one can say:

- (28) (a-xàc'a_A) (a-ph^oós_O)
 ARTICLE-man ARTICLE-woman
 d-i-š-it'
 3sg.HUMAN.O-3sg.MASC.A-kill-AORIST

With the NPs included, this sentence means 'The man killed the woman'. But if the NPs in parentheses are omitted, then *d-i-š-it'* is a well-formed sentence, meaning 'He killed him/her'. (The paradigm of bound pronoun prefixes is very similar to that given in (42) of §15.1.9 for Abaza, another dialect of the same language.) In the first prefix slot (here coding O function) *d-* indicates 3rd singular human, with no gender specified. In the second slot (here showing A function), *-i-* is 3rd singular masculine.

For a language with bound pronouns, the question to ask concerns how they relate to the realization of a CA. There are two possible scenarios:

- (A) It is sufficient to state the CA by a bound pronoun in both MC and RC. That is, realization of the CA by an NP is not required.

- (B) Although the CA will be marked by a bound pronoun in the MC and in the RC (as appropriate) it must also be realized as an NP, somewhere in the relative clause construction.

In an informative discussion of RCs in Navajo, Platero (1974) discusses whether an NP showing the CA may be stated within the RC or within the MC (he finds both alternatives possible). He does mark bound pronouns (for A, S, and O functions) on the verbs of the quoted sentences but pays no attention to how they relate to the CA. Revisiting this topic, Willie (1989: 413) takes account of bound pronouns and indeed shows that a relative clause construction in Navajo may have the CA shown just by a bound pronoun, as in:

- (29) nahacha=gíí t'óó 'ahyóí
 3sgS.hopping=RELATIVE too many
 Those that hop (i.e. grasshoppers) are too many (i.e. There are too many grasshoppers)

The first word here is a fusion of the 3sg subject pronoun, verb root 'hop', and relative clause marker *-igíí*. Another example of the CA being expressed just by a bound pronoun is a reduced version of (34) from Panare.

Navajo is thus seen to be a language of type (A). However, more work is needed on the conditions under which a CA may be realized just by bound pronoun, with no NP in sight.

Each clause in Fijian must include a subject (S/A) pronoun at the beginning of its predicate and, if transitive, an object pronoun immediately after the verb. These pronouns are free words or clitics but they qualify as 'bound pronouns' since they are obligatory and occur at fixed position within the predicate. A clause can consist just of a predicate. This may optionally be followed by NPs providing further information about core arguments (already specified within the predicate by a bound pronoun) and/or NPs realizing peripheral arguments. Note that, in basic structure, the predicate is clause-initial, and NPs can occur in any order after it. Consider:

- (30) [drau rogo-ca ti'o]PREDICATE [(o mudrau)] ([a
 2dualA listen-TR.3sgO CONTINUOUS ART 2dual ART
 we'a-qu)]A
 friend-1sgPOSS
 You two can listen to it [the story I have just recorded], (you two) (my friends)

Final *-ca* on the verb is a fusion of transitive suffix *-ci* and 3sg O pronoun. The A bound pronoun within the predicate, 2dual *drau*, is optionally expanded by a post-predicate NP. This could just involve *o mudrau* (2 dual), where *o* is the

article used in an NP whose head is a pronoun or proper name, or *a we'a-qu* ('my friend') where *a* is the article used in an NP whose head is a common noun. Example (30) has in A function a complex NP, involving the apposition of two simple NPs, *o mudrau* and *a we'a-qu*.

A typical intransitive clause is:

- (31) [drau la'o]_{PREDICATE} ([mai Ositeralia]) ([[o mudrau]]_{2dualS} come FROM Australia ART 2dual ([a we'a-qu]]_S ART friend-1sgPOSS)

You two come (from Australia), (you two) (my friends)

The bound pronoun *drau* is here in S function and is again optionally expanded by *o mudrau a we'a-qu*. There is also a peripheral NP, *mai Ositeralia* 'from Australia', introduced by preposition *mai* 'from'.

Any core or peripheral NP can be topicalized by being moved to a position before the predicate. Topicalizing *o mudrau a we'a-qu* from (31) gives:

- (31') [[o mudrau] [a we'a-qu]]_A [drau la'o] ([mai Ositeralia])

The technique for creating a relative clause construction in Fijian is to front the CA within the RC and then place the RC after the occurrence of the CA in the MC, omitting one of the two consecutive occurrences of the CA. Thus, placing (31') after (30) and omitting one of the occurrences of the CA, *o mudrau a we'a-qu*, we get:

- (32) [drau rogo-ca ti'o]_{PREDICATE} [[o mudrau] [a we'a-qu]]
[[drau la'o]_{PREDICATE} [mai Ositeralia]]_{RC}

You two can listen to it [the story I have just recorded], you two my friends who come from Australia

The common NP is, essentially, the 2dual pronoun. This must be stated in both MC and RC. The important point is that statement of the CA could *not* be confined to bound pronouns. It would be scarcely felicitous to say:

- (33) * [drau rogo-ca ti'o]_{PREDICATE} [[drau la'o]_{PREDICATE} [mai Ositeralia]]_{RC}

You two, who come from Australia, can listen to it [the story I have just recorded]

In a Fijian relative clause construction, the CA must be stated by the appropriate bound pronoun in each clause, and it *must also* be realized as an NP, outside the predicate. In this particular textual example (Dixon 1988a: 350), the CA is realized as *o mudrau a we'a-qu*. It could consist just of a pronoun

(plus article *o*) *o mudrau*. Fijian is thus a language of type (B), in contrast to Navajo, which is of type (A).

We can now survey the various positions for the fullest statement of the CA within a relative clause construction: (a) within the MC; (b) within the RC; (c) either in MC or in RC; (d) in both MC and RC; (e) in a position which could be regarded as in MC or in RC; (f) in neither MC nor RC.

(a) Fullest statement of CA is in MC

This is the most common type cross-linguistically. The RC may follow the CA in the MC or precede it. There are some languages in which either positioning is acceptable.

We should now ask about the possibilities for statement of the CA in the RC. An important matter to take account of is whether the language has bound pronouns and, if so, whether they are included in an RC. Apart from this, there are three main possibilities for statement of the CA within the RC:

- (i) The CA is shown by a relative pronoun in the RC. English is a prototypical example of this, with there being different forms of the relative pronoun according to the function of the CA in the RC (*who* for subject, *whom* for other functions) and according to whether the CA has human or non-human reference (*who/whom* versus *which*), with special forms for a CA referring to a location (*where*) or a time (*when*).
- (ii) The CA is shown by an independent pronoun within the RC.
- (iii) There is no statement of the CA in the RC, simply a 'gap' where it would be. Mangghuer, a Mongolic language from China, is of this profile (Slater 2003: 233–5).

Many languages combine (ii) and (iii). One recurrent pattern is to require the CA to be stated as a pronoun just when it is in possessor function in the RC; this is found in Lao (Enfield 2007: 115). In languages with bound pronouns, these normally mark core arguments in each clause. But if the CA in the RC is not a core argument, then a free pronoun is required. This applies to Panare (Carib, Venezuela; Payne and Payne 1999: 160–1). In (33) the CA is in S function in both clauses and need not be stated in the RC:

- (34) n-u'púma-yaj [Ø_S tē-yaj-něj]_{RC}
 3DIRECT.S-fall-PAST.PERF man go-PAST.PERF-RELATIVE.ANIMATE
 apojs
 man
 The man who left fell down

In fact the NP *apojs* 'man' could be omitted, providing another example where the CA is expressed only by a bound pronoun, similar to (29).

However, when the CA is in indirect object function in the RC, it must be stated by a free pronoun within the RC (here *t-uya* ‘3-DATIVE’):

- (35) *tě-yaj* *apojs* [*arakono* *t-tu-yaj-něj*
 go-PAST.PERF man black.monkey 1.A-give-PAST.PERF-REL.AN
 t-uya]_{RC}
 3-DATIVE

The man to whom I gave the black monkey left (lit. The man I gave a black monkey to him left)

Typically, there may be no statement of the CA in the RC for a function high on the Keenan/Comrie hierarchy, with a pronominal statement possible or required for functions lower on the hierarchy. For example, in Mandarin Chinese—a language with no bound pronouns—there is zero realization for the CA within an RC if it is in a core function (S, A, or O), but pronominal statement is needed for all other functions (Charters 1996). In this language an RC is marked by final element *de*. Compare (36) where the CA *xìn* ‘letter’ is in O function in the RC, and there realized by zero, with (37), where the CA *rén* ‘person’ is in dative function in the RC and requires to be marked by pronoun *tā* ‘him’.

- (36) [[*nǐ* *gěi* *wǒ* *yóu* *lái* \emptyset *de*]_{RC} *xìn* *yībàn* *dōu*]_{CS}
 you to me send come letter REL letter half all
 shì *yīngwén*_{CC}
 COPULA English

The letter you sent me, half was in English

- (37) [[*wǒ* *gěi* *tā* *sòng* *shū* *de*]_{RC} *rén*]_{CS} *shì* [*wǒ*
 I to him give book REL person COPULA my
 didi]_{CC}
 younger.brother

The person I gave the book to was my younger brother

Comrie (1989: 147–8) mentions that in Persian it would be unusual—but not impossible—to state a CA which is in subject function in an RC; when the CA is in O function it may optionally be shown by a pronoun, and when in any other function the pronominal statement is required.

Other languages show different techniques. In Mupun (Nigeria, Chadic branch of Afro-Asiatic; Frajzyngier 1993: 502), a CA which is in O function in the RC will be stated by a pronoun only if it has human reference. In Tariana, the CA can be stated—by a pronoun—within the RC only if it has a different function in that clause from what it has in the MC (Aikhenvald 2003: 539–40).

(b) **Fulllest statement of CA is in RC**

For only a few languages is the fullest realization of the CA exclusively in the RC. Wappo (Yukian, California; Li and Thompson 1976) has no bound pronouns. As a consequence, the CA of a relative clause construction is only realized through an NP, and this is always stated within the RC, there being no statement of the CA within the MC. This language has AOV, SV constituent order, with adjectival modifiers following a head noun. Suffix *-i* is added to an A or S NP in an MC; however, *-i* is not used on an NP in CS function, or on the subject within a subordinate clause—complement clause, adverbial clause, or relative clause.

Consider the relative clause construction:

- (38) [Ø [[ce k'ew]_A ?ew_O ʔ'ohtih]_{RC}]_A ʔi_O pehkhi?
 that.man that man fish catch me look.at
 The man who was catching a fish was looking at me

If the statement of the CA—*ce k'ew* ‘the man’—were in the MC it would take subject marker *-i*. The fact that it does not, shows that it is part of the RC in (38).

In fact, the string of words in (38) could have a different meaning, ‘The fish that the man was catching was looking at me’. The underlying structure for this would be:

- (39) [Ø [[ce k'ew]_A ʔew_O ʔ'ohtih]_{RC}]_A ʔi_O pehkhi?
 fish that man fish catch me look.at
 The fish that the man was catching was looking at me

The statement of the CA here—*ʔew* ‘fish’—could not be analysed as an element of the MC, since it is surrounded by other elements of the RC.

Kobon, spoken in Papua New Guinea (Davies 1981: 30), is similar to Wappo in that the CA is stated in the RC, leading to an ambiguous construction.

- (40) [ñi_A pai_O pak-öp]_{RC:O} Ø_O yad_A nöŋ-bin
 boy girl hit-PERFECT.3sg CA 1sg perceive-PERFECT.1sg

Sentence (40) could mean ‘I know the boy who hit the girl’ or ‘I know the girl who the boy hit’. It can be disambiguated by also including statement of the CA in the MC:

- (41) [ñi_A pai_O pak-öp]_{RC:O} ñi_O yad_A nöŋ-bin
 boy girl hit-PERFECT.3sg boy 1sg perceive-PERFECT.1sg
 I know the boy who hit the girl
- (42) [ñi_A pai_O pak-öp]_{RC:O} pai_O yad_A nöŋ-bin
 boy girl hit-PERFECT.3sg girl 1sg perceive-PERFECT.1sg
 I know the girl who the boy hit

In fact, (40) could also be interpreted as a complement clause construction ‘I saw/know that the boy hit the girl’.

When the CA is stated just in the MC, one can generally recognize which NP is the CA since—in most languages—it has an RC juxtaposed either before or after it. However, if the CA is stated in the RC (and left blank in the MC) it may be impossible, out of context, to identify which NP in the RC is the CA. Another example of this is Navajo where Platero (1974: 209) shows that a sentence glossed as ‘[Ø [boy_A dog_O 3:3:PERFECTIVE:kick-RELATIVE]_{RC}]_S IMPERFECTIVE:3:bark’ can mean either ‘The dog that the boy kicked is barking’ or ‘The boy that kicked the dog is barking’. The former interpretation is preferred, on pragmatic grounds, since boys do not (as a rule) bark. (See also Willie 1989: 423–31.)

The chance of ambiguity when the major statement of the CA is in the RC is undoubtedly one reason why the great majority of languages prefer to state the CA within the MC.

Murinyata (Australia; Walsh 1976: 287–91) has bound pronouns—indicating core arguments—attached to (or fused with) the verb. Adjectival modifiers generally follow the head noun. A case suffix is added to the last word of an NP; ergative *-re* marks A function while absolutive, with zero realization, is used for S and O. It appears that a relative clause construction requires the CA to be stated as an NP, and this is within the RC. There is no specific marking on a relative clause. Consider the following intransitive and transitive sentences.

- (43) mutʷiŋga-ø_S paŋ-anduwi mundakjajya
 old.woman-ABSOLUTIVE 3sg-arrive:PERFECT earlier
 The old woman arrived earlier

- (44) mutʷiŋga-re_A ŋjayi-ø_O pan-ŋi-bard
 old.woman-ERGATIVE 1sg-ABSOLUTIVE 3sgA-1sgO-hit
 The old woman hit me

When (44) is placed as an RC within (43) as MC, we get:

- (45) [Ø [mutʷiŋga-re_A ŋjayi-ø_O
 old.woman old.woman-ERGATIVE 1sg-ABSOLUTIVE
 pan-ŋi-bard]_{RC}]_S paŋ-anduwi mundakjajya
 3sgA-1sgO-hit 3sg-arrive:PERFECT earlier
 The old woman who hit me arrived early

If the statement of CA, *mutʷiŋga* ‘old woman’ were part of the MC, it would not bear any suffix. The fact that *mutʷiŋga* takes ergative marker *-re* in (45) shows that it is a constituent of the RC.

(c) CA can be stated in MC or in RC

A number of languages allow the CA to be stated as an NP in the MC or in the RC (but not in both at the same time). This applies to some languages without bound pronouns such as Hmar, illustrated by (6–7) in §17.2, the Panoan language Shipibo-Konibo (Valenzuela 2002; 2003: 445–86), and Korean (Sohn 1994: 61–8).

Quechua is an example of a language with bound pronouns—here, for S and A functions—which allows for the CA to be stated as an NP in either clause. For example, one can have either of (Cole 1987: 277–9):

- (46) [[nuna_A Ø_O ranti-shaq-n]_{RC} bestya-Ø_{CS} [alli
man horse buy-PERFECT-3 horse-NOM good
bestya-m]_{CC} ka-rqo-n
horse-VALIDATOR be-PAST-3
The horse that the man bought was a good horse

- (47) [[nuna_A bestya-ta_O ranti-shaq-n]_{RC} Ø_{CS} [alli
man horse-ACC buy-PERFECT-3 horse good
bestya-m]_{CC} ka-rqo-n
horse-VALIDATOR be-PAST-3
The horse that the man bought was a good horse

The CA *bestya* ‘horse’ is in O function in the RC and in CS function in the MC. In (46) it is in nominative case (with zero marking), appropriate to CS function, and must thus belong to the MC. In (47) it is in accusative case (shown by suffix *-ta*) and is thus a constituent of the RC.

Other languages which appear to allow the CA to be stated in either clause include Navajo (Platero 1974) and a number from the Tibeto-Burman family, including the Kiranti languages (Bickel 1999).

It appears that in the languages just mentioned there is freedom of statement for the CA—it may be placed within the MC or within the RC. For other languages, the two possibilities are conditioned. Mojave, from the Yuman family, has bound pronouns for core functions, and case marker *-č* on an NP in S or A function; an NP in O function has zero marking. The basic constituent order is AOV, SV. This language has two varieties of relative clause construction, depending on the function of the CA in the RC (Munro 1976: 187–218):

- If the CA is in subject (A or S) function in the RC, the CA is stated in the MC (not in the RC) and, in addition, the verb of the RC bears ‘subject-relative-clause’ prefix *k^w*-. This is illustrated in:

- (48) [hatčəq [Ø_A poš_O k^w-taver]_{RC}]_O ʔ-iyu:-pč
 dog dog cat RELATIVE-chase 1sgA.3sgO-see-TENSE
 I saw the dog that chased the cat

If the stated CA, *hatčəq* ‘dog’, were a constituent of the RC, it would bear subject case suffix -č; the fact that this suffix is missing indicates that *hatčəq* ‘dog’ is here a constituent of the MC.

- If the CA is in non-subject function in the RC, then the CA is stated within the RC (not in the MC) and there is no prefix *k^w*-. This is illustrated in:

- (49) [Ø [hatčəq_O ʔavi:-m ʔ-u:tav]_{RC} -n^y-č]_S
 dog dog rock-WITH 1sgA.3sgO-hit -DEM-SUBJECT
 ø-ən^yəʔi:l^y-pč
 3sgS-be.black-TENSE
 The dog I hit with the rock is black

The zero marking on *hatčəq* ‘dog’ is consistent with its being in O function in the RC.

As in other languages where the CA is stated in the RC, the string of words in (49) is open to a second interpretation, ‘The rock with which I hit the dog is black’ (the underlining would then go on *ʔavi* ‘rock’). Disambiguation would then be in terms of pragmatic context.

Other Yuman languages behave in a similar way; see Gorbet (1976) and Miller (2001) on varieties of Diegueño.

Tibeto-Burman languages typically mark a relative clause construction by an affix on the verb of the RC (the same affix may mark a nominalization, but with different grammatical properties—see Genetti 1992: 425). There are generally a number of relative-clause affixes—which one is used may depend on the function of the CA in the RC, on the reference of the CA (for example, whether animate or inanimate), and on the aspect value of the sentence. For some relative-clause suffixes, the CA is stated in the MC and for others it is in the RC. Good descriptions of individual languages, with comparative assessment, include Mazaudon (1978), Genetti (1992), DeLancey (1999), and Huber (2003).

(d) Statement of CA is split between MC and RC

This can be illustrated for the Australian language Yidiñ. First, a note about the structure of discourse in Yidiñ is in order (Dixon 1977a: 112–18).

In some languages (for instance, Yidiñ’s southerly neighbour Dyirbal) a reply to a question may be simply ‘yes’ or ‘no’. Not so in Yidiñ—here the reply

The predicate follows core arguments in a clause; an A NP most often precedes an O NP but they can occur in the other order. Thus, AOV (or OAV) and SV. The language is rather unusual in that an RC generally comes at the end of its MC, not contiguous with the statement of the CA in that clause. (Recall from Table 17.1 in §17.2.2 that the CA must be in S or O function in both MC and RC.) There are three possibilities for the statement of the CA (Dixon 1977a: 322–41).

- (50) η ay_{uA} buña_O bunja:-ñ [\emptyset]_S
 1sg+NOMINATIVE woman hit-PAST woman
 ja:wurr_{ga}-ñ_{unda}]_{RC:O}
 yaw_n-RELATIVE
 I hit the woman who was yawning

- (51) η ayu_A Ø_O wawa:-l [[miña
1sg+NOMINATIVE mouse see-PAST CLASS:edible.animal
 mugin]_O biju:- η A buga-ñunda]_{RC:O}
 mouse eaglehawk-ERGATIVE eat-RELATIVE
I saw the mouse being eaten by the eaglehawk

(iii) CA stated part in MC and part in RC, as in:

- (52) η ay_A bana_O banji:-li-ñu [bugun_S
 1sg+NOM CLASS:liquid find-GOING-PAST spring
 bayil-ñunda]_{RC:O}
 come.out-RELATIVE
 I went and found a spring coming out (of the ground)

Here the CA *bana* [classifier] *bugun* [specific noun] is in O function in the MC and in S function in the RC. The classifier element is stated in the MC and the specific noun in the RC. Other examples include a specific noun stated in the MC and its adjectival modifier in the RC; or a demonstrative in the MC and specific noun in the RC. Sometimes one word of the CA occurs in both clauses—just specific noun in MC, and then classifier plus specific noun in RC (see example sentences in Dixon 1977a: 328–30).

A similar example from Serbo-Croatian is given by Keenan (1985: 153), from Browne (1973). It is likely that a CA being split between MC and RC would only be found in languages which allow for discontinuous statement of an NP within an MC.

(e) Indeterminate status of CA

There are languages for whose relative clause constructions the CA is stated as an NP but it is impossible to decide whether it is a constituent of the MC or of the RC (or of both simultaneously). This is reported by Asher and Kumari (1997: 56–7) for the Dravidian language Malayalam, and it is illustrated below for Fijian (Dixon 1988a: 251–5).

As mentioned above, the predicate normally comes first in a Fijian clause, and can be followed by core NPs (A and O in either order, for a transitive clause). The following intransitive clause continues comment on the established topic of the discourse ‘our ancestors’:

- (53) [e lailai]_{PREDICATE} [a 'e-dra 'aa'ana]_S
 3sgS be.little ARTICLE CLASSIFIER-3pl food
 They [our ancestors] had little food (lit. Their food was little)

The S argument is here realized by obligatory 3sg bound pronoun *e* within the predicate and by the post-predicate NP *a 'e-dra 'aa'ana*. This consists of common noun *'aa'ana* ‘food’ as head, article *a* (needed whenever the head of the NP is a common noun), and possessor *'e-dra*, which is made up of ‘edible’ classifier element *'e-* and 3pl possessor suffix *-dra*.

Any NP can be topicalized and then fronted before the predicate, as in:

- (54) [a 'e-dra 'aa'ana]_S [e saqa]_{PREDICATE}
 ARTICLE CLASSIFIER-3pl food 3sgS be.boiled

Their food was boiled

The CA must be fronted in an RC. We can have (54) as RC to (53) as MC. This is achieved by placing (54) after the CA in (53). There would then be two consecutive statements of the CA, a 'e-dra 'aa'ana, one of which is omitted, giving:

- (55) [e lailai]_{PREDICATE} [a 'e-dra 'aa'ana]_S [e
 3sgS be.little ART CLASSIFIER-3pl food 3sgS
 saqa]_{RC:S}
 be.boiled

Little of their (our ancestors') food was boiled (lit: Their food which was boiled was little)

The question now facing us is which of MC and RC should the occurrence of the CA in (55) be taken to belong to? In underlying structure, a 'e-dra 'aa'ana functions as S argument for both MC and RC. Should we take (55) as having a 'e-dra 'aa'ana stated in the MC and omitted from the RC, or vice versa? Or should a 'e-dra 'aa'ana in (55) be understood to function as a surface element of both MC and RC? In essence, the question is unimportant. To understand the grammar of the language one needs to know the underlying structure. The status of the CA in a surface structure, such as (55), is truly indeterminate and essentially immaterial.

(f) CA is stated in neither clause

At the beginning of §17.2, it was mentioned that 'head' is used in a variety of ways with respect to relative clause constructions. We can recall that if the CA is stated in the MC the construction is said to be 'externally headed' and if it is stated in the RC the label 'internally headed' is employed. Working in these terms, the description 'headless RC' should mean that the CA is explicitly stated in neither clause. It is sometimes used in this way, but not infrequently in other ways. For instance, Li and Thompson (1976: 452) describe a relative clause construction in which the CA is stated within the RC as 'headless'. And 'condensed relative clause constructions'—such as *I saw what he wanted* and *Mary liked whatever John cooked* in English (see §17.5.3 below)—are described as 'headless relatives' by Comrie and Smith (1977: 14) and by Kuroda (1992: 114–15).

Navajo has been said to have relative clause constructions where the CA is stated in neither clause. But this refers only to the CA being stated as an NP. This is a language with obligatory bound pronouns and, as shown in (29) above, the surface structure of any relative clause construction must have bound pronouns expressing the CA. Similar comments apply for Panare, illustrated in (34), and for Mojave (Munro 1976: 209).

Yidiñ is a language with no bound pronouns. It does have relative clause constructions where the CA is not stated, in surface structure, in either MC or RC. The following sentence comes from a text in which the established topic is *ɲayu* ‘I’; this was stated in an earlier clause and is not repeated here (Dixon 1977a: 328):

- (56) Ø_A Ø_O wawa:-l [Ø_S guwa ɲaɲja:-da
 1sg something see-PAST something west creek-LOC
 ɲina-ɲunda]_{RC:O}
 sit-RELATIVE

[I] saw [something] which was sitting by the creek over to the west

Many languages require the CA to be stated as an NP somewhere in the surface structure of each relative clause construction; they include Lao (Enfield 2007: 116) and Korean (Sohn 1994: 66–7). In others, the CA may be omitted only if the sentence is ‘predictable and stereotypical’, such as ‘Ø who worked’ (for ‘the one who worked’); this is found in Ute (Givón 1980: 200).

17.3 The relative clause (RC)

As set out in §17.1, a canonical relative clause construction involves the RC functioning as syntactic modifier to the CA in the MC. The RC must have the basic structure of a clause—including a predicate, and the core arguments required by the predicate.

In this section we examine the ways in which an RC may be marked, what its internal structure may be, the syntactic functions available to an RC, and finally the types of meaning and meaning contrasts associated with RCs.

17.3.1 Marking of a relative clause

There are a number of ways of marking and recognizing an RC. Each language is likely to combine a number of these.

- (a) By one intonation contour across the relative clause construction.
- (b) By the position of the RC within the MC.
- (c) By prosodic means, such as stress, tone, or creaky voice.
- (d) By an inflection on the verb of the RC.
- (e) By a relative clause marker, generally a clitic or a short grammatical word.
- (f) By a relative pronoun. This will both indicate that we have here a relative clause, and also fill the functional slot for the CA in the RC, providing information concerning the reference, function, etc. of the CA.

Note that in the great majority of instances elements from (c) to (f) have further function(s) within the grammar.

Another important factor is the overall meaning of the relative clause construction. This is dealt with in §17.4.

We can now discuss (a)–(f) in turn.

(a) Intonation contour

There is a habit among grammar-writers of paying little attention to intonation. This is a grave shortcoming, since intonation plays a major role within each grammar. It is always the case that a canonical relative clause construction—consisting of MC and RC—comprises one intonational unit. In contrast, a coordination of clauses may involve two intonational units. If we look back at the Fijian sentence (55), ‘Little of our ancestors’ food was boiled’, this is recognized as a relative clause construction in part because it is pronounced as one intonation unit. The string of words in (55) could represent two coordinated clauses (with the repeated S argument, *a 'e-dra 'aa'ana* ‘our food’, omitted from the second clause). There would then be separate intonation units—one for *e lailai a 'e-dra 'aa'ana* and another for *e saqa*. This would be shown in writing by placing a comma or semicolon after *'aa'ana*. The meaning of the coordinate construction would be quite different—‘Our ancestors had little food, and it was boiled’.

(b) Position

In some languages, the RC always comes at the end of the MC, and just on the matter of position it would then not be possible to recognize a relative clause construction; this applies for (55) from Fijian. (See the discussion in §17.4.) In other languages, an RC occurs adjacent to the CA in the MC. In some languages the RC will precede the CA—this applies for Hmar in (6–7), for Mandarin Chinese in (36–7) and for Quechua in (46–7)—and in others it follows the CA—this applies to Mojave in (48–9). The embedding of an RC within an MC provides evidence that this is a relative clause construction.

(c) Prosody

Sign languages use several production mechanisms simultaneously. An RC may be shown by a mechanism quite different from those employed for indicating lexemes and function markers. For example, in American Sign Language, the CA is stated within the RC, and the RC precedes the MC. The RC is marked by a particular facial expression—raised eyebrows, backward tilt of the head, upper lip raised. This facial expression continues throughout the RC, and stops as the MC begins (Liddell 1980: 137; Baker-Shenk and Cokley 1996: 163; Zeshan 2000: 84).

One of the ways of marking an RC in Igbo (Benue-Congo, Nigeria) is by tone shift. For instance, a noun with inherent LH tones becomes HH when functioning as CA within a relative clause construction (Emenanjo 1978: 199).

In Karajá (Macro-Jê, Brazil; Ribeiro 2006: 18) an RC is marked by stress shift. The verb of an MC is stressed on its final syllable, before tense-aspect clitics are added (these remain unstressed). But in an RC stress shifts to the last syllable of the clitics. In (57) verb *dɔrɔ* is stressed on its final syllable, *dɔ'rɔ*, and is followed by unstressed clitic complex =*de*; the whole word is *dɔ'rɔde*.

- (57) *dori_S ɔ-d-ɔ-rɔ=d-e*
 white 3-CENTRIPETAL-INTR-go.ashore=CENTRIPETAL-IMPERV
 The white man came ashore

The main clause in (57) becomes a relative clause in (58) and stress now shifts to the clitic syllable, producing *dɔrɔ'de*. (Note that the CA is stated within the RC.)

- (58) [*Ø* [*dori_S ɔ-d-ɔ-rɔ=d-e*]_{RC}]_A
 white white 3-CENTRIPETAL-INTR-go.ashore=CENTRIPETAL-IMPERV
 wa-rikok_O
 1-clay.doll
 ɔ-r-i-kōbgra-dã =r-a
 3-CENTRIFUGAL-TR-buy-VERBALIZER=CENTRIFUGAL-PERFV
 The white man who came ashore bought my clay dolls

In Colloquial Burmese (Okell 1969: 18, 59–61, 173, 357, 428), an RC is marked by ‘induced creaky tone’ (shown orthographically by an acute accent) on non-future suffix *-te* or future suffix *-me*. Compare the MC in (59), where the vowel of *-me* has level tone, with the RC in (60) where it has induced creaky tone:

- | | | | | |
|------|----------------|------|---------------------------------|------------|
| (59) | <i>nei-me</i> | (60) | [<i>nei-mé</i>] _{RC} | <i>eiñ</i> |
| | stay-FUTURE | | stay-FUTURE.RELATIVE | house |
| | (He) will stay | | The house (he) will stay in | |

As noted above, almost all markers of relative clauses have further functions in the grammar. Among those for induced creaky tone in Colloquial Burmese are marking a possessive construction, and in numeral compounds. And stress shift in Karajá also functions as a marker of complement clauses and of adverbial clauses.

(d) Verbal inflection in the RC

Many languages have an inflectional system associated with the verb, where for each clause one choice must be made from the system. This typically covers tense, aspect, and mood; in some languages it includes one or more terms which mark a relative clause.

In Dyirbal, each clause may optionally select one or more terms from a set of derivational suffixes—antipassive, reflexive, reciprocal, applicative, ‘start to do’, etc.—and then must choose one term from the inflectional system. This includes:

- tense (future/non-future in southern, and past/non-past in northern, dialects)
- purposive, ‘in order to/as a result of’
- apprehensive ‘lest’
- positive imperative, negative imperative
- relative clause markers *-ŋu* and *-mi*

Examples of relative clauses marked by *-ŋu* are at (15–16) in §17.2.2. In this language, every constituent of an NP takes the case marking which is appropriate to its function in the clause. It can be seen that in (16) the RC modifying dative noun *yara-gu* (‘man-DATIVE’) takes dative case *-gu* on its verb, following the relative clause suffix *-ŋu*. In (61)—from Dixon (1972: 101)—there is an RC modifying the instrumental noun *yugu* ‘stick’. *Yugu* is marked for instrumental case (by allomorph *-ŋgu*) and the verb of the modifying clause also takes instrumental case (allomorph *-rru*) after relative clause suffix *-ŋu*:

- (61) [bala-n jugumbil]_O [ba-ŋgu-l yara-ŋgu]_A
 THERE.ABS-FEM woman THERE-ERG-MASC man-ERG
 balga-n [yugu-ŋgu [ŋaja_A Ø]_O
 hit-PAST stick-INSTRUMENTAL 1sg:NOM stick
 maŋga-ŋu-rru]_{RC}]_{INSTRUMENT}
 pick.up-RELATIVE-INSTRUMENTAL
 The man hit the woman with the stick which I picked up

In Dyirbal, the relative clause markers are mutually exclusive with suffixes showing tense, purposive, apprehensive, and imperative. That is, relative clauses do not mark tense. Instead of this, there is an aspectual-type distinction in relative clauses, according to the suffix used:

- RC marker *-ŋu* indicates that the RC refers to something which is still in progress
- RC marker *-mi* indicates that the RC refers to something which is completed

As mentioned before, RC markers typically have some other function in the language. In Dyirbal suffixes *-ŋu* and *-mi* are also used to indicate possession, as described in §16.4:

- simple genitive suffix *-ŋu* refers to a present possession
- general genitive suffix *-mi* refers to a past possession, or something which a person owns but does not have in their present possession, etc.

In a fair number of languages, a verbal suffix may mark either a relative clause or a nominalization, with a clear syntactic difference between the constructions involved. This applies to languages from several families in South America (see Derbyshire and Pullum 1986b: 19; and Seki 2000 on Kamaiurá, from the Tupí-Guaraní branch of Tupí), and to languages from the Tibeto-Burman group (see, for instance, Genetti 2007 on Dolakha Newar).

(e) Relative clause marker

Many languages have an invariable form which marks a relative clause. Unlike a relative pronoun, it does not fill the CA slot in the RC, nor does it include any information about the reference or function of the CA. Almost all forms which function as a relative marker also have some further role(s) in the grammar—they may mark a complement clause, or a possessive construction, or function as a comparative, or be related to a demonstrative or an interrogative, and so on.

A relative marker will typically occur either at the beginning or at the end of the RC (or at both ends at once), or be added to the beginning or to the end of the verb of the RC. These can be illustrated in turn.

(i) Relative marker at beginning of RC

- Proclitic *she=* in Modern Hebrew (where the RC follows the statement of the CA in the MC), illustrated by (27) in §17.2.2. The same form marks a complement clause.
- Marker *dǝ* in Mupun is placed at the beginning of an RC (which follows the statement of the CA in the MC). It is related to the demonstrative *dě* (Frajzyngier 1993: 498).
- Particle *som* in Danish occurs clause-initially within an RC (which also follows the CA in the MC). It is also a comparative word ‘as’, and indeed this was its only function in Old Norse (Sadock 1972: 59).

(ii) Relative marker at end of RC

- In Mandarin Chinese, relative marker *de* comes at the end of an RC (which precedes the statement of the CA in the MC), as illustrated by (36–7) in §17.2.3. We also find *de* functioning as marker of a possessive construction, as a nominalizer, and as marker of stative clauses (Li and Thompson 1981: 113–23, 623).

- Marker *ká* comes at the end of an RC in Supyire (Gur family, Mali; Carlson 1994: 488, 541); the RC follows the statement of the CA in the MC. It appears that relative marker *ká* is etymologically related to the interrogative question form *ké* ‘where’.
- In Navajo, there is a final clitic on the last word of an RC, as illustrated by (29) in §17.2.3 (where the form $=(\textit{i})gíí$ is used). The relative clause precedes the CA in the MC (in cases where the CA is, in fact, stated in the MC). Clitics marking RCs in Navajo have a range of other uses in the grammar, including functioning as locational postpositions; see Young and Morgan (1987: 20–3).

(iii) Relative markers enclosing the RC

In Tok Pisin, an English-based creole which functions as a national language in Papua New Guinea (Sankoff and Brown 1976: 632), relative marker *ia* is placed immediately after the CA (just before the RC) and also at the end of the RC. Note that the CA is expressed by a noun in the MC and by a corresponding pronoun in the RC. (Each predicate bears the predicate marker *i*.)

- (62) [meri *ia* [em_S *i* yangpela meri *ia*]_{RC}]_S em_S
 woman REL 3sg PRED young woman REL 3sg
 harim *i* stap
 listen PRED CONTIN
 The woman, who was a young girl, was listening

Ia also functions as an emphatic particle in Tok Pisin (and see §17.6).

(iv) Relative marker added to the beginning of the verb in the RC

As mentioned in §17.2.3, if the CA is in subject function in the RC in Mojave, the CA is stated just in the MC, the RC follows the CA, and the verb of the RC bears a ‘subject-relative-clause’ marker, which has the form of prefix *k^w*-. This is illustrated in (48).

(v) Relative marker added to the end of the verb in the RC

In Dhaasanac (Cushitic branch of Afro-Asiatic) an RC follows the statement of the CA in the MC and its verb is followed by *ka*, as in:

- (63) [jigéɲ [ká Ø comii ka]_{RC}]_A ^hé ku
 young HERE young come:PERF RELATIVE.MARKER 3A 2O
 ʔargiyyi
 see:PERF
 The young man who came saw you

Ka also functions as a general determiner, which may occur after verbs, nouns, and modifiers (such as relative clauses). It typically follows a copula subject (Tosco 2001: 225, 283, 289).

Most relative markers are invariable, but there are languages which have a number of forms conditioned by something like the tense-aspect of the RC. For instance, in Amharic, the relative marker is prefixed to the verb of the RC. If the RC refers to past time, the relative marker is *ya-*, and this is appended to the perfect form of the RC verb, as in (64). If the RC refers to present or future time, then relative marker *yamm-* is prefixed to the imperfect form of the RC verb (Amberber 1996: 68–9; see also Leslau 1995: 81).

- (64) [[tilantina
yesterday
ya-matt'ač-iw-in]_{RC}
RELATIVE.MARKER-COME:PERF-3sg,fem.S-DEFINITE:ACCUSATIVE
setiyo]_O aya-h^w-at
woman see:PERF-1sgA-3sg.femO
I saw the woman who came yesterday

The fact that a certain grammatical form in a language has two distinct synchronic functions is generally recognized as such by anyone writing a grammar of that language. Sometimes, however, one comes across a misguided attempt to conflate the two functions. In Lakota the form *wą* marks an NP as indefinite, and is also used after a CA which is stated within an RC. Williamson (1987) concludes that a CA stated within an RC ‘must be indefinite’ (despite the fact that English translations provided clearly indicate that a CA modified by an RC generally has a definite sense; for example ‘I bought the quilt (CA) that a woman made’).

Note that when an RC is shown just by intonation and position, with no explicit prosodic or segmental marker, there may still be grammatical ambiguity. It was mentioned in §17.2.3 that a sentence such as (40) in Kobon (with the CA stated just within the RC) is not only ambiguous with respect to a relative clause interpretation—meaning either ‘I know the girl who hit the boy’ or ‘I know the girl who the boy hit’—it may also have a complement clause interpretation ‘I saw/know that the boy hit the girl’ (Davies 1981: 30).

(f) Relative pronoun

When the fullest statement of the CA is in the MC, the RC may include a relative pronoun. It has two roles—it indicates that this *is* an RC, and it fills the functional slot of the CA in the RC.

There are basically two varieties of relative pronoun. The most common type functions as an NP in the RC. But we also find bound relative pronouns,

which replace that bound pronoun within the RC which refers to the CA. This can be illustrated for Abkhaz, first repeating the simple sentence example from (28) in §17.2.3.

- (28) a-xàc'a_A a-ph^oós_O
 ARTICLE-man ARTICLE-woman
 d-i-š-it'
 3sg.HUMAN.O-3sg.MASC.A-kill-AORIST
 The man killed the woman

If (28) is embedded as RC within 'The man is coming here', with 'the man' as CA, then bound relative pronoun *-z-* 'who (A function)' replaces bound pronoun *-i-* 'he (A function)' as second prefix to verb 'kill' (Hewitt 1979: 156):

- (65) [[a-ph^oós_O də-z-š-it']_{RC}
 ARTICLE-woman 3sg.HUMAN.O-RELATIVE.A-kill-AORIST
a-xàc'a_S d-as-wèit'
 ARTICLE-man 3sg.HUMAN.S-come-PRESENT
 The man who killed the woman is coming here

To the paradigm of bound pronouns in the Abaza dialect—almost identical to those in Abkhaz which were presented at (42) in §15.1.9—we can add bound relative pronouns *y-* in column P1 and *-z-* in P2.

In Abkhaz, the CA is permitted to be in any of those functions in the CA which may be realized in part by a bound pronoun. In fact, Abkhaz (and Abaza) have four bound pronominal slots within their verb, covering almost all syntactic functions. The CA may be A, S, CS, O, or indirect object, instrumental, benefactive, locative, accompaniment, etc. However, there is no bound pronoun for copula complement function, and thus a CC argument within an RC may not function as CA (Hewitt 1979; 1981: 37–9).

In this book, 'pronoun' is used for what is often called 'personal pronoun'. In §15.1, we defined 'pronoun' as 'a small closed system of grammatical words which vary for person'. Pronouns often—but not invariably—also show number, and some may be marked for gender, etc. A bound form such as the relative *-z-* in (65) belongs in the same morphological paradigm as bound pronouns which indicate person, number, gender, and syntactic function. The prefix *-z-* marks that the clause in which it occurs is a relative clause. And it also indicates the function in the RC of the CA; but it does not show person, number, or gender. Nevertheless, because of its paradigmatic disposition, it would be not inappropriate to call it a type of pronoun.

Most 'relative pronouns' are free forms. They substitute for a CA which is generally an NP with common noun as head, and are thus quite different in nature from '(personal) pronouns'. However, the term 'relative pronoun'

is well established, and is useful in order to distinguish these from ‘relative markers’—described under (e)—which simply mark a relative clause as such.

A (non-bound) relative pronoun will include some of the following information concerning the CA for which it substitutes:

- (i) Type of reference of the CA—whether to a person, animal, or thing (there may, for instance, be different forms for human/non-human, or for animate/inanimate), place, time, manner, etc. Sometimes also gender or noun class, or choice from a classifier system.
- (ii) Number; for example singular/plural.
- (iii) Function of the CA in the RC—often by case inflection (if the language has a case system) or marked by an adposition.

The relative pronoun is sometimes in the same surface structure position as the CA would be, but it is generally fronted to the beginning of the RC. As noted in §17.2.2, a CA is often topicalized within its RC and it is thus natural that a relative pronoun (substituting for the CA) should be.

In most of the languages of Europe, a relative pronoun has the same or similar form to a content question word. Interestingly, this is rather rarely found outside Europe. And, although question word and relative pronoun may share the same form, their syntactic properties are different. For example, in English question word *which* modifies (and precedes) a noun—as in [*Which car*]_O *do you*_A *like?*—whereas relative pronoun *which* introduces an RC which follows the noun it modifies—as in *I*_A *know* [*the car* [*which you like*]_{RC}]_O. (Other examples of relative pronoun use in English are (3–5), (8–12), (17–19), (24) above.)

Most European languages have number and gender marked on some of their relative pronouns. And a relative pronoun ‘which’ will typically inflect for case in a similar way to an adjective. English is more limited in that number is not shown and only *who* (subject function)/*whom* (other functions) indicates the syntactic status of the CA. (Even here, *whom* is being replaced by *who*, especially in O function.)

Some languages have relative pronouns whose forms are not identical to those of question words, but are based on them. For example, in Georgian, relative pronouns are formed by suffixing *-c(a)* to content question words (Vogt 1971: 49). In Hungarian, relative pronouns are ‘regularly derived from question-words by means of being prefixed by *a-*, which is historically identical to one form of the demonstrative’ (Kenesei, Vago, and Fenyvesi 1998: 40). For example *ki* ‘who (question word)’ and *a-ki* ‘who (relative pronoun)’. In Yagua, spoken in Peru, relative pronouns are formed by suffixing *-tɨy* to either of the demonstratives, ‘this (animate)’ and ‘this (inanimate)’, or to the 3sg or

3pl (but not 3du) pronouns, or to ‘inanimate’, ‘other’, or ‘someone, anyone’ (Payne and Payne 1990: 342–3). An example of an RC is:

- (66) *riy-tìy=múy* *tuvaqay-su* *sa-imu*
 3pl-RELATIVE=NEGATIVE ear-VERBALIZER 3sg-LOCATIVE
 ... those who didn't listen to him

The relative pronoun here consists of the 3pl pronoun *riy* plus suffix *-tìy*. In a subordinate clause, negation is shown by clitic *=muy* attached to the first constituent.

In Lezgian, from the North-East Caucasian family, a reflexive pronoun may sometimes function as a relative pronoun. This happens when the CA ‘cannot be easily recovered because it is neither an argument of a verb or a noun nor an instrumental, locative or temporal adverbial’. In the following sentence, the CA is the standard of comparison within a comparative construction (Haspelmath 1993: 342):

- (67) [[*wiče-laj* *Šadağ q'aq'an tir*]_{RC} *dağ*]_{CS}
 SELF-SUPERRELATIVE *Šadağ* high COPULA.PARTICIPLE mountain
 hine awa?
 WHERE be.in

Where is the mountain that *Šadağ* [which is 4,243 m] is taller than?

And there are languages in which relative pronouns are independent forms, different from question words, demonstratives, reflexives, and all else. A sample of forms for nominative and oblique cases in Urdu is (Schmidt 2003: 321):

- | | | | | | |
|--|-----------------|-----------------------------------|---|---------------------|-------------------|
| (68) | | NOMINATIVE
SINGULAR AND PLURAL | | OBLIQUE
SINGULAR | OBLIQUE
PLURAL |
| 3rd person pronouns, also
functioning as demonstratives | | vo, ye | | us, is | un, in |
| Question words | ‘who’
‘what’ | kaun
kyā | } | kis | kin |
| Relative pronoun ‘who, which’ | | jō | | jis | jīn |

In languages which have a grammatical category of classifiers, a classifier may be employed as a sort of relative pronoun. Example (69) is from Lao (Enfield 2007: 113–14). The CA, *makø-muangi* ‘mango’ (which includes the ‘fruit’ classifier *makø*) is fully stated in the MC, and then modified by the RC *nuaɟi suk2 kòðn muu1* in which the CA is represented by *nuaɟi*, a numeral classifier referring to ‘one unit (of round things)’.

- (69) khòòj_{5A} siø kin₃ [makø-muang₁
 1sg.POLITE IRREALIS eat CLASSIFIER.FRUIT-mango
 [nuaj_{1s} suk_{2s} kòòn₁ muu₁]_{RC}]O
 CLASSIFIER.UNIT be.ripe before others
 I will eat the mango which is riper than the others

Some languages can have a combination of strategies for marking RCs. In many Berber languages we get:

- (α) If the CA is in S or A function in the RC, then a relative marker enclitic *-n* is added to the verb of the RC.
- (β) If the CA is not in S or A function in the MC, then the RC includes a relative particle.
- (γ) In all other circumstances, a relative particle is used.

Note that (α) and (β) overlap. If the CA is S or A in the RC and not S or A in the MC we get both the verbal ending and the relative particle. Thus, in Zekkara (North Berber, Morocco; data from Alexandra Aikhenvald):

- (70) əsy agmar_O [i-urjaz [əlli
 sell:IMPERATIVE horse TO-man RELATIVE.PARTICLE
 yusa-n]_{RC}]!
 come-RELATIVE.CLITIC
 Sell the horse to the man who comes!

The ways of marking a relative clause are heterogeneous in Wetan (Austronesian, Indonesia; Carpenter 1996). This language marks relative clauses in two ways, depending on whether the CA is an ‘actor’ or ‘non-actor’ within the RC. If it is an actor, the verb of the RC bears a relative marker *mak-* as prefix. If it is a non-actor, a portion of the RC verb is reduplicated.

17.3.2 Structure of a relative clause

An RC can never be in imperative or interrogative mood. Apart from this, it may have virtually all the possibilities open to an MC (as it does in English, for example). But in many languages an RC has less complexity (never more) than an MC.

The definition of an RC, in §17.1, states that it must have the basic structure of a clause—involving a predicate and the core arguments required by that predicate. Going beyond this, there are the following parameters of variation.

(a) Bound pronouns within the RC

In many languages, an RC has the same bound pronoun possibilities as in an MC. But there are exceptions. In Tariana the verb in a positive main clause

will bear a pronominal prefix relating to A function (if the clause is transitive) or to Sa function (if it is 'active intransitive'). An RC is marked by prefix *ka-*, which goes in the slot available to an A/Sa bound pronoun, and replaces it (Aikhenvald 2003: 253).

(b) Markers of core argument NPs within the RC

In most languages, NP core arguments are marked in the same way in an RC as in an MC. As mentioned in §17.2.3, Wappo is an exception in that an NP in S or A function within an MC bears suffix *-i*, but *-i* is not used on a subject NP in an RC (nor in a complement clause or adverbial clause).

There can be case marking restrictions of other types. For example, in Japanese, an NP in S or A function may normally be marked by nominative particle *ga*. But if the subject NP within an MC is the CA of a relative clause construction, only topic particle *wa* may be used, never *ga* (Matsumoto 1988).

(c) Including peripheral arguments within an RC

In most languages, an RC can include NPs in peripheral function, together with constituents indicating time, place, manner, etc. Jarawara is an exception in that an RC may only involve core arguments and the predicate, nothing else. The predicate can be marked by the full range of tense, modality, and evidentiality suffixes, but not of course by mood suffixes (Dixon 2004a: 525–9).

(d) Markers for non-spatial setting

In some languages, an RC may include the same grammatical marking for tense, aspect, modality, etc. as an MC. This applies for the Polynesian language Tuvaluan (Besnier 2000: 63) and the Papuan language Manambu (Aikhenvald 2008a: 469). Other languages have slightly restricted possibilities. For example, in Amharic, the verb of an RC cannot take the compound imperfect aspect (Amberber 1996: 68). In Ute, the immediate and nominal-habitual aspects 'collapse together in relative clauses' (Givón 1980: 185, 191). In Kiranti languages (Tibeto-Burman; Bickel 1999) 'the embedded nature of relative clauses brings with it that they only allow a limited set of tense markers'. For languages with evidentiality specified in the MC, this may be retained in the RC—it is in Quechua and Jarawara—but most often it is missing—as in Tariana.

An RC in Tariana does not include any of the tense and evidentiality markers which characterize an MC. However, it does have its own aspect-type system, indicating whether the event referred to by the RC is simultaneous with, preceding, or following that of the MC (Aikhenvald 2003: 537). Since in Dyirbal the relative marker is a verbal inflection, it replaces tense or other suffixes found in MCs. But, as described in §17.3.1, Dyirbal is like Tariana in that RCs have their own system of temporal contrast—suffix *-ŋu* on the verb

of an RC indicates that it refers to something which is still in progress, while *-mi* relates to something which is completed.

(e) Marking of negation

I know of no language in which an RC may not be negated. This is generally achieved in the same way as in an MC, but in some languages it is done differently. For example, in Mojave (Munro 1976: 65, 213–17) an MC is generally negated with suffix *-mot-*, whereas suffix *-m-* is employed on the verb of an RC (this is also used to negate NPs and complement clauses). Manambu has quite different techniques for negating MCs, on the one hand, and all types of dependent clauses (including RCs), on the other hand (Aikhenvald 2008a: 325–8).

(f) Marking of formality

There are languages with alternative forms for various grammatical markers, relating to degree of formality. For instance, Japanese distinguishes four ‘levels of sentence styles’—informal, polite, superpolite (or hyperpolite), and formal writing (Kuno 1973: 19). These formal distinctions are generally neutralized in RCs (although Harada 1976: 357 states that ‘performative honorifics’ may occur in an RC ‘only in the hyperpolite style’).

17.3.3 Functions of a relative clause

The discussion thus far has focused on constructions involving an MC and one RC. But in most—perhaps in all—languages, more than one RC may be associated with an MC. There can be two or more RCs, whose CAs are in different functions in the MC. An English example is:

- (71) [The man [whos lives across the road]_{RC}]_A gave [that dog [which_O you_A liked so much]_{RC}]_O to [his daughter [whos works in Stroud]_{RC}]

There can be iteration, a sequence of clauses each embedded within another, as in:

- (72) I_A saw [the dog [who_A worried [the cat [who_A chased [the mouse [who_A ate [the cheese [which_O you_A placed in the trap]_{RC}]_O]_{RC}]_O]_{RC}]_O]_{RC}]_O

Munro 1976: 202 provides examples of several types of multiple-RC constructions in Mojave.

An alternative is to have two RCs relating to a single CA in the MC. In English, these are generally conjoined with *and*, as in:

- (73) I_A saw [the man [whos lives next door]_{RC} and [who_O [your father]_A hates]_{RC}]_O

In Tuvaluan, no coordinator (like English *and*) is required. For example (Besnier 2000: 63):

- (74) a ko faafine laa see iloa [me
 BUT FOCUS women THEN NEGATIVE know COMPLEMENTIZER
 ne faafine aa konaa [e toko tolu]_{RC}
 some women WHAT THOSE NON.PAST NUMERAL three
 [e mmai]_{RC}_O
 NON.PAST come

As for the women, [I] don't know where the three women [lit. women who are three], who are coming, are from

Here the two RCs, *e toko tolu* 'are three' and *e mmai* 'are coming', modify noun *faafine* 'women' within the complement clause *me ne faafine aa konaa e toko tolu e mmai* 'the three women who are coming', which is O argument for the verb *iloa* 'know'.

The defining function of an RC is to be syntactic modifier of the CA in the MC. But in some languages RCs have a wider role. It is shown in Chapter 18 that not all languages have complement clause constructions per se. Where these are lacking, the language is likely to have a range of what can be called 'complementation strategies' (see §18.6). One such strategy is to employ an RC, as semantic equivalent of an *ing* complement clause in English. The following example is from Dyirbal:

- (75) ŋaja_A yara_O ŋamba-n [banaga-ŋu]_O
 1sg man hear-PAST return-RELATIVE

This sentence describes hearing the activity of the man returning (as would be described by the complement clause construction *I_A heard [the man returning]*_{COMPLEMENT.CLAUSE:O} in English). But it uses as complementation strategy a relative clause construction, literally 'I_A heard [the man [who was returning]_{RC}]_O'.

Dyirbal has no specific conditional marker, like English *if*, and instead employs a relative clause construction. For example (Dixon 1972: 362):

- (76) [[ŋamu bala]_O ŋinda_A ŋima-ŋu]_{RC}
 trigger THERE:NEUTER you squeeze-RELATIVE
 minba-yirri-ñu
 shoot-REFLEXIVE-FUTURE

If you squeeze the trigger [of a gun], [the gun] will go off (literally: The trigger being squeezed, [the gun] will go off)

This is a remarkably elliptical language. The CA is not stated, but its identity can be inferred, from the overall meaning of the sentence, to be *marrgin*

‘gun’. This functions as S argument of the MC and—modified by inalienably possessed noun *ŋamu* ‘trigger’ and neuter noun marker *bala*—as O argument of the RC.

Relative clause constructions in Dyirbal are also used to indicate temporal linkage. There is no overt marker ‘after’, and the way in which one conveys something like ‘After Bulbu got the money, he returned to the river’ is by employing a relative clause construction—saying, literally, ‘Bulbu, having got the money, returned to the river’.

17.3.4 Meanings of relative clauses

The most important semantic parameter relates to ‘restrictive RCs’, which assist in focusing the reference of the CA, versus ‘non-restrictive RCs’, which apply to a CA that is already fully specified, and provide background information about it. These were illustrated by (4) and (5) in §17.1.

In a fair number of languages, all RCs are of the restrictive type; this applies for Kambera (Austronesian, Indonesia; Klammer 1998: 316), Manambu (Aikhenvald 2008a: 468), and Jarawara (Dixon 2004a: 525). Tosco (2001: 282) states that Dhaasanac employs coordination for what would be expressed by a non-restrictive relative clause construction in other languages.

Some languages appear to make no formal distinction between restrictive and non-restrictive RCs, the interpretation relying entirely on pragmatic matters. Munro (1976: 205–7) reports that in Mojave there is no structural or intonational difference, but that, for a non-restrictive construction, the CA must have definite reference.

Typically, languages with both restrictive and non-restrictive RC types will have some means for distinguishing between them. There are a number of possibilities.

- (a) Position. In Amele (Gum family, Papua New Guinea; Roberts 1987: 49–56), a restrictive RC precedes the CA in the MC, and modifies it. In contrast, a non-restrictive RC follows the CA, and is in apposition to it. Also, it may be extraposed to the end of the MC, whereas a restrictive RC has fixed position.
- (b) Special relative clause marker. In Persian, a restrictive RC includes a special suffix *-i* on the statement of the CA in the MC; this is missing when the RC is non-restrictive (Mahootian 1997: 32–3; Comrie 1989: 139).
- (c) Prosody. In Hausa (Chadic branch of Afro-Asiatic, Ghana; Jaggar 1998, Newman 2000: 541–2), all types of RC are introduced by a relative marker with fixed segmental form. The difference is that for a non-restrictive RC the marker has a ‘distinctive low tone allomorph’.

For Korean (Sohn 1994: 65), the CA in a restrictive RC never receives any stress, but in the non-restrictive variety the CA is frequently

stressed. In Hungarian (Kenesei, Vago, and Fenyvesi 1998: 38–9), when the stress on the CA is ‘at the same level’ as that on the preceding RC, ‘the structure is interpreted as non-restrictive’. But if the stress on the CA ‘is decreased, restrictive interpretation is in order’. Compare (where ‘ indicates primary and , secondary stress):

- (77) [[a 'szorgalmas]_{RC} 'magyarokat]_O
 THE diligent Hungarians:ACCUSATIVE
 'mindenki_A ,megbecsüli
 everyone appreciates

Everyone appreciates the diligent Hungarians

[Non-restrictive reading—all Hungarians are diligent and all are appreciated]

- (78) [[a 'szorgalmas]_{RC} ,magyarokat]_O
 THE diligent Hungarians:ACCUSATIVE
 'mindenki_A ,megbecsüli
 everyone appreciates

Everyone appreciates the diligent Hungarians

[Restrictive reading—some Hungarians are diligent and they are appreciated]

- (d) Intonation. A common way of distinguishing between the varieties of RCs is by employing a distinctive intonation pattern for non-restrictive RCs. (In some languages this may mean that non-restrictive RCs do not satisfy the criteria stated in §17.1 for canonical relative clauses—that MC and RC make up one sentence, which is a single intonation unit.) Intonation is stated as the only (or as a major criterion) for distinguishing between restrictive and non-restrictive RCs in Hebrew (Keenan 1985: 169), in Rumanian (Mallinson 1986: 53–4), in Cairene Egyptian Colloquial Arabic (Gary and Gamal-Eldin 1982: 18), and in the Papuan language Kobon (Davies 1981: 29), as well as in English.

There are other differences between restrictive and non-restrictive RCs in English; we can mention just a couple. A restrictive RC may be marked by a *wh*- relative pronoun or by *that*, whereas *that* is never used (or used very seldom) with the non-restrictive variety. Another difference is that a relative pronoun may be omitted when the CA is not in subject function in the RC, as in *I_A saw [the man [who/that/Ø the managers sacked]_{RC}]_O*. In contrast, a relative pronoun is required in a non-restrictive RC; that is, *who* cannot be omitted from *I_A comforted [John, [who the managers had sacked]_{RC}]_O*.

A different type of semantic distinction can be illustrated from the Papuan language Yimas. Foley (1991: 403–4) distinguishes two types of RC:

- When there is tense marking on the verb of the RC, then the RC denotes a state which holds, or an action which occurs, within the time frame denoted by the tense suffix. For example, ‘the person who is tending the fire (right now, others may do so at other times)’.
- When there is no tense marking, the RC denotes a characteristic action of the referent of the CA, as in ‘the person who tends the fire (as his job)’.

17.4 How to recognize a relative clause construction

In every language, some analytic decisions are easy and others difficult. Which are easy and which difficult vary from language to language.

In two of the languages on which I did intensive fieldwork, it was not at first obvious that there is a relative clause construction. No formal marking is apparent—no relative pronoun, relative clause marker, special verbal inflection, nor special prosodic indicator. One of the languages is pretty analytic, the other highly synthetic. Yet they pose similar problems. It will be instructive to outline the problems posed and solutions found.

In §4.3 there was discussion of relative clauses in Fijian. It is appropriate to here summarize the argumentation involved. In 1985 I worked on the Boumaa dialect of Fijian (Dixon 1988a: 251–5), a language with sparse morphology. There is a very clear set of markers for complement clause constructions—*ni* ‘that’, *me* ‘should’, *dee* ‘lest’, etc. While transcribing and analysing texts, there *seemed to be* a relative clause construction, but with no formal marking at all. On what criteria could I analyse it—within the grammar of Boumaa Fijian—as a relative clause construction? After two months fieldwork there were, in the corpus collected thus far, forty-five putative relative clause constructions, including:

- (79) *mo* *rai-ca* [*a* *kuruse*]_O
 IMPERATIVE:2sgA see-TRANSITIVE:3sgO ARTICLE cross
 au *saa* *tara-a*
 1sgA ASPECT hold-TRANSITIVE:3sgO

This consists of two clauses, whose underlying forms are *mo rai-ca a kuruse!* ‘you look at the cross!’ and *au saa tara-a a kuruse* ‘I am holding the cross’. (Recall that in Fijian the predicate normally comes at the beginning of the clause, but a topicalized NP can be fronted before the predicate.)

Now (79) could be interpreted as a coordinate construction, with the repeated argument, *a kuruse* ‘the cross’, omitted from the second clause. That is, it would mean ‘You look at the cross; I am holding it’. Or *au saa tara-a* could be taken to be an RC to the CA *a kuruse* in the MC, giving ‘You look at the cross

which I am holding!’ It is difficult to choose between these competing analyses for this particular sentence.

Let us examine another example:

- (80) 'eitou mani la'i 'ani-a [a pua'a ni
 3paucalA so GO.AND eat-TRANSITIVE:3sgO ART pig of
 coocoo lailai] 'eirau 'au-ta ti'o
 forest little 2dualA bring-TRANS:3sgO CONT

The coordinate analysis here would give ‘So we (paucal) went and ate the little wild pig; we (two) brought it’. But for coordination in Fijian the order of events follows the order of clauses describing them. The pig could not have been brought in after it was eaten, making the coordination interpretation semantically implausible. In contrast, the relative clause analysis, ‘So we (paucal) went and ate the little wild pig which we (two) had brought’, describes exactly what happened.

Other sentences provide even stronger semantic evidence for a relative clause construction. Repeating (55) from §17.2.3 (which was also (2) in §4.3):

- (55) [e lailai]PREDICATE [a 'e-dra 'aa'ana]_S [e
 3sgS be.little ARTICLE CLASSIFIER-3pl food 3sgS
 saqa]_{RC:S}
 be.boiled

Little of their (our ancestors’) food was boiled (lit: Their food which was boiled was little)

If this were a coordinate construction, the meaning would be ‘Our ancestors had little food, it was boiled’. In fact the message intended to be conveyed is that the ancestors had a great deal of food but few cooking pots and, as a result, only a small portion of their food was boiled, the remainder being roasted. The meaning requires a relative clause interpretation.

Of the six ways of marking a relative clause presented in §17.3.1, those involving prosody or some segmental marking do not apply in Fijian. But for around one-third of the examples I had gathered, a relative clause analysis was required on semantic grounds. And when I went back and listened to the recordings of texts, it became clear that in each case the entire relative clause construction made up one intonation unit. In contrast, if any of (79), (80), or (55) had involved coordination, there would have been an intonational shift (which could be shown by a comma in orthography) at the end of the first clause (after *kuruse*, *lailai*, and *'aa'ana* respectively).

Position is scarcely a relevant criterion in Fijian. In the great majority of instances, an RC (like other ‘long constituents’) comes at the end of the MC and there is thus no difference in ordering between coordinate and relative

clause constructions. However, just occasionally an RC may be placed within the MC, further confirming the analysis.

In the mid-1990s, I was studying Jarawara, deep in the Amazonian jungle. This is a language with rich morphology, presenting a profile quite different from Fijian. Jarawara has an extensive set of types of subordinate clause, all but one with clear formal marking. For example, the ‘because’ clause in (81) is marked by subordinator *kaaro* following the nominalized form of the verb. The exception was a putative relative clause construction, which has no marking at all—no relative pronoun, relative clause marker, special verbal inflection, or prosodic feature. At first I wondered whether it was justified to recognize a relative clause construction in the language. Then, criteria did emerge. MC and RC make up one intonation unit. The CA is stated in the RC and the RC fills the slot which the CA would have in the MC. And, as in Fijian, a relative clause analysis is required on semantic grounds. Consider:

- (81) [\emptyset [jamas ahi na-ba]_{RC}]_S wata-ri
 thing thing(f) be.done AUX-FUT:f exist-NEG:NOMINALIZER
 kaaro
 BECAUSE:f
 [I’ll go and chop trees tomorrow] because there’s nothing else to do

If this were the coordination of two clauses—and it would then have to have an intonation break after *na-ba*—the meaning would be ‘because there will be things to do, they don’t exist’, which is paradoxical. It must be a relative clause construction whose meaning is, literally, ‘because things, which are to be done, don’t exist’—that is ‘because there’s nothing else to do’.

In essence, when in Jarawara something is recognizable as a type of subordinate clause (that is, as not an MC) and there is *no* formal marking, it is taken to be a relative clause. (For fuller discussion, with exemplification, see Dixon 2004a: 525–9.)

17.5 Non-canonical constructions

We can now survey four construction types which—in their separate ways—do not quite accord with the profile of a canonical relative clause construction as set out in §17.1.

17.5.1 The co-relative construction

One sentence may involve two clauses—sharing a CA—which are essentially coordinated, rather than one being embedded within the other. The construction achieves a similar semantic effect to a canonical relative clause construction. These were at one time named ‘co-relatives’, with the name being

gradually recast as ‘corelatives’ and then as ‘correlatives’. The highest concentration of them is in languages spoken in the South Asian sub-continent, from the Indo-Aryan, Dravidian, and also Tibeto-Burman genetic groups.

In some co-relative constructions, each clause could make up a complete sentence on its own, the CA being expressed by a content question word in the first clause and by a corresponding pronoun or demonstrative in the second. The following example is from Dolakha Newar (Tibeto-Burman, Nepal; Genetti 2007: 487):

- (82) [gunān_A [bāmā=e khā]_O ŋen-ai]_{MC},
 WHO:ERG parent=GEN talk listen-3sg-PRES
 [āmūn_A sukha_O sir-ai]_{MC}
 3sgDISTAL:ERG happiness 3sg:PRES

Whoever listens to their parent’s advice, they know happiness

The realization of the CA in the second clause is here a 3rd person ‘distal’ deictic pronoun; it may be omitted. Note that although in (82) the CA has the same function in each clause, this is not always the case. Genetti suggests that the construction in Dolakha Newar is probably a calque on similar construction types in nearby Indo-Aryan languages (see Masica 1972; 1991: 410–11).

In (83), from Malayalam (Dravidian, India; Asher and Kumari 1997: 53), the CA is realized by a content question word plus the noun ‘god’ in the first clause, and by the corresponding remote demonstrative form, again plus noun ‘god’, in the second:

- (83) [[eetə daivam]_S [ellaa vastukka]ilum]
 WHICH god all object:PL:LOC
 unŋ[oo]_{MC},
 be:PRESENT:INTERROGATIVE
 [[aa daivatte]_O praarthikkunnu]_{MC}
 THAT god:ACCUSATIVE pray:PRESENT
 [I] pray to the god who is in every object (lit. Which god is in every object, [I] pray to that god)

The closely related language Tamil (Lehmann 1993: 349–52) differs in that the first clause bears subordinating marker *-oo* (and thus cannot, as it stands, make up a complete sentence). For example:

- (84) [neerru [enta-p paiyan]_S va-nt-aan-oo],
 yesterday WHICH boy come-PAST-3sg.masc.S-SUBORDINATE
 [[anta-p paiyan-ai]_O naan_A inru paar-tt-een]_{MC}
 THAT boy-ACCUSATIVE 1sg today see-PAST-1sg

Today I saw the boy who came yesterday (lit. Which boy came yesterday, that boy I saw today)

A variant on this pattern is where the language has a set of relative pronouns distinct from content question words. In this circumstance, a relative pronoun is used in the first clause (which then could not make up a sentence on its own). This is found in Punjabi (Bhatia 1993: 53–67) and in Bengali, from which the following example is taken (Bhat 2004: 186):

- (85) [jes bajare gie chilo],
 WHO(RELATIVE) market gone had
 [ses ekthane ache]_{MC}
 3sg(CORRELATIVE) HERE be:3sgS

The person who had gone to the market is here (lit. Who had gone to the market, they are here)

Bengali has separate forms for proximate and remote 3rd person pronouns, content question words—such as *ke* ‘who’—relative pronouns—such as *je* ‘who’, in the first clause of (85)—and ‘correlative’ pronouns—such as *se*, in the second clause of (85). See the paradigm in Bhat (2004: 186). (This has some similarities to the paradigm for Urdu partially presented at (68) in §17.3.1, save that Urdu lacks distinct correlative forms.)

Some languages have both a co-relative construction and also a canonical relative clause construction, involving the same set of relative pronouns in each. There is useful discussion of this for Hindi in Mahajan (2000).

17.5.2 ‘Adjoined relative clauses’

In a classic paper, Hale (1976) describes a type of subordinate clause which occurs in many Australian languages and may function as an RC or as an adverbial clause. For example, in Warlpiri:

- (86) [ɲatjulu-rɭu=rna_A yankiri_O pantu-rnu]_{MC},
 1sg-ERGATIVE=AUXILIARY emu spear-PAST
 [kutja=lpa ɲapa_O ɲa-rnu]_{ADJOINED.CLAUSE}
 SUBORDINATOR=AUXILIARY water drink-PAST
 EITHER: I speared the emu which was drinking water
 OR: I speared the emu while it was drinking water

Typically, but not invariably, the adjoined clause ‘is separated from the main clause by a pause’.

If the two clauses share a CA, then a relative clause interpretation is possible. If the two clauses have the same tense value, then a temporal adverbial reading is possible. If both conditions hold, then the sentence is potentially ambiguous—as is (86)—and the appropriate meaning will only be inferable from discourse context. (As an alternative to the temporal sense ‘when’, the construction can be taken as a conditional, ‘if’, provided that there are appropriate tense/modality specifications.)

The same construction type is found in Rembarrnga, an Australian language with a highly synthetic profile. For example (McKay 1988: 13; 1975: 141–8, 330–64):

- (87) [ngalward_O barran-rdiy_h-mər_n]_{MC},
 stone 3plO:3sgA-hit-PAST.CONTIN
 [birri-yud-minj]
 3pl:SUBORD-run-PAST.PUNCTUAL
 EITHER: The stone hit those who ran
 OR: The stone hit them as they ran

Like (86), this construction may have a relative clause interpretation, since there is a shared CA, or an adverbial interpretation, since the two clauses have the same tense. (For this purpose, past punctual, past continuous, and past counterfactual count as identical.) The subordinate marker on the second clause is shown by phonological change; the regular 3pl prefix for S function is *barra-* and to signify that this is a subordinate clause all vowels in the prefix are raised to *i*, giving *birri-*.

Generally, a language will have this ‘adjoined relative clause’ construction instead of (rather than in addition to) the canonical construction.

17.5.3 Condensed (or fused) relative clauses

A number of languages which have a canonical relative clause construction also include a special variant of it. Typical examples from English are:

- (88) (a) [**What(ever)** John says] should be regarded with distrust
 (b) [**Who(ever)** breaks that rule] should be punished
 (c) The cat sleeps [**where(ever)** it chooses]
 (d) You can call me [**when(ever)** you are feeling lonely]

We can compare these with:

- (89) (a) [**Anything which** John says] should be regarded with distrust
 (b) [**Anyone who** breaks that rule] should be punished
 (c) The cat sleeps [**in any place where** it chooses]
 (d) You can call me [**at any time when** you are feeling lonely]

Sentences (88a–d) are rough paraphrases of (89a–d). What we have is a single form doing double duty—as ‘relative pronoun’ and as the sole statement of the CA (shown in bold). Thus, *what(ever)* instead of *anything which*, *who(ever)* instead of *anyone who*, *where(ever)* instead of *in any place where*, and *when(ever)* instead of *at any time when*. Sweet (1891: 81) refers to these as ‘condensed relatives’; a common label is ‘fused relatives’. (Confusingly, they are

called ‘headless relatives’ by Comrie and Smith 1977: 14 and by Kuroda 1992: 114–15.)

Note, though, that there is only an approximate correspondence between the sentences of (88) and those of (89). One should not say, for instance, that (88c) is simply a truncated version of (89c). The ‘condensed relatives’ have their own grammatical structure and meaning.

One interesting point concerns the different possibilities for content question words in English to function as plain relative pronouns and as ‘condensed relatives’. *Who*, *where*, and *when* show both functions, whereas *which* is only a plain relative and *what* is only a condensed one—compare (89a) and (88a).

Condensed relatives marked by *who*, *what*, *where*, and *when* have a fairly generic sense. This is accentuated by adding *-ever*. (Accusative form *whom* tends to retain the archaic linker *-so-* before *-ever*, as in *Whomsoever you told it to has spread the news far and wide*.) Interestingly, if one clause in a condensed relative construction includes *who(ever)* or *what(ever)*, that clause must come first.

There are accounts of condensed relatives in Indo-European languages; for example, Mallinson (1986: 58–60) on Rumanian, and Mahootian (1997: 35) on Persian. Schuh (1998: 273–5) describes similar constructions in the Chadic languages Miya and Hausa. Good discussions of condensed relatives in English are in Jespersen (1927: 52–77) and Huddleston and Pullum (2002: 1068–79).

17.5.4 Relative clauses marked by ‘to’

There is a further, rather unusual, construction which is found in English and in just a few other languages. It has an RC—modifying the CA in the MC—marked by ‘to’, with no relative pronoun. For example:

- (90) (a) John gave Mary [a book [to read]_{RC}]
 (b) They dug [a tunnel [to escape through]_{RC}]
 (c) July is [the best month [to go]_{RC}]
 (d) The Eiffel Tower is [the most important place [to see]_{RC}]

Each of these sentences may be paraphrased by a canonical relative clause construction where the RC includes a modal:

- (91) (a) John gave Mary [a book [which she could/should read]_{RC}]
 (b) They dug [a tunnel [which they could escape through]_{RC}]
 (c) July is [the best month [in which one should go]_{RC}]
 (d) The Eiffel Tower is [the most important place [which one should see]_{RC}]

As with the plain RCs in (89) and condensed RCs in (88), there is only an approximate correspondence between the plain RCs in (91) and the RCs marked by *to* in (90). Each has its own structure and meaning.

The grammar of each language has recurrent motifs. Parallel to the two varieties of RCs in (91) and (90), there are in English two varieties of complement clause constructions. Compare the *THAT* complement clause in (92), where there is a modal in the complement clause, with the Modal *TO* clause in (93):

(92) I told John [that he should do it]_{COMPLEMENT.CLAUSE}

(93) I told John [to do it]_{COMPLEMENT.CLAUSE}

There is thus a congruence (but no exact identity) between a complement clause construction marked by *that* and one shown by *to*—as in (92) and (93)—and also between a relative clause marked by a relative pronoun and one shown by *to*—as in (91) and (90).

For each of the *to* relative clause constructions in (90), the CA is in a function other than subject in the RC, and the reference is to some potential happening (relating to the inclusion of a modal in the corresponding sentence in (91)). If the CA is in subject—S, A, or CS—function in both MC and RC, there is a wider range of possibilities. For example:

(94) Mary is [the best person [to complain about conditions here]_{RC}]

(95) Mary was [the last person [to complain about conditions here]_{RC}]

There is an overtone of potentiality in (94), and we can suggest a corresponding canonical construction with a modal in the RC—*Mary is [the best person [who should/could complain about conditions here]_{RC}]*—although the semantic correspondence is less good than between the sentences of (90) and those of (91). In contrast, (95) refers to something in the past, and for this kind of *to* RC there is no corresponding plain-RC-with-modal. It appears that non-potential RCs marked by *to* typically include, within the NP stating the CA, a modifier such as *only*, *next*, *last*—as in (95)—*first*, *second*, etc. (Huddleston and Pullum 2002: 1067–8).

Quirk et al. (1985: 1265–9) provide a good discussion of *to* (or ‘infinitive’) RCs in English. And Bužarovska (2002) describes a similar construction in Macedonian, ‘purpose *da*-relative clauses’.

17.6 Lines of diachronic development

There are two main questions to be posed. First, how did the relative clause construction in a particular language develop? And secondly, if there is an

explicit marker for an RC, where did this come from? Only summaries of some of the answers to these questions will be presented here.

A relative clause construction can evolve from either outside or inside the grammar, as it were. We can consider two possible paths of development.

SCENARIO A. A clause providing background information can gradually work its way into the core of the grammar. It could first be conjoined to the preceding MC (with which it shares a CA). A schematic example is:

- (96) [John gave some strawberries to Mary;] [Kate had picked them]

Some kinds of co-relative constructions, described in §17.5.1, would be of this type.

The second stage would be for it to become an adjoined clause. Schematically:

- (97) [John gave some strawberries to Mary, [Kate having picked them]]

Warlpiri sentence (86) and Rembarrnga sentence (87), from §17.5.2, are examples of this type.

And finally, at the third stage, it would be ‘attracted’ into the MC, to be modifier within the statement in the MC of the CA. Continuing with the schematic examples:

- (98) [John gave some strawberries [which Kate had picked] to Mary]

Hale (1976) provides an illuminating discussion of this scenario.

There are various ways in which Scenario A could evolve. In a number of languages, case affixes marking NPs in peripheral function appear to have been generalized to mark types of RC. Consider a nominal inflection marking ‘cause’. As a first stage we could have conjoined clauses:

- (99) [John was injured spear_{CAUSAL.CASE}] [the spear pierced him] (that is, ‘John was injured from a spear; the spear pierced him’)

The peripheral NP (relating to the cause) in the first clause, and the second clause (which provides more detail about the cause), could now blend together, yielding an adjoined relative clause construction, schematically:

- (100) [John was injured, [spear pierce him]_{CAUSAL.CASE}] (that is, ‘John was injured, a spear having pierced him’)

The causal case affix, which was on ‘spear’ in (99), would in (100) be added to the verb, ‘pierce’, within the RC. The adjoined clause in (100) could then be attracted into the MC as a canonical RC:

- (101) John was injured [spear [which pierced him]_{CAUSAL.CASE:RC}] (that is, 'John was injured from a spear which pierced him')

Peripheral case affixes may initially be added to the nominalized form of the verb of the RC (see, for instance, Crowley 1983: 378–80 on the Australian language Uradhi) with the nominalizing affix later dropping out. (Example sentences (99–100) are based on what has happened in the Australian language Yidiñ; see Dixon 1977a: 322–41.).

SCENARIO B. A relative clause construction may evolve from the opposite direction, with what was a morphological derivation being expanded so that it becomes a full modifying clause. As the starting point, we can consider a nominalized verb which modifies the head of an NP. In this schematized example, modifier follows head:

- (102) John cooked [the yams gathered]

The yams must have been gathered by some person, in some place, for example, by Kate in the forest. However, the nominalized verb in (102) does not carry with it any core or peripheral arguments. There could then be an historical development whereby these arguments would be included, expanding the nominalization to be a complete clause, a relative clause:

- (103) John cooked [the yams [which Kate had gathered in the forest]]

As is often found, the same verbal form would be used both for nominalization and to mark a relative clause, there being significant grammatical differences associated with the two uses. (This type of development is reported for Tibeto-Burman languages; see, among many other sources, Matisoff 1972; Mazaudon 1978; Genetti 1991; 2007: 387–407; Bickel 1999; Delancey 1999.)

We can now look at the ways in which explicit markers of relative clause constructions evolve. As mentioned before, the great majority of prosodic and segmental forms which define an RC have some other function in the grammar. (Indeed, for every language, most grammatical forms have multiple roles.) The discussion in §17.3.1 of techniques for marking an RC listed many of these further functions.

We saw that one form may be used to mark an RC and also a complement clause and/or an adverbial clause; this applies for stress placement in Karajá, and for the relative marker in Hebrew. In some languages, it is difficult to discover which of several functions was historically prior; it could be that the various functions all came into being more or less simultaneously.

The discussion just above described how case suffixes on NPs were extended to also mark RCs (and sometimes other types of subordinate clause), and how the function of nominalizers was expanded to mark RCs. We saw, under (e)

in §17.3.1, that particle *som* was just used as a comparative word ‘as’ in Old Norse, its function then being broadened to mark RCs in modern Danish. Marking of possession, on a noun or pronoun, has in a number of languages been extended to mark an RC, as in Colloquial Burmese and in Dyirbal.

Among the best-known sources for relative pronouns are content question words and indefinites. In Proto-Indo-European, one set of forms had both interrogative and indefinite meanings. These developed into relative pronouns in the Anatolian, Italic, Germanic, and Balto-Slavic branches. Opinions vary as to whether the relative pronoun use developed from the interrogative sense or from the indefinite sense, or from both at once (Szemerényi 1996: 208–11; Watkins 1998: 66–71). As mentioned in §17.3.1, relative pronouns in Georgian and in Hungarian are based on content question words. The relative marker in Supyire appears to be related to interrogative ‘where’, and that in Lakota is identical to the indefinite marker.

Another common source for relative markers and relative pronouns is demonstratives of various kinds. Examples were mentioned in §17.3.1 for Hungarian (where they combine with interrogative forms), Mupun, and Yagua. Discussing Tok Pisin, Sankoff and Brown (1976: 663) suggest the following development for *ia*—see (62) in §17.3.1—(1) first as a place adverb ‘here’; (2) then extended to be used as a ‘postposed deictic or demonstrative’; (3) further extension to a general ‘bracketing’ use, ‘including topic-comment structures, relativization and cleft sentences’. Its use as a marker of relative clauses appears to have been an innovation within the last hundred or so years.

Most grammatical forms came, originally, out of the lexicon, and relative clause markers are no exception. Thai provides a neat example. Kullavanijaya (2008) describes how *thii* was originally just a lexeme meaning ‘piece of land’. It has developed in many directions—as a preposition ‘at’, as a classifier, as an ordinal number marker, etc. And it has become a relative clause marker. At first, *thii* was restricted to relative clause constructions where the CA referred to a place. Nowadays, *thii* has more general use, as the marker of any RC, whatever the reference of the CA. It appears that *thii* is optional, and that its presence or absence is semantically significant. According to Kuno and Wongkhamthong (1981) ‘a *thii*-less relative clause construction is used to represent the general public’s evaluation’ whereas when *thii* is included it represents ‘a concept which is based on the speaker’s personal evaluation’.

We have said a little about the origins of relative clauses. How about what they may develop into? One possible line of development is that an RC may itself be reinterpreted as an MC. In §17.3.3, it was shown that a relative clause construction in Dyirbal shows a wide range of functions—as complementation strategy, to render a conditional, and to indicate temporal linkage. Just

occasionally, a sentence may consist just of an RC (with no MC present). In one dreamtime legend, the rainbow snake has secreted the only fire in the world on a high mountain ledge. The eagle-hawk asks the satin bird to fly up and seize the fire. But the satin bird refuses, saying:

- (104) *yimba*, *ɲajas* *gayñjin-ja-ɲu*
 no 1sg have.pain-ALL.OVER-RELATIVE

The intransitive verb *gayñjin-* ‘have a pain’ takes derivational suffix *-ja-* ‘applies all over’ plus relative clause inflection *-ɲu*. The literal meaning is ‘No, I who have pains all over’. But the relative clause (plus interjection *yimba* ‘no’) constitutes a full sentence ‘No (I can’t go because) I have pains all over’. It appears that *-ɲu* is in the process of being extended to mark a type of MC, in addition to its primary function of marking an RC.

17.7 Summary

A canonical relative clause construction involves two clauses, a main clause (MC) and a relative clause (RC), which are included within a single intonation unit and make up one sentence. The RC has the essential structure of a clause—with, at the least, a predicate and the core arguments which this requires—but functions as modifier within an NP in the MC.

There is an argument common to the underlying structures of MC and of RC; this is the common argument (CA). It may receive full statement in the MC, or in the RC, or partially or fully in both clauses, or in neither. Languages vary as to what the allowed functions of the CA are, in the RC and in the MC. The possibilities for what can be head of the CA vary—only a common noun, or also a 3rd person pronoun, or also a proper noun, or also any pronoun, etc.

In most languages, a relative clause construction receives a formal marker—a relative pronoun (which also includes information concerning the reference and/or function of the CA in the RC), or a relative marker (which includes no such information), or a special verbal inflection, or some prosodic indicator. The positioning of the RC within the CA may be significant. Above all, there is a particular semantic connection between MC and RC; a relative clause construction is distinguished from other construction types in terms of meaning. A single string of words may constitute either a relative clause construction or some other construction type, but these will have different meanings. Consider the following examples from English:

- (105) [I saw [the official [John helped with the counting]_{RC}]]
 (106) [I saw the official];[John helped with the counting]

The relative clause construction, (105), states that John helped the official with the counting and I saw the official who he helped. In contrast, the unmarked coordination of clauses, in (106), makes two distinct—and possibly unrelated—statements. There would, of course, be a difference of intonation between the sentences.

There is considerable variation in the structure of RCs in different languages. An RC will in most—but not in all—languages include peripheral, as well as core, arguments. In some languages an RC can show the same tense/aspect/modality choices as an MC; in others it is allowed none of these; in a number of languages an RC has its own TAM system, different from that in the MC. In some languages all RCs are ‘restrictive’, providing further information about a referent which is not fully specified. Others also have ‘non-restrictive’ RCs, which merely append background information about some argument whose referent was already fully specified. Restrictive and non-restrictive RCs may receive different formal marking, or they may have the same structure and only be distinguishable on pragmatic grounds.

A language may have one of a number of types of non-canonical relative clause construction, either instead of or as well as the canonical variety. Co-relatives and adjoined relatives essentially involve clausal apposition. In condensed RCs, relative pronoun and the only statement of the CA are fused into one form. And there are, in just a few languages, a type of relative clause introduced by ‘to’.

17.8 What to investigate

There is no requirement, when working in terms of basic linguistic theory, that every language should have something recognizable as a canonical or non-canonical relative clause construction. A language lacking any type of relative clause construction can readily be imagined. The fact is that—working in terms of the characterizations of relative clauses provided in §17.1 and §17.5—no language which I have examined in detail exhibits this lack. Two of the languages on which I have done extensive fieldwork (and written grammars of) do not show any formal marking for an RC. But careful consideration of positioning (for Jarawara) and of intonation and semantics (for both Jarawara and Fijian) enables one clearly to identify a relative clause construction.

In most languages there is some formal marking of an RC—a relative pronoun, relative clause marker, special verbal inflection, or distinctive stress or tone or voice quality. However, in many instances this marking has further function(s) in the grammar. There are in every case some syntactic or other criteria for identifying the various functions.

One recurrent duality involves an RC marker and a complement clause marker. This applies for *that* in English, as shown in:

(107) I_A know [the cat [that_A stole [the fish]_O]_{RC}]_O

(108) I_A know [that [the cat]_A stole [the fish]_O]_{O:COMPLEMENT.CLAUSE}

In (107), *that* introduces an RC, *that stole the fish*, which modifies the CA, *the cat*, which is O argument of the MC. But in (108) *that* introduces a complement clause, *that the cat stole the fish*, which itself functions as O argument in the MC.

Another common double function is as marker of an RC and of an adverbial clause, such as *when* in English:

(109) John_A remembered [the time [when it happened]_{RC}]_O

(110) John_S laughed [when it happened]_{ADVERBIAL.CLAUSE}

In (109), relative clause *when it happened* modifies *the time* within the NP which is O argument for the transitive verb *remembered*. In contrast, *when it happened* in (110) is an adverbial clause modifying the intransitive verb *laughed*. In every instance where a relative clause marker has some other function(s) in the grammar, there will be syntactic or semantic or phonological or pragmatic criteria for distinguishing them.

Once a relative clause construction has been recognized, its various properties should be studied:

- A. What is allowed as the head of an NP which functions as CA?
- B. Where in the relative clause construction is the fullest statement of the CA?
- C. What are the allowed functions for the CA within both RC and MC?
- D. What is the marking of the RC?
- E. How do the structural possibilities within an RC compare to those for an MC?
- F. What are the full syntactic and semantic extents of RCs?

If a language has non-canonical relative clause constructions, these should be fully described. And when a non-canonical coexists with a canonical construction, the relation between these must be studied.

If it is possible to make any reliable statements about the historical origins of a relative clause construction and/or about a marker of a relative clause, these may help explain some aspects of the grammar of the relative clause construction in a currently spoken language.

Sources and notes

Classic accounts of the grammar of relative clause constructions include Keenan (1985), Lehmann (1986), and Comrie (1989). Andrews (2007) is ostensibly an update of Keenan (1985), but the earlier study is to be preferred.

A considerable number of studies focus mainly on surface ordering of CA and RC, etc. These include Dowling (1978) and Mallinson and Blake (1981).

17.2.2 This discussion of relative clause constructions in English relates to the standard variety of the language. In colloquial speech one does hear things like: *Do you remember that guy who we met him at John's party*, with the CA stated in both MC and RC.

17.2.3 Very little is said in this chapter about the surface order of constituents within a relative clause construction. This is dealt with in most of the sources mentioned just above. For example, Keenan (1985: 143) suggests that RCs which follow the CA in the MC 'are almost the only type attested in verb-initial languages'. And, that in languages with constituent order AVO, RCs following the CA 'are the overwhelming norm and are to our knowledge always the dominant or more productive form of RC'. Finally, in verb-final languages, the RC generally precedes the CA within the MC.

17.3.1 A number of languages outside Europe have recently begun using content question words as relative pronouns, but this is in virtually every instance a calque from a language like English or Spanish or Portuguese which is established as lingua franca of the area. Examples include Tariana (Aikhenvald 2003: 543–6).

Bhat (2004: 266–71) has an insightful discussion of the links between question words and relative pronouns. He suggests, 'languages that use the same pronoun as relatives and interrogatives also show affinity between interrogatives and indefinites. On the other hand, languages that have distinct sets of interrogative and indefinite pronouns do not appear to use their interrogative pronouns as relative pronouns.' There are, however, obvious exceptions to these tendencies.

There is discussion of agreement possibilities for relative pronouns in Corbett (1991: 226–42) and in Aikhenvald (2000: 39).

17.3.2 See Haiman (1980: 450) for discussion of a content question which appears to function as RC in the Papuan language Hua.

17.3.4 We stated that a restrictive RC may be marked by a *wh*- relative pronoun or by *that*, whereas *that* is never used (or used very seldom) with the non-restrictive variety. This is the basis for what we can call an example of 'copy editor lore', whereby it is said that—when the CA has non-human reference—every restrictive RC in English must be marked with *that* and every

non-restrictive RC with *which*. I am surely not the only author who has suffered from application of this misguided principle, and had to change oodles of *that*'s back into *which*'s at proof stage.

17.5.1 The term 'correlative construction' was used in traditional grammar with a quite different significance, for constructions such as *As you sow, so shall you reap* (Matthews 1997: 79).

17.5.6 See Dixon (2005: 36–53, 242–7) for discussion of the correspondence between *THAT* and Modal (*FOR*) *TO* complement clause constructions in English.

17.6 Heine and Kuteva (2002: 113–15, 174, 251) provide useful examples of relative markers and pronouns developing from interrogatives and demonstratives, together with references to further literature.

Little reliable information is available on what relative clause markers may develop into. Heine and Kuteva (2002: 254) quote Cristofaro (1998: 64–5), who in turn quotes Givón (1991) as saying that relative clause markers developed into complementizers in Biblical Hebrew. However, other sources on Biblical Hebrew do not attest this directionality of historical development; see, for example, Kautzsch (1910: 487, 491) and Waltke and O'Connor (1990).

17.8 Some discussions of relative clause constructions lack full clarity. As pointed out in §17.2, a great deal of confusion is caused by talking of 'head' rather than of 'common argument'. And a label such as 'subject relative clause' is ambiguous. Does this refer to an RC whose CA is in underlying subject function in the MC, or in the RC, or in both? Or to the fullest statement of the CA being in subject function in the MC, or in the RC?

Complement Clauses and Complementation Strategies

A complement clause is a type of clause which fills an argument slot in the structure of another clause. Most languages have a restricted set of complement-taking verbs (CTVs) which may have either an NP—as in (1) and (3) from English—or a complement clause—as in (2) and (4)—corresponding to a core argument slot.

- (1) I_A remembered [John's birthday]_{NP:O}
- (2) I_A remembered [that John was born on the eighth of August]_{CoCl:O}
- (3) I_A want [an Indian meal]_{NP:O}
- (4) I_A want [to eat a hot curry]_{CoCl:O}

(‘CoCl’ is used as the abbreviation in glosses for ‘complement clause’.)

There was a brief introduction to complement clauses in §3.10. As an illustration of the interaction between grammar and semantics, §1.9 featured a study of the four basic varieties of complement clauses in English, and how they occur with verbs from the THINKING and LIKING types, explaining this in terms of the meanings of the complement clauses and the meanings of the verbs.

The defining characteristics of a complement clause are:

- (I) It has the internal structure of a clause, at least as far as core arguments are concerned.
- (II) It functions as core argument of another clause. The range of functions available to a complement clause always includes O (object in a transitive clause).
- (III) It describes a proposition, which can be a fact, an activity, or a state (not a place or a time).

§18.1 lists the formal markings for a complement clause, and shows how to distinguish them from other construction types. §18.2 then discusses criteria (I)–(III), and §18.3 deals with the varying grammatical properties of

complement clause. §18.4 provides an account of the three recurrent types of complement clause—referring to Fact, Activity, and Potential—while §18.5 investigates the semantic types of CTVs, and which kinds of complement clauses each is likely to occur with.

The majority of the world's languages do have complement clauses. There is a fair degree of variation concerning which verbs take complement clauses, and in which syntactic function. But there are recurrent members of the set of CTVs—prototypically 'see', 'hear', 'know', 'believe', and 'like' (these generally take a complement clause in O function). If a language has an indirect speech construction, then verbs like 'tell' will take a complement clause. When the concept of wanting is coded through a lexical verb, this is also likely to take a complement clause.

A sizeable number of languages lack a complement clause construction. However, they do each have some grammatical mechanisms for coding a proposition which is remembered, wanted, seen, heard, known, believed, liked, etc.—that is, to translate sentences such as (2) and (4). These mechanisms are called complementation strategies. They include serial verb constructions, relative clauses, nominalizations, and such clause linking devices as apposition, clause chaining, and purposives. Complementation strategies are discussed in §18.6.

§18.7 provides a summary of the chapter, and §18.8 is a guide to fieldworkers on matters to investigate concerning complementation.

18.1 Distinguishing complement clauses

Most languages which include in their grammar a complement clause construction have several varieties of complement clauses, each with its own structure, grammatical properties, and meaning. Some verbs occur with more than one kind of complement clause, such as *hear* in English.

- (5) I_A heard [the result]_{NP:O}
- (6) I_A heard [(that) Brazil beat Argentina]_{CoCl:O}
- (7) I_A heard [the game]_{NP:O}
- (8) I_A heard [Brazil('s) beating Argentina]_{CoCl:O}

For (5) and (6), the O argument refers to a fact; this can be shown by an NP like *the result* or by a complement clause introduced by *that*; the *that* may here be omitted. For (7) and (8), the O argument is an activity, which can be shown by an NP such as *the game* or a complement clause marked by *-ing* on the verb and optional 's on the subject. Sentences (5) and (6) would be used when the

speaker just heard the final score; in contrast, sentences (7) and (8) describe the speaker's listening to the full unfolding of the match, perhaps via a radio commentary.

Sentence (6) is an example of a *THAT* complement clause in English. This has the same structure as a main clause with the addition of initial complementizer *that* (which may be omitted when the complement clause is in O function in the main clause, as in (6)). It takes the full set of modal and tense choices open to main clauses. In contrast, (8) involves an *ING* complement clause, shown by suffix *-ing* on its verb and possessive marker *'s* on the subject of the complement clause; the *'s* may be omissible, as it is from (8). Unlike *THAT* clauses, an *ING* clause may not include a modal or tense marking.

Russian distinguishes between reporting a fact and describing an activity but, unlike English, it uses the same complement clause construction, simply employing different complementizers—*čto* for fact and *kak* for activity, as in (Barentsen 1996: 23–4):

- (9) Len_A videl [čto Mardži igraet v kroket]_{CoCl:O}
 Len saw:IMPERV COMP Margie plays:IMPERFV IN croquet
 Len saw that Mary played croquet
- (10) Len_A videl [kak Mardži igraet v
 Len saw:IMPERFV COMP Margie plays:IMPERV IN
 kroket]_{CoCl:O}
 croquet
 Len saw Mary play croquet

A complement clause can be recognized by its grammatical profile—discussed in §18.2—and by formal marking. The latter may involve one or more of:

- (a) A complementizer element, such as *that* in English and *čto* and *kak* in Russian. This typically comes at the beginning of the complement clause and may be a separate word or a clitic (or, sometimes, an affix to the verb).
- (b) A special marker on the subject of the complement clause. As in English—see (8)—complement clauses relating to an activity frequently have possessor marking on their subject. An alternative is for the subject in a complement clause to lack an element which is found on the subject in a main clause. As mentioned under (b) in §17.2.3, in Wappo suffix *-i* is added to an A or S NP within an MC, but not to a subject NP in a complement clause, relative clause, or adverbial clause.
- (c) The verb of a complement clause may be in a special form, such as *-ing* in (8). In Jarawara the predicate comes last in a clause; if the predicate of a complement clause ends in *-a* then this is replaced by *-i* as the

mark of a complement clause—see (33–4) and (48). (In fact, *a* is by far the most common final vowel. If the predicate does end in another vowel, this method of marking is lost. But there are other criteria; for instance, a special form of a singular pronoun as complement clause subject.)

Irish combines (a) and (c). Compare ‘Sheila is young’ as a simple sentence in (11) and as a complement clause in O function to ‘think’ in (12):

- (11) Tá Síle_{CS} óg_{CC}
 be Sheila young
 Sheila is young
- (12) Sílim [go bhfuil Síle_{CS} óg_{CC}]_O
 think:1sgA THAT be Sheila young
 I think that Sheila is young

The complement clause in (12) is marked by initial element *go* ‘that’ and also by using the ‘dependent form’ of ‘be’ which is suppletive *bhfuil*. This is used in complement clauses, and also after clause-initial particles which mark negation or a question (Stenson 1981: 52, 22–4, 130).

There may be other ways of indicating that something is a complement clause. Under (c) in §17.3.1, we noted that in Karajá stress shift marks the verb of a relative clause—see (57) and (58) there. Ribeiro (2006: 37) states that complement clauses ‘are also characterized by stress shift’, but provides little detail.

In some languages, complement clauses and relative clauses are marked in the same way but may be distinguished in terms of their grammatical properties. For example, in Modern Hebrew, *she-* is used to mark both a complement clause, as in (13), and a relative clause, as in (14):

- (13) aní_A yodéa [she-hì yaf-á]_{CoCl:O}
 I know:msg:PRESENT COMP-she beautiful-fsg
 I know that she is beautiful
- (14) raíti [et ha-ishá [she-niytá yaf-á]_{RC}]_O
 see:1sgA:PAST ACC DEF-woman REL-become3fsg:PAST beautiful-fsg
 I saw the woman who became beautiful

Zuckermann (2006: 76–8) provides three criteria to distinguish the two kinds of subordinate clause. First, only a complement clause can be passivized, as in:

- (15) yadúa l-i [she-hì yaf-á]_{CoCl:S}
 know:msg:PRESENT:PASSIVE DATIVE-1sg COMP-she beautiful-fsg
 It is known to me that she is beautiful (lit. That she is beautiful is known to me)

However in (14), *she-niytá yafá* ‘who became beautiful’ behaves like an adjective and cannot be the target of passivization; that is, it is a relative clause. Other criteria are that only a complement clause can be topicalized, and only a relative clause may be coordinated with an adjective.

Different types of complex sentences may have similar surface structures, so that care has to be taken to distinguish them. Turning again to English, we can consider the relative clause construction in (16) and the complement clause construction in (17).

(16) I_A dislike [that man [(who is) painting his front door blue]_{RC}]_O

(17) I_A dislike [that man(’s) painting his front door blue]_{CoCl:O}

In (16) the relative clause (*who is*) *painting his front door blue* serves to identify which man it is that I dislike. There is no indication that I dislike what he is doing (I may quite like blue doors); I just dislike him as a man. But in (17) what I dislike is the activity; I may perfectly well approve of the man in general terms. In their full forms, the two sentences are clearly distinct. However, *who is* may be omitted from a relative clause construction such as (16), and an *ING* complement clause such as that in (17) may omit the *’s* from the subject (indeed, some speakers prefer always to omit *’s*). With these omissions, (16) and (17) have the same form, *I dislike that man painting his front door blue*. They may be distinguishable only by discourse context or by sentence stress—this is likely to go on *man* in (16) but on either *paint* or *blue* in (17).

To has a multiplicity of functions in the grammar of English. It may introduce a variety of complement clause, as in (4), and—as described in §3.11—it can be the short form of the ‘consequence’ clause linker (*in order*) *to*. We can compare the non-embedded subordinate clause structure with (*in order*) *to*, in (18), and the complement clause construction, in (19).

(18) He_S went (*in order*) to swim

(19) He_A wanted [to swim]_{CoCl:O}

Go is an intransitive verb, and *He went* can be a self-contained sentence. It may be—but need not be—extended by an (*in order*) *to* clause, as in (18). In contrast, *want* is a transitive verb and *must* take an O argument; in (19), this is the complement clause *to swim*. Now *in order* can be omitted from (18) and the two sentences would then consist of an identical string of words save for the main clause verb, *went* or *wanted*. They may be distinguished by the possibility of including *in order* in (18) but not in (19), and by the fact that *He went* can be a full clause, whereas *He wanted* is incomplete; it involves a transitive verb which requires an O argument (this can be an NP—for example, *I want a bathe*—or a complement clause).

As mentioned before, languages which lack a complement clause construction will employ some other construction type as a complementation strategy. Two of the possibilities are a relative clause construction and a purposive (*in order*) to linkage.

When a language uses a relative clause construction as a complementation strategy, it is likely to miss a distinction such as that shown between (16) and (17). That is, a given sentence may be ambiguous between a relative clause meaning, as in (16), and a complement clause meaning, as in (17). The ambiguity would be resolved by discourse context and/or by sentence stress.

When clauses linked by something like '(in order) to' are used as a complementation strategy, there may be no apparent distinction between sentences such as (18) and (19), save for the verb in the first clause. In fact, the underlying difference lies in the verbs. Something like 'go' is self-sufficient; it can be followed by a purposive clause but does not have to be. In contrast, a verb like 'want' carries the expectation of a following purposive clause, to specify what is wanted.

18.2 Grammatical criteria

We can now expand on and discuss the three defining criteria for a complement clause, set out at the beginning of this chapter.

(I) It has the internal constituent structure of a clause, at least as far as core arguments are concerned. That is, S, A, and O (or other) arguments, if not omitted by a grammatical rule associated with a particular complement clause construction, should be marked in the same way as in a main clause (allowing for the fact that part of the marking for a type of complement clause may attach to its subject) and have much the same grammatical properties.

A complement clause is an alternative to an NP as exponent of a core slot in clause structure. Grammatical criteria for distinguishing between a complement clause and an NP were outlined in §3.10. This is such an important matter that it repays repetition here, utilizing different examples from English. We can compare (20), with a complement clause as A argument, and (21), with an NP as A:

(20) [John's_A understanding [the rules]_O]_{CoCl:A} impressed Mary_O

(21) [John's understanding of the rules]_{NP:A} impressed Mary_O

The complement clause in (20), *John's understanding the rules*, has similar structure to a main clause, with an A NP, *John* (with possessive 's, one marker of the ING variety of complement clause in English), and an O NP, *the rules*, which immediately follows the verb with no preposition intervening. The verb

of the complement clause is *understand*, with suffix *-ing*, the other marker of this kind of complement clause in English.

In contrast, *John's understanding of the rules*, in (21), is an NP where the nominalization *understanding* is head noun, *John's* is possessive modifier, and *of the rules* is a post-head modifying prepositional phrase (similar to *of the table* in *the legs of the table*).

There are at least five criteria for distinguishing between a complement clause, as in (20), and an NP, as in (21):

- (i) In the complement clause, the O NP, *the rules*, immediately follows the verb, as in a main clause. In the NP, the underlying O must be marked by a preposition, here *of*.
- (ii) In the NP, the possessor *John's* is a modifier of the head noun and can be replaced by another modifier such as the article *the*, giving *The understanding of the rules impressed Mary*. In (20), the subject, *John*, bears *'s*, which is a marker of this variety of complement clause; *John's* cannot here be replaced by *the*; that is, we cannot have **The understanding the rules impressed Mary*.
- (iii) The verb of the complement clause, *understanding*, may be modified by an adverb. As in a main clause, this typically follows the object—for example, *John's understanding the rules thoroughly impressed Mary*. The head noun of the NP, the nominalization *understanding*, can be modified by an adjective, which must precede it—*John's thorough understanding of the rules impressed Mary*.
- (iv) A complement clause is, like every other clause type, negated by *not*; we get *John's not understanding the rules dismayed Mary*. In contrast, an NP—as in (21)—may only be negated by prefixing *non-* to the head noun, giving *John's non-understanding of the rules dismayed Mary*.
- (v) An ING complement clause may include auxiliaries *have (-en)* and *be (-ing)*; one can say *John's having understood the rules impressed Mary*. An auxiliary may not be used with a nominalization, such as we have in (21).

These two structurally different—although superficially similar—sentences have different meanings. Sentence (20) might imply that John at first experienced some difficulty in mastering the rules and the fact that he did now understand them impressed Mary, whereas (21) suggests that what impressed Mary was the nature of the understanding John achieved—he had worked out their intricacies and implications.

What makes (20) and (21) such an intriguing pair of sentences is that the verb *understand* adds *-ing* in being nominalized, the same suffix that marks

the verb of one variety of complement clause. Many verbs in English have different forms in the two circumstances. Compare:

(22)

COMPLEMENT CLAUSE	NP WITH NOMINALIZATION AS HEAD NOUN
John's appreciating the rules	John's/the appreciation of the rules
John's stating the rules	John's/the statement of the rules
John's knowing the rules	John's/the knowledge of the rules
John's using the rules	John's/the use of the rules

Whereas the verb of an *ING* complement clause always carries suffix *-ing*, there are a variety of derivational processes employed for nominalization in English. Those illustrated in (22) are *-(at)ion*, *-ment*, *-ledge*, and plain zero (with *use*). Each pair in (22) shows a semantic difference similar to that between (20) and (21).

One does find, in the literature, statements such as 'it is generally accepted that English complement clauses are simultaneously NPs and S[entence]s' (Stenson 1981: 63). First, they are clauses and not sentences. And secondly, they are certainly not NPs. Sentences (20) and (21) clearly contrast a complement clause and an NP, as alternative fillers of the same argument slot in a main clause.

(II) A complement clause functions as a core argument of a higher clause. In every language in which complement clauses occur they function as O argument; there are often other possibilities as well.

This can first be illustrated for English where a complement clause may fill any core slot. There is rather strict constituent order so that a complement clause can be recognized by its position within the main clause—S and A arguments immediately precede the predicate, and the O argument immediately follows it. This leads to the recognition of the complement clause in (23) as being in S function, that in (24) in A function, and that in (25) in O function.

(23) [That John was an academic]_{CoCl:S} didn't matter (to Mary)

(24) [John's having ignored her mother]_{CoCl:A} annoyed Mary_O

(25) Everybody_A knew [that John had not committed any crime]_{CoCl:O}

In English, most clauses involving a transitive verb can be passivized, with the original O argument being reassigned to S function and the original A argument being moved to the end of the clause, marked with *by* (and being optionally omissible). This applies to many complement clauses in O or A function. For example, the passive of (25) is

(25p) [That John had not committed any crime]_{CoCl:S} was known
(by everyone)

Sentences (26) and (27) have an Activity and a Potential complement clause, respectively, in O slot, with (26p) and (27p) being the corresponding passives.

- (26) [The police]_A had observed [John's taking the money]_{CoCl:O}
 (26p) [John's taking the money]_{CoCl:S} had been observed (by the police)
 (27) [The organizers]_A had already decided [for Mary to lead the parade]_{CoCl:O}
 (27p) [For Mary to lead the parade]_{CoCl:S} had already been decided (on) (by the organizers)

Passivization involving a complement clause in A function can be seen in (24p), relating to (24). The complement clause in underlying A function, from (24), becomes an oblique constituent—which may be omitted—in (24p).

- (24p) Mary_S was annoyed (by [John's having ignored her mother]_{CoCl})

A Fact or Activity complement clause in A function can also undergo passivization, as in:

- (28) [That John always ignored her mother]_{CoCl:A} annoyed Mary_O
 (28p) Mary_S was annoyed ([that John always ignored her mother]_{CoCl})
 (29) [For John to marry Mary]_{CoCl:A} would please Nancy_O
 (29p) Nancy_S would be pleased ([for John to marry Mary]_{CoCl})

There is a general rule in English that a preposition, such as *by*, is dropped when immediately followed by a complementizer, such as *that* or *for* or *to*. For this reason, there is no *by* before *that* in (28p) or before *for* in (29p), in the way that there is before the complement clause in (24p).

In English a complement clause may also function as copula complement (CC), as in (30), or as copula subject (CS), as in (31).

- (30) [The truth]_{CS} is [that John did it]_{CoCl:CC}
 (31) [That John did it]_{CoCl:CS} is true_{CC}, or It_{CS} is true_{CC} [that John did it]_{CoCl:CS}

There are few examples of transitive verbs which may take complement clauses as both A and O arguments. A small number of verbs in English do have these properties; they include *implicate*, *show*, *demonstrate*, *relate to*, *depend on*, and *result from*. For example:

- (32) [John's having carried the log home]_{CoCl:A} shows [that he is a strong fellow]_{CoCl:O}

Many European (and some other) languages are like English in allowing a complement clause to occur in all core functions for transitive (O and A), intransitive (S), and copula (CS and CC) clauses. In others, the possibilities are more limited:

- Modern Hebrew (Zuckermann 2006)—O, S, CS, and CC functions (that is, all but A)
- Akkadian (Deutscher 2006)—O, S, and CS
- Ainu (Onishi 1996)—O and S
- Basque (Curnow 1998)—O and CS
- Tariana (Aikhenvald 2003, 2006), White Hmong (Jarkey 2006), and Panare (Payne and Payne 1999)—O

It was mentioned in §3.2 and §13.1 that some languages have extended intransitive and/or extended transitive clause types including an additional core argument E ('extension' to core). English allows a complement clause in E function within a ditransitive clause; for example, [*The doctor*]_A *promised Mary*_O [*that he would cure her*]_{CoCl:E} and *John*_A *told/showed Mary*_O [*that the house had burned down*]_{CoCl:E}. A complement clause may also be in E function for Modern Hebrew.

It will be seen that the pre-eminent function for complement clauses is O. This is followed by S and/or CS, with A being the least common function. Every language with complement clauses allows them in O function and for the great majority of languages O is the commonest slot for a complement clause. Jarawara is an exception—more than 70 per cent of its complement clauses are in S function, a little less than 30 per cent in O, with just a few in CS function, and in A function. The latter all occur with causative derivations of intransitive verbs, things like 'The man's falling over made me laugh'.

In Jarawara we find complement clauses in S function for verbs like 'begin' and 'finish', for quantity verbs such as 'be two' and 'be much', and for stative verbs such as 'be good' and 'be strong'. One says '[Your talking]_S is good' to translate *You talk well* and '[His paddling]_S was strong' for *He paddled strongly*. In each instance, the main clause verb in Jarawara corresponds to an adverb in English, with the main clause in English being rendered as a complement clause in S function. Verbs of motion may also take an S complement clause, as in:

- (33) [kosi_O weje ni]_S ka-me
 urucuri.palm.fruit carry AUXILIARY:COMP go-return:MASCULINE
 He carried back urucuri palm fruit (lit. His carrying urucuri palm fruit
 went back)

The underlying form of the auxiliary is *na*; final *-a* is raised to *-i* as a mark that this is a complement clause. The last element of the main clause verb is suffix *-ma* ‘return’; its final *-a* is raised to *-e*, indicating masculine gender, to show that the subject of the complement clause is masculine.

However, Jarawara still does conform to the universal norm, having complement clauses in O function for verbs such as *-mita-* ‘hear, smell, taste’, *wati-waha-* ‘remember’, *-wato-* ‘know, understand’, and *-nofa-* ‘want, desire’.

Such is the number and diversity of human languages that one is always likely to encounter some minor modification to a putative generalization. Generally, a complement clause fills a core argument slot in the structure of a higher clause. Almost all of the several hundred complement clauses in my Jarawara corpus are of this type. However, I have recorded three instances where a complement clause appears to function as head of an NP, since it is followed by a modifier (Dixon 2004a: 455–6). When the Jarawara people built me a house, visitors from other villages admired the woven thatched roof, and said:

- (34) [[[jobe baje efe]_S boto ni]_{CoCl} nafi]_S
 thatch(m) palm(m) leaf:m weave AUX:COMP all
 amosa-ka
 be.good-DECM
 All of the palm leaf thatch weaving is good

Here the complex NP *jobe baje efe* ‘palm leaf thatch’ is S argument for *boto -na-* ‘weave thatching’. Raising the final *a* of auxiliary *na* to *i* indicates that *jobe baje efe boto ni* ‘weaving palm leaf thatch’ is a complement clause. This functions as the head of an NP, with modifier *nafi* ‘all’, the whole NP being S argument for intransitive verb *amosa-* ‘be good’.

Note that this does *not* imply that here a complement clause is an NP. It has the structure of a clause, not of an NP, but it functions as head of an NP which fills a core argument slot.

(III) A complement clause must refer to a proposition, something involving at least one participant who is involved in an activity or state. It cannot refer just to a place or a time. In English, *I_A saw [(the place) [where John lives]_{RC}]_{NP:O}* and *I_A know [(the time) [when John came home last night]_{RC}]_{NP:O}* each involves an NP as O argument, with *place* and *time* as heads, each being modified by a relative clause. The head (plus *the*) can be omitted, the NP then consisting just of a headless relative clause, specifying place or time. These do not involve complement clauses.

It is possible to get what appears to be a coordinated set of complement clauses, as in English *Mary_A knows [(that) John is stupid and (that) Tom is*

clever]_{CoCl:O} and *I_A want [John to sing and Mary to dance]*_{CoCl:O}. Many types of complex sentences in English can be explained through ellipsis from a combination of simple clauses; for example *John likes apples and Mary pears* is taken to relate to the underlying *John likes apples and Mary likes pears* with the second occurrence of *likes* omitted. In similar fashion, the coordinated complement clauses just quoted could be said to be reductions from *Mary_A knows [(that) John is stupid]*_{CoCl:O} and *Mary_A knows [(that) Tom is clever]*_{CoCl:O} and from *I_A want [John to sing]*_{CoCl:O} and *I_A want [Mary to dance]*_{CoCl:O} respectively. It is a matter of choice whether one decides to pursue this analysis, or instead to say that a number of compatible complement clauses of the same type may be coordinated as the complex filler for a core argument in the higher clause.

In some languages, a clause which includes a complement-taking verb is simply apposed to a clause including a second verb; this must be distinguished from a complement clause construction. An illustration is provided by Watkins (1984: 235) from Kiowa (Kiowa-Tanoan family, USA):

- (35) à-ʒn mágyá èm-k^hóydé-t'ǝ·
 1sg-think:THAT might 2sg-turn.back-FUT
 I thought that you might turn back (lit. I thought that; you might turn back)

No evidence is provided that (*mágyá*) èm-k^hóydé-t'ǝ· functions as a constituent of the main clause à-ʒn—that it is a complement clause. It is likely that we have here a complementation strategy in which clauses ‘I think that’ and ‘you might turn back’ are apposed, with ‘that’ of the first clause referring to the second clause. This kind of strategy is discussed under (a) in §18.6.4.

18.3 Grammatical parameters

There are a number of parameters of variation for complement clauses. We can here list the major ones, commenting further on some of them in §18.4, relating to types of complement clause.

(a) **Marking of core arguments.** In most complement clauses, A, O, S, CS, and CC arguments are marked in the same way as in a main clause. An exception is when a particular complement clause type is shown by a special marker on its subject (A, S, or CS). In a number of languages, a complement clause which describes an activity takes possessive marking on its subject; this applies for ING clauses in English. Our discussion of sentences (20) and (21) showed that although an ING clause includes ‘s on its subject, it is of a quite different structure from an NP which involves possession.

In Jarawara, the 1sg or 2sg subject of a complement clause is in possessive form. This does not indicate that the complement clause is a type of NP since other subjects (1nsg, 2nsg, and 3rd person) are shown as in a main clause, with no possessive marker (Dixon 2004a: 446–61; 2006b).

(b) **Use of bound pronouns.** If a language has bound pronouns, there is a strong expectation that these will be included in a complement clause as they are in a main clause (generally, on the verb). Tariana provides an example of this (Aikhenvald 2006).

An interesting question concerns how a complement clause in a core argument function is referred to by the bound pronoun in the main clause which relates to that core argument. Typically, a complement clause (whatever the persons of its participants) is cross-referenced as 3sg, the default term in the pronoun system. Bilua (Papuan type, Solomon Islands; Obata 2003: 213–22) has a gender contrast in 3sg, and a complement clause in O function triggers 3sg feminine enclitic on the main clause verb, since this is the default form.

Rukai (Austronesian, Taiwan; Zeitoun 2007: 414–36) has bound pronouns in subject function. Consider:

- (36) o-iriho'o-irao o-'ongolo-nga-nomi vavan
 DYNAMIC-know-1sg.SUBJ DYNAMIC-drink-ALREADY-2plSUBJ wine
 I know that you (plural) drank wine

At first sight this appears just to involve two clauses in apposition, similar to (35) in Kiowa. That is, it would be a complementation strategy ('I know. You (plural) drank wine.') rather than a complement clause construction. However, there is an alternative form of (36):

- (37) o-iriho'o-irao-nomi o-'ongolo-nga-nomi
 DYNAMIC-know-1sgSUBJ-2plSUBJ DYNAMIC-drink-ALREADY-2plSUBJ
 vavan
 wine
 I know that you (plural) drank wine

Here, the subject pronoun from the second clause is copied onto the end of the verb of the first clause (it must be retained in the second clause). This syntactic integration shows that (37)—and thus also (36)—should be regarded as a rather special type of complement clause construction, rather than just an appositional strategy.

(c) **Inclusion of peripheral constituents.** All varieties of complement clauses in English may include peripheral constituents referring to time, place, etc., in the same way as a main clause. In contrast, the single complement clause

type in Jarawara only includes the predicate and core arguments, nothing else. Other languages fall between these extremes. (Further work is needed on this topic.)

(d) Tense, aspect, and modality specification. Many languages have several varieties of complement clause, some of which may take the same choice of tense and/or aspect and/or modality—and evidentiality, in languages which have this category—as a main clause, while others lack these. In English, for instance, a *THAT* clause has the full set of TAM choices, while *ING* and *TO* complement clauses only allow auxiliaries *have* (*-en*) and *be* (*-ing*), not permitting tense inflection or modal verbs. How then does one refer to past time within such a complement clause? What happens is that *have* (*-en*) corresponds to both *have* (*-en*) and to past time in a *THAT* clause or in a main clause. Suppose that I have been daydreaming and then say:

- (38) (a) I_A imagined [that Mary writes books]_{CoCl:O}
 (b) I_A imagined [that Mary wrote a book]_{CoCl:O}
 (c) I_A imagined [that Mary has written a book]_{CoCl:O}
 (d) I_A imagined [that Mary had written a book]_{CoCl:O} (before she met John)

If an *ING* complement clause were used in place of a *THAT* clause, we would get:

- (39) (a) I_A imagined [Mary('s) writing books]_{CoCl:O}
 (b/c/d) I_A imagined [Mary('s) having written a book]_{CoCl:O}

All of past tense, in (38b), *have* (*-en*) plus present, in (38c), and *have* (*-en*) plus past, in (38d) are rendered in the same way within an *ING* complement clause. (An example involving a *TO* complement clause is in Dixon 2005: 50–1.)

A *THAT* clause may include any of the modals—*can*, *must*, *will*, and so on. These are not used in *ING* and *TO* complement clauses. However, there are ‘semi-modal’ verbs which have very similar meanings to some of the modals—*be able to* corresponds to *can*, *have* (*got*) *to* for *must*, and *be going to* for *will*. The semi-modals are regular intransitive verbs which may be employed in every kind of complement clause. A *THAT* clause can involve either *can* or *be able to*, as in:

- (40) (a) I_A imagined [that Mary can write books]_{CoCl:O}
 (b) I_A imagined [that Mary is able to write books]_{CoCl:O}

An *ING* or *TO* clause may only use *be able to*:

- (41) (a/b) I_A imagined [Mary('s) being able to write books]_{CoCl:O}

A modal and the corresponding semi-modal do not have exactly the same meaning. Generally, a semi-modal may carry an ‘unconditional’ sense—*Mary is able to write books (easily, scarcely having to try at all)*. In contrast, a modal may indicate ability, etc, subject to certain specifiable conditions—*Mary can write books (if she puts her mind to it, but most of the time she is too lazy to bother)*. This contrast is maintained in a THAT but lost from an ING or TO complement clause. (See Dixon 2005: 172–88 for discussion of modals, semi-modals, and their meanings.)

Valentine (2001b: 671) states for Nishnaabemwim (or Ojibwe; Algonquian, Southern Ontario) that ‘verbs in complement clauses allow a rather free range of tense and aspect preverbs’. For Tariana, independent tense and evidentiality specification is found in purpose and interrogative complement clauses, but not in Fact clauses marked by *-ka* (Aikhenvald 2003, 2006). Bilua has ‘finite’ complement clauses which include a tense suffix but this must have future reference; it can be either the near future or the future tense marker (but not present, recent past, remote past, or historical tenses). There are also ‘non-finite’ complement clauses which include no tense marker and can have any temporal reference.

The verb in a main clause for Jarawara may take a multiplicity of suffixes; there are six ‘echelons’ for what I call ‘miscellaneous suffixes’ followed by a choice from a system of eleven tense-modal markers and then one from a system of fourteen mood markers. The verb in a complement clause may only include suffixes from the first four echelons (‘upstream’, ‘do first’, ‘soon’, ‘do quickly’, etc.) but not from echelons 5 or 6 (‘do/happen all night’, ‘again’, ‘also’, and clausal negator *-ra*, etc.), nor anything from the tense-modal or mood systems (Dixon 2004a, 2006b).

Some languages mark the time of a complement clause as it relates to that of the main clause. Quechua (Weber 1989: 289) has two Fact complementizers: *-sha* is used if the time of the event of the complement clause is earlier than or at the same time as the event referred to in the main clause, as in (42a), and *-na* is used if it is later, as in (42b).

- (42) (a) mana musya-shka-:-chu [achka
 NOT know-perfect-1-NEG many
 chaya-mu-sha-a-ta]_{CoCl:O}
 arrive-afar-comp-3-object
 I did not know that so many arrived
- (b) mana musya-shka-:-chu [achka
 NOT know-perfect-1-NEG many
 chaya-mu-na-a-ta]_{CoCl:O}
 arrive-afar-comp-3-object
 I did not know that so many were going to arrive

(e) **Negation.** In most languages, negation can apply either to the main clause or to the complement clause in a complement clause construction. This may be illustrated from English:

- (43) (a) Mary_A didn't notice [that John was working]_{CoCl:O}
 (b) Mary_A noticed [that John wasn't working]_{CoCl:O}

In (43a), John may or may not have been working; Mary just didn't notice what he was doing. In (43b), Mary did notice that John was not working—perhaps it was when he should have been working, but he had left his desk and was watching football on the TV in the common room.

It is also the case the negation may apply to all kinds of complement clause in Modern Hebrew (Zuckermann 2006), in White Hmong (Jarkey 2006), and in Tariana (Aikhenvald 2003, 2006). Complement clauses in Akkadian may be negated but require a different negator (the same as in other kinds of subordinate clause) from that used in main clauses (Deutscher 2006). Jarawara may only negate a clause through the sixth echelon 'miscellaneous suffix' *-ra*. As mentioned under (d), a complement clause cannot include this suffix and thus negation can only apply to a main clause. It would not be possible in Jarawara to distinguish between (43a) and (43b), except through circumlocution.

One needs to investigate whether complement clauses behave, with respect to negation, like main clauses or like some or all other kinds of subordinate clauses.

(f) **Derivational processes.** In the great majority of languages, the verb of a complement clause is open to the same inventory of derivational processes as a main clause verb. Even in Jarawara, where the verb of a complement clause accepts only a limited selection of suffixes, it may take causative prefix *ma-* or applicative *ka-*. This is a question which needs to be investigated for each language.

(g) **Position in main clause.** A complement clause will generally occur at the same place in the surface structure of the main clause as would an NP in the same function. However, there is a prevalent tendency to shift a 'heavy constituent'—such as a complement clause—to the end of the main clause. This applies in English to complement clauses in A or S function. In place of (28) and (23) one is likely to encounter (28e) and (23e), in which the complement clause is extraposed to clause-final position. Dummy pronoun *it* now fills the A or S argument slot before the verb.

- (28) [That John always ignored her mother]_{CoCl:A} annoyed Mary_O
 (28e) It_A annoyed Mary_O [that John always ignored her mother]_{CoCl:A}

(23) [That John was an academic]_{CoCl:S} didn't matter (to Mary)

(23e) It_S didn't matter (to Mary) [that John was an academic]_{CoCl:S}

And in (31) of §18.2, a complement clause in CS function is extraposed.

Stenson (1981: 67) states that, in Irish, 'extraposition of complements, especially subject, further distinguishes them from non-sentential NPs'. Fijian is a predicate-initial language. A, O, S, and peripheral arguments, if realized by NPs, can occur in any order after the predicate. But if one of the core slots is filled by a complement clause, this should be placed in clause-final position (Dixon 1988a: 273).

Or a complement clause may have positional restrictions, when compared with a corresponding NP. In Q'eqchi' Maya (Kockelman 2003: 29) 'one cannot prepose a full-clause complement into focus position, as one can an NP, even though both are cross-referenced on the main verb in exactly the same way'. Similarly, in Fijian any NP can be topicalized and fronted to appear before the predicate, but this option is not available for complement clauses.

(h) Same and different subjects, and ellipsis. In many complement clause types there is no restriction on whether main clause and complement clause have the same or different subjects. But constraints do sometimes apply. For instance, in White Hmong, for a Fact clause its subject can be the same as or different from the subject of the main clause, but for a 'potential' complement clause, the subjects must be different, as in Jarkey (2006: 124):

(44) kuv_A nyiam [kom nws_S mus]_{CoCl:O}
 1sg like to 3sg go
 I like him to go

Indonesian (Hill 1997a; Wayan Pastika, personal communication) is interesting in that the verb 'decide' is generally used with a Fact complement clause, as in (45), when the two clauses have different subjects, and with a 'potential' clause, as in (46), when subjects are the same.

(45) saya_A memutuskan [bahwa dia harus pergi]_{CoCl:O}
 1sg decided THAT 3sg MUST go
 I decided that he must go

(46) saya_A memutuskan [untuk pergi]_{CoCl:O}
 1sg decided TO go
 I decided to go

There is discussion in §18.5.2 of 'same subject' and 'different subject' possibilities for 'want'.

Rukai (Zeitoun 2007: 421–2) has an unusual variety of construction in which the complement clause includes causative prefix *pa-*. For example:

- (47) 'ooho-lra-ine [pa-noa
 DYNAMIC.FIN:order-1sgNOM-3sgOBL CAUS-DYNAMIC.NON.FIN:go
 taipake]_{CoCl}
 Taipei
 I ordered him to go to Taipei

If there were no causative derivation, 'go' would have a different subject from 'order'. When causative is applied to 'go' the two verbs have the same subject, literally 'I order him, (I) make-go (him)'.

In a 'same subject' construction, the occurrence of the subject in the complement clause may be optionally or obligatorily omissible. In English, nothing can be omitted from a THAT clause, but a repeated subject may be omitted from an ING clause—I_A *remember* [(my) *winning the prize*]_{CoCl:O}. And the subject of a Modal TO complement must be omitted if it is the same as the main clause subject—compare I_A *longed* [for Mary/her to win]_{CoCl:O} and I_A *longed* [to win]_{CoCl:O}, rather than *I_A *longed* [for me to win]_{CoCl:O} (this could only be said as a sort of 'echo').

(j) **'Raising' a pronoun from complement clause to main clause.** There are a number of situations which have been described as 'raising'. One of the clearest is found in Jarawara. Consider a transitive complement clause which is in S function within the main clause:

- (48) [sina_O otaa hisi ni]_{CoCl:S} ø hawa
 snuff(f) 1sg.excA sniff AUX:COMP 3sgS be.finished
 to-ha-ke
 AWAY-AUX-DECf
 We have finished sniffing snuff (lit. our sniffing snuff is finished)

In Jarawara, 3sg has zero realization. The complement clause in (48) is cross-referenced within the main clause as 3sg and there is thus nothing in the S pronominal slot for the main clause (ø indicates an empty slot).

The raising rule is as follows. If a pronoun is in A function within a transitive complement clause, which is itself in S function within the main clause, this pronoun may be raised into the S pronominal slot in the main clause (effectively, filling an empty slot). Thus:

- (48r) [sina_O hisi ni]_{CoCl:S} otaa hawa to-ha-ke
 snuff(f) sniff AUX:COMP 1sg.exc be.finished AWAY-AUX-DECf
 We have finished sniffing snuff (lit. our sniffing snuff is finished)

These sentences are simply stylistic variants, there being no difference in meaning.

Interestingly, raising is only possible from A to S, as in (48r), not from S to S, or from S to A, or from A to A. (Jarawara has an entirely nominative–accusative profile for the marking of core arguments, with S and A there treated in the same way. These facts of raising show the importance of distinguishing between S (intransitive subject) and A (transitive subject), even in a language of this type.)

It has been suggested that ‘raising’ is involved for some complement clause constructions in English; however, this is a matter for debate. Consider a situation where a certain NP has one function in the main clause and another in the complement clause. Compare:

(49) Mary_A persuaded John_O [that he_A should hit Fred_O]_{COCL:SECOND.O}

(50) Mary persuaded John to hit Fred

In (49), *John* is O argument of the main verb *persuade* and the coreferential pronoun *he* is A argument for the complement clause verb *hit*. In (50), *John* has both argument functions simultaneously. It is unprofitable to put forward the limited view that an NP can have only one function and then to try to decide which of the two functions to assign to *John* in (50). A reflexive pronoun ending in *-self* can only be used when coreferential with another NP in the same clause. Note that one can say *Mary forced herself to hit Fred* (the *herself* is in the same clause as *Mary*) and also *Mary forced John to hit himself* (the *himself* is in the same clause as *John*); this shows that in (50) *John* is functioning both as O for the first clause and as A argument for the second one. The underlying structure of (50) is:

(50u) Mary_A persuaded John_O [John_A to hit Fred_O]_{COCL:SECOND.O}

One of the two successive occurrences of *John* is then omitted from surface structure.

When *John* is replaced by a pronoun, the underlying structure becomes:

(51u) Mary_A persuaded him_O [he_A hit Fred_O]_{COCL:SECOND.O}

And this comes out in surface structure as:

(51) Mary persuaded him to hit Fred.

The fact that the object form *him* is used (rather than the subject form *he*) has been taken as an indication that pronoun *he* has been raised from complement clause into main clause. But *him/he* in (51) has syntactic function in both clauses, just like *John* in (50). It has to have just one form in surface structure,

and the fact that *him* is used in preference to *he* does not affect this bifunctionality. In all linguistic analysis, one must take care always to distinguish function from form.

(A possible explanation for the choice of *him* rather than *he* in a structure like (51) is that the *to*-plus-verb sequence may be omissible when able to be understood from context, and the abbreviated sentence is then grammatical. A dialogue might comprise: *Is John going to go to the market this week? Yes, I finally persuaded him* (sc. to go).)

Pronoun *otaa* 'us' has syntactic function in only one clause for (48) in Jarawara. After it has been raised into the S pronominal slot in the main clause, in (48r), it still functions as A argument in the complement clause. This is a true example of 'raising'. In contrast, *John* in (50) and *he/him* in (51) have syntactic function in both clauses. Nothing has been 'raised'; it is just that only one of the functions is realized in surface structure.

18.4 Types and meanings

In terms of semantics, there are three recurrent types of complement clause, each of which can have a number of subtypes. Their typical properties will now be outlined.

I. Fact type

- Generally refers to the fact that something took place.
- Typically, shows similar structure to a main clause, with full possibilities for negation, tense–aspect marking, etc., and for bound pronominal reference in a language which has this.
- Its subject may or may not be identical to the subject of the main clause; but if it is the same, it is unlikely to be omitted.
- The time reference of a Fact complement clause is generally independent of that in the main clause, and the two clauses may show different tense–aspect values.
- Typically, marked as a complement clause by a complementizer element (similar to English *that*). This may be omissible under certain conditions, the complement clause then being recognized as such perhaps by its position within the main clause. There may be a preference (or even a requirement) that a Fact complement clause—as a 'heavy constituent'—be extraposed to the end of the main clause—see (g) in §18.3. (Other varieties of complement clause may be extraposable, but not so commonly as the Fact type.)

Most frequently, a complementizer form has additional functions in the grammar; for example, *that* in English is also a marker of a relative clause, and a nominal demonstrative. Complementizers have often developed from a demonstrative, or from a verb such as ‘be like’ or ‘say’. In Mokilese (Austronesian; Harrison 1976: 266–8), the grammatical element *pwa* functions as a clause linker ‘because, so that’ and as the marker of a Fact complement clause. It has the same form as the verb *pwa* ‘say’ and is probably historically derived from it. The complementizer *pwa* is sometimes omitted, this being particularly common after the verb *pwa* ‘say’.

In Akkadian (Deutscher 2000, 2006) complementizer *kīma* ‘that’ also functions as ‘a preposition and adverbial conjunction with a range of meanings: “as”, “like”, “instead of”, “when”, “because”’. In Tariana (Aikhenvald 2003, 2006), the complementizer *-ka* also marks a type of sequential clause. When a complementizer does not have any other function in the grammar, it frequently shows a transparent etymology. For example, in Dolakha Newar (Genetti 2006) complementizer *khā* is plainly derived from noun *khā* ‘talk, matter, news’. The task for a grammarian is to provide clear criteria for distinguishing between the various functions of a given form.

There may be a subtype of Fact complement clause relating to the reliability of the information provided. In Jacalteco (Craig 1977: 267–8), complementizer *chubil* indicates a high degree of certainty, as in (52b) and *tato* a degree of doubt, as in (52a).

- (52) (a) xal naj_A [tato chuluj [naj
 said CLASS.MASC THAT will.come CLASS.MASC
 presidente]_S]_{CoCl:O}
 president
 He said that the president is going to come
- (b) xal [naj alca]_A [chubil chuluj [naj
 said CLASS.MASC mayor THAT will.come CLASS.MASC
 presidente]_S]_{CoCl:O}
 president
 The mayor said that the president is going to come

The information in (52a) is attributed to an anonymous reporter (shown just by masculine classifier *naj*) and is, as a consequence, regarded as less than fully reliable, hence use of complementizer *tato*. But in (52b) a person in high authority provides the report, indicating its reliability, and *chubil* is employed.

In Modern Greek (Joseph and Philippaki-Warbuton 1987: 182, 22–3), a verb such as ‘see’ may take a complement clause marked by complementizer

óti, indicating a sure fact, or by ‘subjunctive’ complementizer *na*, indicating a possible fact. This difference is nicely brought out with a negated main clause:

- (53) (a) δen ton $i\delta a$ na $koli(m)b\acute{a}$
 NEGATIVE he:ACCUSATIVE saw:1sgA COMP swim:3sg:PRESENT
 I did not see him swim
- (b) δen ton $i\delta a$ $\acute{o}ti$
 NEGATIVE he:ACCUSATIVE saw:1sgA COMP
 $koli(m)b\acute{u}se$
 swim:3sg:IMPERFECT:PAST
 I did not see that he was swimming (lit. I did not see him that he was swimming)

In (53a) ‘the implication is that he may have swum but he may not have’, while in (53b) ‘the implication is that he did swim but the act was not witnessed by the speaker’.

A complement clause is not marked for mood, as a main clause may be, but many languages have interrogative complement clauses as a subtype of the Fact variety. Illustrating from English, a complement clause can relate to a polar interrogative—for example, *Mary asked whether/if dinner was ready*—or to a content question—*John asked who was cooking dinner*. (There may also be an interrogative version of a Potential clause; see III.)

In Tzotzil (Mayan, Mexico; Robinson 1999: 98; Haviland 1981: 342–3), a declarative complement clause may be marked by *ti*, as in (54a), and an interrogative one by complementizer *mi*, as in (54b).

- (54) (a) mi $av-a'i-\emptyset$ $[ti$ $i-\emptyset-jatav$ li
 QUESTION 2:A-hear-3:O THAT COMPL-3:A-flee ART
 $'antze]_{CoCl:O}$
 woman
 Did you hear that the woman fled?
- (b) mi $av-a'i-\emptyset$ $[mi$ $i-\emptyset-jatav$ li
 QUESTION 2:A-hear-3:O WHETHER COMPL-3:A-flee ART
 $'antze]_{CoCl:O}$
 woman
 Did you hear whether the woman fled?

Completive *ti* also functions as an article within an NP (referring to something distant or remote), while *mi* is also the marker of a polar interrogative, as in the main clauses of (54a/b).

II. Activity type

- Generally, refers to some ongoing activity, relating to its extension in time.
- Typically, has some structural similarities to a noun phrase, although it must retain crucial characteristics of a clause in order to be analysed as a complement clause. The subject may be marked like a possessor in an NP. The verb may have a special form, but this must be distinguishable from a verbal nominalization (which is a noun, and functions as head of an NP). The similarities and differences between an Activity complement clause and an NP whose head is a deverbal noun were brought out in the discussion of the English sentences (20) and (21) in §18.2.
- Its subject may or may not be the same as the subject of the main clause; if it is, it may be omissible.
- Typically, has available less specification of tense and/or aspect and/or modality and/or negation than a main clause (partly, through use of a special verb form). An Activity complement clause may well have different time reference from that of the main clause. In a language where it cannot include a grammatical tense marker, the time reference has to be shown by lexical means; for example, in English *I_A do remember [seeing him last Wednesday]*_{CoCl:O}.
- May not be able to include the same bound pronominal elements as does a main clause (for a language which has bound pronouns).

Finnish has two subtypes of Activity complement clauses—a ‘present participle’, used when the activity described by the complement clause is at the same time or later than that of the main clause, and a ‘past participle’ used when it precedes it (Sands 2000).

III. Potential type

- Generally, refers to the potentiality of the subject of the complement clause (which is almost always the same as some argument in the main clause) becoming involved in an activity.
- Typically—having satisfied the requirement to be analysed as a complement clause—has less structural similarity to a main clause than a Fact complement clause, and has less structural similarity to an NP than an Activity clause.
- In some languages, must have the same subject as its main clause, and statement of the subject must (or may) be omitted from the complement clause.
- Generally, lacks the tense-aspect and similar choices available to a main clause. And, where a main clause includes bound pronouns, may lack these.

- As a rule, has implicit reference to the same time as, or a later time than, that of the main clause.
- The verb generally has a special form (sometimes called ‘infinitive’; see (g) in §2.5 for difficulties connected with using this term). Alternatively, the verb may be marked in a similar way to dative (or some other) case on an NP.

The number of distinct types of complement clause constructions varies from language to language. Illustrating from languages already mentioned in this chapter:

- **One.** Jarawara (Dixon 2004a, 2006b) has a single complement clause construction, whose formal properties most resemble those outlined for the Activity type, while its meaning encompasses all of Fact—as in (34)—Activity—as in (33) and (48)—and Potential (‘I want him to go’). The verb takes a special form, which has some similarities to—and some differences from—a nominalization, and 1sg and 2sg subjects (but not the non-singulars or 3rd person) are expressed as possessives.
- **Two.** Both Akkadian (Deutscher 2006) and Irish (Stenson 1981) have two construction types, one for Fact and the other for both Activity and Potential.
- **Three.** Tariana has declarative and interrogative Fact complement clause constructions, plus a Potential type. (Activity is shown by the complementation strategy of nominalization—see §18.6.3.)
- **Five.** White Hmong (Jarkey 2006) has two Fact complement clauses, one occurring after verbs of speaking and thinking and the second after other verbs. There is an Activity and two Potential constructions—one for direct and the other for indirect expression of intention/will. The first roughly has the import of a Modal TO construction in English (‘want people to go’) and the other of a THAT construction including a modal (‘told people that they should go’).

Fijian also has five complement clause constructions—one Potential, one Activity, and three Fact—definite, uncertain, and interrogative (Dixon 1988a: 130–5, 267–85).

Larger inventories are reported for some languages. Zuckermann (2006) recognizes six types for Modern Hebrew, and Burridge (2006) has seven for Pennsylvania German (spoken by Mennonite Anabaptists in Canada), including three Potential types.

Yimas (Lower Sepik family, Papuan area, New Guinea; Foley 1991: 384–402) constitutes an exception to the typical schema set out above. Complement clauses involve a verb marked with *-ru*, which also forms nominalizations. There is a choice of four complementizers, which follow the complement

clause: *-mpwi* (lit. 'talk') for a complement clause referring to speech or language, as in 'I tried to tell them to buy betel nut'; *-wampuj* ('heart') for desire, as in 'he feels like eating sago'; *-nti* ('act') for action, as in 'I'm tired of building houses'; and *-wal* ('custom') for customary activity, as in 'smoking tobacco is bad'.

18.5 Semantic types of verbs and varieties of complement clause

Each type of complement clause, in a given language, has a certain meaning. Each verb has its meaning. Which type of complement clause can be used with a given complement-taking verb depends on the interrelation of these two semantic parameters. To examine each verb in isolation would be a tiresome business. It is also unnecessary.

As discussed in §1.11 (and briefly in §3.3), words making up the lexicon of any language fall naturally into a number of 'semantic types', which have a common meaning component and share certain grammatical properties. There was illustration in §1.9 of how verbs from the THINKING and LIKING types co-occur with the four major types of complement clause in English. Each semantic type in a language has—by virtue of its shared semantic content—a predilection for occurring with certain complement clause varieties.

Every language has a large open class of 'Primary verbs' which can make up a complete sentence by choosing appropriate NPs and/or (bound or free) pronouns to fill their argument slots. There is then a division into Primary-A, which cannot take a complement clause as one argument, and Primary-B, which can.

- **Primary-A.** Argument slots may *not* be filled by complement clauses, only by NPs/pronouns. This covers semantic types such as MOTION ('walk', 'follow'), REST ('stand', 'put'), AFFECT ('kick', 'knit'), GIVING ('give', 'lend'), CORPOREAL ('drink', 'shiver', 'cure'), among others. (For a comprehensive list of verb types in English, see Dixon 2005: 485–91.)
- **Primary-B.** All argument slots may be filled by NPs/pronouns, but one (very occasionally, two) argument slot(s) may alternatively be filled by a complement clause. Each semantic type within Primary-B has its own semantic profile and, associated with this, a principled strategy for which varieties of complement clauses it may occur with. These semantic types are surveyed in §18.5.1.

English also has a set of 'Secondary verbs'. Cross-linguistically, we can recognize a number of 'Secondary concepts' which may be realized either through grammar or lexicon. At the underlying level they serve to modify a Primary

verb. Ideas such as ‘can’, ‘try’, ‘want’, and ‘make’ may be coded through a verbal affix in a language with rich morphology, and through a lexical verb (or a grammatical form such as a modal) in a language like English which has relatively sparse morphology. We thus have:

- **Secondary verbs.** In underlying structure, one argument slot must always be filled by a complement clause. There can be ellipsis, leaving just an NP in the argument slot in surface structure. For example, *I want to eat an apple* may be shortened to *I want an apple*. But *I want an apple* will only relate to *I want to eat an apple* in certain circumstances. In an art class *I want an apple* would be said when an apple was requested as an object to be drawn. (And for William Tell the import would be different again.) The semantic types within Secondary verbs, and the complement clauses they tend to occur with, are the topic of §18.5.2.

The discussion which follows of semantic types associated with the two kinds of CTVs is exemplified from English, which has a fairly typical set of complement clauses. (There are further complement clause types in English, which are not taken account of here. For example, the Judgement TO construction, as in *I_A know [Mary to be clever]_{CoCl:O}*. Full details are in Dixon 1991, 2005.)

A preliminary comment is in order here. Most verbs in English consist just of a verb root. Others are a kind of inherently phrasal verb, involving verb and preposition in their lexical form; for example, *refer to*, *decide on*, *think about*. The argument which follows one of these phrasal verbs behaves like a direct object (O function). For example, it can become the S argument within a passive derivation; compare *They_A had thought about [John’s having nominated Mary]_{CoCl:O}* and *[John’s having nominated Mary]_{CoCl:S} had been thought about (by them)*. It is thus appropriate to treat *refer to*, *decide on*, *think about*, and others of this nature, as transitive verbs. (Further arguments in favour of this analysis are in Dixon 1991: 13–14, 271–4; 2005: 14–15, 290–3.)

18.5.1 Primary-B semantic types

We can now briefly survey the main semantic types of Primary-B verbs, and their subtypes.

Attention

(a) verbs such as ‘see’, ‘hear’, ‘notice’, ‘smell’, ‘show’

- Prototypically take an Activity complement clause, describing the perception of a continuous activity; for example *I_A noticed [Mary(’s) weeding the garden]_{CoCl:O}*.

- May also take a Fact complement clause, for the perception that some activity is completed, or of some state; for example *I_A noticed [that Mary had weeded the garden]_{CoCl:O}* and *I_A saw [that John was incompetent]_{CoCl:O}*.

(b) ‘recognize’, ‘discover’, ‘find’

- Are expected to take a Fact complement clause; for example *I_A discovered [that Mary had resigned]_{CoCl:O}*.

Thinking

(a) ‘think (of/about/over)’, ‘consider’, ‘imagine’, ‘dream (of/about)’

- Depending on the senses of a verb ‘think’ in a particular language, it may relate to a Fact complement clause, such as *John_A thinks [that Mary is clever]_{CoCl:O}*, or to an Activity one, as in *John_A is thinking about [(Mary’s) weeding the garden]_{CoCl:O}*. (As noted in §18.2, there is a rule of English grammar that a preposition is omitted before complementizer *that*, *for*, or *to*.)

(b) ‘assume’, ‘suppose’

- Generally restricted to a Fact complement clause.

(c) ‘remember’, ‘forget’

- Similar to set (a). One can remember (or forget) just the fact that something happened, or else the details of the activity involved; for example *I_A remembered [that I had visited Paris]_{CoCl:O}* (but couldn’t recall anything I did there) and *I_A remembered [visiting Paris]_{CoCl:O}* (and had a clear recollection of every part of the holiday). English is perhaps unusual in also permitting a Potential complement clause, as in *I_A remembered [to lock the door]_{CoCl:O}*. The difference in meaning between a *to* clause after *remember* and a *that* clause which includes modal *should* is brought out in:

(55) I remembered that I should lock the door (but then decided not to, as a way of asserting my distaste for the rules)

(56) I remembered to lock the door (but then Mary snatched the key and pushed it down a grating, so I couldn’t)

And see (3–6) in §1.9.

(d) ‘know’, ‘understand’; (e) ‘believe’, ‘suspect’

- Generally take a Fact or Potential complement clause—*He_A knew [that the door should be locked]_{CoCl:O}* and *He_A knew [to lock the door]_{CoCl:O}*.

The verb ‘know’ may also have a second sense ‘know about’, and can then take an Activity clause.

Deciding—‘decide (to)’, ‘resolve’, ‘plan’, ‘choose’

- May often take either a Fact or a Potential complement clause; for example I_A *decided* [*that Mary/I should lead the parade*]_{CoCl:O}, I_A *decided* [*for Mary to lead the parade*]_{CoCl:O}, and I_A *decided* [*to lead the parade*]_{CoCl:O} (note that here the repeated subject is omitted from the second clause). We saw under (h) in §18.3 that in Indonesian a Fact complement is likely to be used if the two clauses have different subjects, and a Potential clause if the subjects are the same—see (45) and (46).

Liking

(a) ‘like’, ‘love’, ‘prefer’, ‘regret’, and ‘fear’

- Most frequently relate to an Activity complement clause, as I_A *like/fear* [*John’s getting drunk*]_{CoCl:O}. They may also be used with a Fact clause, as in I_A *like (it)* [*that John gets drunk*]_{CoCl:O} or I_A *fear* [*that John may get drunk*]_{CoCl:O}. Note the optional inclusion of *it* (or an NP such as *the fact*) before a Fact clause in English with *like*, *love* and *prefer* (but not with *fear*). English also allows a Potential complement clause with *like*, *love*, *prefer*, and *fear* (but not with *dislike* and *regret*); for example I_A ’d *like* [*to go*]_{CoCl:O} and I_A *fear* [*to go*]_{CoCl:O}.

(b) ‘enjoy’

- Refers to a pleasant perception which is extended in time and expects an Activity complement clause.

English has a further semantic type, which I dub ANNOYING verbs, that take a complement clause in A function; for instance, [*That John stays out late at night*]_{CoCl:A} *annoys Mary*_O. Some verbs in this type have similar meanings to LIKING verbs, but with core arguments reversed; compare $Mary_A$ *likes* [*John’s singing*]_{CoCl:O} and [*John’s singing*]_{CoCl:A} *pleases Mary*_O. Verbs with meanings and syntactic profiles similar to *annoy* and *please* in English are rather rare across the world’s languages.

Speaking

Most languages have both direct speech—*John said* ‘I’ll go’—and also indirect speech— $John_A$ *said* [*that he would go*]_{CoCl:O}. However, some languages lack indirect speech, using only the direct variety. Indirect speech involves a complement clause, generally of the Fact kind. We can hypothesize that

a language which lacks a complement clause construction, relying instead on complementation strategies, will not have indirect speech. Direct speech should be regarded—save in exceptional cases—not as a complementation strategy (nor as any other kind of complementation) but instead as a grammatical mechanism quite distinct from any form of complementation.

When a language does feature indirect speech, there is a fair range of verbs of speaking and they tend to select different varieties of complement clauses.

- (a) ‘say’, ‘inform’, ‘tell’ (one sense)
 - Are generally confined to a Fact clause.
- (b) ‘report’
 - May take a Fact or an Activity clause.
- (c) ‘describe’, ‘refer to’
 - Typically take an Activity clause.
- (d) ‘promise’, ‘threaten’
 - Generally take a Potential complement clause, which may be in E (indirect object) slot.
- (e) ‘order’, ‘command’, ‘persuade’, ‘tell’ (one sense)
 - Also generally take a Potential clause.

In English, sets (d) and (e) may alternatively take a Fact complement clause, so long as this includes a modal element which describes the potentiality; for example, *I_A persuaded John_O [to go]_{CoCl:E}* and *I_A persuaded John_O [that he should go]_{CoCl:E}*.

There may be a verb *tell* (about) with a wide range of meaning, corresponding to several of the Speaking semantic subtypes. In English we have (a) *I_A told Mary_O [that it was late]_{CoCl:E}*, and (c) *I_A told Mary_O about [Brazil’s having scored four goals]_{CoCl:E}*, and (e) *I_A told Mary_O [to go]_{CoCl:E}*.

Indirect—and direct—speech constructions may also involve some verbs from the THINKING type. For example, *Mary thought ‘John is a fool’* and *Mary_A thought [that John was a fool]_{CoCl:O}*.

We sometimes find that two Primary-B verbs, which differ only in that one has a positive and the other a negative meaning, take different types of complement clause. In White Hmong (Jarkey 2006: 132) *nyiam* ‘like’ may take Fact, Activity, or Potential complement clauses, while *ntxub* ‘hate’ is restricted to the Fact variety (showing the wider pragmatic possibilities of ‘like’ in this language). In English, verbs like *persuade* and *encourage* take a Modal to complement clause while their antonyms, *dissuade* and *discourage*, take the rather different FROM ING type. Compare *I persuaded/encouraged Tom to go* with *I dissuaded/discouraged Tom from going*. This can extend to complementation strategies, discussed in §18.6. For example, in Dyirbal *gigal* ‘tell to do, let do’

takes the purposive complementation strategy, while *jabil* ‘tell not to do, refuse to allow’ requires the relative clause strategy.

In some languages there are intransitive verbs with adjective-type meanings—and in others there are adjectives themselves—which take a complement clause. Most of these relate, semantically, to some of the Primary-B (or Secondary) types, and have similar complement-taking properties. English has, for example, pairs such as:

ADJECTIVE	VERB	ADJECTIVE	VERB
unsure (of/about)	doubt	fond (of)	enjoy
sorry (about)	regret	eager (for)	want
afraid (of)	fear		

It is unusual to encounter a noun (or NP) which can govern a complement clause. However, in English (and in Modern Hebrew, see Zuckermann 2006: 80), we find sentences like:

- (57) (a) John_A told me_O [the news]_{CoCl:E}
 (b) John_A told me_O [that Fred had broken his leg]_{CoCl:E}
 (c) John_A told me_O [[the news] [that Fred had broken his leg]_{CoCl}]_E

The message which John purveys can be realized by an NP in (57a), or by a complement clause in (57b), or by both together in (57c). Sentence (57c) has a complex O argument, involving an NP and a complement clause in apposition. (Note that it would be unhelpful to suggest that the clause *that Fred had broken his leg* is a complement clause (or a relative clause) to *the news*.)

18.5.2 Secondary semantic types

Some concepts are always expressed by lexemes—for example ‘eat’ and ‘rub’—while others are always part of the grammar—pronouns, and markers of syntactic function. Between these, there lies a set of ‘Secondary concepts’, which are realized by grammatical forms in some language and by lexemes (Secondary verbs) in others. They include ‘not’, ‘can’, ‘must’, ‘begin’, ‘try’, ‘want’, ‘hope’, ‘plan’, ‘make’, and ‘help’. A Secondary concept will always modify the meaning of a Primary verb to which it is linked, either as a grammatical element or as a lexeme through a complement clause construction (discussed below) or a complementation strategy (dealt with in §18.6).

We can summarize the ways in which a Secondary concept may be realized:

- (i) As an affix to a verb. For example, Macushi, a Carib language from South America, has verbal suffixes which include *-yonpa* ‘try’, *-pia’iti*

‘begin’, and *-aretika* ‘finish’ (Abbott 1991: 120–1). Fijian has a prefix *via-* ‘want to’. As mentioned in §1.11, ‘begin’ is shown by verbal suffix *-yarra-* in Dyirbal—*baŋga-* is ‘paint’ and *baŋga-yarra-* ‘begin to paint’.

- (ii) As a Secondary sense of an affix, often one from a TAM system; for example, the ‘intention’ modality suffix in Jarawara (with feminine form *-(ha)bone* and masculine form *-(hi)bona*) can be used with the sense ‘should’ or ‘want to’.
- (iii) As a word modifying the verb, or modifying the whole clause; for example *not*, and modals *can*, *must*, etc. in English.
- (iv) As a lexical verb. This may occur in a complement clause construction with the Primary verb which it semantically modifies. An example from English is:

(58) John_A began [to write a detective story]_{CoCl:O}

Here the Secondary verb *begin* is predicate of the main clause, while Primary verb *write* occurs in the complement clause which is in O function for *begin*. But, at the semantic level, ‘begin’ is here a modifier to ‘write’. Sentence (58) is not about an activity of beginning, it is about an activity of writing, with *begin* specifying a phase of this activity.

A complement clause verb (and the associated complementizer) may be omissible if they would be understood by the addressee, on the basis of the context in which the utterance occurs and information which speech act participants share. One can say *John_A began [to write a detective story]_{CoCl:O}* or *John_A began [to read a detective story]_{CoCl:O}* or *John_A began [to typeset a detective story]_{CoCl:O}*, and so on. Any of these can be shortened to *John began a detective story*, if the addressee(s) can supply the omitted complement clause verb—if they know that John is a writer of detective stories (and the speaker is describing this aspect of John’s life, not what he does to relax in the evenings), or if they know that all John ever does with respect to detective stories is to read them, or if they know that he is a typesetter (and the speaker is describing a current work task). That is, a speaker will only say something like *John began a detective story* if they consider that the addressee(s) should be able to infer the semantic content of the unstated complement clause verb.

This point can be further demonstrated by a conjunction of two abbreviated sentences, such as:

(59) John_A began [a detective story]_O and Mary_A (began) [a historical novel]_O.

Either John and Mary are authors, each beginning to write a work belonging to their respective genres, or they are both readers, or both typesetters, and so on. This abbreviated conjunction could *not* be used if John were beginning

to write and Mary to read or typeset, or vice versa. They *must* be engaged in the same sort of activity if the underlying full complement clause construction is to be abbreviated. That is, in the conjunction *John_A began [to X a detective story]_{CoCl:O}* and *Mary_A began [to Y a historical novel]_{CoCl:O}*, the verbs X and Y in the complement clauses must have the same meaning, as a condition for this conjunction to be reduced to *John began a detective story and Mary (began) a historical novel*. Similar examples and arguments can be given for *finish*, *want*, *try*, and indeed for all other Secondary verbs. In summary, a Secondary verb always provides semantic modification for a Primary verb, which is either explicitly stated or understood from the context.

There are three types of Secondary verbs:

Secondary-A. These do not add any argument to those of the Primary verb being modified. Both *John wrote a detective story* and *John began to write a detective story* involve just two NP arguments, *John* and *a detective story*.

Secondary-B. One argument is added to those of the Primary verb being modified—compare *Mary will go* and *John plans for Mary to go*. The subjects of the two clauses are often identical, and the second occurrence is then omitted. One says *John plans to go*, rather than **John plans for himself to go* (this would only be likely to be heard as an ‘echo’).

Secondary-C. Again, an argument is added—compare *Mary went* and *John forced Mary to go*. With this type of Secondary verb, it is unlikely that main and complement clauses will have the same subject. If they do, there can be no omission—one must say *John forced himself to go*, not **John forced to go*.

It is now time to examine these three main types of Secondary verbs, and note which kinds of complement clause construction each is likely to occur with.

Secondary-A

Since there is no addition to the semantic roles of the Primary verb which is being modified, a Secondary-A concept is particularly likely to be realized by a verbal affix or a grammatical modifier. However, all subtypes can be shown as lexical verbs.

(i) **NEGATORS** such as ‘not’, ‘never’, ‘don’t’ A handful of languages code negation through an intransitive verb which takes a Fact complement clause in S function. They include *sega* ‘it is not the case’, and negative imperative ‘*ua* ‘it is not the case that (someone) should’ in Fijian (Dixon 1988a: 281–2), and *ta?* ‘it is not the case’ in Shuswap (Salish, British Columbia; Kuipers 1974: 81). Makah

(Wakashan, British Columbia; Davidson 2002: 155–6) has three negative verbs, *wiki* ‘it is not the case’, *wi·ya* ‘it is never the case’, and *yubut* ‘it is not the case that (someone) is able to’. For example:

- (60) *wi·ya-id* [wa·
 BE.NEVER.THE CASE.-INDICATIVE:1pl say:PERFECTIVE
 ʔu·]CoCl:S
 DEMONSTRATIVE
 We never say this (lit. That we say this is never the case)

(ii) *MODAL-TYPE*, such as ‘can’, ‘should’, ‘must’ Often realized as a verbal affix or—as in English—by a modal verb. May be an intransitive verb taking a complement clause in S function; generally a Potential clause, sometimes also a Fact one.

In Fijian *dodonu* ‘must, be necessary’ only accepts a Potential complement clause, while *rawa* ‘can, be able to’ may occur with a Fact clause, as in (61a), or a Potential one, as in (61b).

- (61) (a) *e rawa [ni la’o [o Mika]NP:S]CoCl:S*
 3sgS can THAT go ARTICLE Mika
 Mika can go
 (b) *e rawa [me la’o [o Mika]NP:S]CoCl:S*
 3sgS can TO go ARTICLE Mika
 Mika can go

Whereas (61a) simply states that there is nothing to impede Mika from going, (61b) has stronger import—literally, he is able and willing to go (Dixon 1988a: 283). Note that the complement clause is coded in the main clause by 3sg subject pronoun *e*.

Irish is another language in which Modal concepts are coded as complement-taking verbs—see Stenson (1981: 65–5, 86).

(iii) *BEGINNING TYPE*, such as ‘begin’, ‘start’, ‘continue’, ‘stop’, ‘cease’, ‘finish’ If realized as a lexeme, this is generally a transitive verb, taking a complement clause as O argument. The complement clause is most commonly of the Activity type, as in English *He_A began/continued/finished [washing the clothes]*CoCl:O. May also describe the potentiality of getting into, or continuing with, or ceasing from an activity, as in *He_A began/continued/ceased [to wash the clothes]*CoCl:O. (See Dixon 1991: 172–9, 2005: 177–83 for discussion of the different complement clause possibilities for these verbs in English.)

(iv) *TRYING-TYPE*, such as ‘try’, ‘attempt’ Again, may be realized by a transitive verb. The prototypical complement clause—in O function—is Potential, as in *John_A tried/attempted [to eat the pie]*CoCl:O (but wasn’t able to get a bite

of it, since other people were pushing him away). In English, the verb *try* also has the sense of ‘testing, tasting’ and can then take an Activity clause, as *John_A tried [eating the pie]_{CoCl:O}* (he took a mouthful and then decided he didn’t like it).

Note that when a Secondary-A concept is realized as a Secondary verb which takes a complement clause in O function, the main and complement clauses must have the same subject. As stated above, a Secondary-A verb may not add an argument to those of the verb it is semantically modifying.

Secondary-B

Includes verbs such as ‘want’, ‘wish (for)’, ‘hope (for)’, ‘intend’, ‘plan (for)’, ‘pretend’.

Some languages have a desiderative verbal affix, as in Awa Pit (Barbacoan, Colombia and Ecuador; Curnow 1997a: 166–8):

- (62) tuk-shi-s
 suck-DESIDERATIVE-CONJUNCT
 I want to smoke

Plainly, this involves the person doing the wanting being the same as the person who undertakes the wanted activity.

In most languages, Secondary-B concepts are expressed by transitive verbs. The complement clause in O function to such a verb is typically of the Potential type, but some Secondary-B verbs may also accept a Fact clause.

In English, the word *want* is directly pragmatic, referring to something which could be achieved, and is limited to a Potential complement clause. *Wish*, in contrast, may have wistful overtones, referring through a Fact clause (which includes a modal) to something that could not be realized, as in *I_A wish [that I could have talked with Aristotle]_{CoCl:O}*.

The complement clause of ‘hope (for)’ may refer to the potentiality of something happening in the future, or to the fact of something which has already happened but concerning which the speaker does not yet have information. For the latter sense a Fact complement clause will be appropriate; for example, in English *I_A hope [that John did lock the door last night]_{CoCl:O}*.

‘Pretend’ behaves pretty much like other Secondary-B verbs, except that it often refers to the present or past. As a consequence, a Fact complement clause need not include a modal; for example, in English *I_A pretended [that I was a preacher]_{CoCl:O}*.

Many languages use the same construction irrespective of whether main and complement clauses have the same or different subjects—for example *I want to go first* and *I want Mary to go first* in English. This also applies for Modern Hebrew (Zuckermann 2006), Jarawara (Dixon 2004a, 2006b),

Dolakha Newar (Genetti 2006), and Akkadian (Deutscher 2006). In other languages, different construction types are required.

Verb *xav* ‘want’ in White Hmong (Jarkey 2006) occurs in a Potential complement clause construction only if the subjects of the two clauses are different:

- (63) *yog koj_A xav [kom tsis txho xa nyiaj*
 IF 2sg want TO NEG IRREALIS:NEG send money
*tuaj]*_{CoCl:O} ...
 come

If you don’t want (people) to send money...

But if the subjects are the same a serial verb construction must be employed (this is a complementation strategy, rather than a complement clause construction):

- (64) [*cov menyuam*]_S [*xav ua:si*]_{PREDICATE} [*nram*
 COLLECTIVE.CLASS child want play down
pas-dej]
 pond-water

The children want to play down (at) the pond

And there are languages which require different complementation strategies for same and different subjects. In Tariana (Aikhenvald 2006), a serial verb construction strategy is used for same subject and a nominalization strategy for different subject. In Dyirbal (§18.6.4 and Dixon 2006d) verbs of wanting are all intransitive and their S can be coreferential with S or O of a following verb in purposive inflection. There is no simple way of expressing a different-subject intention, such as ‘I wanted John to go’. Depending on the circumstances, one could be specific and say something like ‘I told John to go’.

In Jacalteco (Mayan family; Guatemala; Craig 1977: 234–8) ‘would like’ (literally ‘stomach wants’)—with main and complement clause having the same subject—takes a Potential clause if the verb of the complement clause refers to an activity, but a Fact clause if the verb refers to a state. Thus, literally ‘I would like [to see him]_{CoCl}’ but ‘I would like [that I be rich]_{CoCl}’. When the subjects of the two clauses differ, only a Fact complement clause may be used; we get, literally ‘I would like [that you sleep]_{CoCl}’. (See also Dixon 1995: 215 on Kamaiurá.)

Secondary-C

Includes verbs such as ‘make’, ‘cause’, ‘force’, ‘let’, and ‘help’.

This kind of Secondary concept can be realized by a verbal affix which increases the valency of the verb by one, adding a ‘causer/helper’ role. Or

it may be realized through a Secondary verb, taking a complement clause as O or E argument. The complement clause is typically of the Potential variety.

If the complement clause is in E function, its subject is likely to be coreferential with the O argument of the main verb, as in *I forced him to say sorry*. It is unlikely to be the same as the main clause subject. In the unlikely event that it is, ellipsis may not be allowed.

18.6 Complementation strategies

All languages have a set of ‘complement-taking verbs’ (CTVs); typical members include ‘see’, ‘think’, ‘know’, and ‘like’. There may be a full set of complement clause constructions relating to these verbs; this applies to familiar languages such as English, and also to Pennsylvania German (Burridge 2006) and Modern Hebrew (Zuckermann 2006). And there are languages whose grammars have no instance of a clause filling a core argument slot in a higher clause—there are no complement clauses whatsoever. In place of these, other grammatical resources are employed, what we can call ‘complementation strategies’. Dyirbal is of this ilk; it uses three kinds of strategy—serial verb constructions, relative clause constructions, and purposive clause linking. In between these extremes, there are languages with a fair array of complement clause constructions which also employ one or more strategies. For example, White Hmong has five varieties of complement clause constructions and also employs, as a strategy, a serial verb construction. We saw just above that *xav* ‘want’ takes a Potential complement clause when the wanter and the person who undertakes the wanted activity are different, as in (63), but a serial verb strategy, exemplified in (64), when they are the same.

Rather few of the grammars which deal with complementation provide explicit criteria for recognizing a putative complement clause as an argument of the verb in the main clause. Some of the phenomena which have been called complement clauses are undoubtedly not this, but rather complementation strategies. Since most grammarians do not explicitly distinguish between complement clauses and complementation strategies (a distinction introduced in Dixon 1995), it is difficult to provide a full account of the latter. However, all of the contributions to Dixon and Aikhenvald (2006) do make this distinction, as do a number of other recent studies (including Deutscher 2000, Zeitoun 2007, Jany 2007, and Aikhenvald 2008a). These enable a start to be made on a cross-linguistic study.

The sections which follow discuss serial verb constructions, relative clauses, nominalizations, and clauses linked together within one sentence.

18.6.1 Serial verb construction strategy

In a serial verb construction (SVC), two (or sometimes more) verbs function together like a single predicate and are conceived of as describing a single action. The most common variety of SVC is asymmetrical, with a Major member (covering a wide range of verbs) and a Minor member (one of a small set of verbs); see Aikhenvald and Dixon (2006). A language with SVCs typically has a number of distinct asymmetrical varieties, where the Minor set for each variety is semantically homogeneous. Recurrent varieties include Direction (for example, ‘go’, ‘come’, ‘return’) and Association (‘be with’, ‘be together’). They also include all kinds of Secondary verbs except negators.

There may be a variety of asymmetrical SVC where the Minor member is a Secondary-A verb such as ‘can’ or ‘must’ or ‘begin’ or ‘stop’ or ‘try’. An example from Tariana (Aikhenvald 2003: 433) is, with the serial verbs indicated by subscript ‘_{sv}’:

- (65) [wa-rapa wa-thaka]_{sv} whas
 1pl-dance 1pl-stop we
 We stopped dancing for a while

Dyirbal has a set of verbs which feature as Minor member of an SVC. The two verbs in the SVC must agree in final inflection (tense, purposive, imperative, etc.) and in transitivity. For example:

- (66) ŋaja_A bala-m_O [muñama-n jaŋga-ñu]_{sv}
 1sg THERE:ABSOLUTIVE-EDIBLE finish.off-PAST eat-PAST
 I’ve eaten up all the vegetable food

There are (in my corpus) about forty verbs which can be minor members of an SVC such as that in (66). They include ‘be the first to do’, ‘stop doing’, ‘do again’, ‘do too much’, ‘do quickly’, ‘do well’, and ‘do badly’ (see Dixon 2006d: 346).

There may be a variety of asymmetrical SVC whose Minor member is a Secondary-B verb such as ‘want’, ‘plan’, ‘intend’ (also found in Tariana). The general rule is that all verbs in an SVC should have the same subject. This naturally applies to Secondary-A verbs, but it means that Secondary-B verbs may be restricted to the same subject sense when used in an SVC (as illustrated by (64) from White Hmong).

Secondary-C verbs such as ‘make’, ‘cause’, ‘force’, ‘let’, and ‘help’ are sometimes found in a special subtype of SVC—the ‘switch function’ subtype—where the O argument of the Minor verb (the Secondary-C item) is identical to the A or S argument of the Major verb. An example from North-East Ambae (Austronesian, Vanuatu; Hyslop 2001: 303) is:

- (67) mo [vai ngire dolegi ra=mo inu=e]_{SV}
 REALIS make 3nsg all 3nsgA=REALIS drink=3sgO
 He made all of them drink it

SVCs are typically used with Secondary verbs and just occasionally with Primary-B verbs. This is often by analogy with Secondary verbs; for example, ‘order’ may be coded in a similar way to ‘make’, and ‘know how to’ in a similar way to ‘can, be able to’. Further work is needed to explore the possibilities of Primary-B verbs as Minor members within an asymmetrical SVC.

18.6.2 Relative clause strategy

If a language does not have an appropriate complement clause construction, it may code an Activity by means of a relative clause strategy. Dyirbal uses this strategy, with CTVs of the ATTENTION and THINKING types. An example from a story is:

- (68) ŋaja_A bura-n [[gayu-ŋga]_{RC} ñalŋga [wanda-ŋu]_{RC}]_O
 1sg see-PAST cradle-LOCATIVE child hang-REL
 I saw the child hanging in a cradle (lit. I saw a child which was hanging in a cradle)

Note that in Dyirbal a relative clause is marked by verbal suffix *-ŋu*, in place of a TAM ending. Word order is quite free in this language and here the relative clause *gayu-ŋga wanda-ŋu*, ‘which was hanging in a cradle’, is discontinuous within the O NP.

Now in English there is a distinction between a relative clause construction, as in (69a), and a complement clause construction, as in (69b).

- (69) (a) I_A saw [the child [(who was) hanging in a cradle]_{RC}]_O
 (b) I_A saw [the child(’s) hanging in a cradle]_{CoCl:O}

Sentence (69a) states that I saw a child, with the relative clause specifying which child it was I saw. In contrast, (69b) states that I saw a happening, a child’s hanging in a cradle.

In Dyirbal this distinction cannot be made. Sentence (68) is potentially ambiguous between the two readings, (69a) and (69b). However, there is no significant difference in meaning between (69a)—relating to seeing a child, who is hanging in a cradle—and (69b)—relating to seeing a child hanging in a cradle—so that little is lost. The specific import of (68) will be inferred from the discourse context in which it occurs.

Whereas a complement-taking verb such as ‘see’ can take a complement clause in a language which has complement clauses, in a language lacking

complement clauses it may carry the *expectation* of occurring in an appropriate complementation strategy.

In §18.1 we gave a pair of sentences—(16) and (17)—which are syntactically like (69a/b) but have *dislike* as the verb in the main clause; as a consequence, their meanings are significantly different. However, in Dyirbal verbs from the *LIKING* type enter into a purposive complementation strategy, not a relative clause strategy. The likelihood of confusion is thus avoided.

A relative clause strategy is also employed in Akkadian (Deutscher (2000: 141–3, 2006: 172–3) and in Manambu (Aikhenvald 2008a: 500–1).

18.6.3 Nominalization strategy

‘Nominalization’ is used to describe a process (and its result) by which something with the properties of a nominal can be derived from a verb or adjective. English is replete with deverbal nominalizations—*appreciation*, *statement*, *knowledge*, and *use* were mentioned at (22) in §18.2. There are also some nominalizations of verb-plus-object, such as *wife-bashing*.

Languages lacking a full range of complement clause constructions will often employ some kind of nominalization as a complementation strategy. For North-East Ambae, Hyslop (2001: 392) contrasts the verb ‘ask’ occurring with a Fact-type complement clause, in (70), and with a nominalization strategy, in (71).

- (70) na=ni hui [huri vo go=mo
 1sgA=IRREALIS ask COMP say 2sgA=REALIS
 domi-gine=o mwerehilogo]CoCl:O
 think-APPLICATIVE=3sgO how
 I’ll ask what (how) you think about it

- (71) na=ni hui [na domi=mu]NP:O
 1sgA=IRREALIS ask ACCUSATIVE think=3sgPOSSESSIVE
 I’ll ask your thoughts

In (70), the complement clause functions as O argument of ‘ask’; it includes a realis marker, showing that it is a clause. The O constituent in (71) has the structure of an NP, with accusative marker preceding the nominalized verb, *domi=mu*, ‘your thoughts’, which is NP head.

Kham (Tibeto-Burman, Watters 2002: 331–41) uses nominalization as its complementation strategy (and has no complement clauses). Generally, the implicit subject of the nominalized verb should be identical to the subject of the main clause. The same-subject sense of ‘want’ takes this strategy, as in:

- ŋa-pəĩ-zya
1sg-want-CONT

Nominalization is recognized as a complementation strategy for Akkadian (Deutscher 2000, 2006), Tariana (Aikhenvald 2003, 2006), Rukai (Zeitoun 2007: 419–21), and Manambu (Aikhenvald 2008a: 502–3). The CTVs with which it occurs differ from language to language, but ‘know’ appears to recur.

Bontkes (1985: 197–202) describes an appositional construction in Suruí (Tupí family, Brazil), where the clause with the CTV comes last:

- (73) *été awuru sádé a-peyare; ewe íkin o-or*
 THEN dog IMPERV 3-bark THAT see 1sg-come
é
 SENTENCE.MARKER

Then I came and saw the dog barking (lit. Then the dog was barking; I came and saw that)

Here the ‘that’ of the second clause refers to ‘the dog was barking’. Note that the first clause in (73) can include any tense–aspect marker.

The appositional strategy is particularly suited for use with CTVs which in other languages occur with Fact and Activity complement clauses, especially verbs like ‘know’ and ‘see’. This strategy is described for Akkadian (Deutscher 2006), Chimariko (isolate, California; Jany 2007), and Manambu (Aikhenvald 2008a: 497–500).

(b) **Clause chaining.** Some languages have a device of ‘clause linking’, whereby a number of clauses referring to distinct but related events are placed in sequence. There is one main clause (often the last in the chain), marked for the full set of verbal categories. Other (‘medial’) clauses may just mark whether they have the same or different arguments to a following clause. Just occasionally, clause chaining can function as a complementation strategy, as in Manambu (Ndu family, Papua New Guinea; Aikhenvald 2008a: 497):

- (74) [*amæy wa-lə-k*] [a-wuk]
 mother see-3fem.sg-COMPL.DIFFERENT.SUBJECT IMPERATIVE-obey
 Obey what mother said! (lit. Mother having said, obey)

There can be other clause chaining strategies, such as sequential subordinate clauses in Tariana (Aikhenvald 2006). For Matses (Panoan, Peru), Fleck (2006) describes a further kind of complementation strategy which he terms ‘adverbialization’. This is used with Primary-B verbs of ATTENTION, THINKING, SPEAKING, and LIKING, and with Secondary-A verbs of BEGINNING and TRYING. For example ‘I stopped working’ is rendered by ‘While working I stopped’, and ‘I don’t know how to hunt alone’ is, literally, ‘While I am hunting alone I don’t know.’

(c) **Purposive linking.** In §18.1 we compared the two English sentences:

- (18) He_S went (in order) to swim
 (19) He_A wanted [to swim]_{CoCl:O}

He went is an intransitive clause which can stand alone or be linked to another clause by (*in order*) *to*, as in (18). In contrast, *want* is a transitive verb for which an O argument must be stated—this can be an NP or, as in (19), a complement clause. The underlying structures of (18) and (19) are totally different.

These two sentences are translated into Dyirbal by:

- (75) *bayiṣ* *yanu* *juwi-gu*
 THERE:MASCULINE go:PAST swim-PURPOSIVE
 He went to swim
- (76) *bayiṣ* *walḡgarra-ñu* *juwi-gu*
 THERE:MASCULINE want-PAST swim-PURPOSIVE
 He wanted to swim

Walḡgarra- ‘want to do something to satisfy persistent emotional worry or desire’ is an intransitive verb, just like *yanu* ‘go’. Unlike (18) and (19) in English, (75) and (76) in Dyirbal have exactly the same underlying and surface structures.

The difference is that whereas *yanu* ‘go’ is a Primary-A verb, so that *bayi yanu* ‘he went’ is a self-contained sentence, *walḡgarra-* ‘want’ is a Secondary-B verb which carries the expectation of being linked to a following verb in purposive inflection, as a complementation strategy (since the language lacks a Potential complement clause construction). That is, one could not have a sentence consisting just of *bayi walḡgarra-ñu* ‘he wanted’.

Dyirbal works in terms of an entirely ergative syntax. That is, any two clauses which are linked—as in (75–6)—must have a shared argument which is in S or O function in each. In (76) *bayi* ‘he’ is in S function in each clause. In (77) *bayi* is in S function in the first clause and (although ellipsed) in O function in the second clause.

- (77) *bayiṣ* *walḡgarra-ñu* *yibi-ḡgu*_A *bura-li*
 THERE:MASCULINE want-PAST woman-ERGATIVE see-PURPOSIVE
 He wanted to be seen by the woman (that is: He wanted the woman to see him)

A Secondary-B verb such as ‘want’ typically has its subject identical with the subject of the activity it relates to: ‘I want to do’ or ‘I want to do something to X’ (these are much more frequent than ‘I want X to do something to me’). In English the A argument of *want* is likely to be identical with the S or A argument (which is then omitted) of the complement clause which is the O argument of *want*. This is in keeping with the S/A syntactic pivot of English.

But Dyirbal has an S/O pivot. It is in keeping with this that verbs of wanting in this language are intransitive, with their S argument being identical with the S or O argument (which is then omitted) of the following purposive clause. If it is the A argument of the purposive clause which is identical with the S argument of the WANTING verb, then the purposive clause must be recast as an antipassive, as in:

- (78) $\eta a\eta a_s$ $wal\eta garra-\tilde{n}u$ $wuju-gu$ $\eta a\eta ga-nay-gu$
 1sg want-PRESENT food-DATIVE eat-ANTIPASSIVE-PURPOSIVE
 I want to eat some food.

The verb *\eta a\eta ga-* ‘eat’ is transitive. Suffix *-nay-* derives an antipassive stem for which the underlying A argument (‘I’) goes into S function, and the underlying O (‘food’) now takes dative case.

The purposive complementation strategy is generally used with CTVs which in languages with complement clause constructions take a Potential complement clause—Secondary-B verbs of WANTING, which are all intransitive, and also with Primary-B verbs of LIKING, which are transitive. Similar to WANTING verbs, those from the LIKING type expect a following clause marked as purposive. Once, in the field, I asked whether one could say a sentence such as:

- (79) $\eta a\eta a_A$ $[bala-m$ $\eta arrinyji]_O$ $\eta uymi-nyu$
 1sg THERE-EDIBLE orange like-PRESENT
 I like oranges

Speakers were not happy with such a sentence, and told me that a verb such as ‘eat’, in purposive inflection, should be added (as the complementation strategy):

- (80) $\eta a\eta a_A$ $[bala-m$ $\eta arrinyji]_O$ $\eta uymi-nyu$ $\eta a\eta ga-ygu$
 1sg THERE-EDIBLE orange like-PRESENT eat-PURPOSIVE
 I like to eat oranges (lit. I like oranges, to eat them)

Here ‘I’ is the A argument and ‘oranges’ the O argument for both ‘like’ and ‘eat’. That is, one cannot just say that one likes something, it is necessary to state what one wants to do with the something.

Languages which lack a Potential complement clause construction—illustrated in (19)—may use purposive clause linking, similar to (76–8), as a complementation strategy. Whereas a clause involving any (intransitive or transitive) verb may be—but need not be—linked to a following purposive clause (for example, ‘He went to bathe’ or ‘He took the car to get it repaired’), a verb like ‘want’ may *carry the expectation* of a following purposive clause, as a complementation strategy (something like ‘He wanted to bathe’ or ‘He wanted to get the car repaired’).

Interestingly, verbs which take a purposive complementation strategy are often intransitive and extended intransitive. This applies for *wal\eta garra-* ‘want’ in Dyirbal and for *d\eta l* ‘want, desire, like, love’ in the Australian language Djambarrpuynngu (Wilkinson 2004). They can, however, be transitive, as in Akkadian (Deutscher 2006) and Manambu (Aikhenvald 2008a: 501). The

purposive strategy is attested with verbs such as ‘remember’, ‘like’, ‘promise’, ‘threaten’, ‘persuade’, ‘tell (to do)’, ‘order’, ‘want’, and ‘try’; it is not attested with ATTENTION verbs.

In summary, in a language which lacks a complement clause construction, a complement-taking verb carries the expectation of entering into one or more complementation strategies, some of the varieties of which were surveyed above.

As mentioned before, Dyirbal lacks any complement clause construction, but instead employs three complementation strategies. These function as follows (fuller details are in Dixon 2006c):

- **Serial Verb Construction strategy.** Secondary-A verbs of BEGINNING and TRYING (some are transitive, such as ‘finish doing’, ‘try’, and some intransitive, such as ‘stop doing’, ‘pretend to do’).
- **Relative clause strategy.** ATTENTION and THINKING verbs (all transitive). Some SPEAKING verbs (mostly transitive, such as ‘tell’, ‘answer’, ‘grumble at’, and one intransitive ‘call out in fright’). Some LIKING verbs (all intransitive, such as ‘be jealous about’, ‘be shy and ashamed’).
- **Purposive strategy.** Some SPEAKING verbs (mostly transitive, such as ‘ask’, ‘tell to do, let do’, and one intransitive, ‘promise to come’). Some LIKING verbs (some transitive, such as ‘like’, ‘dislike’, and some intransitive, such as ‘show off, act proud’). WANTING verbs (all intransitive).

18.7 Summary

In every language there are one or more grammatical processes for relating the action or state described by one verb—which can, in essence, be any verb—to an argument of another verb—from a restricted set of ‘complement-taking verbs’ (CTVs).

Cross-linguistically, the most common CTVs have meanings such as ‘see’, ‘hear’, ‘know’, ‘believe’, and ‘like’ and often also ‘tell’ (these are Primary-B verbs). Their object argument may refer to an object (through a noun or pronoun) or to an activity or state (through a complement clause).

A number of universal ‘Secondary concepts’—including ‘can’, ‘begin’, ‘try’, ‘want’, and ‘make’—have varying realization across languages; they may be an affix to a verb, or a modifier to a verb or a clause, or a lexical verb (a Secondary verb). If realized as verbs, they are also CTVs, and will enter into the same sort of grammatical relations with another verb as do Primary-B verbs.

A complement clause construction involves a CTV as predicate of the main clause with a complement clause filling one of its core argument slots. Some

languages lack complement clause constructions (or else, have only a limited set of them) and then employ alternative grammatical means to express the relationship between a CTV and another verb. Such ‘complementation strategies’ include serial verb constructions, relative clause constructions, nominalizations, and a variety of kinds of clause linking. These may involve clausal apposition (with one argument of a clause which has a CTV as predicate being something like a demonstrative which refers to a following clause, as in ‘I saw that. You ate the mango’), clause chaining, or a clause with a CTV being linked to a following clause with purposive marking. Some languages work entirely in terms of complement clauses, others only through complementation strategies, while a further set combine the two.

Complement clauses are likely always to function as O argument for some verbs. Depending on the language, they may also be in S and/or A and/or CS and/or CC and/or E function. There are three recurrent varieties of complement clause, each with its typical meaning and grammatical properties (including whether main and complement clause subjects can be the same or different). Each of the varieties has typical co-occurrences with CTVs (but note that there are many variations on this scheme).

- Fact complement clauses are often found with Primary-B verbs such as ‘think (of/about/over)’, ‘imagine’, ‘dream (of/about)’, ‘assume’, ‘remember’, ‘forget’, ‘know’, ‘understand’, ‘believe’, ‘recognize’, ‘discover’, ‘say’, ‘inform’, and ‘report’; and with Secondary verbs such as ‘not’, ‘can’, and ‘wish’.
- Activity complement clauses are often used with Primary-B verbs such as ‘see’, ‘hear’, ‘like’, ‘fear’, ‘enjoy’, and ‘describe’. Also with Secondary verbs such as ‘begin’ and ‘continue’.
- Potential complement clauses tend to be used with Primary-B verbs such as ‘promise’, ‘threaten’, ‘order’, and ‘persuade’. And with such Secondary items as ‘should’, ‘try’, ‘want’, and ‘make’.

The various complementation strategies do not have direct semantic correspondence with the varieties of complement clauses. Tentative associations with CTVs are (note that much more work is needed on complementation strategies):

- Serial verb constructions—typically with Secondary verbs.
- Relative clause constructions—typically referring to an Activity, with ‘see’, ‘hear’, ‘discover’, ‘think of’, ‘dream about’.
- Nominalization and Clause Chaining—may be almost equally acceptable with any verb.

- Apposition—typically with verbs which take Fact and Activity complement clauses, such as ‘know’ and ‘see’.
- Purposive linking is particularly suitable for verbs which may relate to Potential, such as ‘remember’, ‘like’, ‘promise’, ‘threaten’, ‘persuade’, ‘tell (to do)’, ‘order’, ‘want’, and ‘try’; it is unlikely to be used with ATTENTION verbs.

18.8 What to investigate

I Complement clauses. Does the language you are working on have complement clauses? These have the structure of a clause but function as an argument of the main verb (the complement-taking verb or CTV). Provide a thorough description of each variety of complement clause, including the following.

- (a) Its meaning—could be a variety of Fact, Activity, or Potential, or something else. (If a verb can occur with two or more complement clauses, it may be useful to contrast the different meanings involved.)
- (b) The grammatical criteria for identifying it as a complement clause; that is, the ways in which it functions as an argument of the main verb. (For example, how does a complement clause behave under passivization?)
- (c) What are its functional possibilities within the main clause? (Some of O, A, S, CS, CC, E, and perhaps also a non-core function.) If the language has bound pronouns, how is a complement clause in a core argument slot coded in terms of the bound pronoun paradigm which relates to that slot?
- (d) What is the position of the complement clause within the main clause? Can (or must) it be extraposed to the end of the main clause? If extraposition is optional, can you say what motivates it?
- (e) State how it is marked as a complement clause, e.g. complementizer, special affix on verb, special marking on subject. For each complementizer, state what other functions (if any) this form has in the grammar.
- (f) Its internal structure, compared with that of a main clause (and of other varieties of subordinate clause). This includes:
 - How are core arguments marked (as NPs, and as bound pronominals in a head-marking language)? Do the core arguments have the same grammatical possibilities in the complement clause as in a main clause?
 - Can it include peripheral arguments? How are these marked?
 - Can it be negated? How is negation marked?

- How are tense, aspect, and modality marked (plus any other categories shown in a main clause)?
 - Can the complement clause verb undergo the same kinds of derivational process as a main clause verb?
- (g) Does the complement clause have any other (perhaps unusual) properties? For example, may it be modified by an adjective?
 - (h) Are there constraints on coreferentiality between subject of complement clause and subject (or some other argument) in main clause? If subject identity (or some other argument identity) is required, then can or must the token of this argument in one of the clauses be omitted?
 - (i) Can an NP (or a pronominal) argument have simultaneous (and different) functions in main clause and in complement clause, as does *John* in sentence (49) and (50) from the end of §18.3?
 - (j) Is there any ‘raising’ of an NP or pronoun from the complement clause into the main clause? (Be sure to explain very carefully what you regard as ‘raising’.)

II Complementation strategies. List the strategies employed. They may include:

- (a) Serial verb construction
- (b) Relative clause construction
- (c) Nominalization (nominalized verb functions as head of NP). Summarize ways in which a nominalization behaves the same as, and ways in which it behaves differently from, an underived noun.
- (d) Linked clauses: (i) Apposition; (ii) Clause chaining; (iii) Purposive linking. (Or any other?)

For each strategy, first of all give a general characterization of the construction type (when not being used as a complementation strategy), then describe how it is also used as a complementation strategy.

III Primary-B complement-taking verbs. For each CTV, state the types of complement clauses it takes (and in what syntactic function) and/or what complementation strategies it enters into. Make sure that you include the following Primary-B types (for whichever of these lexemes the language has).

ATTENTION, such as: (a) ‘see’, ‘hear’, ‘notice’, ‘smell’, ‘show’; (b) ‘recognize’, ‘discover’, ‘find’.

THINKING, such as: (a) ‘think (of/about/over)’, ‘consider’, ‘imagine’, ‘dream (of/about)’; (b) ‘assume’, ‘suppose’; (c) ‘remember’, ‘forget’; (d) ‘know’, ‘understand’; (e) ‘believe’, ‘suspect’.

DECIDING, such as ‘decide (to)’, ‘resolve’, ‘plan’, ‘choose’.

LIKING, such as: (a) 'like', 'love', 'prefer', 'regret'; (b) 'fear'; (c) 'enjoy'.

SPEAKING (if the language has indirect speech), such as: (a) 'say', 'inform', 'tell' (one sense); (b) 'report'; (c) 'describe', 'refer to'; (d) 'promise', 'threaten'; (e) 'order', 'command', 'persuade', 'tell' (another sense).

There are likely to be other verbs, besides those just listed, which take complement clauses or enter into complementation strategies. These should be listed.

Are there any verbs which may take complement clauses in two functional slots (as *show* does in English, illustrated in (32) from §18.2)?

Are there any adjectives which may take complement clauses as a peripheral constituent? If so, say how they behave similarly to, and how differently from, complement-taking verbs.

IV Secondary concepts. For each secondary concept, say how it is realized. Typical possibilities are: as an affix or some other morphological process to a verb; as a secondary sense of some affix whose primary meaning deals with tense or aspect, etc.; as a modifier to a verb; as a modifier to a clause; as a lexical verb.

For those which are lexical verbs state what complement clauses they take, or what complementation strategies they enter into. The major secondary concepts include:

SECONDARY-A (the Secondary concept provides no addition to the semantic roles associated with the verb to which it is related):

(a) Negators 'not', 'don't'; (b) Modal-type, such as 'can', 'should', 'must', 'might'; (c) Beginning-type, such as 'begin', 'start', 'continue', 'stop', 'cease', 'finish'; (d) Trying-type, such as 'try', 'attempt'.

SECONDARY-B—'want', 'wish (for)', 'hope (for)', 'intend', 'plan (for)', 'pretend'.

SECONDARY-C—'make', 'cause', 'force', 'let', and 'help'.

V Historical inferences. Are you able to say anything about how each type of complement clause evolved (perhaps from an earlier complementation strategy)? Or can you say anything about how a particular construction type came to be used as a complementation strategy?

If you are familiar with how complementation is dealt with in a number of languages which are all spoken in the same geographical area, can you perceive any ways in which the languages have influenced each other in this respect? That is, has there been areal diffusion of varieties of complement clause constructions or of complementation strategies?

Sources and notes

This chapter is a thorough revision of Dixon (2006a). The other chapters in Dixon and Aikhenvald (2006) provide high-quality accounts of complement clause varieties and complementation strategies in eleven languages from across the world.

Noonan (1985) was for some years the standard text on complementation. His 2007 version is only slightly revised and, unfortunately, fails to take account of much of the recent literature. For instance, there is no mention of the distinction between complement clause varieties and complementation strategies—see, among others, Dixon (1995) and Deutscher (2000). Noonan's (1985, 2007) account is in a number of ways complementary to that presented here and repays study. Other useful publications include Ransom (1986) and Dik and Hengeveld (1991). There are detailed accounts of complementation in English in Dixon (1991a, 2005).

18.2. A further criterion for distinguishing between a complement clause and an NP—illustrated by our discussion of (20) and (21)—is that plural marker *-s* may be added to a nominalization, but never to the verb within a complement clause. This does, of course, only apply for a nominalization which is 'countable', such as *craving*. One can say [*Your craving(s) for sweet things*]_{NP:A} *worried* [*the doctor*]_O, whereas in the complement clause construction [*Your craving sweet things*]_{CoCl:A} *worried* [*the doctor*]_O, it is not permissible to add plural *-s* to *craving*, since it is here a verb.

As an addition to the discussion of criterion (I), it can be noted that the post-Bloomfieldian and Chomskian schools fail adequately to distinguish between clause and sentence; 'sentence' tends to be used when 'main clause' is intended; see §2.5. In general linguistic terminology, a sentence has an obligatory main clause and an optional number of subordinate clauses. Several sources refer to 'sentence-like complement clauses', meaning 'main-clause-like'. Craig (1977) talks of 'complement sentences' rather than 'complement clauses'. Noonan (1985: 42, 2007: 52) begins his seminal article by saying, 'by complementation we mean the syntactic situation that arises when a notional sentence or predication is an argument of a predicate'. At the least, the word 'simple' should be included before 'sentence'.

The post-war American tradition failed to distinguish between argument slot in clause structure, and the class of items that may fill it. They employed an oversimplified structural formula:

sentence consists of NP and VP

There was then no alternative to saying that a complement clause was a kind of NP. A theoretically more adequate characterization is:

sentence consists of main clause and optional subordinate clauses
 main clause consists of a predicate, a number of core arguments and a
 number of optional peripheral arguments
 corresponding to an argument slot in clause structure there may be an NP
 or a complement clause.

Criterion (II) was usefully discussed in Rosenbaum (1967), an early work on complement clauses.

There is a possible counter-example to the statement that a complement clause may in every language be used in O function. Fleck (2006) analyses Matses (Panoan family, Peru) as having just one verb, 'want', which takes a complement clause, and it is then in E function. Other CTVs in Matses employ a variety of complementation strategies.

18.4. Dimmendaal (1989) disuses complementizers in Hausa which distinguish 'predetermined, determined or future truth', 'doubtfulness', 'indetermination', and so on.

18.5. In their central meanings, Primary-A verbs may not take a complement clause as a core argument. However, a number of these verbs in English have metaphorical or idiomatic senses which may involve complement clauses; for example, [*That he had been passed over for promotion*]_{CoCl:A} *stung* [*John's pride*]_O. And some Primary-A verbs in English can take an optional oblique argument which may be a complement clause; for example, [*Mary*]_S *cried over* [*the policeman's having shot her dog*]_{CoCl}.

18.5.1. Güldermann and von Roncador (2002) and Aikhenvald (2004b: 132–42) describe ways of reporting what someone has said, as alternatives to complement clause constructions.

It appears that Goemai (Chadic branch of Afro-Asiatic, Nigeria) has no direct speech, a most unusual characteristic. Hellwig (2006) states that if one wants to quote what someone said then it must be done through reported (or indirect) speech, which is a complementation strategy. A set of logophoric pronouns facilitates this. The reported speech copies precisely what was said (including any errors).

Genetti (2006) put forward evidence for a direct quote in Dolakha Newar (Tibeto-Burman, Nepal) having the status of a 'complement clause', although it can involve several clauses or even just an interjection (but it is most commonly a single clause). One significant factor is that whereas in other languages the quote margin must either follow or precede direct speech—as in English *'I'll go,* *Mary told John* or *Mary told John 'I'll go'*—in Dolakha Newar the direct quote complement clause frequently intrudes into the middle of the quote clause—for example, literally, 'speaker-ERGATIVE_A [direct speech]_{CoCl:O}

say'. Another piece of supporting evidence is that the complete sentence may have prosodic unity, with no intonation breaks before or after the direct speech (just like an NP in O function).

18.5.2. A few Primary-A verbs referring to motion (prototypically 'come' and 'go', illustrated in §1.11) have alternative realization as verbal affixes in a few languages.

Noonan (1985: 132; 2007: 144)) mentions Shuswap and Fijian as languages where negation is realized by a complement-taking verb. But his Fijian example comes from an old source which apparently glosses *ena* as 'future'. All modern grammarians of Fijian would say that future is marked just by *na* with the initial *e* being the 3sg subject pronoun which cross-references the complement clause in S function, as in (61a/b).

Make and *let* in English appear to differ from other Secondary-C verbs in that they omit *to* from a Potential complement clause when in active voice; compare *They forced John to go* and *They made John go*. However, the *to* has to be included in the passive; one must say *John was made to go (by them)*, and not **John was made go (by them)*. The omission of *to* from an active clause with *make* and *let* is a minor matter of realization, and by no means sets them apart as a distinct set of CTVs.

18.6. Englebretson (2003) suggests that the grammar of colloquial Indonesian conversation involves no complement clause constructions, only complementation strategies. However, in terms of the theoretical stance he takes, Englebretson might well come to a similar conclusion about colloquial speech in many (perhaps all) other languages. He does include some interesting data, although a number of details of analysis are arguable.

18.6.1. Detailed discussion of SVCs in a range of languages are in the contributions to Aikhenvald and Dixon (2006).

18.6.4. Deutscher (2000: 55; 2006) describes how, in Akkadian, a linked clause construction has developed into a bona fide complement clause construction. In judicial statements the form *kīma* at first had the meaning 'as', so that one could say 'I_A prove-him_O, *kīma* (as) a slave'; this developed into 'I_A prove-him_O, *kīma* he is a slave'. It was then generalized to sentences like 'I_A prove-him_O, *kīma* he ran away'. Finally, the *kīma* clause was reanalysed as O argument of 'prove': 'I_A prove [*kīma* (that) he ran away]_{C_OCl:O}', where *kīma* is now the marker of a Fact complement clause, filling the O slot within the main clause. Another early use of *kīma* was as a clause linker 'because' as in 'I informed the governor *kīma* (because) the barley was not collected'; this use of *kīma* was also reanalysed as complementizer, giving 'I informed the governor [*kīma* (that) the barley was not collected]_{C_OCl}'; the construction type was then extended to 'know', 'see', and 'hear'. There were thus two complementation

strategies of different kinds—one involving *kīma* with the meaning ‘as’ and the other having *kīma* with the meaning ‘because’—which converged to become a Fact-type complement clause construction with *kīma* as the complementizer ‘that’.

Heine and Kuteva (2002) list a variety of diachronic lexical and grammatical sources for complementizers; these undoubtedly point to further kinds of complementation strategies, which could develop into complement clause constructions.

Glossary

Definitions are provided for a number of technical terms which recur in these volumes. For some entries there is reference to the chapter or section in which they are discussed. Note that Chapters 1–9 are in Volume 1 while Chapters 10–18 comprise Volume 2. Complementary terms are cross-references by ‘Compl.’

ABLATIVE: marker indicating movement away from the referent of the noun phrase to which it is attached.

ABSOLUTIVE: case inflection marking intransitive subject (S) and transitive object (O). Compl. ergative. §3.9, §13.2, §13.5.4.

ACCUSATIVE: case inflection marking transitive object (O). Compl. nominative. §3.9, §13.2, §13.5.4.

ACTIVE/STATIC: label covering split-S and fluid-S systems.

ADJECTIVE: class of words which typically refer to properties and have two main roles: (a) make a statement that something has a certain property through functioning in intransitive predicate slot or copula complement slot; and (b) help to specify the referent of the head noun in an NP by functioning as modifier to it. §3.6, §4.5, §6.1, §8.3.2, Chapter 12.

ADPOSITION: a marker of a (predominantly peripheral) grammatical relation which is realized as a separate phonological word or as a clitic, not as an affix. §5.4.

AFFINAL: kinship relation which involves a link by marriage. §1.3, §16.1.

AFFIX: a bound form added to a root or stem. §5.4.

AFFIXATION: morphological process which involves adding an affix to a root or stem. §3.13.

AGGLUTINATIVE: a type of language whose words are readily segmentable into a sequence of morphemes, each of which typically conveys one piece of information. §5.5.

AGREEMENT: when two words (for example, noun and modifying adjective within an NP) are marked for the same grammatical category. §5.6.

AIRSTREAM MECHANISM: a system for initiating a flow of air which will facilitate speech; see pulmonic, glottalic. §7.2.

ALIENABLE POSSESSION: when the possessed does not have an inherent connection with the possessor. §1.3, §16.5.

ALLATIVE: marker indicating movement towards the referent of the noun phrase to which it is attached.

ALLOMORPH: one of several alternative forms of a morpheme. §5.2.

ALLOPHONE: one possible pronunciation of a phoneme. §7.1.

AMBITRANSITIVE: verb which can function in both a transitive and an intransitive clause; of type S = A or S = O. §3.3, §13.3.

- ANALYTIC:** language whose words generally each have a small number of grammatical components. Compl. synthetic. §5.5.
- ANAPHORA:** a pronoun or demonstrative referring to something which was explicitly stated earlier in the discourse, such as *he* in *John came in and he sat down*. §15.3.
- ANTIPASSIVE:** valency-reducing derivation which puts underlying A argument into derived S function, and places underlying O argument in a peripheral function. §3.20.
- APPLICATIVE:** valency-increasing derivation which prototypically operates on an intransitive clause, putting underlying S argument into A function and introducing a new O argument (which may have been in peripheral function in the underlying clause). §3.20.
- ARCHIPHONEME:** unit resulting from the neutralization of a phonological contrast in a certain environment. §7.2.
- ARGUMENT, CORE:** an obligatory argument for a specific verb, which must be either stated or understood from the context. §3.2, §3.9, §5.6, §13.2.
- ARGUMENT, PERIPHERAL:** non-core argument, which is optional; typically include instrument, accompaniment, recipient, beneficiary, time, place, manner. §3.9, §5.6.
- ARTICLE:** a type of determiner, whose prototypical role is to mark an NP as definite or indefinite. The label is used in special ways for particular languages; for instance the tradition in Fijian linguistics is to use ‘article’ for the first word of an NP, which is *a* or *na* if the NP head is a common noun and *o* if the head is a proper name or pronoun. §3.4, §3.18.
- ARTICULATORS:** an active articulator (for example, tongue tip) is brought into contact with—or into approximation with—a passive articulator (for example, the teeth). §7.2.
- ASPECT:** term used for composition (perfective/imperfective), sometimes also for boundedness, completion, etc. §3.15.
- ASSIMILATION:** a process by which one sound changes to become more similar to a neighbouring sound, for example *-nb-* becoming *-mb-*.
- ATELIC:** an event which is unbounded and has no definite end-point. Compl. telic. §3.15.
- AUGMENTED:** pronoun paradigm in which one or more further participants are added to each term in a minimal paradigm. Compl. minimal. §3.7, §15.1.2.
- AUXILIARY:** a grammatical form (sometimes called an auxiliary verb) which occurs together with a lexical verb. It typically inflects for some non-spatial setting categories, in place of the verb inflecting for these categories.
- AVERSIVE:** case which is added to a noun or pronoun referring to something for fear of which the action described by the verb of the clause takes place or should take place. For example, ‘Come away from the fire for fear of the flying sparks.’
- BENEFICIARY:** peripheral argument referring to someone who will benefit from an action, as in *John wrote the letter [for Mary]*_{BENEFICIARY}.
- BOUNDEDNESS (or telicity):** grammatical category indicating whether or not an activity has a definite end-point; see telic, atelic. §3.15.
- BOUND FORM:** form which cannot occur alone but must be attached to some other form, e.g. *un-* in English. Compl. free form. §5.2.

- CASE:** a system of nominal inflections, marking the syntactic function of an NP in its clause. §1.5, §1.10, §13.2.
- CATAPHORA:** a pronoun or demonstrative referring to something which is explicitly stated earlier in the discourse, such as *he* in *After he stopped smoking, John lived to a ripe old age*. §15.3.
- CAUSAL:** peripheral argument whose referent is responsible for a state or activity, as in *John is sick [from eating rotten meat]_{CAUSAL}*.
- CAUSATIVE:** valency-increasing derivation which prototypically operates on an intransitive clause, putting underlying S argument into O function and introducing a 'causer' as A argument. §3.20.
- CIRCUMFIX:** a type of affix made up of one part which precedes the root or stem (like a prefix) and one part which follows (like a suffix). §5.2.
- CLASSIFIERS:** a set of (free or bound) forms which serve to categorize most of the nouns of a language, typically in terms of shape, composition, arrangement, or function/use. §3.16.
- CLAUSE:** the description of some activity, state or property. Consists of an obligatory predicate which requires certain core arguments and may also have peripheral arguments. §3.2.
- CLITIC:** a surface element part-way between a word and an affix in its properties. It is typically a separate grammatical word which is attached to a contiguous phonological word. §5.4, §10.5.
- COGNATES:** forms which are historically related; that is, go back to a single original form.
- COMITATIVE:** an affix (generally derivational, sometimes inflectional) added to a form with reference X, giving the meaning 'with (accompanied by) an X' or 'having an X'. Compl. privative.
- COMMON ARGUMENT:** an argument shared, in their underlying structures, by main clause and relative clause within a relative clause construction. Chapter 17.
- COMPARATIVE CONSTRUCTION:** typically involves comparing two participants (the comparee and the standard) in terms of some property (the parameter) this being marked by an index. §3.23.
- COMPLEMENTARY DISTRIBUTION:** the occurrence of each of two or more items (sounds or forms) in mutually exclusive environments.
- COMPLEMENT CLAUSE:** clause which fills a (normally core) argument slot in a higher clause. §1.9, §3.10, Chapter 18.
- COMPLEMENTIZER:** grammatical form which marks a complement clause. Chapter 18.
- COMPLEMENT-TAKING VERB:** a verb which may have a complement clause filling one of its (generally, core) argument slots. Chapter 18.
- COMPLETION:** grammatical category covering perfect and imperfect. §3.15.
- COMPOSITION:** grammatical category covering perfective and imperfective. §3.15.
- COMPOUNDING:** morphological process which joins two roots to form one stem. §3.13.
- CONCORD:** when two words (for example, noun and modifying adjective within an NP) are marked for the same grammatical category. §5.6.
- CONJUGATION:** a class of verbs all of which take the same inflectional allomorphs.

- CONJUNCT:** grammatical element showing that the subject is 1st person in a statement and 2nd person in a question. Compl. disjunct. §15.1.10.
- CONSANGUINEAL:** kinship relation which does not involve marriage but is entirely through descent (a 'blood relation'). §1.3, §16.1.
- CONSTITUENT:** anything which fills a slot in a syntactic structure. §5.6.
- CONSTITUENT ORDER:** the order in which phrasal constituents occur within a clause (often mis-termed 'word order'). §2.4, §5.6.
- CONSTRUCTION:** type of clause (or, sometimes, phrase) with specified properties. §5.6.
- CONTENT INTERROGATIVE:** question which enquires concerning a core or peripheral argument (including time, place, and manner), or predicate, or some action or state or property. A word defining such a question. §3.7.
- CONTINUOUS:** see durative.
- COPULA CLAUSE:** indicating a relational meaning between CS (copula subject) and CC (copula complement) functions. §3.2, Chapter 14.
- COPULA COMPLEMENT (CC):** the argument in a copula clause which is shown to be in a specified relation to the copula subject (typically, may be realized as a plain NP, an NP marked with a preposition, a possessive clause, an adjective, or a complement clause). Chapter 14.
- COPULA SUBJECT (CS):** that argument in a copula clause which is topic for the discourse in which it occurs (generally realized by an NP or a complement clause). Chapter 14.
- CORE ARGUMENT:** an obligatory argument for a specific verb, which must be either stated or understood from the context. §3.2, §3.9, §5.6, §13.2.
- COVERB:** word (generally non-inflecting) which may be combined with an inflecting verb to form a complex verbal lexeme. §1.11.
- DATIVE:** a case which typically marks the beneficiary of 'give', the addressee of 'tell', and the person to whom something is shown for 'show'.
- DECLARATIVE:** choice from a mood system used in a statement. §3.2.
- DEICTIC REFERENCE:** pointing to some participant, activity, or place within the context of speaking. §15.2.
- DEMONSTRATIVE:** grammatical element whose primary function is to point to an object in the situation of discourse; may also have anaphoric and/or cataphoric functions. §3.7, §§15.2–3.
- DERIVATION:** optional morphological process which applies to a root or stem and derives a stem; may or may not change word class. §3.13–14, §5.3.
- DETERMINER:** grammatical modifier within an NP, typically including demonstratives and articles.
- DIPHTHONG:** vowel phoneme which has two or more phonetic components. §4.9.
- DIRECT SPEECH:** verbatim quotation of what was said.
- DISJUNCT:** grammatical element showing that the subject is not 1st person in a statement and not 2nd person in a question. Compl. conjunct. §15.1.10.
- DISSIMILATION:** change by which one sound becomes more dissimilar to some neighbouring sound.
- DURATIVE** (also called continuous or progressive): an event seen as unfolding over a period of time. Compl. punctual. §3.15.

- ENCLITIC: clitic which is attached to the end of a word. §5.4, §10.5.
- ERGATIVE: case inflection marking transitive subject (A). Compl. absolutive. §3.9, §13.2, §13.5.4.
- EVIDENTIALITY: grammatical system providing information about the evidence on which a report is based. §1.5, §3.15.
- EXCLUSIVE: non-singular 1st person pronoun, referring to speaker and one or more other people who do not include the addressee. Compl. inclusive. §15.1.2.
- EXTENDED INTRANSITIVE: clause type with two core arguments, in S (intransitive subject) and E (extension to core) functions. Verb which occurs in the predicate of such a clause. §3.2, §13.1.
- EXTENDED TRANSITIVE (or ditransitive): clause type with three core arguments, in A (transitive subject), O (transitive object), and E (extension to core) functions. Verb which occurs in the predicate of such a clause. §3.2, §13.1.
- EXTENT: grammatical category covering punctual and durative. §3.15.
- FLUID-S: system where some verbs may have their S argument marked like A (Sa) or like O (So) with a difference in meaning. §3.9, §13.2, §13.5.4.
- FOCAL CLAUSE: that clause in a linking construction which carries the mood of the sentence. §3.11.
- FOCUS: an argument accorded prominence within a clause. §3.21.
- FORMAL MARKEDNESS: if a term in a grammatical system has zero realization (or a zero allomorph) it is said to be formally unmarked. Other terms in the system are formally marked. §5.7.
- FREE FORM: a form which constitutes a grammatical word without any morphological processes having to be applied. §5.2.
- FUNCTIONAL LOAD of a contrast: the extent to which that contrast is utilized within that language.
- FUNCTIONAL MARKEDNESS: a term in a grammatical system which is employed in neutral or unspecified circumstances (or when a contrast is neutralized) is said to be functionally unmarked. Other terms in the system are functionally marked. §5.7.
- FUSIONAL: a type of language whose words involve a number of grammatical elements fused together (that is, not segmentable in surface structure). §5.5.
- GENDER: small system of noun classes one of whose semantic distinctions is masculine/feminine. §1.5, §1.10, §3.16.
- GENITIVE: marker of an intra-NP possessive relation, which is added to the possessor item. Compl. pertainive. §1.10, §16.2.
- GLOTTALIC AIRSTREAM MECHANISM: air movement initiated at the glottis. §7.2.
- GRAMMATICAL WORD: a unit on the hierarchy of grammatical units (just below phrase) defined on grammatical criteria. Generally (but not necessarily always) coinciding with phonological word. Chapter 10.
- HEAD: obligatory nucleus of a phrase which determines the grammatical profile of the whole phrase (for example, gender of a noun phrase). §3.4, §5.6, §16.8, §17.2.
- HETERORGANIC: sequence of sounds which have different place of articulation, for example *-nb-*.

HOMORGANIC: sequence of sounds which have the same place of articulation, for example *-mb-*.

IDEOPHONE: word class which often has special phonology (often involving inherent reduplication and onomatopoeia). Typically relating to manner, colour, sound, smell, action, state, or intensity. §8.3.

IMPERATIVE: choice from a mood system used in a direct command. §1.5, §3.2.

IMPERFECT: something which began in the past and is still continuing. Compl. perfect. §3.15.

IMPERFECTIVE: focusing on the temporal make-up of an event. Compl. perfective. §3.15.

INALIENABLE POSSESSION: when the possessed has an inherent connection with the possessor, and cannot be given away. §1.3, §16.5.

INCLUSIVE: non-singular 1st person pronoun, referring to speaker and one or more other people who do include the addressee. Compl. exclusive. §15.1.2.

INDIRECT SPEECH: a report of what someone else has said (often cast into the reporter's own words).

INFLECTION: morphological process which obligatorily applies to a root or derived stem of a certain word class, producing a grammatical word. §3.13, §5.3.

INSTRUMENTAL: case inflection marking the referent of the NP to which it is attached as weapon, tool, or material used in the activity described by the verb. §4.3, §13.2.1.

INTERJECTION: a conventionalized cry, typically indicating the speaker's emotional response to something that has happened to them, or something which they have observed or become aware of. §10.7.

INTERNAL CHANGE: morphological process which involves changing a vowel (or, less frequently, a consonant) in the middle of a word, for instance, from *take* /teik/ to *took* /tuk/ in English. §3.13.

INTERROGATIVE: choice from a mood system used in a (content or polar) question. A content interrogative word. §3.2, §3.7.

INTONATION: type of prosody realized by pitch, generally applying over clause or sentence. §7.6.

INTRANSITIVE: clause type with one core argument, in S (intransitive subject) function. Verb which occurs in the predicate of such a clause. §3.2, §5.6, Chapter 13.

IRREALIS: referring to something that didn't happen (but could have happened) or which might happen. Compl. realis. §3.15.

ISOLATING: a type of language most of whose words consist of one morpheme. §5.5.

LABILE: older name for ambitransitive.

LANGUAGE: in the technical sense of linguists, a number of forms of speech are said to constitute a single language if they are mutually intelligible.

LENITION: the replacement of a sound by another sound that has the same place of articulation but a weaker manner of articulation (involving less muscular tension).

LEXEME (OR LEXICAL ITEM): a root or underlying form. §10.2.

LOCATIVE: marker indicating position of rest at, on, or near the referent of the noun phrase to which it is attached.

- LOGOPHORIC PRONOUN:** used in a complement clause, this refers back to the subject of the matrix clause. §15.3.4.
- MARKEDNESS:** see formal markedness, functional markedness. §5.7.
- MINIMAL:** pronoun paradigm in which ‘me and you’ is a term on a par with 1st person singular and 2nd person singular (and, in some languages, 3rd person singular). Compl. augmented. §3.7, §15.1.2.
- MODALITY:** one of a number of choices (within irrealis) referring to some aspect of the future. §3.15.
- MODAL VERB:** a verb which indicates a modality.
- MOOD:** grammatical system indicating the pragmatic function of a sentence, covering indicative (for a statement), interrogative (for a question), and imperative (for a command). §3.2.
- MORA:** unit between phoneme and syllable, variously defined. §7.6.
- MORPHEME:** the minimum meaningful unit of speech. §5.2.
- MORPHOLOGICAL PROCESS:** process which applies to a root, forming a stem. §3.13.
- MORPHOLOGY:** that part of grammar which studies the structure of words. Compl. syntax. §3.13, §5.2.
- NEUTRALIZATION:** when a certain grammatical or phonological contrast may not apply in a certain environment, it is then said to be neutralized. §5.7, §7.2, §15.1.3.
- NOMINAL HIERARCHY:** hierarchy of items which can be head of an NP, according to how likely they are to be in A rather than in O function. §3.9; §13.5.4.
- NOMINALIZATION:** morphological derivation which forms a noun stem from a verb or adjective root or stem. §3.14
- NOMINATIVE:** case inflection marking intransitive subject (S) and transitive subject (A). Compl. accusative. §3.9, §13.2, §13.5.4.
- NON-CANONICAL MARKING OF CORE ARGUMENTS:** when most of the instances of a core argument receive a certain marking, but there are a minority of instances which attract a different marking, this is termed non-canonical. §13.6.
- NON-SPATIAL SETTING:** covers the range of parameters which describe the setting for an activity or state other than those referring to spatial location. It typically includes evidentiality, reality, degree of certainty, phase of activity, completion, boundedness, extent, composition (some of the last three, and more besides, may be called aspect), and tense. §3.15.
- NOUN:** word class whose primary function is as head of an NP; many of its members refer to concrete objects. §3.3, §8.3.1, Chapter 11.
- NOUN CLASSES:** grouping of all the nouns of a language into a number of small classes which comprise a small closed grammatical system. Noun class membership must be marked somewhere outside the noun itself. Also see gender. §1.9, §3.16.
- NOUN INCORPORATION:** the incorporation of a noun (generally in underlying S or O function) into a verb to create a compound stem.
- NOUN PHRASE (NP):** a constituent which can fill an argument slot in clause structure. It has a noun or pronoun or demonstrative, etc. as head. §3.4, §5.6, §11.4.
- NP:** see noun phrase.

- NUMBER:** grammatical system one of whose terms is singular. There will be one or more further terms. §1.4; §13.7.
- PASSIVE:** valency-reducing syntactic derivation which puts underlying O argument into derived S function and places underlying A argument in a peripheral function. §3.20.
- PAUCAL NUMBER:** referring to 'a few' (more than two).
- PERFECT:** a past action which is completed but still has present relevance. Compl. imperfect; §3.15.
- PERFECTIVE:** an event regarded as a whole, without regard for its temporal constituency. Compl. imperfective; §3.15.
- PERIPHERAL ARGUMENT:** a non-core argument, which is optional. Typically includes instrument, accompaniment, recipient, beneficiary, time, place, manner. §3.9, §5.6.
- PERIPHERAL PLACE OF ARTICULATION:** cover term for sounds made at the front or back of the mouth, covering bilabial and dorso-velar.
- PERSON:** speech act participants; always including 1st person (speaker) and 2nd person (addressee), and sometimes also 3rd person (neither speaker nor addressee). §15.1.1.
- PERTENSIVE:** marker of an intra-NP possessive relation, which is added to the possessed item. Compl. genitive. §16.2.
- PHASE OF ACTIVITY:** whether beginning, continuing, ending, etc. §3.15.
- PHONEME:** the minimum segmentable unit of phonology. §7.1.
- PHONETICS:** articulatory and/or acoustic study of the sounds of speech.
- PHONOLOGICAL WORD:** a unit on the hierarchy of phonological units (just above syllable) defined on phonological criteria. Generally (but not necessarily always) coinciding with grammatical word. Chapter 10.
- PHONOLOGY:** description of the phonetic contrasts which are used to distinguish between distinct words in a given language. Chapter 7.
- PHONOTACTICS:** statement of which consonants and vowels may correspond to each structural slot in syllable (and word) structure. §7.4.
- PHRASE:** a constituent which can fill a slot in clause structure—noun phrase in an argument slot and verb phrase in predicate slot. §3.4.
- PIVOT:** a topic which is recognizable as such by its grammatical properties. §3.21.
- POLAR QUESTION:** question enquiring whether or not a proffered statement is correct. Can be answered by 'yes' or 'no' in languages which have such words (not all do). §3.2.
- POLARITY:** grammatical system whose terms are positive and negative. §3.12.
- POLYSYNTHETIC:** highly synthetic. §5.5.
- POSSESSIVE PHRASE:** a type of NP which is included within a larger NP and indicates the possessor with respect to the head of the larger NP, which is the possessed. §3.4, Chapter 16.
- POSTPOSITION:** an adposition which follows the constituent for which it provides grammatical marking. §5.4.
- PRAGMATICS:** the practical consequences of the use of a given portion of language.
- PREDICATE:** the central (and obligatory) structural element of a clause, generally realized by a verb phrase (with verb as head). It determines the number and type of core arguments required in the clause. §2.5, §3.2, §11.5.

PREFIX: an affix which precedes a root or stem.

PREPOSITION: an adposition which precedes the constituent for which it provides grammatical marking. §5.4.

PRIMARY VERBS: referring directly to an activity or state. Compl. secondary verbs. §1.11, §18.5.

PRIVATIVE: an affix (generally derivational, sometimes inflectional) added to form with referent X, giving the meaning ‘without an X’. Compl. comitative.

PROCLITIC: clitic which is attached to the beginning of a word: §5.4, §10.5.

PROGRESSIVE: see durative.

PRONOUN: small closed class of grammatical items which relate to person (and usually also to number). Can be free forms or bound forms. §3.7; §15.1.

PROSODY: a system of phonological contrasts which has scope over a sequence of segments. §7.5.

PROTO-LANGUAGE: putative single ancestor language for a group of modern languages that are held to be genetically related, each having developed by regular changes from the proto-language.

PULMONIC AIRSTREAM MECHANISM: air movement initiated in the lungs. §7.2.

PUNCTUAL: an event which happens more or less instantaneously. Compl. durative. §3.15.

REALIS: referring to something that has happened or is happening. Compl. irrealis. §3.15.

REALITY: grammatical category covering realis and irrealis. §3.15.

RECIPROCAL: clause describing several instances of an activity such that what is A argument in one instance is O argument in another. §3.22.

REDUPLICATION: morphological process which involves repeating all or part of a root (or stem or full word) either before, after, or in the middle of it. §3.13.

REFLEXIVE: clause in which underlying A and O arguments have the same reference. §3.22.

RELATIVE CLAUSE: clause which modifies the head of an NP. Relative clause and main clause share, in their underlying structures, a common argument. Chapter 17.

ROOT: unanalysable lexical element.

S = A AMBITRANSITIVE: the S argument, when the verb is used intransitively, corresponds to the A argument, when it is used transitively. §3.3, §13.3.

S = O AMBITRANSITIVE: the S argument, when the verb is used intransitively, corresponds to the O argument, when it is used transitively. §3.3, §13.3.

SECONDARY CONCEPTS: provide modification for a primary verb. May be realized as an affix or as a verb (a secondary verb). Compl. primary verb. §1.11; §18.5.

SEMANTIC ROLES: the types of participant involved with verbs of a certain semantic type. §1.9, §3.3, §13.5.1.

SEMANTICS: study of the meaning relations conveyed by the grammatical systems and lexical contrasts of a language.

SEMANTIC TYPE: a set of words with similar meanings and grammatical properties. §1.9, §1.11, §3.3, §8.3, §12.4, §13.5.12, §18.5.

SENTENCE: no simple definition is feasible—see §3.11.

SERIAL VERB CONSTRUCTION: has a predicate consisting of two (or more) verbs, each of which could make up a predicate on its own, and whose combination is conceived of as describing a single action; there must be a single subject applying to the whole. §18.6.1.

SHIFTER: grammatical item whose reference changes depending on who is speaking (pronouns) or what the place or time is. §3.7.

SPLIT-S: system where the S argument for some verbs is marked like A (Sa) and for other verbs S is marked like O (So); also called active/stative. §3.9, §13.2, §13.5.4.

STATIVE/ACTIVE: label covering split-S and fluid-S systems.

STEM: the nucleus of a word, to which an inflectional process applies, forming a word.

STRESS (or accent): a contrastive prosody generally having scope over a word, characterized by some or all of: loudness, vowel quality, pitch, and length. §7.6.

SUBGROUP: set of languages within a language family which descend from a single ancestor language, this being itself a descendant of the proto-language for the whole language family.

SUBTRACTION: morphological process which involves deleting something from a root. §3.13.

SUFFIX: an affix which follows a root or stem.

SUPPLETION: when a lexeme has two forms which are not cognate (as *go* and *went* in English).

SUPPORTING CLAUSE: that clause in a linking construction which does not carry the mood of the sentence. §3.11.

SYLLABLE: a phonological unit centred on a nucleus (typically a vowel) which may be preceded and/or followed by one or more consonants. §1.4, §6.3, §7.4.

SYNCHRONIC DESCRIPTION: description of a language system at one point in time, without taking account of historical changes.

SYNTAX: study of the organization and interrelation of the components of a grammar above the level of word.

SYNTHETIC: language whose words generally each have a large number of grammatical components. Compl. analytic. §5.5.

TELIC: an event which is bounded and has a definite end-point. Compl. atelic. §3.15.

TENSE: grammatical category, with shifting reference, which refers to time. §1.5, §1.7, §1.10, §3.15.

TOPIC: an argument which occurs in a succession of clauses in a discourse and binds them together. §3.21.

TRANSITIVE: clause type with two core arguments, in A (transitive subject) and O (transitive object) functions. Verb which occurs in the predicate of such a clause. §3.2, §5.6, Chapter 13.

TRIPARTITE MARKING: when each of transitive subject (A), intransitive subject (S), and transitive object (O) receives a distinct surface marking. §3.9, §13.2.

UNMARKED: see formal markedness, functional markedness.

VALENCY: the number of core arguments a verb requires.

- VALENCY-CHANGING:** derivations which may increase valency (causative, applicative) or decrease it (passive, antipassive, some varieties of reflexive and reciprocal, etc.). §3.20.
- VERB:** word class whose primary function is as head of a predicate. Most of its members refer to actions and states. §3.3, §8.3.3, Chapter 11.
- VERBALIZATION:** morphological derivation which forms a verb stem from a noun or adjective root or stem. §3.14.
- VERBLESS CLAUSE:** similar to a copula clause but with the predicate slot left blank. It indicates a relational meaning between verbless clause subject and verbless clause complement. Chapter 14.
- VERBLESS CLAUSE COMPLEMENT (VCC):** the argument in a verbless clause which is shown to be in a specified relation to the verbless clause subject (typically, may be realized as a plain NP, an NP marked with a preposition, a possessive clause, an adjective, or a complement clause). Chapter 14.
- VERBLESS CLAUSE SUBJECT (VCS):** that argument in a verbless clause which is topic for the discourse in which it occurs (generally realized by an NP or a complement clause). Chapter 14.
- VERB PHRASE:** a constituent which can fill the predicate slot within a clause. Typically has a verb as its head. §3.4, §5.6.
- VOWEL HARMONY:** prosody applying over a phonological stretch (typically, a phonological word) whereby all vowels within that stretch agrees in some feature, e.g. front/back.
- WORD:** the result of applying optional derivational processes to a root, and then any obligatory inflectional process to the resulting stem. Subtypes: phonological word, grammatical word. Unit at the intersection of morphology and syntax. §3.1, Chapter 10.
- WORD ORDER:** the order in which words must or may occur in a phrase, in a clause, or in a sentence. (This label is often misleadingly used for (phrasal) constituent order.) §2.4, §5.6.
- YES/NO QUESTION:** see polar question.
- ZERO:** when one term in a grammatical system has no explicit marking it is said to have zero realization (\emptyset). For example, in English a noun with singular number reference receives zero marking (for instance *horse*- \emptyset) whereas one with plural reference is marked by orthographic *-s* (*horse-s*). §3.13, §5.3.
- ZERO ANAPHORA:** when anaphora is shown simply by leaving a gap. Compare anaphoric *he* in *John came in and he sat down* with anaphoric \emptyset in *John came in and \emptyset sat down*. §15.3.
- ZERO DERIVATION:** a word-class-changing derivation with zero marking. Compare noun *hospital* and verbalization *hospital-ize*, marked by *-ize*, with noun *market* and verbalization *market*- \emptyset , with zero marking. §3.5, §3.13, §11.3.

References

- Abbott, Miriam. 1991. 'Macushi', pp. 23–160 of Derbyshire and Pullum 1991.
- Abraham, R. C. 1959. *Language of the Hausa people*. London: University of London Press.
- Adelaar, Willem. 1977. *Tarma Quechua: Grammar, texts, dictionary*. Lisse: The Peter de Ridder Press.
- 2002. 'Copula clauses and verbless clauses in Quechua and in Aymara'. Handout for presentation at RCLT Local Workshop on 'Copula clauses and verbless clauses'.
- 2004. *The languages of the Andes*. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra Y. 1995. 'Split ergativity in Berber languages', *St Petersburg Journal of African Studies* 4: 39–68.
- 1996. 'Words, phrases, pauses and boundaries: Evidence from South American Indian languages', *Studies in Language* 20: 487–517.
- 1998. 'Warekena', pp. 225–439 of Derbyshire and Pullum 1998.
- 1999a. 'Multiple marking of syntactic function and polysynthetic nouns in Tariana', *Papers from the Regional Meeting of the Chicago Linguistic Society* 35, 2: 235–48.
- 1999b. 'The Arawak language family', pp. 65–106 of Dixon and Aikhenvald 1999.
- 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.
- 2002a. *Language contact in Amazonia*. Oxford: Oxford University Press.
- 2002b. 'Typological parameters for the study of clitics, with special reference to Tariana', pp. 42–78 of Dixon and Aikhenvald 2002a.
- 2003. *A grammar of Tariana, from northwest Amazonia*. Cambridge: Cambridge University Press.
- 2004a. 'The adjective class in Tariana', pp. 97–124 of Dixon and Aikhenvald 2004.
- 2004b. *Evidentiality*. Oxford: Oxford University Press.
- 2006. 'Complement clause types and complementation strategies in Tariana', pp. 178–203 of Dixon and Aikhenvald 2006.
- 2008a. *The Manambu language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- 2008b. 'Versatile cases'. *Journal of Linguistics* 44: 565–603.
- and Dixon, R. M. W. 2006. Editors of *Serial verb constructions: A cross-linguistic typology* (*Explorations in linguistic typology*, Vol. 3). Oxford: Oxford University Press.
- Forthcoming. 'Explaining associations between intransitive subject (S) and transitive object (O)'. To appear in *Word*.
- and Onishi, Masayuki. 2001. Editors of *Non-canonical marking of subjects and objects*. Amsterdam: John Benjamins.
- and Green, Diana. 1998. 'Palikur and the typology of classifiers', *Anthropological Linguistics* 40: 429–80.

- Akmajian, Adrian, and Anderson, Stephen. 1970. 'On the use of the fourth person in Navajo, or Navajo made harder', *International Journal of American Linguistics* 36: 1–8.
- Allen, W. S. 1956. 'Structure and system in the Abaza verbal complex', *Transactions of the Philological Society for 1956*: 127–76.
- 1957. 'Aspiration in the Hārauti nominal', pp. 68–86 of *Studies in linguistic analysis*, by J. R. Firth et al. Oxford: Blackwell.
- 1964. 'Transitivity and possession', *Language* 40: 337–43.
- 1972. *Sandhi: The theoretical, phonetic and historical basis of word-junction in Sanskrit*. The Hague: Mouton.
- Alpher, Barry. 1991. *Yir-Yoront lexicon: Sketch and dictionary of an Australian language*. Berlin: Mouton de Gruyter.
- Alvarez, José. 1994. *Estudios de lingüística Guajira*. Maracaibo: República de Venezuela, Ministerio de la Cultura.
- Amberber, Mengistu. 1996. 'Amharic'. RCLT internal document, ANU.
- Ameka, Felix K. 1992a. Editor of special issue on 'Interjections', *Journal of Pragmatics* 18: 101–301.
- 1992b. 'Interjections: The universal yet neglected part of speech', pp. 101–18 of Ameka 1992a.
- 1994. 'Interjections', pp. 1712–15 of *The international encyclopedia of language and linguistics*, edited by R. E. Asher and J. M. Y. Simpson. Oxford: Pergamon Press.
- 1996. 'Body parts in Ewe grammar', pp. 780–840 of Chappell and McGregor 1996.
- 2004. 'Grammar and cultural practices: The grammaticalization of triadic communication in West African languages', *Journal of West African Languages* 30: 5–28.
- Andersen, Torben. 1988. 'Ergativity in Pāri, a Nilotic OVS language', *Lingua* 75: 289–324.
- Anderson, Stephen. 1976. 'On the notion of subject in ergative languages', pp. 1–23 of *Subject and topic*, edited by Charles N. Li. New York: Academic Press.
- Anderson, Stephen R., and Keenan, Edward L. 1985. 'Deixis', pp. 259–308 of Shopen 1985: Vol. 3.
- Andrade, Manuel J. 1933. 'Quileute', pp. 149–292 of Boas 1933.
- Andrews, Avery D. 2007. 'Relative clauses', pp. 206–36 of Shopen 2007: Vol. 2.
- Anschutz, A. 1997. 'How to choose a possessive noun phrase construction in four easy steps', *Studies in Language* 21: 1–35.
- Arms, David G. 1974. 'Transitivity in Standard Fijian'. PhD dissertation, University of Michigan.
- Arnott, D. W. 1970. *The nominal and verbal systems of Fula*. Oxford: Clarendon Press.
- Asher, R. E. 1968. 'Existential, possessive, locative and copulative sentences in Malayalam', pp. 88–111 of Verhaar, Part 2, 1968.
- 1985. *Tamil*. London: Croom Helm.
- and Kumari, T. C. 1997. *Malayalam*. London: Routledge.
- Aspinion, Robert. 1953. *Apprenons le berbère: Initiation aux dialectes chleuhs*. Paris: Éditions Félix Moncho.
- Austin, Peter. 1978. 'A grammar of the Diyari language of north-east South Australia'. PhD thesis, ANU.

- 1981a. *A grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.
- 1981b. 'Case marking in southern Pilbara languages', *Australian Journal of Linguistics* 1: 211–26.
- Avrorin, V. A. 1959. *Grammatika nanajskogo jazyka*, 1. Moscow and Leningrad: AN.
- Backhouse, Anthony E. 1984. 'Have all the adjectives gone?', *Lingua* 62: 169–86.
- 2004. 'Inflected and uninflected adjectives in Japanese', pp. 50–73 of Dixon and Aikhenvald 2004.
- Baker-Shenk, Charlotte, and Cokley, Dennis. 1996. *American sign language: A teacher's resource text on grammar and culture*. Washington, DC: Gallaudet University Press.
- Ball, Martin J. 1993. Editor of *The Celtic languages*. London: Routledge.
- Bally, Charles. 1950. *Linguistique générale et linguistique française*, 3rd edition. Berne: Francke.
- Barentsen, Adrian. 1996. 'Shifting points of orientation in Modern Russian: Tense selection in "reported perception"', pp. 15–55 of *Reported speech: Forms and functions of the verb*, edited by Theo A. J. M. Janssen and Wim van der Wurff. Amsterdam: John Benjamins.
- Bargery, G. P. 1934. *A Hausa-English dictionary and English-Hausa vocabulary*. London: Oxford University Press.
- Bauer, Laurie, and Huddleston, Rodney. 2002. '19. Lexical word-formation', pp. 1621–721 of Huddleston and Pullum 2002.
- Bazell, C. E. 1953. *Linguistic form*. Istanbul: Istanbul Press.
- Beaumont, Clive H. 1980. *The Tigak language of New Ireland*. Canberra: Pacific Linguistics.
- Beck, David. 2000. 'The syntax, semantics, and typology of adjectives in Upper Necaxa Totonac', *Linguistic Typology* 1: 213–50.
- Benjamin, Geoffrey. 1976. 'An outline of Temiar grammar', pp. 129–87 of *Austroasiatic studies*, Vol. 1, edited by Phillip N. Jenner, Laurence C. Thompson, and Stanley Starosta. Honolulu: University of Hawaii Press.
- Benveniste, Emile. 1971a. *Problems of general linguistics*, translated by Mary E. Mack. Coral Gables, Fla.: University of Miami Press. [Original: *Problèmes de linguistique générale*. 1966. Paris: Éditions Gallimard.]
- 1971b. 'The linguistic functions of "to be" and "to have"', pp. 163–79 of Benveniste 1971a.
- Besnier, Niko. 2000. *Tuvaluan: A Polynesian language of the Central Pacific*. London: Routledge.
- Bhat, D. N. S. 1994. *The adjectival category: Criteria for differentiation and identification*. Amsterdam: John Benjamins.
- 2004. *Pronouns*. Oxford: Oxford University Press.
- Bhatia, Tej K. 1993. *Punjabi: A cognitive-descriptive grammar*. London: Routledge.
- Bickel, Balthasar. 1999. 'Nominalization and focus constructions in some Kiranti languages', pp. 271–96 of *Topics in Nepalese linguistics*, edited by Yogendra P. Yadava and Warren G. Glover. Kathmandu: Royal Nepal Academy.

- Bing, Janet M. 1991. 'Color terms and lexical classes in Krahn/Wobé', *Studies in African Linguistics* 22: 277–96.
- Blackwell, Aleka A. 2000. 'On the acquisition of the syntax of English adjectives', *Chicago Linguistic Society Proceedings* 36, 2: 361–75.
- Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. Urbana: University of Illinois Press.
- 1933. *Language*. New York: Holt, Rinehart and Winston.
- Boas, Franz. 1911a. Editor of *Handbook of American Indian languages*, Part 1 (Smithsonian Institution, Bureau of American Ethnology, Bulletin 40). Washington, DC: Government Printing Office.
- 1911b. 'Introduction', pp. 1–83 of Boas 1911a.
- 1911c. 'Tsimshian', pp. 283–422 of Boas 1911a.
- 1922. Editor of *Handbook of American Indian languages*, Part 2 (Smithsonian Institution, Bureau of American Ethnology, Bulletin 40). Washington, DC: Government Printing Office.
- 1933. Editor of *Handbook of American Indian languages*, Part 3. New York: Columbia University Press.
- and Deloria, Ella. 1941. *Dakota grammar*. Washington, DC: US Government Printing Office.
- and Swanton, John R. 1911. 'Siouan (Dakota)', pp. 875–965 of Boas 1911a.
- Bochner, H. 1984. 'Inflection within derivation', *Linguistic Review* 3: 411–21.
- Bolinger, Dwight. 1967. 'Adjectives in English: Attribution and predication', *Lingua* 18: 1–34.
- Bontkes, Carolyn. 1985. 'Subordinate clauses in Surui', pp. 189–207 of *Porto Velho workpapers*, edited by David L. Fortune. Brasília: Summer Institute of Linguistics.
- Borer, Hagit. 1984. 'Restrictive relatives in Modern Hebrew', *Natural Language and Linguistic Theory* 2: 219–60.
- Borgman, Donald M. 1990. 'Sanuma', pp. 15–248 of Derbyshire and Pullum 1990.
- Boxwell, Maurice. 1967. 'Weri pronoun system', *Linguistics* 29: 34–43.
- Breen, J. G. 1976. 'Gugadj and Gog-Nar: A contrastive sketch', pp. 151–62 of *Languages of Cape York*, edited by Peter Sutton. Canberra: Australian Institute of Aboriginal Studies.
- Bright, William. 1957. *The Karok language*. Berkeley and Los Angeles: University of California Press.
- Broschart, Jürgen. 1997. 'Why Tongan does it differently: Categorical distinctions in a language without nouns and verbs', *Linguistic Typology* 1: 143–65.
- Browne, W. 1973. 'Relative pronoun = relative + pronoun? Evidence from Serbo-Croatian'. Paper presented at Department of Linguistics, University of Cambridge.
- Bruce, Les. 1984. *The Alamlak language of Papua New Guinea (East Sepik)*. Canberra: Pacific Linguistics.
- Bugenhagen, Robert D. 1994. 'The exponents of semantic primitives in Mangap-Mbula', pp. 87–108 of *Semantic and linguistic universals: Theory and empirical findings*, edited by Cliff Goddard and Anna Wierzbicka. Amsterdam: John Benjamins.

- Burridge, Kate. 2006. 'Complement clause types in Pennsylvania German', pp. 49–71 of Dixon and Aikhenvald 2006.
- Buse, J. E. 1965. 'Problems of morphology and classification illustrated from Rarotongan', *Lingua* 15: 32–47.
- Butt, John, and Benjamin, Carmen. 2004. *A new reference grammar of modern Spanish*, 4th edition. London: Hodder Arnold.
- Bužarovska, Eleni. 2002. 'The purpose-modification continuum: Purposive *da*-relative clauses in Macedonian', *Southwest Journal of Linguistics* 21: 67–97.
- Bybee, Joan, Perkins, Revere, and Pagliuca, William. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Carlin, Eithne B. 2004. *A grammar of Trio, a Cariban language of Suriname*. Frankfurt am Main: Peter Lang.
- Carlson, Robert. 1994. *A grammar of Supyire*. Berlin: Mouton de Gruyter.
- Carpenter, Kathie. 1996. 'Subordination by reduplication in Wetan', *Studies in Language* 20: 37–51.
- Ceria, Verónica G., and Sândalo, Filomena. 1995. 'A preliminary reconstruction of Proto-Waikurúan with special reference to pronominals and demonstratives', *Anthropological Linguistics* 37: 169–91.
- Černý, Václav. 1971. 'Some remarks on syntax and morphology of verb in Avar', *Archiv Orientální* 39: 46–56.
- Chafe, Wallace. 1990. 'Uses of the defocusing pronominal prefixes in Caddo', *Anthropological Linguistics* 32: 57–68.
- Chao, Yuen-Ren. 1946. 'The logical structure of Chinese words', *Language* 22: 4–13.
- 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chappell, Hilary. 2001. 'Mandarin demonstratives'. Handout for RCLT Local Workshop on 'Demonstratives and related items'.
- and McGregor, William. 1996. Editors of *The grammar of inalienability: A typological perspective on body part terms and the part-whole relation*. Berlin: Mouton de Gruyter.
- Charters, Helen. 1996. 'Relative clauses in Mandarin'. Handout for a presentation in the Workshop on 'Relative clauses', ANU.
- Churchward, C. Maxwell. 1941. *A new Fijian grammar*. Sydney: Australasian Medical Publishing Company.
- 1953. *Tongan grammar*. London: Oxford University Press.
- Clark, Eve V., and Clark, Herbert H. 1979. 'When nouns surface as verbs', *Language* 55: 767–811.
- Clements, George N. 1975. 'The logophoric pronoun in Ewe: Its role in discourse', *Journal of West African Languages* 10: 141–77.
- Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North-Holland.
- 1987. 'The structure of internally headed relative clauses', *Natural Language and Linguistic Theory* 5: 277–302.

- Comrie, Bernard. 1981. *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- 1989. *Language universals and linguistic typology, syntax and morphology*, 2nd edition. Oxford: Basil Blackwell.
- 1997. 'The typology of predicate case marking', pp. 39–50 of *Essays on language function and language type*, edited by Joan Bybee, John Haiman, and Sandra A. Thompson. Amsterdam: John Benjamins.
- 2003. 'Recipient person suppletion in the verb "give"', pp. 265–81 of *Language and life: Essays in memory of Kenneth L. Pike*, edited by Mary Ruth Wise, Thomas N. Headland, and Ruth M. Brend. Dallas: SIL International and University of Texas at Arlington.
- and Smith, Norval. 1977. 'Lingua Descriptive Studies: Questionnaire', *Lingua* 42: 1–72.
- Conklin, Harold C. 1962. 'Lexicographical treatment of folk taxonomies', pp. 119–41 of *Problems in lexicography*, edited by Fred W. Householder and Sol Saporta. Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics.
- Cook, Eung-Do. 1984. *A Sarcee grammar*. Vancouver: University of British Columbia Press.
- Cook, Kenneth W. 1978. 'The mysterious Samoan transitive suffix', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4: 53–66.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- 2004. 'The Russian adjective: A pervasive yet elusive category', pp. 199–222 of Dixon and Aikhenvald 2004.
- Corominas, Joan, and Pascual, Joan. 1980–91. *Diccionario crítica etimológico castellano e hispánico*. Madrid: Editorial Gredos.
- Coulmas, Florian. 1982. 'Some remarks on Japanese deictics', pp. 209–21 of *Here and there: Cross-linguistic studies in deixis and demonstration*, edited by Jürgen Weissenborn and Wolfgang Klein. Amsterdam: John Benjamins.
- Craig, Colette G. 1977. *The structure of Jacalteco*. Austin: University of Texas Press.
- Crapo, Richley H., and Aitken, Percy. 1986. *Bolivian Quechua reader and grammar-dictionary*. Ann Arbor: Karoma.
- Crazzolaria, J. P. 1933. *Outlines of a Nuer grammar*. Vienna: Anthropos.
- Creissels, Denis. 2005. 'A typology of subject and object markers in African languages', pp. 43–70 of *Studies in African linguistic typology*, edited by F. K. Ekhard Voeltz. Amsterdam: John Benjamins.
- Cristofaro, Sonia. 1998. 'Grammaticalization and clause linkage: A typological approach with particular reference to Ancient Greek', pp. 59–88 of *The limits of grammaticalization*, edited by Anna G. Ramat and Paul J. Hopper. Amsterdam: John Benjamins.
- Crowell, Thomas H. 1979. 'A grammar of Bororo'. PhD dissertation, Cornell University.

- Crowley, Terry. 1978. *The middle Clarence dialects of Bandjalang*. Canberra: Australian Institute of Aboriginal Studies.
- 1983. 'Uradhi', pp. 306–428 of Dixon and Blake 1983.
- Culy, Christopher. 1994. 'Aspects of logophoric marking', *Linguistics* 32: 1055–94.
- Cummins, George M., III. 1998. 'Definiteness in Czech', *Studies in language* 22: 567–96.
- Curnow, Timothy J. 1997a. 'A grammar of Awa Pit (Cuaiquer), an indigenous language of south-western Colombia'. PhD thesis, ANU.
- 1997b. 'Awa Pit'. RCLT Internal Document, ANU.
- 1998. 'Basque: Grammar summary'. RCLT internal document, ANU.
- 2002. 'Copula and verbless clauses in Awa Pit'. Handout for presentation at RCLT Local Workshop on 'Copula clauses and verbless clauses'.
- Cysouw, Michael. 2003. *The paradigmatic status of person marking*. Oxford: Oxford University Press.
- Czaykowska-Higgins, Ewa. 1998. 'The morphological and phonological constituent of words in Moses-Columbia Salish (Nxaʔamxín)', pp. 154–95 of *Salish languages and linguistics: Theoretical and descriptive perspectives*, edited by Ewa Czaykowska-Higgins and M. Dale Kinkade. Berlin: Mouton de Gruyter.
- Daguman, Josephine S. 2004. 'A grammar of Northern Subanen'. PhD thesis, La Trobe University.
- Dahlstrom, Amy. n.d. 'Grammar of Fox'. MS.
- Dai, John Xiang-Ling. 1998. 'Syntactic, phonological and morphological words in Chinese', pp. 103–34 of *New approaches to Chinese word formation: Morphology, phonology and the lexicon in Modern and Ancient Chinese*, edited by Jerome L. Packard. Berlin: Mouton de Gruyter.
- Davidson, Matthew. 2002. 'Studies in Southern Wakashan (Nootkan) grammar'. PhD dissertation, State University of New York at Buffalo.
- Davies, John. 1981. *Kobon*. Amsterdam: North-Holland.
- Deane, P. 1987. 'English possessives, topicality, and the Silverstein hierarchy', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13: 65–76.
- DeLancey, Scott. 1999. 'Relativization in Tibetan', pp. 231–48 of *Topics in Nepalese linguistics*, edited by Yogendra P. Yadava and Warren W. Glover. Kathmandu: Royal Nepal Academy.
- Dench, Alan C. 1995. *Martuthunira: A language of the Pilbara region of Western Australia*. Canberra: Pacific Linguistics.
- 1998. *Yingkarta*. Munich: Lincom Europa.
- Derbyshire, Desmond C. 1979. *Hixkaryana*. Amsterdam: North-Holland.
- 1985. *Hixkaryana and linguistic typology*. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- and Pullum, Geoffrey K. 1986–98. Editors of *Handbook of Amazonian languages*, Vol. 1, 1986a; Vol. 2, 1990; Vol. 3, 1991; Vol. 4, 1998. Berlin: Mouton de Gruyter.
- — 1986b. 'Introduction', pp. 1–28 of Derbyshire and Pullum 1986a.
- Deutscher, Guy. 2000. *Syntactic change in Akkadian: The evolution of sentential complementation*. Oxford: Oxford University Press.

- Deutscher, Guy. 2006. 'Complement clause types and complementation strategies in Akkadian', pp. 159–77 of Dixon and Aikhenvald 2006.
- Devitt, Dan. 1990. 'The diachronic development of semantics in copulas', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 16: 103–15.
- Dickens, Patrick. n.d. 'Juù'hoan grammar'. Typescript.
- Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins.
- Dik, Simon C., and Hengeveld, Kees. 1991. 'The hierarchical structure of the clause and the typology of perception verb complements', *Linguistics* 29: 231–59.
- Dimmendaal, Gerrit J. 1989. 'Complementizers in Hausa', pp. 87–110 of *Current papers in Chadic linguistics*, edited by Zygmunt Frajzyngier. Amsterdam: John Benjamins.
- . 2000. 'Number marking and noun categorization in Nilo-Saharan languages', *Anthropological Linguistics* 42: 214–61.
- Di Sciullo, Anna-Maria, and Williams, Edwin. 1987. *On the definition of word*. Cambridge, Mass.: MIT Press.
- Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
- . 1977a. *A grammar of Yidj. Cambridge*: Cambridge University Press.
- . 1977b. 'Some phonological rules in Yidin', *Linguistic Inquiry* 8: 1–34.
- . 1977c. 'Where have all the adjectives gone?', *Studies in Language* 1: 19–80. [Reprinted as pp. 1–62 of Dixon 1982.]
- . 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
- . 1981. 'Warrgamay', pp. 1–144 of Dixon and Blake 1981.
- . 1982. *Where have all the adjectives gone? And other essays in semantics and syntax*. Berlin: Mouton.
- . 1988a. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- . 1988b. '"Words" in Fijian', pp. 65–71 of *Lexicographical and linguistic studies: Essays in honour of G. W. Turner*, edited by T. L. Burton and Jill Burton. Cambridge: D. S. Brewer.
- . 1989. 'Subject and object in universal grammar', pp. 91–118 of *Essays on grammatical theory and universal grammar*, edited by Doug Arnold, Martin Atkinson, Jacques Durand, Claire Grover, and Louisa Sadler. Oxford: Clarendon Press.
- . 1991a. *A new approach to English grammar, on semantic principles*. Oxford: Clarendon Press.
- . 1991b. *Words of our country: Stories, place names and vocabulary in Yidiny, the Aboriginal language of the Cairns-Yarrabah region*. St Lucia: University of Queensland Press.
- . 1994. *Ergativity*. Cambridge: Cambridge University Press.
- . 1995. 'Complement clauses and complementation strategies', pp. 175–220 of *Grammar and meaning: Essays in honour of Sir John Lyons*, edited by F. R. Palmer. Cambridge: Cambridge University Press.
- . 1999a. 'Arawá', pp. 293–306 of Dixon and Aikhenvald 1999.
- . 1999b. 'Semantic roles and syntactic functions: The semantic basis for typology', *Papers from the Regional Meeting of the Chicago Linguistic Society* 35, 2: 323–41.

-
- 2000. 'Categories of the noun phrase in Jarawara', *Journal of Linguistics* 36: 487–510.
- 2002a. *Australian languages: Their nature and development*. Cambridge: Cambridge University Press.
- 2002b. 'The eclectic morphology of Jarawara, and the status of word', pp. 125–52 of Dixon and Aikhenvald 2002a.
- 2002c. 'Copula clauses in Australian languages: A typological perspective', *Anthropological Linguistics* 44: 1–26.
- 2003. 'Demonstratives: A cross-linguistic typology', *Studies in language* 27: 61–112.
- 2004a. *The Jarawara language of southern Amazonia*. Oxford: Oxford University Press.
- 2004b. 'Adjective classes in typological perspective', pp. 1–49 of Dixon and Aikhenvald 2004.
- 2004c. 'The small adjective class in Jarawara', pp. 177–98 of Dixon and Aikhenvald 2004.
- 2005. *A semantic approach to English grammar*. Oxford: Oxford University Press. [Revision of Dixon 1991a, with three chapters added.]
- 2006a. 'Complement clauses and complementation strategies in typological perspective', pp. 1–48 of Dixon and Aikhenvald 2006.
- 2006b. 'Complement clause type and complementation strategy in Jarawara', pp. 93–114 of Dixon and Aikhenvald 2006.
- 2006c. 'Complementation strategies in Dyirbal', pp. 263–79 of Dixon and Aikhenvald 2006.
- 2006d. 'Serial verb constructions: Conspectus and coda', pp. 338–50 of Aikhenvald and Dixon 2006.
- 2006e. 'Grammatical diffusion in Australia: Free and bound pronouns', pp. 67–93 of *Grammars in contact: A cross-linguistic typology*, edited by Alexandra Y. Aikhenvald and R. M. W. Dixon. Oxford: Oxford University Press.
- 2008. 'Comparative constructions: A cross-linguistic typology', *Studies in Language* 32: 787–817.
- and Aikhenvald, Alexandra Y. 1999. Editors of *The Amazonian languages*. Cambridge: Cambridge University Press.
- — 2002a. Editors of *Word: A cross-linguistic typology*. Cambridge: Cambridge University Press.
- — 2002b. 'Word: A typological framework', pp. 1–41 of Dixon and Aikhenvald 2002a.
- — 2004. Editors of *Adjective classes: A cross-linguistic typology*. Oxford: Oxford University Press.
- — 2006. Editors of *Complementation: A cross-linguistic typology*. Oxford: Oxford University Press.
- and Blake, Barry J. 1979–2000. Editors of *Handbook of Australian Languages*, Vol. 1, 1979; Vol. 2, 1981; Vol. 3, 1983; Canberra: Australian National University Press and Amsterdam: John Benjamins; Vol. 4, 1991; Vol. 5, 2000; Melbourne: Oxford University Press.

- Donaldson, Tamsin. 1980. *Ngiyambaa: The language of the Wangaaybuwan of New South Wales*. Cambridge: Cambridge University Press.
- Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
- Douglas, Wilfrid H. 1981. 'Watjarri', pp. 196–272 of Dixon and Blake 1981.
- Dowling, Bruce T. 1978. 'Some universals of relative clause structure', pp. 375–418 of *Universals of human language*, Vol. 4: *Syntax*, edited by Joseph H. Greenberg. Stanford, Calif.: Stanford University Press.
- Dresher, Bezalel E. 1994. 'The prosodic basis of the Tiberian Hebrew system of accents', *Language* 70: 1–52.
- Dryer, Matthew. 1983. 'Indirect objects in Kinyarwanda revisited', pp. 129–40 of *Studies in relational grammar*, edited by David M. Perlmutter. Chicago: University of Chicago Press.
- Duranti, Alessandro, and Byarushengo, Ernest R. 1977. 'On the notion of "direct object"', pp. 45–72 of *Haya grammatical structure*, edited by Ernest B. Byarushengo, Alessandro Duranti, and Larry M. Hyman. Los Angeles: University of Southern California Department of Linguistics.
- Durie, Mark. 1985. *A grammar of Acehnese, on the basis of a dialect of North Aceh*. Dordrecht: Foris.
- Dutton, Tom E. 1996. *Koiari*. Munich: Lincom Europa.
- Eades, Diana. 1979. 'Gumbaynggir', pp. 244–361 of Dixon and Blake 1979.
- Emenajo, E. Nlue. 1978. *Elements of modern Igbo grammar: A descriptive approach*. Ibadan: Oxford University Press.
- Emeneau, M. B. 1961. 'Brahui demonstrative pronouns', *Journal of the Asiatic Society* (Bengal) 3: 1–5. [Reprinted as pp. 17–22 of M. B. Emeneau, *Dravidian studies: Selected papers*. Delhi: Motilal Banarsidass, 1994.]
- Enfield, N. J. 2003. 'Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis', *Language* 79: 82–117.
- 2004. 'Adjectives in Lao', pp. 323–47 of Dixon and Aikhenvald 2004.
- 2007. *A grammar of Lao*. Berlin: Mouton de Gruyter.
- England, Nora. 2004. 'Adjectives in Mam', pp. 125–46 of Dixon and Aikhenvald 2004.
- Englebretson, Robert. 2003. *Searching for structure: The problem of complementation in colloquial Indonesian conversation*. Amsterdam: John Benjamins.
- Enrico, John. 2003. *Haida syntax*, 2 vols. Lincoln: University of Nebraska Press.
- Escalante, Fernando. 1990. 'Voice and argument structure in Yaqui'. PhD dissertation, University of Arizona.
- Evans, Nicholas D. 1995. *A grammar of Kayardild, with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
- 1996. 'The syntax and semantics of body part incorporation in Mayali', pp. 65–109 of Chappell and McGregor 1996.
- and Osada, Toshiki. 2005. 'Mundari: The myth of a language without word classes', *Linguistic Typology* 9: 351–90.
- Faverey, Margot, Johns, Brenda, and Wouk, Fay. 1976. 'The historical development of locative and existential copula constructions in Afro-English creole languages', pp. 88–95 of *Papers from the parasession on diachronic syntax*, edited by Stanford

- B. Steever, Carol A. Walker, and Salikoko S. Mufwene. Chicago: Chicago Linguistic Society.
- Feeling, Durbin. 1975. *Cherokee–English dictionary*. Talequah: Cherokee Nation of Oklahoma.
- Ferrar, H. 1967. *A French reference grammar*, 2nd edition. Oxford: Oxford University Press.
- Ferris, Connor. 1993. *The meaning of syntax: A study in the adjectives of English*. London: Longman.
- Filimonova, Elena. 2005. Editor of *Clusivity: Typology and case studies of the inclusive/exclusive distinction*. Amsterdam: John Benjamins.
- Firth, J. R. 1957. *Papers in linguistics, 1934–1951*. London: Oxford University Press.
- Fleck, David W. 2006. 'Complement clause type and complementation strategies in Matses', pp. 224–44 of Dixon and Aikhenvald 2006.
- 2008. 'Coreferential fourth-person pronouns in Matses', *International Journal of American Linguistics* 74: 279–311.
- Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- 1991. *The Yimas language of New Guinea*. Stanford, Calif.: Stanford University Press.
- Forchheimer, Paul. 1953. *The category of person in language*. Berlin: de Gruyter.
- Ford, Lysbeth J. 1990. 'The phonology and morphology of Bachamal (Wogait)'. MA thesis, ANU.
- 1998. 'A description of the Emmi language of the Northern Territory of Australia'. PhD thesis, ANU.
- Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.
- 1988. 'Eskimo orientation systems', *Meddelelser om Grønland, Man and Society* 11: 1–30.
- Fortune, Reo. 1942. *Arapesh*. Publications of the American Ethnological Society, Vol. 19.
- Fox, Barbara A. 1987. 'The Noun Phrase Accessibility Hierarchy revisited', *Language* 63: 856–70.
- Frachtenberg, Leo J. 1922a. 'Coos', pp. 297–429 of Boas 1922.
- 1922b. 'Siuslawan (Lower Umpqua)', pp. 431–619 of Boas 1922.
- Frajzyngier, Zygmunt. 1993. *A grammar of Mupun*. Berlin: Dietrich Reimer.
- Franklin, Karl J. 1971. *A grammar of Kewa, New Guinea*. Canberra: Pacific Linguistics.
- and Franklin, Joice. 1978. *A Kewa dictionary, with supplementary grammatical and anthropological materials*. Canberra: Pacific Linguistics.
- Gak, V. G. 1990. 'Slovo [Word]', pp. 464–7 of *Lingvističeskij enciklopedičeskij slovar'* [*Linguistic encyclopaedic dictionary*], edited by V. N. Yartseva. Moscow: Sovetskaya Enciklopediya.
- Galloway, Brent D. 1993. *A grammar of Upriver Halkomelem*. Berkeley and Los Angeles: University of California Press.
- Gardiner, Alan H. 1927. *Egyptian grammar: Being an introduction to the study of hieroglyphs*. Oxford: Clarendon Press.

- Gary, Judith Olmsted, and Gamal-Eldin, Saad. 1982. *Cairene Egyptian Colloquial Arabic*. Amsterdam: North-Holland.
- and Keenan, Edward L. 1977. 'On collapsing grammatical relations in universal grammar', pp. 83–120 of *Syntax and semantics*, Vol. 8: *Grammatical relations*, edited by Peter Cole and Jerrold M. Sadock. New York: Academic Press.
- Genetti, Carol. 1992. 'Semantic and grammatical categories of relative clause morphology in the languages of Nepal', *Studies in Language* 15: 405–27.
- 2006. 'Complement clause types and complementation strategy in Dolakha Newar', pp. 137–58 of Dixon and Aikhenvald 2006.
- 2007. *A grammar of Dolakha Newar*. Berlin: Mouton de Gruyter.
- and Hildebrandt, Kristine. 2004. 'The two adjective classes in Manange', pp. 74–96 of Dixon and Aikhenvald 2004.
- George Madugu, Isaac S. 1976. 'Yoruba adjectives have merged with verbs, or are they just emerging?', *Journal of West African Languages* 11: 85–102.
- Gerzenstein, Ana. 1994. *Lengua Maká: Estudio descriptivo*. Buenos Aires: Instituto de Lingüístico, Universidad de Buenos Aires.
- Giridhar, P. P. 1980. *Angami grammar*. Mysore: Central Institute of Indian Languages.
- 1994. *Mao Naga grammar*. Mysore: Central Institute of Indian Languages.
- Givón, Talmy. 1970. 'Notes on the semantic structure of English adjectives', *Language* 46: 816–37.
- 1980. *Ute reference grammar*. Ignacio, Colo.: Ute Press.
- 1991. 'The evolution of dependent clause morpho-syntax in Biblical Hebrew', pp. 257–310 of *Approaches to grammaticalization*, Vol. 2, edited by Elizabeth C. Traugott and Bernd Heine. Amsterdam: John Benjamins.
- Glasgow, Kathleen. 1994. *Burarra–Gun–Nartpa dictionary with English finder list*. Darwin: Summer Institute of Linguistics.
- Gorbet, Larry. 1976. *A grammar of Diegueño nominals*. New York: Garland.
- Gordon, Raymond G., Jr. 2005. Editor of *Ethnologue: Languages of the World*, 15th edition. Dallas: SIL International.
- Gragg, Gene. 1976. 'Oromo of Welledda', pp. 166–95 of *The non-Semitic languages of Ethiopia*, edited by M. Lionel Bender. East Lansing: African Studies Center, Michigan State University.
- Gray, Louis H. 1934. *Introduction to Semitic comparative linguistics*. New York: Columbia University Press.
- Greenberg, Joseph H. 1988. 'The first person inclusive dual as an ambiguous category', *Studies in Language* 12: 1–18.
- Gregor, Douglas B. 1980. *Celtic: A comparative study of the six Celtic languages, Irish, Gaelic, Manx, Welsh, Cornish, Breton, seen against the background of their history, literature, and destiny*. Cambridge: The Oleander Press.
- Grimes, Charles E. 1991. 'The Buru languages of eastern Indonesia'. PhD thesis, ANU.
- Guedes, Marymarcia. 1991. *Subsídios para uma análise fonológica do Mbyá*. Campinas: Editora da Unicamp.
- Guillaume, Antoine. 2008. *A grammar of Cavineña*. Berlin: Mouton de Gruyter.
- Güldemann, Tom. 2003. 'Logophoricity in Africa: An attempt to explain and evaluate the significance of its modern distribution', *Sprachtypologie und Universalienforschung* 56: 366–87.

- and von Roncador, Manfred. 2002. *Editors of Reported discourse: A meeting ground for different linguistic domains*. Amsterdam: John Benjamins.
- Haag, Marcia. 1998. 'Word-level evidence for lexical categories in Salishan languages', *International Journal of American Linguistics* 64: 379–93.
- Haas, Mary R. 1941. *Tunica*. New York: J. J. Augustin.
- 1969. '“Exclusive” and “inclusive”': A look at early usage', *International Journal of American Linguistics* 35: 1–6.
- Hagège, Claude. 1974. 'The “adjective” in some African languages', *Studies in African Linguistics*, Supplement 5: 125–33.
- Haiman, John. 1980. *Hua: A Papuan language of the eastern highlands of New Guinea*. Amsterdam: John Benjamins.
- 1983. 'Iconic and economic motivation', *Language* 59: 781–819.
- 1985. *Natural syntax*. Cambridge: Cambridge University Press.
- Hajdú, Péter. 1963. *The Samoyed peoples and languages*. Bloomington: Indiana University Press.
- Hale, Kenneth. 1976. 'The adjoined relative clause in Australia', pp. 78–105 of *Grammatical categories in Australian languages*, edited by R. M. W. Dixon. Canberra: Australian Institute of Aboriginal Studies.
- Hall, T. Alan, and Kleinhenz, Ursula. 1999. *Editors of Studies on the phonological word*. Amsterdam: John Benjamins.
- Halliday, M. A. K., and Hasan, Ruqaiya. 1976. *Cohesion in English*. London: Longman.
- Harada, S. I. 1976. 'Honorifics', pp. 499–561 of *Syntax and semantics*, Vol. 5: *Japanese generative grammar*, edited by Masayoshi Shibatani. New York: Academic Press.
- Harris, Alice C. 1991. 'Mingrelian', pp. 313–94 of *The indigenous languages of the Caucasus*, Vol. 1: *The Kartvelian languages*, edited by Alice C. Harris. Delmar, NY: Caravan.
- Harris, Zellig S. 1946. 'From morpheme to utterance', *Language* 22: 161–83.
- Harrison, Sheldon P. 1976. *Mokilese reference grammar*. Honolulu: University Press of Hawaii.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.
- 2001. 'Non-canonical marking of subjects and objects', pp. 55–83 of Aikhenvald, Dixon, and Onishi 2001.
- Dryer, Matthew S., Gill, David, and Comrie, Bernard. 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Haviland, John. 1979. 'Guugu Yimidhirr', pp. 1–180 of Dixon and Blake 1979.
- 1981. *Sk'op sotz'leb: El Tzotzil de San Lorenzo Zinacantán*. Mexico City: Universidad Nacional Autónoma de México.
- Hayward, Richard J. 1990. 'Notes on the Zayse language', pp. 210–355 of *Omoti language studies*, edited by Richard J. Hayward. London: School of Oriental and African Studies.
- Hazlewood, David. 1850. *A Feejeean and English and an English and Feejeean dictionary... and a compendious grammar of the Feejeean language...* Vewa: Wesleyan Missionary Press.
- Heath, Jeffrey. 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- Hébert, Yvonne M. 1983. 'Noun and verb in a Salishan language', *Kansas Working Papers in Linguistics* 8, 2: 36–81.

- Heffernan, J. A. 1984. *Papunya Luritja language notes*. Papunya: Literature Production Centre.
- Heine, Bernd. 1997a. *Possession: Cognitive sources, forces and grammaticalization*. Cambridge: Cambridge University Press.
- 1997b. *Cognitive foundations of grammar*. New York: Oxford University Press.
- and Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hellwig, Birgit. 2006. 'Complement clause type and complementation strategies in Goemai', pp. 49–71 of Dixon and Aikhenvald 2006.
- Henadeerage, Kumara. 1997. 'Relative clauses in Sinhala (Sri Lanka)'. Presentation at the ANU.
- Henderson, John. 2002. 'The word in Eastern/Central Arrernte', pp. 100–24 of Dixon and Aikhenvald 2002a.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Hercus, Luise A. 1982. *The Bāgandji language*. Canberra: Pacific Linguistics.
- 1994. *A grammar of the Arabana-Wangkanguru language, Lake Eyre Basin, South Australia*. Canberra: Pacific Linguistics.
- Hewitt, B. G. 1979. 'The relative clause in Abkhaz (Abžui dialect)', *Lingua* 47: 151–88.
- 1981. *Abkhaz*. Amsterdam: North-Holland.
- Hill, Deborah. 1992. 'Longgu grammar'. PhD thesis, ANU.
- 1997a. 'Indonesian grammar summary'. RCLT internal document, ANU.
- 1997b. 'Longgu'. RCLT internal document, ANU.
- Himmelman, Nikolaus P. 1996. 'Demonstratives in narrative discourse: A taxonomy of universal uses', pp. 205–54 of *Studies in anaphora*, edited by Barbara Fox. Amsterdam: John Benjamins.
- Hinds, John. 1986. *Japanese*. London: Croom Helm.
- Hock, Hans H. 1985. 'Sanskrit double-object constructions: Will the real object please stand up', *Praci-Bhasha-Vijnan: Journal of Indian Linguistics* 12: 50–70.
- Hockett, Charles F. 1944. Review of *Linguistic interludes and Morphology: The descriptive analysis of words* (1944 edition), both by Eugene A. Nida, in *Language* 20: 252–5.
- 1958. *A course in modern linguistics*. New York: Macmillan.
- Hoffmann, Carl. 1963. *A grammar of the Margi language*. London: Oxford University Press.
- Hoffmann, J. 1903. *Mundari grammar*. Calcutta: Bengal Secretariat Press.
- 1905. *A Mundari grammar with exercises*. Calcutta: Catholic Orphan Press. [Appears to be identical to 1903 grammar, with text reset, plus 127 pp. of exercises.]
- 1930–79. *Encyclopaedia Mundarica*, 16 vols. Patna: Superintendent, Government Printing, Bihar and Orissa.
- Hoiijer, Harry. 1933. 'Tonkawa: An Indian language of Texas', pp. 1–148 of Boas 1933.
- Holisky, DeeAnn. 1987. 'The case of the intransitive subject in Tsova-Tush (Batsbi)', *Lingua* 71: 103–32.
- Hopper, Paul J., and Thompson, Sandra A. 1980. 'Transitivity in grammar and discourse', *Language* 56: 251–99.

- and Traugott, Elizabeth C. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huang, Chenglong. 2002. 'Copula clauses and verbless clauses in Mandarin Chinese'. Handout for presentation at RCLT Local Workshop on 'Copula clauses and verbless clauses'.
- Huber, Brigitte. 2003. 'Relative clauses in Kyirong Tibetan', *Languages of the Tibeto-Burman Area* 26: 1–14.
- Huddleston, Rodney, and Pullum, Geoffrey K. 2002. Chief authors of *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hudson, Joyce. 1978. *The core of Walmatjari grammar*. Canberra: Australian Institute of Aboriginal Studies.
- Hyman, Larry M. 1977. 'The syntax of body parts', pp. 99–117 of *Haya grammatical structure*, edited by Ernest R. Byarushengo, Alessandro Duranti, and Larry M. Hyman. Los Angeles: Department of Linguistics, University of Southern California.
- Alford, Denny Keith, and Akpati, Elizabeth. 1970. 'Inalienable possession in Igbo', *Journal of West African Languages* 7: 85–101.
- and Comrie, Bernard. 1981. 'Logophoric reference in Gokana', *Journal of African Languages and Linguistics* 3: 19–37.
- and Duranti, Alessandro. 1982. 'On the object relation in Bantu', pp. 217–39 of *Syntax and semantics*, Vol. 15: *Studies in transitivity*. New York: Academic Press.
- Hyslop, Catriona. 2001. *The Lolovoli dialect of the North-East Ambae language, Vanuatu*. Canberra: Pacific Linguistics.
- 2004. 'Adjectives in North-East Ambae', pp. 263–282 of Dixon and Aikhenvald 2004.
- Ikoro, Suanu. 1996a. *The Kana language*. Leiden: CNWS.
- 1996b. 'Kana'. RCLT internal document, ANU.
- 1996c. 'Igbo'. RCLT internal document, ANU.
- Imedadze, Natela, and Tuite, Kevin. 1992. 'The acquisition of Georgian', pp. 39–109 of *The crosslinguistic study of language acquisition*, Vol. 3, edited by Dan I. Slobin. Hillsdale, NJ: Lawrence Erlbaum.
- Ingram, David. 1978. 'Typology and universals of personal pronouns', pp. 213–47 of *Universals of human language*, Vol. 3: *Word structure*, edited by Joseph H. Greenberg. Stanford, Calif.: Stanford University Press.
- Isačenko, Alexander V. 1974. 'On "have" and "be" languages', pp. 43–77 of *The Slavic Forum: Essays in linguistics and literature*, edited by Michael S. Flier. The Hague: Mouton.
- Jacob, Judith M. 1968. *Introduction to Cambodian*. London: Oxford University Press.
- Jacobsen, William, Jr. 1979. 'Noun and verb in Nootkan', pp. 83–155 of *The Victorian conference on Northwestern languages*, edited by Barbara S. Efrat. British Columbia Provincial Museum, Heritage Record No. 4.
- 1980. 'Inclusive/exclusive: A diffused pronominal category in native western North America', pp. 204–27 of *Papers from the parasession on pronouns and anaphora*, edited by Jody Kreiman and Almerindo E. Ojeda. Chicago: Chicago Linguistic Society.

- Jaggat, Philip J. 1998. 'Restrictive vs. non-restrictive relative clauses in Hausa: Where morphosyntax and semantics meet', *Studies in African Languages* 27: 199–238.
- Jany, Carmen. 2007. 'Is there any evidence for complementation in Chimariko?', *International Journal of American Linguistics* 73: 94–113.
- Jarkey, Nerida. 2006. 'Complement clause types and complementation strategy in White Hmong', pp. 115–36 of Dixon and Aikhenvald 2006.
- Jastrow, Otto. 1997. 'The Neo-Aramaic languages', pp. 334–77 of *The Semitic languages*, edited by Robert Hetzron. London: Routledge.
- Jauncey, Dorothy. 1997. 'A grammar of Tamambo, the language of western Malo, Vanuatu'. PhD thesis, ANU.
- Jelinek, Eloise, and Demers, Richard A. 1994. 'Predicates and pronominal arguments in Straits Salish', *Language* 70: 697–736.
- Jespersen, Otto. 1924. *The philosophy of grammar*. London: George Allen and Unwin.
- . 1927. *A modern English grammar on historical principles*, Part III: *Syntax*, second volume. Heidelberg: Winter.
- . 1933. *Essentials of English grammar*. London: Allen and Unwin.
- Jeyapal, V. Y. 1987. *Karbi grammar*. Mysore: Central Institute of Indian Languages.
- Johanson, Lars. 2006. 'Adjectives and nouns in South Siberian and other Turkic languages', pp. 57–78 of *Exploring the eastern frontiers of Turkic in South Siberia*, edited by Marcel Erdal and Irina Nevskaya. Wiesbaden: Harrassowitz.
- Johansson, Stig, and Lysvåg, Per. 1987. *Understanding English grammar*, Part II: *A closer view*. Oslo: Universitetsforlaget.
- and Oksefjell, S. 1998. 'A friend of my mother(s): On the use of genitive vs. common case in postmodifying *of*-phrases', *Studia Anglica Posnaniensia* 33: 153–64.
- Jones, Daniel. 1956. *Everyman's English pronouncing dictionary*, 11th edition. London: Dent and New York: Dutton.
- Jones, Morris, and Thomas, Alan R. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.
- Joseph, Brian D., and Philippaki-Warbuton, Irene. 1987. *Modern Greek*. London: Croom Helm.
- Jungbluth, Konstanze. 2001. 'Binary and ternary deictic systems in speech and writing. Evidence from the use of demonstratives in Spanish', *Philologie im Netz* 15: 1–24.
- Kachru, Yamuna. 1968. 'The copula in Hindi', pp. 33–59 of Verhaar, Part 2, 1968.
- Kautzsch, E. 1910. Editor of *Gesenius's Hebrew Grammar*, 2nd English edition, revised by A. E. Cowley. Oxford: Clarendon Press.
- Keenan, Edward L. 1972. 'Relative clause formation in Malagasy', pp. 169–89 of Peranteau, Levi, and Phares 1972.
- . 1975. 'Logical expressive power and syntactic variation in natural language', pp. 406–21 of *Formal semantics of natural languages*, edited by Edward L. Keenan. Cambridge: Cambridge University Press.
- . 1985. 'Relative clauses', pp. 141–70 of Shopen 1985: Vol. 2.

- and Comrie, Bernard. 1977. 'Noun phrase accessibility and universal grammar', *Linguistic Inquiry* 8: 63–99.
- 1979. 'Data on the Noun Phrase Accessibility Hierarchy', *Language* 55: 333–51.
- Kenesei, István, Vago, Robert M., and Fenyvesi, Anna. 1998. *Hungarian*. London: Routledge.
- Khaidakov, S. M. 1966. 'The dialect divisions of Lak', *Studia Caucasica* 2: 9–18.
- Kimball, Geoffrey D. 1991. *Koasati grammar*. Lincoln: University of Nebraska Press.
- Kinkade, M. Dale. 1983. 'Salish evidence against the universality of "noun" and "verb"', *Lingua* 60: 25–40.
- Kiparsky, Paul. 1998. 'Partitive case and aspect', pp. 265–307 of *The projection of arguments: Lexical and compositional factors*, edited by Miriam Butt and Wilhelm Gender. Stanford, Calif.: CSLI.
- Kita, Sotaro. 2003. Editor of *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum.
- Klamer, Marian. 1998. *A grammar of Kambara*. Berlin: Mouton de Gruyter.
- Kleinschmidt, S. 1851. *Grammatik der grönländischen Sprache mit teilweisem Einschluss des Labradordialekts*. Berlin: G. Reimer.
- Kockelman, Paul. 2003. 'The interclausal relations hierarchy in Q'eqchi' Maya', *International Journal of American Linguistics* 69: 25–48.
- Koehn, Edward, and Koehn, Sally. 1986. 'Apalai', pp. 33–127 of *Derbyshire and Pullum 1986a*.
- König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.
- Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.
- Koshal, Sanyukta. 1979. *Ladakhi grammar*. Delhi: Motilal Banarsidass.
- Krámský, Jiří. 1969. *The word as a linguistic unit*. The Hague: Mouton.
- Kratochvíl, František. 2007. 'A grammar of Abui, a Papuan language of Alor'. PhD thesis, University of Leiden.
- Krishnamurti, Bh. 2001. 'Telugu demonstratives'. Handout for RCLT Local Workshop on 'Demonstratives and related items'.
- Kroeger, Paul D. 1999. *The Salish language family: Reconstructing syntax*. Lincoln: University of Nebraska Press.
- Kroeger, Paul. 1998. 'Nouns and verbs in Tagalog: A reply to Foley'. Available on web at <<http://www.sultry.arts.usyd.edu.au/LFG98/austro/download/kroeger.doc>>.
- Kruspe, Nicole. 2004a. *A grammar of Semelai*. Cambridge: Cambridge University Press.
- 2004b. 'Adjectives in Semelai', pp. 283–305 of Dixon and Aikhenvald 2004.
- Kuipers, Aert H. 1967. *The Squamish language: Grammar, texts, dictionary*. The Hague: Mouton.
- 1974. *The Shuswap language: Grammar, texts, dictionary*. The Hague: Mouton.
- Kullavanijaya, Pranee. 2008. 'A historical study of *thii* in Thai', pp. 445–67 of *The Tai-Kadai languages*, edited by Anthony V. N. Diller, Jerrold A. Edmondson, and Yongxian Luo. London: Routledge.

- Kumar, Rajesh, and Subbarao, Karumuri V. 2005. 'Postposition incorporation and internally-headed relative clauses in Hmar', pp. 129–38 of *Contemporary issues in Nepalese linguistics*, edited by Yogendra Yadava, Govinda Bhattarai, Ram Raj Lohani, Balaram Prasain, and Krisha Parajuli. Kathmandu: Linguistic Society of Nepal.
- Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: MIT Press.
- and Wongkhamthong, Preya. 1981. 'Relative clauses in Thai', *Studies in Language* 5: 195–226.
- Kuroda, S.-Y. 1992. *Japanese syntax and semantics: Collected papers*. Dordrecht: Kluwer.
- Kuteva, Tania. 2001. *Auxiliation: An enquiry into the nature of grammaticalization*. Oxford: Oxford University Press.
- Langendoen, D. Terence. 1967. 'The copula in Mundari', pp. 75–100 of Verhaar, Part 1, 1967.
- LaPolla, Randy J. 2003. *A grammar of Qiang, with annotated texts and glossary*. Berlin: Mouton de Gruyter.
- 2005. 'The inclusive–exclusive distinction in Tibeto-Burman languages', pp. 291–311 of Filimonova 2005.
- and Huang, Chenglong. 2002. 'The copula and existential verbs in Qiang'. Presentation at RCLIT Local Workshop on 'Copula clauses and verbless clauses'.
- — 2004. 'Adjectives in Qiang', pp. 306–22 of Dixon and Aikhenvald 2004.
- Launey, Michel. 1994. *Une grammaire omniprédicative: Essai sur la morphosyntaxe du nahuatl classique*. Paris: Éditions CNRS.
- Lee, Jennifer. 1987. *Tiwi today: A study of language change in a contact situation*. Canberra: Pacific Linguistics.
- Leeding, Velma J. 1989. 'Anindilyakwa phonology and morphology'. PhD thesis, University of Sydney.
- Leer, Jeff. 1987. 'Aleut anaphora'. MA thesis, University of Chicago.
- Lehiste, Ilse. 1962. 'Acoustic studies of boundary signals', pp. 178–87 of *Proceedings of the 4th International Congress of Phonetic Sciences, Helsinki*, edited by A. Sovijärvi and P. Aalto. The Hague: Mouton.
- Lehmann, Christian. 1986. 'On the typology of relative clauses', *Linguistics* 24: 663–80.
- Lehmann, Thomas. 1993. *A grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.
- Leslau, Wolf. 1995. *Reference grammar of Amharic*. Wiesbaden: Harrassowitz.
- Levin, Beth, and Rappaport Hovav, Malka. 1995. *Unaccusativity: At the syntax–lexical semantics interface*. Cambridge, Mass.: MIT Press.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- 1999. 'Deictics and demonstratives: Field guide for 1998', pp. 29–40 of 'Manual' for the 1999 field season, edited by David Wilkins. Nijmegen: Language and Cognition Group, Max Planck Institute.
- Levy, Paulette. 2004. 'Adjectives in Papantla Totonac', pp. 147–76 of Dixon and Aikhenvald 2004.
- Lewis, Geoffrey. 2000. *Turkish grammar*, 2nd edition. Oxford: Oxford University Press.

- Li, Charles N., and Thompson, Sandra A. 1976. 'Strategies for signaling grammatical relations in Wappo', *Papers from the Regional Meeting of the Chicago Linguistic Society* 12: 450–8.
- 1977. 'A mechanism for the development of copula morphemes', pp. 419–44 of *Mechanism of syntactic change*, edited by Charles N. Li. Austin: University of Texas Press.
- 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley and Los Angeles: University of California Press.
- Li, Feng. 1993. 'The copula in Classical Chinese declarative sentences', *Journal of Chinese Linguistics* 21: 277–311.
- Li Xinyin. 1996. 'Chinese: grammar summary'. RCLT internal document, ANU.
- Lichtenberk, Frantisek. 1985. 'Possessive constructions in Oceanic languages and in Proto-Oceanic', pp. 93–140 of *Austronesian linguistics at the 15th Pacific Science Congress*, edited by Andrew Pawley and Lois Carrington. Canberra: Pacific Linguistics.
- 2000. 'Inclusory pronominals', *Oceanic Linguistics* 39: 1–32.
- 2005. 'Inclusive–exclusive in Austronesian: An opposition of equals', pp. 261–89 of Filimonova 2005.
- Liddell, Scott K. 1980. *American sign language syntax*. The Hague: Mouton.
- Liddicoat, Tony. 1993. 'Possession in Jersey Norman French (with reference to Standard French)'. Handout for presentation in Workshop on 'Possession', ANU.
- Lindström, Eva. 2000. 'Some uses of demonstratives in spoken Swedish', pp. 107–28 of *Corpus-based and computational approaches to discourse anaphora*, edited by Simon Botley and Anthony N. McEnery. Amsterdam: John Benjamins.
- Longacre, Robert E. 1964. *Grammar discovery procedures*. The Hague: Mouton.
- Lynch, John. 1973. 'Verbal aspects of possession in Melanesian languages', *Oceanic Linguistics* 12: 69–102.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- 1977. *Semantics*, 2 vols. Cambridge: Cambridge University Press.
- McCarthy, John J. 1982. 'Prosodic structure and expletive infixation', *Language* 58: 574–90.
- McConvell, Patrick. 1990. 'Gurindji'. MS. Northern Territory University.
- McGuckin, Catherine. 2002. 'Gapapaiwa', pp. 297–321 of *The Oceanic languages*, edited by John Lynch, Malcolm Ross, and Terry Crowley. Richmond, Surrey: Curzon.
- McKay, Graham R. 1975. 'Rembarnga: A language of Central Arnhem Land'. PhD thesis, ANU.
- 1978. 'Pronominal person and number categories in Rembarnga and Djeebbana', *Oceanic Linguistics* 17: 27–37.
- 1988. 'Figure and ground in Rembarnga complex sentences', pp. 7–36 of *Complex sentence constructions in Australian languages*, edited by Peter Austin. Amsterdam: John Benjamins.
- McKay, Graham R. 1996. 'Body parts, possession marking and nominal classes in Ndjébbana', pp. 293–326 of Chappell and McGregor 1996.

- Mahajan, Anoop. 2000. 'Relative asymmetries and Hindi correlatives', pp. 201–29 of *The syntax of relative clauses*, edited by Artemis Alexiadou, Paul Law, André Meinunger, and Chris Wilder. Amsterdam: John Benjamins.
- Mahootian, Shahrzad. 1997. *Persian*. London: Routledge.
- Mallinson, Graham. 1986. *Rumanian*. London: Croom Helm.
- Mallinson, Graham, and Blake, Barry J. 1981. *Language typology: Cross-linguistic studies in syntax*. Amsterdam: North-Holland.
- Malotki, Eckehart. 1982. 'Hopi person deixis', pp. 223–52 of *Here and there: Cross-linguistic studies in deixis and demonstration*, edited by Jürgen Weissenborn and Wolfgang Klein. Amsterdam: John Benjamins.
- Mamiani, Luis Vincencio. 1877. *Arte de grammatica da lingua brasilica da nação Kiriri*, 2nd edition. Rio de Janeiro: Bibliotheca Nacional. [1st edition, Lisbon, 1699.]
- Maring, Joel M. 1967. 'Grammar of Acoma Keresan'. PhD dissertation, Indiana University.
- Martin, Jack. 1993. "'Inalienable possession" in Creek (and its possible origin)', *International Journal of American Linguistics* 59: 442–52.
- Masica, Colin. 1972. 'Relative clauses in South Asia', pp. 198–204 of Peranteau, Levi, and Phares, 1972.
- . 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Matisoff, James A. 1972. 'Lahu nominalization, relativization and genitivation', pp. 237–57 of *Syntax and semantics I*, edited by John P. Kimball. New York: Seminar Press.
- . 1973. *The grammar of Lahu*. Berkeley and Los Angeles: University of California Press.
- Matsumoto, Yoshiko. 1988. 'Semantics and pragmatics of noun-modifying constructions in Japanese', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 14: 166–75.
- Matthews, P. H. 1991. *Morphology*, 2nd edition. Cambridge: Cambridge University Press.
- . 1997. *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.
- . 2002. 'What can we conclude?', pp. 266–81 of Dixon and Aikhenvald 2002a.
- Matthews, Stephen, and Yip, Virginia. 1994. *Cantonese: A comprehensive grammar*. London: Routledge.
- Mazaudon, Martine. 1978. 'La Formation des propositions relatives en tibétain', *Bulletin de la Société de Linguistique de Paris* 73: 401–14.
- Meillet, Antoine. 1964. *Introduction à l'étude comparative des langages indo-européennes*. University, Ala.: University of Alabama Press.
- . 1970. *General characteristics of the Germanic languages*, translated by William P. Dismukes. Coral Gables, Fla.: University of Miami Press.
- Meinhof, Carl. 1930. *Der Karanadialekt des Hottentottischen*. Berlin: Dietrich Reimer.
- Meira, Sérgio. 1999. 'A grammar of Tiriyo'. PhD dissertation, Rice University.
- Mejlanova, U. A. 1960. *Morfologičeskaja i sintaksičeskaja xarakteristika padežej lezgin-skogo jazyka* [Morphological and syntactic characteristics of Lezgian cases]. Maxačkala: Dagestanskij Filial Akademii Nauk SSSR.

- Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland.
- Meyerstein, Zlata P. 1972. 'Czech deictics: Pronouns or articles?', *Linguistics* 91: 17–30.
- Michael, Ian. 1970. *English grammatical categories, and the tradition to 1800*. Cambridge: Cambridge University Press.
- Migeod, F. W. H. 1914. *A grammar of the Hausa language*. London: Kegan Paul.
- Mikkelsen, Line. 2005. *Copula clauses: Specification, predication and equation*. Amsterdam: John Benjamins.
- Miller, Amy. 2001. *A grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter.
- Miller, Wick R. 1965. *Acoma grammar and texts*. Berkeley and Los Angeles: University of California Press.
- 1996. 'Sketch of Shoshone, a Uto-Aztecan language', pp. 693–720 of *Handbook of North American Indians*, Vol. 17: *Languages*, edited by Ives Goddard. Washington, DC: Smithsonian Institution.
- Milner, G. B. 1956. *Fijian grammar*. Suva: Fiji Government Press.
- Mithun, Marianne. 1984. 'The evolution of noun incorporation', *Language* 60: 847–94.
- 1987. 'The grammatical nature and discourse power of demonstratives', *Berkeley Linguistics Society Proceedings* 13: 184–94.
- 1991. 'Active/agentive case marking and its motivations', *Language* 67: 510–46.
- 1996. 'Multiple reflections of inalienability in Mohawk', pp. 633–49 of Chappell and McGregor 1996.
- 1999. *The languages of native North America*. Cambridge: Cambridge University Press.
- 2000. 'Noun and verb in Iroquoian languages: Multicategorisation from multiple criteria', pp. 397–420 of *Approaches to the typology of word classes*, edited by Petra M. Vogel and Bernard Comrie. Berlin: Mouton de Gruyter.
- Moliner, Maris. 1984. *Diccionario de uso del español*. Madrid: Editorial Gredos.
- Monserrat, Ruth. 2000. 'A lingua do povo Mýky'. PhD dissertation, Universidade Federal de Rio de Janeiro.
- Mosel, Ulrike. 2004. 'Complex predicates and juxtapositional constructions in Samoan', pp. 263–96 of *Complex predicates in Oceanic languages: Studies in the dynamics of binding and boundedness*, edited by Isabelle Bril and Françoise Ozanne-Rivière. Berlin: Mouton de Gruyter.
- and Hovdhaugen, Even. 1992. *Samoan reference grammar*. Oslo: Scandinavian Universities Press.
- Mous, Maarten. 1993. *A grammar of Iraqw*. Hamburg: Helmut Buske.
- Munro, Pamela. 1976. *Mojave Syntax*. New York: Garland.
- Nababan, P. W. J. 1981. *A grammar of Toba-Batak*. Canberra: Pacific Linguistics.
- Nagaraja, K. S. 1999. *Korku language: Grammar, texts and vocabulary*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Nakayama, Toshihide. 2001. *Nuuchahnulth (Nootka) morphosyntax*. Berkeley and Los Angeles: University of California Press.
- Nater, H. F. 1984. *The Bella Coola language*. Ottawa: National Museum of Canada.

- Nedjalkov, Igor V. 1992. 'Functions of Evenki verbal suffixes with variable morpheme ordering (comparative approach)', *Languages of the World* 4: 20–41.
- 1997. *Evenki*. London: Routledge.
- Nespor, Marina, and Vogel, Irene. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Newman, John. 2002. Editor of *The linguistics of sitting, standing and lying*. Amsterdam: John Benjamins.
- Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.
- Newman, Stanley. 1967. 'Yokuts', *Lingua* 17: 182–99.
- 1968. 'Zuni equivalents of English "to be"', pp. 60–70 of Verhaar, Part 2, 1968.
- Ng, Eve. 2004. 'What constitutes a case of grammaticalization? Evidence for the development of copulas from demonstratives in Passamaquoddy', pp. 281–98 of *Up and down the cline: The nature of grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon. Amsterdam: John Benjamins.
- Nguyễn, Đình-Hoà. 1987. 'Vietnamese', pp. 777–96 of *The world's major languages*, edited by Bernard Comrie. London: Croom Helm.
- Nichols, Johanna. 1980. 'Control and ergativity in Chechen', *Papers from the Regional Meeting of the Chicago Linguistic Society* 16: 259–68.
- 1982. 'Ingush transitivization and detransitivization', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 8: 445–62.
- 1983. 'Switch-reference in the Northeast Caucasus', pp. 245–65 of *Switch-reference and universal grammar*, edited by John Haiman and Pamela Munro. Amsterdam: John Benjamins.
- 1986. 'Head-marking and dependent-marking grammar', *Language* 62: 56–119.
- 1988. 'On alienable and inalienable possession', pp. 557–609 of *In honor of Mary Haas*, edited by William Shipley. Berlin: Mouton de Gruyter.
- 1994. 'Ingush', pp. 79–145 of *The indigenous languages of the Caucasus*, Vol. 4: *North east Caucasian languages*, Part 2, edited by Ricks Smeets. Delmar, NY: Caravan Books.
- Nikiforidou, Kiki. 1991. 'The meanings of the genitive: A case study in semantic structure and semantic change', *Cognitive linguistics* 2: 149–205.
- Noonan, Michael. 1985. 'Complementation', pp. 42–141 of *Shopen 1985: Vol. 2*.
- 1992. *A grammar of Lango*. Berlin: Mouton de Gruyter.
- 2007. 'Complementation', pp. 52–150 of *Shopen 2007: Vol. 2*.
- Nordlinger, Rachel. 1998. *A grammar of Wambaya*. Canberra: Pacific Linguistics.
- and Sadler, Louisa. 2004. 'Nominal tense in cross-linguistic perspective', *Language* 80: 776–806.
- Ọmọrúyí, Thomas O. 1986. 'Adjectives and adjectivalization processes in Èdó', *Studies in African Linguistics* 17: 83–302.
- Obata, Kazuko. 2003. *A grammar of Bilua, a Papuan language of the Solomon Islands*. Canberra: Pacific Linguistics.
- Okell, John. 1969. *A reference grammar of Colloquial Burmese*. London: Oxford University Press.

- Olawsky, Knut J. 1999. *Aspects of Dagbani grammar, with special emphasis on phonology and morphology*. Munich: Lincom Europa.
- . 2002. 'The copula in Urarina'. Handout for presentation at RCLT Local Workshop on 'Copula clauses and verbless clauses'.
- . 2006. *A grammar of Urarina*. Berlin: Mouton de Gruyter.
- Onishi, Masayuki. 1994. 'A grammar of Motuna (Bougainville, Papua New Guinea)'. PhD thesis, ANU.
- . 1996. 'Ainu grammar summary'. RCLT internal document, ANU.
- . 1997. 'Bengali: Grammar summary'. RCLT internal document, ANU.
- . 2001a. 'Introduction: Non-canonically marked subjects and objects: parameters and properties', pp. 1–51 of Aikhenvald, Dixon, and Onishi 2001.
- . 2001b. 'Non-canonically marked A/S in Bengali', pp. 113–47 of Aikhenvald, Dixon, and Onishi 2001.
- Osam, E. Kweku. 1999. 'Adjectives in Akan', *Afrika und Übersee*, 32: 189–211.
- Overall, Simon. 2007. 'A grammar of Aguaruna'. PhD thesis, La Trobe University.
- Owens, Jonathan. 1985. *A grammar of Harar Oromo (northeastern Ethiopia), including a text and glossary*. Hamburg: Helmut Buske.
- Packard, Jerome L. 2000. *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.
- Palmer, F. R. 1967. 'Bilin', *Lingua* 17: 200–9.
- Paris, Catherine. 1985. 'Relations antanciennes et valence verbale en avar: Effacement d'actant et mise au jour du sens', *Actances* 1: 135–53.
- Parsons, F. W. 1960. 'An introduction to gender in Hausa', *African Language Studies* 1: 117–36.
- Pawley, Andrew. 1973. 'Some problems of Proto-Oceanic grammar', *Oceanic Linguistics* 12: 103–88.
- Payne, Doris L., and Barshi, Immanuel. 1999. Editors of *External possession*. Amsterdam: John Benjamins.
- and Payne, Thomas E. 1990. 'Yagua', pp. 249–474 of Derbyshire and Pullum 1990.
- Payne, Thomas E., and Payne, Doris L. 1999. 'Panare: A Cariban language of Central Venezuela'. RCLT internal document, ANU.
- Pensalfini, Robert J. 1997. 'Jingulu grammar, dictionary and texts'. PhD dissertation, MIT.
- Peranteau, Paul M., Levi, Judith N., and Phares, Gloria C. 1972. Editors of *The Chicago Which Hunt: Papers from the Relative Clause Festival*. Chicago: Chicago Linguistic Society.
- Perlmutter, David M. 1978. 'Impersonal passives and the unaccusative hypothesis', *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4: 157–89.
- Piau, Julie. 1985. 'The verbal syntax of Kuman'. MA thesis, ANU.
- Pike, Kenneth L. 1947. *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.
- Platero, Paul R. 1974. 'The Navajo relative clause', *International Journal of American Linguistics* 40: 202–46.

- Pope, M. K. 1934. *From Latin to Modern French*. Manchester: Manchester University Press.
- Poppe, Nikolai N. 1960. *Buriat grammar*. Bloomington: Indiana University Press.
- Post, Mark. 2008. 'Adjectives in Thai: Implications for a functionalist typology of word classes'. *Linguistic Typology* 12: 339–81.
- Poulos, George. 1990. *A linguistic analysis of Venda*. Pretoria: Via Africa.
- Press, Margaret L. 1979. *Chemehuevi: A grammar and lexicon*. Berkeley and Los Angeles: University of California Press.
- Prista, Alexander da R. 1966. *Essential Portuguese grammar*. New York: Dover.
- Pustet, Regina. 2001. 'Lakota'. RCLT internal document.
- 2003. *Copulas: Universals in the categorization of the lexicon*. Oxford: Oxford University Press.
- Quesada, J. Diego. 2000. *A grammar of Teribe*. Munich: Lincom Europa.
- Quirk, Randolph, and Greenbaum, Sidney. 1973. *A university grammar of English*. London: Longman.
- Leech, Geoffrey, and Svartvik, Jan. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Ramirez, Henri. 1992. *Le Bahuana: Une nouvelle langue de la famille Arawak*. Paris: Amerindia.
- Ramstedt, G. J. 1939. *A Korean grammar*. Oosterhout: Anthropological Publications.
- Randal, Scott. 1998. 'A grammatical sketch of Tennen', pp. 219–72 of *Surmic languages and cultures*, edited by Gerrit J. Dimmendaal and Macro Last. Cologne: Rüdiger Köppe.
- Rangan, K. 1979. *Purki grammar*. Mysore: Central Institute of Indian Languages.
- Rankin, Robert L. 1976. 'From verb to auxiliary to noun classifier and definite article: Grammaticalization of the Siouan verbs "sit", "stand", "lie"', pp. 273–83 of *Proceedings of the 1976 Mid-America Linguistic Conference*, edited by R. L. Brown, K. Houlihan, L. Hutchinson, and A. MacLeish. Minneapolis: University of Minnesota.
- 2004. 'The history and development of Siouan positionals, with special attention to polygrammaticalization in Dhegiha', *Sprachtypologie und Universalienforschung* 57: 202–27.
- Boyle, John, Graczyk, Randolph, and Koontz, John. 2002. 'Synchronic and diachronic perspective on "word" in Siouan', pp. 180–204 of Dixon and Aikhenvald 2002a.
- Ransom, Evelyn. 1986. *Complementation: Its meaning and forms*. Amsterdam: John Benjamins.
- Refsing, Kirsten. 1986. *The Ainu language: The morphology and syntax of the Shizunai dialect*. Aarhus: Aarhus University Press.
- Reh, Mechthild. 1996. *Anywa language: Description and internal reconstructions*. Cologne: Rüdiger Köppe.
- Rehg, Kenneth L. 1981. *Ponapean reference grammar*. Honolulu: University of Hawaii Press.
- Rennison, John R. 1997. *Koromfe*. London: Routledge.

- de Reuse, Willem P. 1994. *Siberian Yupik Eskimo: The language and its contacts with Chukchi*. Salt Lake City: University of Utah Press.
- Ribeiro, Eduardo R. 2006. 'Subordinate clauses in Karajá', *Boletim do Museu Paraense Emílio Goeldi, Ciências Humanas*, 1, 1: 17–47.
- Rice, Keren. 1989. *A grammar of Slave*. Berlin: Mouton de Gruyter.
- 1990. 'Predicting rule domains in the phrasal phonology', pp. 289–312 of *The phonology–syntax connection*, edited by Sharon Inkelas and Draga Zec. Chicago: University of Chicago Press.
- Roberts, John R. 1987. *Amele*. London: Croom Helm.
- Robins, R. H. 1957. 'Vowel nasality in Sundanese: A phonological and grammatical study', pp. 87–103 of *Studies in linguistic analysis*, by J. R. Firth et al. Oxford: Blackwell.
- Robinson, Stuart P. 1999. 'Tzotzil (dialect of Zinacantan): Grammar summary'. RCLT internal document, ANU.
- Rodrigues, Aryan D. 1997. 'Nominal classification in Karirí'. *Opción* 13: 65–79.
- Rose, Françoise. 2003. 'Morphosyntaxe de l'émérillon, langue tupi-guarani de Guyane française'. PhD thesis, Université Lumière Lyon 2.
- Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, Mass.: MIT Press.
- Rosetti, Alexandru. 1947. *Le Mot: Esquisse d'une théorie générale*, 2nd edition. Copenhagen: Munksgaard, and Bucharest: Institutul de Linguistică Română.
- Routenaa, Judy. n.d. 'Kamula grammar essentials (W. Province, PNG)'. MS. SIL, Ukarumpa, Papua New Guinea.
- Rowan, O., and Burgess, E. 1979. *Gramática Parecis*. Brasília: Summer Institute of Linguistics.
- Rubino, Carl. 1998. 'Ilocano'. RCLT internal document, ANU.
- Russell, Kevin. 1999. 'The "word" in two polysynthetic languages', pp. 203–21 of Hall and Kleinhenz 1999.
- Sadiqi, Fatima. 1986. *Studies in Berber syntax: The complex sentence*. Würzburg: Königshausen und Neumann.
- Sadock, Jerrold M. 1972. 'A conspiracy in Danish: Relative clause formation', pp. 59–62 of Peranteau, Levi, and Phares 1972.
- 1999. 'The nominalist theory of Eskimo: A case study in scientific self-deception', *International Journal of American Linguistics* 65: 383–406.
- Saeed, John I. 1993. *Somali reference grammar*, 2nd revised edition. Kensington, Md.: Dunwoody Press.
- 1999. *Somali*. Amsterdam: John Benjamins.
- Saltarelli, Mario. 1988. *Basque*. London: Routledge.
- Sandler, Wendy. 1999. 'Cliticisation and prosodic words in a sign language', pp. 223–54 of Hall and Kleinhenz 1999.
- 2000. 'The medium and the message: Prosodic interpretation of linguistic content in Israeli Sign Language', *Sign Language and Linguistics* 2: 187–215.
- Sands, Kristina. 2000. 'Complement clauses and grammatical relations in Finnish'. PhD thesis, ANU.

- Sankoff, Gillian, and Brown, Penelope. 1976. 'The origins of syntax in discourse: A case study of Tok Pisin relatives', *Language* 52: 631–66.
- Sapir, Edward. 1917. Review of C. C. Uhlenbeck, *Het identificeerend karakter der possessieve flexie in talen van Noord-Amerika*, in *International Journal of American Linguistics* 1: 86–90. [Reprinted as pp. 75–80 of *The collected works of Edward Sapir*, Vol. 5: *American Indian languages—1*, edited by William Bright. Berlin: Mouton de Gruyter, 1996.]
- 1921. *Language: An introduction to the study of speech*. New York: Harcourt Brace.
- 1922. 'The Takelma language of south western Oregon', pp. 1–296 of Boas 1922.
- 1930–1. 'The Southern Paiute language', *American Academy of Science, Boston Proceedings*, 65, 1–3.
- 1933. 'La Réalité psychologique des phonèmes', *Journal de psychologie normale et pathologique* 30: 247–65. [English version, 'The psychological reality of phonemes', pp. 46–60 of *Selected writings of Edward Sapir in language, culture and personality*, edited by David G. Mandelbaum. Berkeley and Los Angeles: University of California Press, 1949.]
- Sari, Nirmala. 1984. 'Banjarese verbal morphology'. MA thesis, ANU.
- Sasse, Hans-Jürgen. 1999. 'Cayuga: Grammar summary'. RCLT internal document, ANU.
- Schachter, Paul. 1973. 'Focus and relativization', *Language* 49: 19–46.
- 1985. 'Parts-of-speech systems', pp. 3–61 of Shopen 1985: Vol. 1.
- and Otnes, Fe T. 1972. *Tagalog reference grammar*. Berkeley and Los Angeles: University of California Press.
- Schebeck, Bernhard. 2001. *Dialect and social groupings in Arnhem Land*. Munich: Lincom Europa.
- Schegloff, Emanuel A., Jefferson, Gail, and Sacks, Harvey. 1977. 'The preference for self-correction in the organization of repair in conversation', *Language* 53: 361–82.
- Schmidt, Ruth L. 2003. 'Urdu', pp. 286–350 of *The Indo-Aryan languages*, edited by George Cardona and Dhanesh Jain. London: Routledge.
- Schuh, Russell G. 1998. *A grammar of Miya*. Berkeley and Los Angeles: University of California Press.
- Schultze-Berndt, Eva. 2000. 'Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language'. PhD thesis, Katholieke Universiteit Nijmegen.
- Schütz, Albert J. 1985. *The Fijian language*. Honolulu: University of Hawaii Press.
- Seiler, Hansjakob. 1983. *Possession as an operational dimension of language*. Tübingen: Gunter Narr.
- Seki, Lucy. 1990. 'Kamaiurá (Tupí-Guaraní) as an active-stative language', pp. 367–91 of *Amazonian linguistics: Studies in lowland South American languages*, edited by Doris L. Payne. Austin: University of Texas Press.
- 2000. *Gramática da Kamaiurá, língua Tupí-Guaraní de Alro Xingu*. Campinas: Editora da Unicamp.
- Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

- Shirjaev, E. N. 1979. 'Nazyvnoe predlozhenie' ['Naming sentence'], p. 152 of *Russkij jazyk. Enciklopedija* [Russian language: Encyclopaedia], by F. P. Filin. Moscow: Sovetskaja Enciklopedija.
- Shopen, Timothy. 1985. Editor of *Language typology and syntactic description*, 3 vols. Cambridge: Cambridge University Press.
- 2007. Editor of *Language typology and syntactic description*, 3 vols., 2nd edition. Cambridge: Cambridge University Press.
- Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
- Slater, Keith W. 2003. *A grammar of Mangghuer: A Mongolic language of China's Qinghai-Gansu Sprachbund*. London: RoutledgeCurzon.
- Smeets, Ineke. 1989. 'A Mapuche grammar'. PhD thesis, University of Leiden.
- Smith, Ian, and Johnson, Steve. 2000. 'Kugu Nganhcara', pp. 355–489 of Dixon and Blake 2000.
- Sneddon, James N. 1996. *Indonesian reference grammar*. Sydney: Allen and Unwin.
- Sohn, Ho-Min. 1994. *Korean*. London: Routledge.
- 1999. *The Korean language*. Cambridge: Cambridge University Press.
- 2004. 'The adjective class in Korean', pp. 223–41 of Dixon and Aikhenvald 2004.
- Sokolov, S. N. 1967. *The Avestan language*. Moscow: Nauka.
- Sprigg, R. K. 1955. 'The tonal system of Tibetan (Lhasa dialect) and the nominal phrase', *Bulletin of the School of Oriental and African Studies* 17: 133–53.
- Stassen, Leon. 1997. *Intransitive predicates*. Oxford: Clarendon Press.
- Stebbins, Tonya N. 2001. 'Sm'algyax (Coast Tsimshian)'. Handout for RCLT Local Workshop on 'Demonstratives and related items'.
- Forthcoming. *Mali (Baining) grammar*. Canberra: Pacific Linguistics.
- Stenson, Nancy. 1981. *Studies in Irish syntax*. Tübingen: Gunter Narr.
- Stumme, Hans. 1899. *Handbuch des Schilhischen von Tazerwalt*. Leipzig: J. C. Hinrichs.
- Sumbuk, Kenneth K. 1999. 'Morphosyntax of Sare'. PhD thesis, University of Waikato, New Zealand.
- Sutton, P. n.d. 'The Flinders Island language'. MS.
- Svorou, Soteria. 1993. *The grammar of space*. Amsterdam: John Benjamins.
- Swadesh, Morris. 1938. 'Nootka internal syntax', *International Journal of American Linguistics* 9: 77–102.
- 1946. 'Chitimacha', pp. 312–36 of *Linguistic structures of Native America*, by Harry Hoijer et al. New York: Viking Fund.
- Swanton, John R. 1911a. 'Haida', pp. 205–82 of Boas 1911a.
- 1911b. 'Tlingit', pp. 159–204 of Boas 1911a.
- Sweet, Henry. 1875/6. 'Words, logic and grammar', pp. 470–83 of *Transactions of the Philological Society for 1875/6*.
- 1891. *A new English grammar, logical and historical*. Oxford: Clarendon Press.
- 1898. *A short historical English grammar*. London: Oxford University Press.
- Szemerényi, Oswald J. L. 1996. *Introduction to Indo-European linguistics*. Oxford: Clarendon Press.
- Takeuchi, Lone. 1999. *The structure and history of Japanese: From Yamatokotoba to Nihongo*. London: Longman.

- Tamura, Suzuko. 2000. *The Ainu language*. Tokyo: Sanseido. [English translation of article 'Ainu-go' ['The Ainu language'] from *Gengogaku-dai-jiten* [*The Sanseido Encyclopaedia of Linguistics*], Vol. 1, Part 1, 1988.]
- Teeter, Karl V. 1964. *The Wiyot language*. Berkeley and Los Angeles: University of California Press.
- Terrill, Angela. 2003. *A grammar of Lavukaleve*. Berlin: Mouton de Gruyter.
- Thalbitzer, William. 1911. 'Eskimo', pp. 967–1069 of Boas 1911a.
- Thiesen, Wesley. 1996. *Gramática del idioma bora*. Pucallpa, Peru: Instituto Lingüístico de Verano.
- Thomas, David. 1955. 'Three analyses of the Ilocano pronoun system', *Word* 11: 204–8.
- 1971. *Chrau grammar*. Honolulu: University of Hawaii Press.
- Thompson, Chad. 1996. 'On the grammar of body parts in Koyukon Athabaskan', pp. 651–76 of Chappell and McGregor 1996.
- Thornell, Christina. 1997. *The Sango language and its lexicon*. Lund: Lund University Press.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.
- Topping, Donald M. 1973. *Chamorro reference grammar*. Honolulu: University of Hawaii Press.
- Tosco, Mauro. 1999a. 'Dhaasanac (Cushitic)'. RCLT internal document, ANU.
- 1999b. 'Somali (Cushitic)'. RCLT internal document, ANU.
- 2001. *The Dhaasanac language*. Cologne: Rüdiger Köppe.
- Travis, Catherine. 2002. 'Copula clauses in Spanish'. Handout for presentation at RCLT Local Workshop on 'Copula clauses and verbless clauses'.
- Trubetzkoy, N. S. 1969. *Principles of phonology*, translated by Christiane A. M. Baltaxe. Berkeley and Los Angeles: University of California Press.
- Tucker, A. N., and Mpaayei, J. T. O. 1955. *A Maasai grammar with vocabulary*. London: Longmans, Green.
- Ullmann, Stephen. 1957. *The principles of semantics*. Oxford: Blackwell and Glargow: Jackson.
- Valentine, J. Randolph. 2001a. 'Being and becoming in Ojibwe', *Anthropological Linguistics* 43: 431–70.
- 2001b. *Nishnaabemwin reference grammar*. Toronto: Toronto University Press.
- Valenzuela, Pilar. 2002. *Relativization in Shipibo-Konibo: A typologically-oriented study*. Munich: Lincom Europa.
- 2003. 'Transitivity in Shipibo-Konibo grammar'. PhD dissertation, University of Oregon.
- van den Berg, René. 1997. 'Spatial deixis in Muna (Sulawesi)', pp. 197–220 of *Referring to space: Studies in Austronesian and Papuan languages*, edited by Gunter Senft. Oxford: Clarendon Press.
- van Driem, George. 1993. *A grammar of Dumi*. Berlin: Mouton de Gruyter.
- van Eijk, Jan. 1997. *The Lilloet language: Phonology, morphology, syntax*. Vancouver: UBC Press.
- and Hess, Thom. 1986. 'Noun and verb in Salish', *Lingua* 69: 319–31.

- van Wyk, E. B. 1967. 'Northern Sotho', *Lingua* 17: 230–61.
- 1968. 'Notes on word autonomy', *Lingua* 21: 543–57.
- Variar, K. M. Prabhakara. 1979. *Studies in Malayalam grammar*. Madras: University of Madras.
- Velázquez-Castillo, Maura. 1996. *The grammar of possession: Inalienability, incorporation and possessor ascension in Guaraní*. Amsterdam: John Benjamins.
- Vendryes, Joseph. 1925. *Language: A linguistic introduction to history*, translated by Paul Radin. London: Routledge.
- Verhaar, John. 1967–72. Editor of *The verb 'be' and its synonyms: Philosophical and grammatical studies*, Part 1 (1967), Part 2 (1968), Part 3 (1968), Part 4 (1969), Part 5 (1972). Dordrecht: Reidel.
- Veselinova, Ljuba. 1999. 'Suppletion in the paradigms of copula verbs', pp. 65–83 of *Cognition and function in language*, edited by Barbara A. Fox, Dan Jurafsky, and Laura A. Michaelis. Stanford, Calif.: CSLI Publications.
- Vesper, Don R. 1968. 'A generative grammar of Kurukh copula', pp. 112–48 of Verhaar, Part 2, 1968.
- Vincennes, St Louis de, and Dallet, J. B. 1960. *Initiation à la langue berbère (Kabylie)*. Fort-National, Algeria: Fichier de Documentation Berbère.
- Vogt, Hans. 1940. *The Kalispel language: An outline of the grammar, with texts, translations and dictionary*. Oslo: I Kommissjon hos Jacob Dybwad.
- 1971. *Grammaire de la langue géorgienne*. Oslo: Universitetsforlaget.
- Walsh, Michael J. 1976. 'The Murinypata language'. PhD thesis, ANU.
- Waltke, Bruce K., and O'Connor, M. 1990. *An introduction to Biblical Hebrew syntax*. Winona Lake, Ind.: Eisenbrauns.
- Watanabe, Honoré. 2003. *A morphological description of Sliammon, Mainland Comox Salish, with a sketch of syntax*. Osaka: Endangered Languages of the Pacific Rim.
- Waterson, Natalie. 1956. 'Some aspects of the phonology of the nominal forms of the Turkish word', *Bulletin of the School of Oriental and African Studies* 18: 578–91.
- Watkins, Calvert. 1998. 'Proto-Indo-European: Comparison and reconstruction', pp. 25–73 of *The Indo-European languages*, edited by Anna G. Ramat and P. Ramat. London: Routledge.
- Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln: University of Nebraska Press.
- Watters, David E. 2002. *A grammar of Kham*. Cambridge: Cambridge University Press.
- 2006. 'The conjunct-disjunct in Kaike', *Nepalese Linguistics* 22: 300–19.
- Waugh, Linda R. 1977. *A semantic analysis of word order: Position of the adjective in French*. Leiden: Brill.
- Weber, David J. 1983. *Relativization and nominalized clauses in Huallaha (Huánuco) Quechua*. Berkeley and Los Angeles: University of California Press.
- 1989. *A grammar of Huallaga (Huánuco) Quechua*. Berkeley and Los Angeles: University of California Press.
- Weinreich, Uriel. 1954. 'Stress and word structure in Yiddish', pp. 1–27 of *The field of Yiddish*, 1, edited by Uriel Weinreich. New York.
- Wells, Rulon S. 1947. 'Immediate constituents', *Language* 23: 81–117.

- Welmers, W. E. 1973. *African language structures*. Berkeley and Los Angeles: University of California Press.
- and Welmers, Beatrice F. 1968. *Igbo: A learner's manual*. Los Angeles: the authors.
- ——— 1969. 'Noun modifiers in Igbo', *International Journal of American Linguistics* 35: 315–22.
- Wetzer, Harrie. 1992. '“Nouny” and “verby” adjectivals: A typology of predicate adjectival constructions', pp. 223–62 of *Meaning and grammar: Cross-linguistic perspectives*, edited by Michel Kefer and Johan van der Auwera. Berlin: Mouton de Gruyter.
- 1996. *The typology of adjectival predication*. Berlin: Mouton de Gruyter.
- Wiesemann, Ursula. 1986. 'The pronoun systems of some Jê and Macro-Jê languages', pp. 359–80 of *Pronominal systems*, edited by Ursula Wiesemann. Tübingen: Gunter Narr.
- Wilkins, David. 1999. '1999 demonstratives questionnaire: “THIS” and “THAT” in comparative perspective', pp. 1–24 of *'Manual' for the 1999 field season*, edited by David Wilkins. Nijmegen: Language and Cognition Group, Max Planck Institute.
- Wilkinson, Melanie. 2004. 'Complementation strategies in Djambarrupyu'. Handout for presentation at International Workshop on 'Complement clauses and complementation strategies' at RCLT.
- Willett, Thomas L. 1991. *A reference grammar of Southeastern Tepehuan*. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Williams, Corinne J. 1980. *A grammar of Yuwaalaraay*. Canberra: Pacific Linguistics.
- Williamson, Janis S. 1987. 'An indefiniteness restriction for relative clauses in Lakota', pp. 168–90 of *The representation of (in)definiteness*, edited by Eric J. Reuland and Alice G. B. ter Meulen. Cambridge, Mass.: MIT Press.
- Williamson, Kay. 1965. *A grammar of the Kolokuma dialect of Ijò*. Cambridge: Cambridge University Press.
- Willie, Mary Ann. 1989. 'Why there is nothing missing in Navajo relative clauses', pp. 407–37 of *Athabaskan linguistics: Current perspectives on a language family*, edited by Eung-Do Cook and Keren Rice. Berlin: Mouton de Gruyter.
- Wonderly, William L. 1951. 'Zoque II: Phonemics and morphophonemics', *International Journal of American Linguistics* 17: 105–23.
- Woodbury, Anthony C. 1986. 'Interactions of tense and evidentiality: A study of Sherpa and English', pp. 188–202 of *Evidentiality: The linguistic coding of epistemology*, edited by Wallace Chafe and Johanna Nichols. Norwood, NJ: Ablex.
- 2002. 'The word in Cup'ik', pp. 79–99 of Dixon and Aikhenvald 2002a.
- Xu Weiyaun. 1988. 'Capturing the adjective in Chinese'. MS. ANU.
- Yoshimura, Chikao. 1993. 'Possession in Tialo, Central Sulawesi'. Handout for presentation in Workshop on 'Possession', ANU.
- Young, Robert W., and Morgan, William, Sr. 1987. *The Navajo language: A grammar and colloquial dictionary*. Albuquerque: University of New Mexico Press.
- Zaenen, A., Maling, J., and Thráinsson, H. 1985. 'Case and grammatical functions: The Icelandic passive', *Natural Language and Linguistic Theory* 3: 441–83.
- Zandvoort, R. W. 1975. *A handbook of English grammar*. London: Longmans.

-
- Zeitoun, Elizabeth. 2007. *A grammar of Mantaauran (Rukai)*. Taipei: Institute of Linguistics, Academia Sinica.
- Zeshan, Ulrike. 2000. 'American sign language (ASL)'. RCLT internal document.
- . 2002. 'Towards a notion of "word" in sign languages', pp. 153–79 of Dixon and Aikhenvald 2002a.
- Žirmunskij, V. M. 1966. 'The word and its boundaries', *Linguistics* 27: 65–91.
- Zuckermann, Ghil'ad. 2006. 'Complement clause types in Hebrew', pp. 72–82 of Dixon and Aikhenvald 2006.
- Zwicky, Arnold M., and Pullum, Geoffrey K. 1983. 'Cliticization versus inflection: English *n't*', *Language* 59: 502–13.

Author Index

The indices cover both volume 1 (shown by 1:) and volume 2 (shown by 2:), followed by page numbers.

- Abbott, Miriam 1: 91; 2: 93–4, 400
 Abercrombie, David 1: 287
 Abraham, R. C. 2: 114
 Adelaar, Willem 2: 181, 184, 204, 286
 Aikhenvald, Alexandra Y. 1: 87, 180–1, 236, 238, 241, 255; 2: 31, 406
 on cases 1: 87; 2: 55, 120, 221
 on clitics 1: 222; 2: 21, 23
 on evidentiality 1: 56, 87, 180, 260, 319; 2: 113, 244, 259–60, 419
 on nominal classification 1: 158, 180, 238, 263, 318; 2: 243, 246, 259, 277, 368
 See also entries in the Index of languages on Baniwa of Içana, Bare, Manambu, Tariana, Tucano, Warekena and Zekkara.
 Aissen, Judith L. 1: 327
 Aitken, Percy 2: 271
 Akmajian, Adrian 2: 204
 Akpati, Elizabeth 2: 278
 Alford, Denny Keith 2: 278
 Allen, W. S. 1: 279; 2: 10, 12, 214, 259, 391, 305
 Alpher, Barry 2: 59, 97, 104
 Alvarez, José 2: 286
 Amberber, Mengistu 1: 180; 2: 344, 349
 Ameka, Felix 1: 54; 2: 36, 254, 280, 285, 288
 Andersen, Torben 2: 142
 Anderson, Stephen 2: 204, 213, 242–5
 Anderson, Victoria 1: 180
 Andrade, Manuel J. 1: 179; 2: 60
 Andrews, Avery D. 2: 368
 Andrews, Edna 1: 241
 Anschutz, A. 2: 312
 Arms, David G. 2: 108, 145
 Arnott, D. W. 2: 86
 Asher, R. E. 2: 91, 174, 180, 233, 251, 336, 357
 Ashton, E. O. 1: 85
 Aspinion, Robert 1: 54; 2: 89
 Atattürk, Kemal 1: 20, 55
 Austin, Peter 2: 157, 169, 181
 Avrorin, V. A. 2: 285
 Baker-Shenk, Charlotte 2: 339
 Ball, Martin J. 2: 12
 Bally, Charles 2: 5
 Bani, Ephraim 1: 55
 Barentsen, Adrian 2: 372
 Bargery, G. P. 2: 114
 Barnes, Janet 1: 55–6, 181
 Barshi, Immanuel 2: 311
 Baudouin de Courtenay, Jan 1: 180, 209
 Bauer, Laurie 1: 185, 241; 2: 61
 Bazell, C. E. 1: 241–2, 261; 2: 12
 Bazin, Louis 1: 55
 Beaumont, Clive H. 2: 80
 Beck, David 2: 84, 90
 Benjamin, Carmen 2: 176
 Benjamin, Geoffrey 2: 78
 Benveniste, Emile 1: 88, 241; 2: 174, 188, 203, 301, 312
 Bergslund, Knut 1: 241
 Berlin, Brent 1: 90, 256
 Besnier, Niko 2: 322, 349, 351
 Bhat, D. N. S. 2: 88, 200, 257, 260, 358, 368
 Bhatia, Tej K. 2: 158, 303, 358
 Bickel, Balthasar 2: 333, 349, 363
 Bing, Janet M. 2: 76
 Blackwell, Aleka A. 2: 76
 Blake, Barry J. 1: 308; 2: 368
 Blandford, F. G. 1: 90
 Bloch, Bernard 1: 288
 Bloomfield, Leonard 1: 85–88, 328; 2: 2, 10–11, 18–19, 32, 35, 37
 Boas, Franz 1: 84; 2: 64, 198, 243–4, 278–9, 281, 282, 288
 Bochner, H. 2: 17
 Bolinger, Dwight 1: 89–90; 2: 71
 Bontkes, Carolyn 2: 409
 Borer, Hagit 2: 325
 Borgman, Donald M. 2: 193
 Bowers, Claire 1: 241
 Boxwell, Maurice 2: 198
 Boyle, John 1: 353; 2: 456
 Breen, J. G. 2: 291
 Bright, William 2: 64
 Broschart, Jürgen 2: 54
 Brown, Cecil H. 1: 262

- Brown, D. Richard 1: 181
 Brown, Penelope 2: 343, 364
 Browne, W. 2: 336
 Bruce, Les 2: 86
 Bugenhagen, Robert D. 2: 231
 Burgess, E. 2: 205
 BurrIDGE, Kate 1: 18, 55; 2: 393, 405
 Burzio, Luigi 1: 91
 Buse, J. E. 2: 70
 Butt, John 2: 176
 Bužarovska, Eleni 2: 361
 Byarushengo, Ernest R. 2: 157
 Bybee, Joan 1: 259, 262; 2: 301
- Canger, U. 1: 326
 Carlin, Eithne B. 2: 197, 311
 Carlson, Robert 1: 55; 2: 236, 251, 343
 Carpenter, Kathie 2: 348
 Cawdrey, Robert 1: 179
 Ceria, Verónica G. 2: 243
 Černý, Václav 2: 158
 Chafe, Wallace 2: 205
 Chalker, Sylvia 1: 91
 Chao, Yuen-Ren 1: 88, 287; 2: 4, 30–1
 Chappell, Hilary 2: 233, 310
 Charney, Jean Ormsbee 1: 180
 Charters, Helen 2: 330
 Churchill, Winston S. 1: 64, 66
 Churchward, C. Maxwell 1: 90, 179–80;
 2: 25, 54, 233
 Clark, Eve V. 2: 61
 Clark, Herbert H. 2: 61
 Clements, George N. 2: 253
 Cokley, Dennis 2: 339
 Cole, Peter 2: 85, 333
 Comrie, Bernard 1: 55, 86; 2: 157–8, 170,
 254, 259
 on relative clauses 1: 89; 2: 320, 326, 330,
 337, 352, 369, 368
 Conklin, Harold C. 2: 196, 198, 258
 Cook, Eung-Do 1: 179, 220–1; 2: 242
 Cook, Kenneth W. 2: 126
 Corbett Greville G. 1: 87, 180; 2: 85, 368
 Corominas, Joan 2: 183
 Coulmas, Florian 2: 231, 233
 Craig, Colette Grinevald 1: 55, 91, 179;
 2: 292, 304, 390, 404, 418
 Crapo, Richley H. 2: 271
 Crazzolaria, J. P. 2: 177
 Creissels, Denis 2: 212, 259
 Crisp, Simon 2: 158, 213
 Cristofaro, Sonia 2: 369
 Croft, William 1: 241
 Crowell, Thomas H. 2: 81
- Crowley, Terry 2: 87, 363
 Culy, Christopher 2: 253
 Cummins, George M. III 2: 236
 Curnow, Timothy J. 2: 87, 178, 231, 274,
 300–1, 379, 403
 Cysouw, Michael 2: 258
 Cytowic, Richard E. 1: 90
 Czaykowska-Higgins, Ewa 2: 72, 77
- Daguman, Josephine S. 2: 114, 165, 237
 Dahl, Otto Christian 1: 54
 Dahlstrom, Amy 2: 251
 Dai, John Xiang-Ling 2: 35
 Dallet, J. B. 2: 167
 Davidson, Matthew 2: 61, 402
 Davies, John 2: 331, 344, 353
 Day, Christopher 1: 179
 Deane, P. 2: 312
 DeLancey, Scott 2: 334, 363
 Deloria, Ella 1: 84; 2: 199, 282–3, 288
 Demers, Richard A. 2: 38
 Dench, Alan C. 1: 250, 267, 276; 2: 8, 157
 Derbyshire, Desmond C. 2: 92, 271, 316,
 341
 Deutscher, Guy 2: 379, 385, 390, 393, 405–6,
 408–12, 418, 420
 Devitt, Dan 2: 188
 Di Sciullo, Anna-Maria 2: 35
 Dickens, Patrick 2: 260
 Diessel, Holger 1: 90; 2: 233, 236, 238, 242,
 245, 259–60
 Diffloth, Gerard 1: 90
 Dik, Simon C. 2: 418
 Dimmendaal, Gerrit J. 2: 97, 419
 Dirr, A. 1: 91
 Doke, C. M. 1: 302
 Donaldson, Tamsin 1: 180; 2: 213
 Donohue, Mark 2: 61, 82
 Douglas, Wilfrid H. 1: 262; 2: 291
 Dowling, Bruce T. 2: 368
 Dresher, Bezalel E. 2: 9
 Dryer, Matthew S. 1: 348; 2: 157
 Du Bois, John W. 1: 89
 Dunn, Michael John 1: 181
 Duponceau, Peter S. 1: 241, 260
 Duranti, Alessandro 2: 157
 Durie, Mark 1: 330; 2: 205, 258
 Dutton, Tom E. 2: 136, 194
- Eades, Diana 1: 90; 2: 113
 Eckert, Paul 1: 262
 Elkin, A. P. 1: 317
 Emenanjo, E. Nolue 2: 268, 340
 Emeneau, M. B. 2: 250

- Enfield, N. J. 1: 55, 83; 2: 82, 113, 236, 259, 329, 338, 347
 England, Nora C. 1: 82, 181; 2: 91
 Englebretson, Robert 2: 420
 Enrico, John 2: 140
 Escalante, Fernando 2: 211
 Evans, Nicholas D. 1: 55, 82, 179; 2: 36, 44–5, 49, 61, 208, 311
 Faverey, Margot 2: 176
 Feeling, Durbin 2: 77
 Fenyvesi, Anna 2: 11, 346, 353
 Ferrar, H. 2: 238
 Ferris, Connor 2: 114
 Filimonova, Elena 2: 258
 Firth, J. R. 2: 35
 Fleck, David W. 1: 262; 2: 204, 410, 419
 Fodor, István 1: 55
 Foley, William A. 1: 55, 63, 82, 181; 2: 9, 23–4, 169, 178, 199, 353, 393
 Forchheimer, Paul 2: 258
 Ford, Lysbeth J. 2: 87, 90, 215
 Fortescue, Michael 2: 243, 275, 304
 Fortune, Reo 2: 85
 Fox, Barbara A. 2: 322
 Frachtenberg, Leo J. 2: 37, 42, 64
 Frajzyngier, Zygmunt 1: 179; 2: 79, 82–3, 169, 231, 330, 342
 Franklin, Joice 2: 10
 Franklin, Karl J. 2: 10
 Freeland, L. S. 1: 84
 Fries, Charles C. 1: 89, 283
 Gak, V. G. 2: 35
 Galloway, Brent D. 2: 61
 Gamal-Eldin, Saad 2: 292, 353
 Gardiner, Alan H. 2: 183
 Gary, Judith Olmsted 2: 157, 292, 353
 Genetti, Carol 2: 95, 165, 175, 258
 on relative clauses 2: 319, 334, 342, 357, 363
 on complementation 2: 390, 402, 419
 George Madugu, Isaac S. 2: 74, 92
 Gerzenstein, Ana 2: 286
 Gill David 1: 348; 2: 445
 Giridhar, P. P. 2: 42, 54, 267
 Givón, Talmy 2: 63, 301, 338, 349, 369
 Glasgow, Kathleen 2: 157
 Glass, Amee 1: 262
 Goddard, Cliff 1: 262, 308
 Gorbet, Larry 2: 334
 Gordon, Lynn 1: 54, 181
 Gordon, Raymond G, Jr. 1: xiii, 347
 Graczyk, Randolph 1: 353; 2: 456
 Gragg, Gene 2: 227
 Gray, Louis H. 1: 301; 2: 9, 201
 Green, Diana 2: 244
 Greenbaum, Sidney 2: 245
 Greenberg, Joseph H. 1: 65, 72–3, 90–1, 241, 254; 2: 258
 Gregor, Douglas B. 2: 12
 Grimes, Barbara 1: 90
 Grimes, Charles E. 2: 270
 Guedes, Marymarcia 2: 8
 Guillaume, Antoine 2: 165, 181
 Guirardello, Raquel 1: 179
 Güldemann, Tom 2: 253, 419
 Haag, Marcia 2: 61
 Haas, Mary R. 1: 84; 2: 91, 201, 221–2, 258
 Hackett, Dorothy 1: 262
 Hagège, Claude 2: 81
 Haiman, John 2: 84, 91, 199, 243, 273, 285–6, 293, 368
 Hajdú, Péter 1: 110; 2: 42
 Hale, Horatio 1: 197
 Hale, Kenneth 1: 17, 55; 2: 358, 362
 Hall, T. Alan 2: 27
 Halle, Morris 1: 288
 Halliday, M. A. K. 2: 260
 Halpern, A. M. 1: 55
 Hansen, K. C. 1: 262
 Hansen, L. E. 1: 262
 Harada, S. I. 2: 350
 Harkins, Jean 1: 308
 Harris, Alice C. 2: 168
 Harris, Zellig S. 2: 4
 Harrison, Sheldon P. 2: 101, 390
 Hasan, Ruqaiya 2: 260
 Haspelmath, Martin 1: 180, 241, 260, 262; 2: 148, 258, 347
 Haviland, John 1: 90, 308; 2: 158, 260, 391
 Hayward, Richard J. 2: 170, 237, 243
 Hazlewood, David 2: 25
 Heath, Jeffrey 1: 181; 2: 89, 201, 209
 Hébert, Yvonne M. 2: 38, 61
 Heffernan, J. A. 2: 211
 Heine, Bernd 1: 180; 2: 188, 301–2, 310, 369, 421
 Hellwig, Birgit 2: 419
 Henadeerage, Kumara 2: 326
 Henderson, E. J. A. 1: 283
 Henderson, John 2: 9, 23, 25
 Hengeveld, Kees 2: 188, 418
 Hercus, Luise A. 2: 182, 194
 Hess, Thom 2: 38, 51, 53, 56

- Hewitt, B. G. 2: 326, 344
 Hildebrandt, Kristine 2: 95
 Hill, Archibald A. 1: 65, 90
 Hill, Deborah 1: 55; 2: 252, 386
 Himmelmann, Nikolaus P. 2: 238
 Hinds, John 2: 259
 Hinton, Leanne 1: 90
 Hock, Hans H. 1: 90; 2: 137
 Hockett, Charles F. 1: 146, 180; 2: 4, 69
 Hoffmann, Carl 2: 197
 Hoffmann, J. 2: 43–4
 Hoijer, Harry 2: 60, 215
 Holisky, Dee Ann 1: 91; 2: 121, 141, 157
 Hopper, Paul J. 1: 89; 2: 143, 149, 156, 409
 Horne, Kibble M. 1: 241
 Hosokawa, K. 1: 56
 Hovdhaugen, Even 2: 45
 Huang, Chenglong 2: 79, 82–3, 114, 165, 173
 Huber, Brigitte 2: 334
 Huddleston, Rodney 1: 185; 2: 61, 360
 Hudson, Joyce 1: 262; 2: 8
 Hurch, Bernhard 1: 180, 262
 Hyman, Larry M. 1: 288; 2: 157, 254, 278, 311
 Hymes, Dell 1: 343, 348–9
 Hyslop, Catriona 2: 45, 82, 113, 406, 408

 Ikoro, Suanu 1: 55; 2: 136, 178, 192, 254, 299
 Imedadze, Natela 2: 251
 Ingram, David 2: 258
 Isačenko, Alexander V. 2: 299, 301

 Jacob, Judith M. 2: 231
 Jacobsen, William, Jr. 1: 179; 2: 43, 51, 53, 258
 Jaggar, Philip J. 2: 352
 Jake, Janice 1: 66, 90
 Jakobson, Roman 1: 70, 88, 90, 241, 287
 Jany, Carmen 2: 405, 410
 Jarkey, Nerida 2: 379, 385–6, 393, 398, 404
 Jastrow, Otto 2: 11
 Jauncey, Dorothy 2: 217, 299
 Jefferson, Gail 2: 36
 Jelinek, Eloise 2: 38
 Jerry, Bessie 1: 20
 Jespersen, Otto 1: 87–8; 2: 68, 114, 250, 360
 Jeyapul, V. Y. 2: 269
 Johanson, Lars 2: 88
 Johansson, Stig 2: 71, 312
 Johns, Brenda 2: 176
 Johnson, Steve 2: 213
 Jones, Daniel 1: 273, 326; 2: 28
 Jones, Morris 2: 302

 Joos, Martin 1: 344, 349, 353–4
 Joseph, Brian D. 2: 10, 26, 178, 201, 390
 Jungbluth, Konstanze 2: 260

 Kachru, Yamuna 2: 174, 177
 Kautzsch, E. 2: 369
 Kay, Paul 1: 256
 Keenan, Edward L. 1: 55; 2: 157, 204, 242–5
 on relative clauses 1: 89; 2: 320, 324, 326, 330, 336, 350, 361
 Keesing, Roger M. 1: 262
 Kenesei, István 2: 11, 346, 353
 Key, Harold 1: 180, 262
 Khaidakov, S. M. 1: 55; 2: 242
 Kibrik, Aleksandr E. 1: 262
 Kilian-Hatz, Christa 1: 362
 Kilby, David 2: 158
 Kimball, Geoffrey D. 1: 82, 180; 2: 83, 200, 233, 316
 Kinkade, M. Dale. 2: 38
 Kiparsky, Paul 2: 158
 Kita, Sotar 2: 260
 Klammer, Marian 1: 55; 2: 352
 Kleinhenz, Ursula 2: 27
 Kleinschmidt, S. 2: 44
 Klokeid, Terry 1: 55
 Kockelman, Paul 2: 386
 Kodzasov, Sandro V. 1: 262
 Koehn, Edward 2: 274
 Koehn, Sally 2: 274
 König, Christa 2: 120
 Koontz, John 1: 353; 2: 456
 Kornfilt, Jaklin 2: 271
 Koshal, Sanyukta 2: 291
 Krámský, Jiří 2: 34
 Kratochvíl, František 2: 240
 Krishnamurti, Bh. 1: 308; 2: 324
 Kroeber, Paul D. 2: 61, 224
 Kroeger, Paul 2: 56
 Kruspe, Nicole 1: 82; 2: 187
 Kruszewski, Nikolaï 1: 180
 Kuipers, Aert H. 1: 288; 2: 61, 401
 Kulemeka, Andrew T. 1: 302
 Kullavanijaya, Pranee 2: 364
 Kumar, Ann 1: 90
 Kumar, Rajesh 2: 317
 Kumari, T. C. 2: 91, 336, 357
 Kuno, Susumu 2: 326, 350, 364
 Kuroda, S.-Y. 2: 337, 360
 Kuteva, Tania 2: 188, 301, 369, 421

 Laanest, Arvo 1: 181
 Ladefoged, Peter 1: 87, 274, 258

- Lakoff, George 1: 55
 Landar, Herbert 1: 240
 Lang, Adrienne 1: 262
 Langacker, Ronald W. 1: 56
 Langendoen, D. Terence 2: 178
 LaPolla, Randy J. 2: 79, 82–3, 114, 165, 258
 Launey, Michel 2: 61
 Lawler, John M. 1: 330
 Lee, Jennifer 1: 55, 205; 2: 199
 Leech, Geoffrey 2: 456
 Leeding, Velma J. 2: 291
 Leer, Jeff 2: 271
 Lehiste, Ilse 2: 9
 Lehmann, Christian 2: 368
 Lehmann, Thomas 2: 357
 Leslau, Wolf 1: 181; 2: 364
 Levi, Judith N. 2: 455
 Levin, Beth 2: 158
 Levinson, Stephen C. 2: 249, 261
 Levy, Paulette 2: 77, 80, 84, 88
 Lewis, Geoffrey 2: 64
 Li Xinyin 2: 245
 Li, Charles N. 2: 69, 183, 268, 331, 337, 342
 Li, Feng 2: 183
 Lichtenberk, Frantisek 2: 205, 259
 Liddell, Scott K. 2: 339
 Liddicoat, Tony 2: 273
 Lindau, Mona 1: 288
 Lindström, Eva 2: 238
 Longacre, Robert E. 2: 34
 Loogman, Alfons 1: 85
 Lorimer, D. L. R. 1: 90
 Lynch, John 1: 351; 2: 277
 Lyons, John 1: 86, 234; 2: 5, 18, 69, 190, 229, 246, 261
 Lysvåg, Per 2: 71
 Lytkin, V. I. 1: 181

 Maddieson, Ian 1: 87, 262, 274, 288
 Mahajan, Anoop 2: 358
 Mahootian, Shahrzad 2: 352, 360
 Majid, Asifa 1: 55
 Maling, J. 2: 148, 158
 Malinowski, Bronislaw 1: 291, 298
 Mallinson, Graham 2: 201, 353, 360, 368
 Malotki, Eckehart 2: 340
 Mamiani, Luis Vincencio 2: 277
 Maring, Joel M. 2: 217, 258
 Marsh, James Lewis 1: 262
 Martin, Jack 2: 280, 287

 Martin, Laura 1: 65, 90
 Masica, Colin 2: 357
 Matisoff, James A. 1: 64, 81; 2: 243, 363
 Matsumoto, Yoshiko 2: 349
 Matthews, P. H. 1: 86, 241; 2: 1–2, 7, 16, 19, 160, 369
 Matthews, Stephen 2: 180, 192
 Maybury-Lewis, David 1: 328
 Mazaudon, Martine 2: 334, 363
 McCarthy, John J. 2: 18
 McConvell, Patrick 2: 197
 McGregor, William 2: 310
 McGuckin, Catherine 1: 54; 2: 281–2, 290
 McKay, Graham R. 2: 196, 293–4, 359
 Meillet, Antoine 1: 25, 88; 2: 9
 Meinhof, Carl 2: 259
 Meira, Sérgio 2: 92, 311
 Mejlanova, U. A. 2: 158
 Merlan, Francesca C. 1: 82, 277; 2: 198, 291
 Meyerstein, Zlata P. 2: 236
 Michael, Ian 2: 247
 Migeod, F. W. H. 2: 114
 Mikkelsen, Line 2: 187
 Miller, Amy 2: 324
 Miller, Jim 1: 91
 Miller, Wick R. 2: 244, 258
 Milner, G. B. 2: 24, 37, 65, 108, 112
 Mitchell, T. F. 1: 180
 Mithun, Marianne 1: 89, 91; 2: 60, 140, 198, 204–5, 257, 284, 311
 on demonstratives 2: 239, 245
 on Tuscaraora 2: 239, 293, 300
 Moliner, Maris 2: 183
 Monserrat, Ruth 2: 268
 Morgan, William, Sr. 2: 343
 Morphy, Frances 1: 288
 Mosel, Ulrike 2: 45, 61
 Mous, Maarten 2: 200, 216, 221, 242
 Mpaayei, J. T. O. 2: 86
 Munro, Pamela 2: 167, 333, 337, 350, 352
 Muravyova, Irina A. 1: 262
 Murray, L. 1: 80

 Nababan, P. W. J. 2: 82
 Nagaraja, K. S. 2: 42, 50
 Nakayama, Toshihide 2: 51–2
 Nater, H. F. 2: 61
 Nedjalkov, Igor V. 2: 16, 303
 Nespor, Marina 2: 11, 23, 27, 34–5
 Newman, John 1: 308; 2: 188
 Newman, Paul 1: 60; 2: 227, 352

- Newman, Stanley 2: 2, 22, 103
 Ng, Eve 2: 188
 Nguyễn, Đình-Hòa 2: 78
 Nichols, Johanna 1: 90; 2: 88, 96, 148, 158, 285, 310–11
 Nida, Eugene A. 1: 81, 85, 179
 Nikiforidou, Kiki 2: 263, 265
 Noonan, Michael 1: 54, 91, 288; 2: 125, 231, 280, 286, 288, 303, 418, 420
 Nordlinger, Rachel 2: 53, 217, 221
 Obata, Kazuko 2: 382
 Ochs, Elinor 1: 55
 O'Connor, M. 2: 369
 Ohala, John J. 1: 90
 Okell, John 2: 340
 Oksefjell, S. 2: 312
 Olawsky, Knut J. 2: 92, 178, 182, 299
 Omorey, Thomas O. 2: 109–10
 Onishi, Masayuki 1: 55, 128, 330; 2: 149, 158, 178, 195, 244, 292, 300, 379
 Osada, Toshiaki 2: 44–5, 49, 61
 Osam, E. Kweku 2: 114
 Osborne, C. R. 1: 205
 Otanes, Fe T. 2: 52–3, 56
 Otsuka, Yuko 1: 180
 Overall, Simon 2: 251
 Owens, Jonathan 2: 167, 299
 Packard, Jerome L. 2: 31, 35
 Pagliuca, William 1: 259; 2: 301
 Palmer, F. R. 1: 56, 288; 2: 84, 88
 Palmer, H. E. 1: 90
 Pāpini 1: 45; 2: 68
 Paris, Catherine 2: 158
 Parsons, F. W. 2: 107
 Pascual, Joan 2: 183
 Pastika, Wayan 2: 386
 Pawley, Andrew 2: 277
 Payne, David L. 1: 180
 Payne, Doris L. 1: 241; 2: 311
 on Panare 2: 246, 279, 288, 292, 329, 347, 379
 on Yagua 2: 200, 241, 272, 292
 Payne, Thomas E. 1: 327
 on Panare 2: 246, 279, 288, 292, 329, 347, 379
 on Yagua 2: 200, 241, 272, 292
 Pensalfini, Robert J. 2: 167
 Peranteau, Paul M. 2: 448
 Perkins, Revere 1: 259; 2: 301
 Perlmutter, David M. 2: 155
 Pesetsky, David 1: 91
 Phares, Gloria C. 2: 448
 Philippaki-Warbuton, Irene 2: 10, 26, 178, 201, 390
 Piau, Julie 2: 200
 Pike, Kenneth L. 1: 88–9, 201, 283, 287–8; 2: 25
 Plaatje, Solomon T. 1: 326
 Plank, Frans 1: 55
 Platero, Paul R. 2: 327, 332–3
 Platt, John T. 1: 262
 Pope, M. K. 2: 227
 Poppe, Nikolai N. 2: 85, 88
 Post, Mark 2: 65
 Poulos, George 2: 87
 Press, Margaret L. 2: 81
 Prista, Alexander da R. 2: 22
 Protagoras 1: 155
 Pullum, Geoffrey K. 1: 81, 94, 106; 2: 35, 342, 360–1
 Pustet, Regina 2: 187, 198, 288, 300
 Quesada, J. Diego 2: 90
 Quirk, Randolph 2: 245, 361
 Ramirez, Henri 2: 205
 Ramstedt, G. J. 2: 83
 Randal, Scott 2: 167, 172
 Rangan, K. 2: 203
 Rankin, Robert 1: 180; 2: 35, 243
 Ransom, Evelyn 2: 418
 Rappaport Hovav, Malka 2: 158
 Refsing, Kirsten 1: 84; 2: 226
 Reh, Mechthild 1: 180; 2: 178
 Rehg, Kenneth L. 2: 200, 231
 Reichard, Gladys A. 1: 90
 Reichmann, W. J. 1: 257–9, 262
 Rennison, John R. 2: 178
 Reuse, Willem P. de 2: 291
 Ribeiro, Eduardo R. 2: 340, 373
 Rice, Keren 2: 12, 199, 279, 287
 Roberts, John R. 1: 53, 56, 181; 2: 84, 199, 279, 288, 293, 352
 Robins, R. H. 1: 90; 2: 10
 Robinson, Stuart P. 2: 391
 Rodrigues, Aryon D. 2: 277
 Rose, Françoise 2: 114
 Rose, Phil 1: 90
 Rosenbaum, Peter S. 2: 419
 Rosetti, Alexandru 2: 31, 34
 Roth, Walter E. 1: 181
 Routenaa, Judy 2: 114
 Rowan, O. 2: 205
 Rubino, Carl 1: 262; 2: 321

- Rumsey, Alan 1: 262, 288, 308
 Rusell, Kevin 2: 18
- Sacks, Harvey 2: 36
 Sadiqi, Fatima 2: 89
 Sadler, Louisa 2: 53, 221
 Sadock, Jerrold M. 2: 44, 242
 Saeed, John I. 2: 66, 114, 200
 Saltarelli, Mario 2: 84, 203, 213, 251
 Sândalo, Filomena 2: 243
 Sandler, Wendy 2: 12
 Sands, Kristina 2: 392
 Sankoff, Gillian 2: 343, 364
 Sapir, Edward 1: 84–8, 226, 265–6
 on morphological processes 1: 138–9, 146
 on possession 1: 54; 2: 310
 on Southern Paiute 1: 84, 265; 2: 99, 197
 on Takelma 1: 84; 2: 89, 218
 on word 2: 1, 31, 35
 on word classes 2: 37–8, 54
 Sari, Nirmala 2: 134
 Sasse, Hans-Jürgen 2: 217, 251
 Saussure, Ferdinand de 1: 25, 87
 Schachter, Paul 2: 37, 52–3, 56, 62, 69, 81, 321
 Schebeck, Bernhard 2: 196
 Schegloff, Emanuel A. 2: 36
 Schleicher, August 1: 180
 Schmidt, Ruth L. 2: 347
 Schuh, Russell G. 2: 201, 360
 Schultze-Berndt, Eva 2: 217
 Schütz, Albert J. 2: 108, 112
 Seiler, Hansjakob 2: 310–11
 Seki, Lucy 2: 70, 78, 80, 213, 342
 Shibatani, Masayoshi 2: 98
 Shirjaev, E. N. 2: 187
 Shopen, Timothy 1: 86, 318
 Siewierska, Anna 1: 246; 2: 258
 Silverstein, Michael 1: 179
 Sinclair, John 1: 307
 Skorik, P. Ja. 1: 181
 Slater, Keith W. 2: 179, 329
 Smeets, Ineke 2: 229, 242
 Smith, Ian 2: 213
 Smith, Norval 2: 259, 337, 360
 Sneddon, James Neil 1: 181, 2: 181, 231, 241, 268, 292
 Sohn, Ho-min 1: 55, 77, 82, 103; 2: 333, 338, 352
 Sokolov, S. N. 2: 85
 Sprigg, R. K. 2: 10
 Staden, Miriam van 1: 55
 Stankiewicz, Edward 1: 180, 209
 Stassen, Leon 2: 187
 Stebbins, Tonya N. 2: 114, 120, 157, 217
 Stenson, Nancy 1: 128; 2: 373, 386, 393, 402
 Stumme, Hans 2: 246
 Subbarao, Karumuri V. 2: 317
 Sumbuk, Kenneth K. 2: 114
 Sutton, P. 2: 218
 Svartvik, Jan 2: 456
 Svorou, Soteria 2: 311
 Swadesh, Morris 1: 88, 179, 215, 240, 287, 299; 2: 37, 43, 56, 90
 Swanton, John R. 2: 64
 Sweet, Henry 1: 135; 2: 1–2, 19, 36, 61, 227, 315, 359
 Szemerényi, Oswald J. L. 2: 364
- Takeuchi, Lone 2: 94
 Tamura, Suzuko 1: 179; 2: 166, 205, 226, 240, 290
 Taumoeofalau, Malenaite 1: 179–80
 Teeter, Karl V. 2: 286–7
 Tepljashina, T. I. 1: 181
 Terrill, Angela 2: 261
 Thalbitzer, William 2: 44
 Thiesen, Wesley 2: 201
 Thomas of Erfurt 2: 68
 Thomas, Alan R. 2: 302
 Thomas, David 2: 258, 302
 Thompson, Chad 2: 281–2, 301, 310
 Thompson, Sandra A. 1: 89; 2: 69, 143, 149, 156, 183, 268, 331, 337, 342
 Thomsen, Marie-Louise 1: 84
 Thornell, Christina 2: 183
 Thráinsson, H. 2: 148
 Timberlake, Alan 2: 170
 Topping, Donald M. 2: 78
 Tosco, Mauro 2: 299, 301, 344, 352
 Toulmin, Stephen E. 1: 1
 Traugott, Elizabeth C. 2: 409
 Travis, Catherine 2: 168, 175, 183
 Trubetzkoy, Nicolai S. 1: 70, 85–8, 235–6, 272, 287; 2: 8–11;
 Tucker, A. N. 2: 86
 Tuite, Kevin 2: 251
- Ullmann, Stephen 2: 34
 Vago, Robert M. 2: 11, 346, 352
 Valentine, J. Randolph 2: 181, 321, 384
 Valenzuela, Pilar 2: 165, 322, 333
 van den Berg, René 2: 244, 252
 van Driem, George 2: 187
 van Eijk, Jan P. 2: 38, 51, 53, 56, 61, 241, 244

- van Wyk, E. B. 2: 6, 17
 Variar, K. M. 2: 91, 183
 Vázolyi, Eric G. 1: 262
 Velazquez-Castillo, Maura 2: 302, 310–11
 Vendryes, Joseph 2: 6–7, 10
 Verhaar, John 2: 187
 Veselinova, Ljuba 2: 187
 Vesper, Don R. 2: 178
 Vincennes, St Louis de 2: 187
 Vitale, Anthony J. 1: 85
 Voegelin, Charles F. 1: 88
 Voeltz, F. K. Erhard 1: 302
 Vogel, Irene 2: 11, 23, 27, 34–5
 Vogt, Hans 1: 241; 2: 53, 61, 346
 von Roncador, Manfred 2: 419

 Walsh, Michael J. 2: 332
 Waltke, Bruce K. 2: 369
 Watanabe, Honoré 2: 61
 Waterson, Natalie 2: 11
 Watkins, Calvert 2: 364
 Watkins, Laurel J. 2: 217, 283, 381
 Watters, David E. 1: 82–3; 2: 63, 125, 222, 319, 408
 Waugh, Linda R. 2: 114
 Weber, David J. 2: 251, 384, 409
 Weiner, Edmund 1: 91
 Weinreich, Uriel 2: 10, 34
 Wells, Rulon S. 2: 2
 Welmers, Beatrice F. 2: 81, 84, 114
 Welmers, William E. 2: 81, 84, 114
 Werth, Elizabeth 1: 90
 Westermann, Diedrich 1: 180

 Wetzter, Harrie 2: 114
 Whorf, Benjamin Lee 1: 180
 Wiesemann, Ursula 2: 201
 Wierzbicka, Anna 1: 308
 Wilkins, David 2: 259–60
 Wilkinson, Melanie 2: 412
 Willett, Thomas L. 2: 243
 Williams, Corrinne J. 2: 180
 Williams, Edwin 2: 35
 Williamson, Janis S. 2: 344
 Williamson, Kay 2: 84
 Willie, Mary Ann 2: 327, 332
 Wonderly, William L. 2: 8
 Wongkhomthong, Preya 2: 364
 Woodbury, Anthony C. 2: 19, 21, 23, 222
 Wordick, F. J. F. 1: 288
 Wouk, Fay 2: 176

 Xu Weiyaun 2: 69, 78–9, 81, 83

 Yip, Virginia 2: 180, 192
 Yoshimura, Chikao 2: 272
 Young, Robert W. 2: 343

 Zaenen, A. 2: 148
 Zandvoort, R. W. 2: 245
 Zeitoun, Elizabeth 1: 83; 2: 382, 387, 405
 Zeshan, Ulrike 2: 12, 339
 Zhu Deni 1: 179
 Žirmunskij, V. M. 2: 1
 Zuckermann, Ghil'ad 2: 373, 379, 385, 393, 399, 403, 405
 Zwicky, Arnold M. 2: 35

Language Index

- Abaza 2: 213, 305, 326, 345
 Abkhaz 2: 135, 326, 345
 Abui 2: 240
 Acehnese 1: 347, 359; 2: 135, 205, 268, 272
 Acoma 2: 217–18, 258
 Adjnyamathanha 2: 213
 African languages 1: 155, 302; 2: 29, 74, 256, 259, 301
 Aguaruna 2: 251
 Ainu 1: 84; 2: 96, 205, 207, 225, 379
 copula clauses 1: 101, 179; 2: 166–8
 demonstratives 2: 240–1, 250
 possession 2: 290, 293, 301
 Akan 1: 10–11, 55; 2: 75–6, 114, 191, 277
 Akkadian 2: 379, 385, 390, 393, 405, 408–12, 420
 Alamlak 2: 86
 Aleut 2: 271, 292, 305
 Algonquian languages 1: 227; 2: 96, 204
 Alutor 1: 253, 262
 Amele 2: 84, 199, 352
 possession 1: 5–6, 11–12, 54; 2: 278–9, 283, 285, 288, 293, 298
 tense and polarity 1: 47, 56, 162–3, 181
 American sign language 2: 339
 Amharic 1: 140, 163, 180–1; 2: 213, 344, 349
 Angami 2: 267, 270, 272
 Anindilyakwa 2: 291
 Anywa 1: 140, 186; 2: 178
 Apalai 2: 274, 311
 Arabana 2: 182
 Arabic 1: 142, 180, 196; 2: 84, 86, 292, 321, 35
 also see Classical Arabic
 Arapesh 2: 86
 Arawak language family 1: 15, 259, 303; 2: 96, 204–5, 241, 311
 Archi 2: 151
 Arrente 2: 9, 23, 25
 Athapaskan 2: 204
 Australian languages 1: 45, 165, 180, 223, 296, 308; 2: 31, 69, 97, 157, 182, 283, 291, 399
 phonology 1: 210–11, 250, 270, 283, 288; 2: 8
 pronouns 2: 194, 196, 213, 219, 257–9
 Avar 2: 151, 212
 Avestan 2: 85
 Awa Pit 2: 87, 178, 231–2, 272, 274, 310–11, 403
 Axininca Campa 1: 159, 189
 Aymara 2: 197
 Baagandji 2: 194
 Bahnar 1: 90
 Bahuana 2: 205
 Balto-Finnic languages 1: 165
 Bandjalang 2: 87
 Baniwa of Içana 2: 158, 276
 Banjarese 2: 134
 Bantu languages 1: 226, 263; 2: 5, 137, 157
 Bardi 1: 231–2, 241, 246
 Bare 2: 9
 Basque 2: 84, 97, 180, 203, 213, 239, 250–1, 277, 379
 Batsbi (Tsova-Tush) 1: 73, 91; 2: 121, 141, 157
 Bella Coola 2: 454
 Bengali 1: 17, 55, 161, 180; 2: 8, 130, 150, 202, 234, 244, 292, 358
 Berber 1: 52; 2: 63, 91, 106
 Biblical Hebrew 2: 9, 369
 Bilin 2: 84, 88
 Bilua 2: 382, 384
 Blackfoot 2: 311
 Bolivian Quechua 2: 271
 Bora 2: 201
 Bororo 2: 81
 Boumaa Fijian, *see* Fijian
 Brahui 2: 260
 Burarra (Gun-Nartpa) 2: 157
 Buriat 2: 85, 88
 Burmese 2: 135
 Buru 2: 270
 Burushaski 1: 69–70, 90; 2: 97, 141
 Caddo 2: 205
 Caddoan languages 2: 140
 Cairene Egyptian Colloquial Arabic 2: 292, 353
 Cambodian, *see* Khmer
 Cantonese 2: 180, 192
 Carib languages 2: 92–4, 197, 311

- Cashinawa 1: 123–4
 Cavineña 2: 165, 181
 Cayuga: 2: 60, 217–8, 250–1
 Celtic languages 2: 12
 Central Pomo 2: 140
 Chadic languages 2: 197
 Chamorro 2: 78
 Chechen 2: 8, 148, 151
 Chemehuevi 2: 81
 Cherokee 1: 65, 90; 2: 77–8
 Chimariko 2: 410
 Chipweyan 1: 115, 179; 2: 191
 Chinese, *see* Mandarin Chinese, Cantonese
 Chitimacha 2: 90
 Chrau 2: 302
 Chukchi 1: 165, 181
 Classical Arabic 2: 9
 also see Arabic
 Classical Armenian 2: 141
 Classical Greek 2: 263, 265
 also see Greek
 Classical Nahautl 2: 61
 Coast Timsian 2: 217
 Colloquial Arabic 1: 142, 196; 2: 292, 353
 Colloquial Burmese 2: 340, 364
 Colloquial Czech 2: 236
 Colloquial Welsh 2: 183, 302
 Comanche 1: 159, 180
 Coos 2: 37, 42, 64
 Cree 2: 18
 Creek 2: 278, 280, 283–7, 293
 Cup'ik Eskimo 2: 19, 21, 23, 44
 Cushitic languages 2: 300
- Dagbani 2: 92, 299
 Dakota 1: 84, 141, 180; 2: 278, 281–4, 288–9
 also see Lakota
 Danish 2: 342, 364
 Dëne Suñiné, *see* Chipewyan
 Dhaasanac 2: 299–300, 343–4, 352
 Dhalanji 2: 156
 Diegueño 2: 334
 Diyari 2: 169, 181
 Dhamindhung 2: 217
 Djambarrpuynu 2: 412
 Djapu (a Yolngu dialect) 1: 280, 288
 Djeebbana, *see* Ndjébbana
 Djingulu 2: 167–8
 Dolakha Newar 2: 165, 175, 258, 342, 357,
 390, 404, 414
 Dravidian languages 1: 305; 2: 234, 260
 Dumi 2: 187
 Dutch 2: 6
- Dyrbal 1: 19–20, 61, 68–9, 82–3, 96, 133,
 144–8, 172–7; 221–2; 2: 22, 206–7, 238–9
 adjectives 1: 26; 2: 63, 85–8, 99–104
 antipassive 1: 148, 165–8, 172–3, 188–9,
 234
 avoidance style (Jalnguy) 1: 215–16,
 293–5, 308
 cases and genitive 1: 44–5, 126–8, 188–91,
 213
 complementation strategies 2: 351, 398,
 400–13
 constituent order and word order
 1: 22–4, 233–4
 demonstratives 2: 229, 237, 242, 244, 248,
 250
 derivational suffixes 1: 41, 219–20;
 2: 14–17
 lexicon 1: 22–3, 256, 283, 290–1, 300–8,
 344
 locational suffixes 1: 16–17, 69; 2: 320–1
 noun classes 1: 27–31, 55, 238; 2: 85–6
 phonology 1: 198–200, 210–12, 265–6,
 272–3, 278, 283; 2: 30
 possession 1: 5–6; 2: 268–9, 272, 275
 reduplication 1: 140, 294; 2: 56, 87
 relative clauses 2: 319–24, 334, 341–2, 349,
 351–2, 364, 407–8
 structure/word class correspondence
 1: 110; 2: 41
 transitivity 2: 121–2, 131, 134, 150, 157
- East Tucanoan 2: 89
 Èdo 2: 109–10, 126
 Egyptian 2: 183
 Egyptian Colloquial Arabic 1: 142, 180;
 2: 321
 Emerillon 2: 114
 Emmi 2: 87, 90, 97–8
 Enga 1: 257, 262
 English 1: 22–3, 76, 79, 107–8, 118–22, 137–9,
 172–6, 225, 241, 255, 283, 301, 305;
 2: 27–30
 adjectives 1: 52–3, 113, 240, 303; 2: 70–1,
 76, 80, 90, 99, 104, 114
 clause linking: 1: 133–7, 228
 comparatives 1: 177–9, 201–2
 complement clauses 1: 30–34, 55, 97,
 129–32; 2: 370–90, 394–401, 408, 411,
 417–20
 constituent order 1: 37–8, 72–3
 copula constructions 1: 100–1; 2: 163,
 170–4, 177, 180, 183, 186–7
 definiteness 1: 160–1, 180

English (*cont.*)

- demonstratives 2: 223–8, 231–6, 245–52, 260
 - derivational processes 1: 149–52, 185, 217–19; 2: 16, 46–9
 - irregular forms 1: 15, 140, 205–6
 - number marking 1: 143, 158, 184–5, 302; 2: 55
 - orthography 1: 67, 90, 281
 - passive 1: 166–8, 231
 - phrasal verbs 1: 36–7, 290
 - phonology 1: 9, 264–6, 272, 277
 - possession 1: 5–6, 229; 2: 262–6, 295–9, 302, 311
 - pronouns 1: 10, 18, 47, 115; 2: 200–4, 207, 210
 - relative clauses 1: 23; 2: 313–6, 319–20, 324–5, 329, 346–51, 359–62, 367
 - semantic types 1: 42–3, 50, 104–6, 394–411
 - time specification 1: 118–22, 154
 - transitivity 1: 98–105; 2: 118–9, 124, 127–34, 144, 147, 149, 154, 147
 - word 2: 2–5, 17–26, 36
 - word classes 1: 24–7; 2: 39–40, 61
- Eskimo 1: 65, 90; 2: 19, 21, 23, 44–5, 204, 243, 275, 291, 304
- Estonian 1: 51, 164–5; 2: 8, 149–50, 170
- Evenki 2: 16, 303
- Ewe 1: 5–6, 54; 2: 253, 278, 280, 285, 288
- Fijian 1: 23–4, 59–60, 82–3, 90, 108, 141, 199–200, 255, 286–8, 301–3, 307–9; 2: 23–4, 161–4
- adjectives, *see* distinguishing word classes
 - complement clauses 2: 386, 393, 400–2, 420
 - constituent order 1: 74–5, 91, 233
 - demonstratives 2: 229–33, 240, 242, 251
 - distinguishing word classes 2: 37–8, 42, 46, 52–6, 65, 71–2, 78–83, 100, 108–12
 - number words 1: 191–3, 213
 - phonology 1: 66–7, 197–9, 278, 281–2
 - possession 1: 303; 2: 270–8, 284–8, 298, 307–10
 - pronouns 2: 191–2, 202, 208, 215, 218
 - relative clauses 1: 186–8, 213; 2: 321, 327–9, 336, 339, 353–6
 - transitivity 2: 131–2, 135, 143–7, 157–8
 - sentence 1: 133, 179
 - word 2: 4, 8–9, 20–26, 36

- Finnish 1: 159; 2: 86, 180, 298, 392
- adjective class 1: 63, 68, 88
 - cases 1: 12, 223; 2: 148, 158, 167
- Finno-Ugric languages 2: 277
- Flinders Island language 2: 218
- Fox 2: 250–1
- French 1: 160, 270; 2: 4–7, 114, 180, 227, 278, 298–9, 320
- demonstratives 2: 231–2, 238, 242
 - gender 1: 12, 155–7, 163
 - pronouns 1: 18, 202, 204
- Fula 2: 86
- Gaagujju 2: 182
- Gala 1: 21; 2: 201
- Gapapaiwa 1: 6, 54; 2: 281–5, 290
- Georgian 1: 45; 2: 239, 250–1, 364
- German 1: 68, 160, 224; 2: 4–5, 41, 227, 239, 292, 298–9, 316
- gender 1: 12, 155, 163
 - phonology 1: 224, 267, 272
 - see also* Pennsylvania German
- Germanic languages 1: 15; 2: 27, 364
- Girramay: dialect of Dyirbal, q. v.
- Goemai 2: 419
- Gog-Nar: 2: 291
- Gokana 2: 251
- Gola 1: 161, 180
- Gooniyandi 2: 182
- Greek 1: 76, 155, 160, 225–6; 2: 5, 10, 26, 68, 177, 390
- cases and genitive 2: 45–6, 225
 - gender 1: 155; 2: 68
 - pronouns 2: 190, 201
 - see also* Classical Greek
- Greenlandic Eskimo 2: 275, 304
- Guajiro 2: 286
- Guaraní 2: 8, 140, 302, 310
- Gugada (dialect of the Western Desert language) 1: 260
- Gugadj 2: 291
- Gumbaynggirr 1: 60, 90; 2: 113, 182
- Gurindji 2: 197–8
- Guugu Yimidhirr 2: 142, 158, 182
- Haida 2: 5–6, 54, 64, 140
- Hanis 2: 37, 42, 64
- Hanunóo 2: 140
- Harar Oromo, *see* Oromo
- Hārāuti 2: 210
- Hausa 1: 60; 2: 234, 246, 251, 352, 360, 419
- adjectives 2: 74–6, 107, 114
- Haya 2: 311

- Hdi 1: 15, 179
 Hebrew 1: 12, 163–4; 2: 353, 363
 see also Modern Hebrew
 Hindi 1: 165; 2: 135, 162, 174, 177, 358
 Hixkaryana 2: 92–3, 239, 270–1, 316
 Hmar 2: 317–8, 333, 339
 Hopi 2: 240
 Hua 2: 84, 91–2, 199, 243, 272–3, 285, 293, 368
 Huichol 2: 135
 Hungarian 1: 12, 20, 55, 2: 5, 11, 15, 63, 88, 181, 346, 353, 364
 Icelandic 2: 148–50
 Igbo 2: 53, 109; 2: 269–70, 277–8, 340
 adjectives 2: 63, 74–5, 81, 84, 114
 Ijo 2: 64
 Ilocano 2: 197, 258, 321
 Indo-Aryan 2: 141, 234, 357
 Indo-European 1: 8, 15, 155–6, 224; 2: 4, 41, 97, 213, 277, 301, 360, 364
 Indonesian 1: 178, 181; 2: 181, 231–3, 241, 268, 292, 386, 420
 Ingush 2: 88, 148, 151
 Iranian 2: 141
 Iraqw 2: 200, 216, 221, 292
 Irish 2: 373, 381, 393, 402
 Iroquoian 2: 60
 Israeli sign language 2: 12
 Italian 2: 21
 Jacalteco 1: 18, 55, 118, 179; 2: 92, 304, 404
 Jalnguy, *see* Dyirbal avoidance style
 Japanese 2: 94–5, 98–9, 169, 277, 298, 304, 349–50
 pronouns and demonstratives 2: 201–3, 231–3, 239, 259
 Jarawara 1: 23, 60, 82–3, 117, 137, 140, 152, 233, 255, 290, 324; 2: 17, 23–4, 84, 238
 adjective class 2: 74–6, 86, 114
 complementation 2: 372, 379–89, 393, 400, 403
 copula clauses 2: 165, 170, 173–4, 180, 184
 gender 1: 62–3, 156–7, 238
 lexicon 1: 301–2, 302, 306, 308
 number system 1: 11, 159
 phonology 1: 278; 2: 23, 29
 pronouns 2: 205, 215, 221
 possession 1: 230; 2: 278, 281–5, 293, 296–300, 304, 311–12
 relative clauses 2: 319, 321, 349, 352, 356
 sentence 1: 75, 91, 133, 179
 transitivity 2: 124–5, 132–5, 144–7, 154, 157
 Jersey Norman French 2: 272–3
 Jirrbal: dialect of Dyirbal, q.v.
 Ju'hoan 2: 260
 Kabardian 1: 288
 Kabyle 2: 167–8
 Kaiké 2: 222–3
 Kalispel 2: 53
 Kalkatungu 2: 135
 Kamaurá 2: 70, 78, 80, 213, 221, 293, 342, 404
 Kambera 1: 24, 55; 2: 352
 Kamula 2: 75–6, 114
 Kana 2: 136, 178, 192, 254, 299
 Kanada 2: 200, 260
 Karajá 2: 201, 340, 363, 373
 Karbi 2: 169, 272
 Karok 2: 64
 Kayardild 1: 82–3; 2: 36
 number system 1: 10–11, 55; 2: 191
 pronouns 1: 115, 179; 2: 208
 Kazakh 2: 11
 Kewa 2: 10
 Kham 1: 82–3; 2: 63, 74, 125, 319, 408–9
 Khmer 1: 69; 2: 202, 231–3, 242, 277
 Kinyarwanda 2: 442
 Kiowa 2: 217–8, 283, 331–2
 Kiranti 2: 333, 349
 Kiriri 2: 277
 Koasati 1: 82–3, 152, 180, 233; 2: 93, 200, 316
 Kobon 2: 331, 343, 353
 Koiari 2: 136, 194
 Kolokuma dialect of Ijo 2: 64
 Koran (or !Oro) 2: 259
 Korean 1: 65, 73, 90, 265; 2: 277
 adjectives 2: 63, 77, 82–3, 99, 103
 relative clauses 2: 320, 333, 338, 352
 state verbs 1: 19–20, 55
 Korku 2: 42, 50
 Koromfe 2: 178
 Koyukon Athabaskan 2: 281–4, 289, 292, 301, 310
 Krahn/Wobé 2: 76
 Kresh 1: 163, 181
 Kugu Muminh (or Kugu Nganhcara) 2: 213
 Kuman 2: 200
 Kurdish 1: 165, 181
 Kurukh 2: 178
 Kwakiutl 2: 244
 Ladakhi 2: 291
 Lahu 1: 64, 81, 83; 2: 243

- Lak 1: 12, 15–16, 55; 2: 242
 Lakota 2: 140, 199, 344, 364
 also see Dakota
 Lango 1: 6, 11–12, 54, 279, 288; 2: 125, 231–2, 239
 possession 2: 278, 280, 283–8, 298, 303
 Lao 1: 83; 2: 113, 236, 259, 329, 338, 347
 Lardil 1: 17, 55; 2: 182
 Latin 1: 79–80, 97, 155, 240, 282; 2: 85, 124, 150, 190, 213, 265, 299
 cases and prepositions 1: 9, 45–6, 96, 165, 224–5, 299; 2: 123, 150, 265
 constituent order and word order 1: 37–8, 71–2
 copula clauses 1: 101, 254; 2: 160, 183
 fusional character 1: 43–4, 49, 55, 117, 144–6, 217, 220; 2: 58, 216
 word 2: 2–5, 10, 15, 17
 word classes 1: 25–6, 52, 102, 110, 194; 2: 38–41, 63, 68
 Latvian 2: 298
 Lavukaleve 2: 261
 Lezgian 1: 152, 189; 2: 151, 158, 347
 Lillooet 2: 51, 53, 241, 244
 Longgu 1: 10–11, 55; 2: 252
 Luiseno 1: 50, 56
 Luritja (dialect of the Western Desert language) 2: 211
 Lushootseed 2: 51, 53

 Maasai 1: 73; 2: 86
 Mabuiag: dialect of West(ern) Torres Strait language, q.v.
 Macedonian 2: 361
 Macushi 1: 73, 91; 2: 93–4, 399
 Madi-Madi (dialect of Wemba-Wemba) 2: 220–1
 Maká 2: 286
 Makah 2: 51, 53, 401–2
 Malagasy 1: 18–19, 55; 2: 320, 324
 Malay 1: 83
 Malayalam 2: 91–2, 180, 183, 260, 336, 357
 Mali (Baining) 2: 120, 157
 Mam 1: 82–3, 168, 181; 2: 64, 90
 Mamu: dialect of Dyirbal
 Manambu 1: 83, 164, 181, 277, 306–8; 2: 13–14, 362: 408–12
 copula clauses 2: 169, 173, 176–7, 182, 184, 300
 pronouns 2: 199, 201, 206–7
 relative clauses 2: 316, 349–50, 352
 transitivity 2: 124, 145, 154, 157

 Manange 2: 95
 Mandarin Chinese 1: 5–7, 110–11, 179, 227; 2: 42–3, 173, 183
 adjectives 1: 52; 2: 63, 69–70, 78–9, 81, 83, 88, 95
 demonstratives 2: 233, 246–7
 phonology 1: 267, 277, 279, 288
 possession 2: 268–9, 272, 278
 relative clauses 2: 330, 339, 342
 word 2: 4–5, 30–31, 35
 Mangap-Mbula 2: 231–2
 Mangarayi 2: 198, 291
 Mangghuer 2: 179, 329
 Mantauran (Rukai) 1: 83; 2: 382, 387, 409
 Mao Naga 2: 42, 54
 Mapuche 2: 229, 242
 Margi 2: 197
 Maricopa 1: 5–6, 54, 163, 181
 Martuthunira 1: 250, 267, 276; 2: 157, 182
 Matses 2: 204, 410, 419
 Mayali 2: 311
 Mayan languages 2: 80, 141
 Mbyá (variety of Guaraní) 2: 8
 Mende 2: 212
 Menomini 1: 328
 Middle English 1: 51
 Mingrelian 2: 168
 Miya 2: 201, 360
 Modern Hebrew 2: 325, 342
 complementation 2: 373–4, 379, 385, 393, 399, 403, 405
 see also Hebrew
 Modern Standard Arabic 2: 84
 Mohawk 2: 284
 Mojave 2: 81, 167–8, 333–4, 337, 339, 343, 350, 352
 Mokilese 2: 101, 399
 Moses-Columbia Salish (Nxaʔamxcín) 2: 22, 27
 Motuna 1: 330; 2: 178, 195, 300
 Muna (Sulawesi) 2: 244, 252
 Mundari 2: 43–4, 49, 50, 178
 Mupun 2: 79, 82–3, 169, 231–3, 330, 352, 364
 Murinypata 2: 332
 Muskogean 2: 96
 M̐yky 2: 269–70, 272

 Nakkara 2: 135
 Nanai 2: 283
 Navajo 2: 327, 332–3, 337, 343
 Ndjébbana 2: 293
 Nenets 1: 110; 2: 42

- Ngajan: dialect of Dyirbal, q.v.
 Ngandi 1: 163, 181
 Ngiyambaa 1: 161–2, 180; 2: 213
 Nhangu 2: 196
 Nilo-Saharan languages 2: 97
 Nishnaabemwin: dialect of Ojibwe, q.v.
 Njangumarta 2: 182
 Nootka
 possession 1: 5–6, 11–12, 54
 structure/word class correspondence
 1: 111, 179; 2: 37, 45, 51–2, 56–9, 61, 103
 North-east Ambae 2: 45, 82, 113, 406, 408
 North-eastern Neo-Aramaic language 2: 11
 Northern Sotho 2: 5, 17
 Northern Subanen 2: 75–6, 114, 165–6, 237, 248
 Nuer 2: 177
 Nunggubuyu 2: 63, 89, 97–9, 201–2, 209
 Nuuchahnulth, *see* Nootka
 Nyawaygi 1: 277, 284–6, 300, 303

 Oceanic languages 1: 313; 2: 8, 45, 54, 70, 101, 202, 205, 259,
 possession 2: 276–7, 284, 288, 290
 Ojibwe 2: 181, 321, 384
 Old English 1: 62; 2: 3, 220, 227
 Old Norse 2: 342, 364
 Olgolo 1: 9, 54
 Oroco 2: 167–8, 227, 299–300

 Paamese 2: 135
 Palikur 2: 244
 Panare 2: 246, 278–9, 285, 288, 292, 327, 329, 337, 379
 Papantla Totonac 2: 77, 80, 84, 88
 Papuan languages 1: 308, 313; 2: 198–99
 Parecis 1: 140; 2: 205
 Pări 2: 142
 Passamaquoddy 2: 188
 Patjtjamalh 2: 215, 221
 Pennsylvania German 1: 18, 55; 2: 393, 405
 Persian 2: 330, 352, 360
 Pitjantjatjara (dialect of the Western Desert language) 2: 8
 Pitta-Pitta 1: 165; 2: 182
 Polish 2: 135, 170
 Pomoan languages 2: 204
 Ponapean 2: 200, 231–2, 239
 Portuguese 1: 18, 67, 160, 237–8, 281; 2: 21–2, 129, 176, 260
 Punjabi 2: 152, 158, 303, 358
 Purki 2: 203

 Quechua 1: 9, 45; 2: 60, 85, 97, 181, 184, 197, 204, 271, 409
 demonstratives 2: 239, 250–1
 relative clauses 2: 333, 339, 349
 Qeqchi Maya 2: 386
 Qiang 2: 79, 114, 165
 Quileute 1: 137, 179; 2: 60

 Rarotongan 2: 70
 Rembarnga 2: 359, 362
 Rotuman 1: 69, 90
 Rukai (Mantauran) 1: 83; 2: 382, 387, 409
 Rumanian 2: 201, 353, 360
 Russian 1: 225, 281, 290; 2: 4, 88, 135, 292, 316, 372
 copula clauses 2: 162, 167, 170, 181, 187
 genders 1: 12, 156
 tense and aspect 1: 162, 181

 Sahaptin 2: 97
 Salinan 1: 227
 Salish(an) languages 1: 190; 2: 96, 241
 distinguishing noun from verb 2: 22, 27, 38, 43, 51–56, 59–61
 Samoan 1: 141, 210; 2: 13, 44–5, 61, 126
 Sango 2: 183
 Sanskrit 1: 45–6; 2: 68, 137, 190
 phonology 1: 8, 10, 12, 272, 279–80
 Sanuma 2: 193
 Sarcee 1: 220–1; 2: 242
 Sare 2: 75, 114
 Semelai 1: 83, 187; 2: 82
 Semitic 1: 139, 143, 301, 350; 2: 9, 189, 201, 310
 Serbo-Croatian 2: 336
 Shih 2: 246
 Shipibo-Konibo 2: 166, 322, 333
 Shoshone 2: 244
 Shuswap 2: 401, 420
 Sierra Miwok 1: 84–5; 2: 97
 Sign languages 2: 12, 339
 Sinhala 2: 326
 Siouan (languages) 1: 84; 2: 35, 64, 96, 197, 243
 Siriono 1: 140
 Siuslawan (Lower Umpqua) 2: 64
 Slave 2: 199, 278–9, 283, 287, 298
 Slavic languages 1: 86, 153–4
 Somali 2: 66, 75–6, 114, 200, 301
 Sorowahá 2: 259
 Southeastern Tepehuan 2: 243
 Southern Paiute 1: 84, 265; 2: 99, 197
 Southern Sotho 2: 5

- Spanish 1: 67, 281; 2: 63, 70, 129, 148, 213, 216, 260, 300
 copula clause 2: 168–9, 175–6, 178, 183
 Sranan 2: 176
 Sumerian 1: 84, 159, 180; 2: 162
 Sundanese 2: 10
 Supyire 1: 12, 55; 2: 236, 246, 250–1, 343, 364
 Suruí 2: 209
 Swahili 1: 12, 85, 157, 177; 2: 86, 239, 277
- Tachelhit 1: 5–6, 12, 54
 Tagalog 2: 37, 52–3, 56, 59
 Takelma 1: 84, 218; 2: 63, 89–90
 Tamambo 2: 217, 299
 Tamil 2: 108, 174, 177, 233, 242, 250–1, 260, 302, 320, 357
 Tarascan 2: 97
 Tariana 1: 20, 83, 164, 181, 290; 2: 54, 84, 241, 293
 adjectives 1: 194–6, 213; 2: 83, 88–9
 complement clauses 2: 379, 382–6, 390, 393, 406, 409–10
 copula clauses 2: 163–4, 169, 179
 evidentials and tense 1: 15, 18–19, 44
 relative clauses 2: 321, 330, 348–9
 pronouns 2: 205, 210–11, 221
 transitivity 2: 124, 126, 145, 147
 word 2: 8, 17, 21, 33
 Tarma Quechua 2: 181, 184
 Tawala 2: 135
 Telugu 1: 69, 307–8; 2: 108, 135, 224, 242, 247, 260
 Temiar 2: 78
 Tennes 2: 167–8, 172
 Teribe 2: 64, 90
 Thai 2: 5, 63–5, 113, 119, 202, 239, 300
 Tialo 2: 272, 298
 Tibetan 2: 10
 Tibeto-Burman languages 2: 222, 258, 319, 333–4, 342, 363
 Tigak 2: 80
 Tiriyo (or Trio) 2: 92–3, 197–8, 311
 Tiwi 1: 39, 40, 55, 116, 205, 227; 2: 199
 Tjaitjala (dialect of Wemba-Wemba) 2: 219
 Tlingit 2: 64
 Toba-Batak 2: 82
 Tok Pisin 1: 21, 320; 2: 343, 364
 Tongan 1: 160, 180, 284, 288; 2: 54
 transitivity 1: 99–100, 179; 2: 116–8, 154, 161
 Tonkawa 1: 140; 2: 60, 215
- Totonac 2: 77, 80, 84, 88, 90
 Trio (or Tiriyo) 2: 92–3, 197–8, 311
 Trumai 1: 100, 179
 Tsimshian 2: 96, 278–9, 285–6, 292
see also Coast Tsimshian
 Tsova-Tush (Georgian name for Batsbi) 2: 121, 157
 Tswana 2: 5
 Tübatulabal 1: 88
 Tucano 1: 13, 15, 55, 47, 56
 Tucanoan languages 1: 15, 20, 210, 221; 2: 89, 205
 Tukang Besi 2: 61, 82
 Tunica 1: 84; 2: 64, 91, 201, 221–2, 287, 293
 Tupuri 2: 81
 Turkic languages 1: 226; 2: 11, 88
 Turkish 1: 9, 12, 20, 55, 161, 227; 2: 64, 271
 word 2: 10, 12, 15, 17, 23–4
 Tuscarora 2: 239, 241, 293, 300
 Tuvaluan 2: 322, 349, 351
 Tuyuca 1: 13–14, 55, 162, 181
 Tyaddywurr (dialect of Wemba-Wemba) 2: 219
 Tzotzil 1: 69–70, 90, 304, 308; 2: 391
- Udmurt 1: 164, 181
 Ungarinjin 1: 260, 262
 Upper Necaxa Totonac 2: 84, 90
 Uradhi 2: 363
 Urarina 2: 178, 181
 Urdu 2: 347, 358
 Ute 2: 301, 338, 349
- Venda 2: 87
 Veps 1: 164
 Vietnamese 1: 148, 226; 2: 58, 63, 78, 234, 277
- Waga-Waga 2: 182
 Waikurúan languages 2: 243
 Wakashan languages 1: 37, 43, 45, 51–56, 59–60
 Walmatjari 2: 8, 182
 Wambaya 2: 217
 Wappo 2: 331, 349, 372
 Wardaman 1: 82–3, 277; 2: 182
 Warekena 1: 191–6, 213, 245; 2: 8–9, 126, 156–7; 275, 393, 320–1
 Warlpiri 1: 282, 306; 2: 135, 213, 218–9, 358, 362
 Warray 2: 182
 Warrgamay 1: 197, 203, 238, 277, 300; 2: 24, 237, 248

- Watjarri 2: 213, 241
 Welsh 2: 320, 326
 Wemba-Wemba 2: 219–21
 Weri 2: 198
 West Greenlandic (Eskimo) 2: 275,
 304–5
 Western Desert language 1: 260, 262,
 307; 2: 8, 182, 211, 258
 West(ern) Torres Strait language 1: 12–14,
 23, 53, 55
 Wetan 2: 348
 White Hmong 2: 379, 385–6, 393, 398,
 404–6
 Wirangu 2: 182
 Wiyot 2: 286–7
 Worora 1: 260
 Wunambal 2: 191

 Xhosa 2: 5

 Yagua 2: 200, 241, 272, 274, 292, 298, 346,
 364
 Yana 1: 227
 Yaqui 2: 211
 Yawuru 1: 52, 56; 2: 63, 151
 Yiddish 2: 10, 17

 Yidiñ 1: 50–1, 118, 133, 146–8, 170, 179–80,
 255; 2: 101, 207–8, 259
 classifiers 1: 18, 157; 2: 86
 demonstratives 2: 226, 236, 242, 245
 lexicon 2: 297–9, 307–8
 phonology 1: 206–9, 213, 273, 288
 possession 2: 278, 280, 283, 288, 296, 298,
 302
 transitivity 2: 128, 139, 144, 157
 Yimas 1: 23, 55, 82–3, 162, 181; 2: 9, 23–4,
 169, 178, 199, 353–4, 393–4
 adjectives 2: 63, 74, 76
 Yingkarta 2: 8
 Yir-Yoront 2: 69, 104, 182
 Yokuts 2: 2, 22, 97
 Yoruba 2: 74, 92, 169
 Yuma 1: 140
 Yuman languages 1: 5, 26, 55; 2: 120, 333–4
 Yurok 1: 70, 90
 Yuwaalaraay 2: 180

 Zayse 2: 170, 237, 243
 Zekkara 2: 348
 Zoque 2: 8
 Zulu 2: 5
 Zuni 2: 96, 103

Subject Index

Note that entries in the glossary (which appears in both volumes) are not included in this index.

- A, S and O core arguments 1: 76–7, 98–100, 122–5, 228–9; 2: 116–23, 129–33, 138–40, 147–55, 161–2, 166–70, 292–4
- ablative 1: 132, 145–7, 224, 231–2, 282; 2: 291
- absolutive 1: 76, 122–3; 2: 116–23, 137, 145–52, 165, 167
- accent, *see* stress
- accessibility hierarchy 2: 320–4
- accompaniment 1: 126
- accusative 1: 9, 76, 122–3, 161; 2: 116–23, 147–52, 162, 167–8, 172
- active articulator 1: 269
- active voice 1: 167, 240
- active/stative, *see* split-S marking, fluid-S marking
- Activity type of complement clause 2: 382–421
- addition type of clause linking 1: 134–6
- adjective class 1: 52–3, 112–14, 194–6, 243–5, 304–5; 2: 62–114
- criteria for recognition 2: 70–73
- distinguished from noun class 2: 69, 84–8, 106–7
- distinguished from verb class 2: 77–83, 105–6
- grammatical properties 2: 63–5
- semantic content 2: 73–6
- adjoined relative clause 2: 358–9
- adjunct 1: 101–2
- adposition 1: 73, 127, 224–5, 231–3
- adverb 1: 109, 301; 2: 76, 82, 88
- AFFECT semantic type 1: 104; 2: 127–33, 147, 394
- affective case 2: 151
- affinal kin 1: 6–7
- affix 1: 221–5
- affix(ation) 1: 141–4, 217–8, 269
- AGE semantic type 1: 114; 2: 73–6, 79, 82–5, 104
- Agent semantic role 1: 99, 104–5; 2: 127–33, 153
- agglutinating language 1: 226–7
- agreement 1: 220, 230
- airstream mechanisms 1: 271
- alienable possession 1: 5–7, 11–12, 230; 2: 277–312
- allative 1: 132, 227–8
- allomorph 1: 179, 185
- allophone 1: 180, 264–6
- alphabet 1: 264–5
- also see* orthography
- alternative syntactic frame 1: 98–9, 105–6
- ambitransitive 1: 77–8, 103–4; 165, 305; 2: 100, 124–6, 143–7, 154–7, 300
- analytic language 1: 226–8
- analysis, linguistic 1: 182–99, 243–7
- anaphora 1: 332; 2: 247–61
- ANNOYING semantic type 2: 129, 397
- antipassive 1: 165–8, 172–4, 207–8; 2: 237
- apical place of articulation 1: 267, 276
- applicative 1: 165, 168–71; 2: 123, 186
- apposition, *see* verbless clause, complementation strategies
- archiphoneme 1: 272
- argument, *see* core argument, peripheral argument
- argument identity 1: 175–7
- article 1: 160–1; 2: 51, 55
- articulator 1: 268
- aspect 1: 86, 154, 162–5; 2: 52–4, 177–8, 181–5, 188, 301
- motivating split marking 2: 141
- aspiration 1: 250, 271–2; 2: 9–10
- associated motion affixes 1: 50–1, 180
- atelic 1: 153
- ATTENTION semantic type 1: 104; 2: 127–30, 146–52, 385–413
- Attribution semantic relation 1: 101; 2: 159, 171–84, 188
- augmented term in pronoun system 1: 115; 2: 196–9, 253
- auxiliary 1: 96, 125, 130, 154, 166
- aversive 1: 127
- avoidance style 1: 214–16, 293–5
- base 2: 37
- BEGINNING semantic type 2: 402–17
- benefactive semantic relation 2: 159, 171–84

- beneficiary 1: 108, 126
bilateral opposition 1: 70, 235–6
binarism 1: 70–1
blood relation 1: 6–7
body part terms 1: 5–7, 22–3, 55, 303
bound form 1: 145, 217
bound pronoun 1: 39–40, 55, 82, 116–7, 125–6, 159; 2: 169, 209–23
boundedness 1: 153
- cardinal vowels 1: 173–4
case 1: 12–13, 43–5, 85, 125–6, 164–5, 224–5; 2: 55, 88
cataphora 2: 247–61
causal 2: 291, 362–3
causative 1: 165, 168–78; 2: 17, 165, 186
certainty, degree of 1: 153
changing valency 1: 165–71
circumfix 1: 141
classifiers 1: 18, 87, 157–8; 2: 55, 86, 248
clause 1: 75–6, 93–102, 132–7, 228
clause linking 1: 94–5, 133–7; 2: 352, 374–5, 410–14
clause structure 1: 97–102, 110–12, 254
click 1: 271, 282; 2: 29
clitic 1: 221–5; 2: 20–2, 215–8, 254
co-existing phonological systems 1: 283
Cogitator semantic role 1: 104–5; 2: 127–30
cohesiveness 2: 14–15
COLOUR semantic type 1: 53, 114, 194, 196, 304; 2: 73–6, 79, 92–5, 104, 114
colour terms 1: 256, 291–2
comitative 1: 145–8
command 1: 95–6
common argument in a relative clause construction 1: 105, 246; 2: 313–69
comparative construction 1: 62, 113, 177–9, 210; 2: 64–6, 71–2, 77, 82, 88–91, 265, 320, 342, 347, 364
complement clauses 1: 27, 30–1, 91, 94, 128–32, 201–3; 2: 171, 185, 361, 370–411
 grammatical criteria for 2: 375–81
 grammatical parameters for 2: 384–9
 types and meanings 2: 388–94
complement-taking verbs 2: 253, 370–424
complementation strategies 1: 83; 2: 351–2, 405–15
 apposition strategy 2: 409–15
 clause chaining strategy 2: 410–14
 nominalization strategy 2: 408–14
 purposive strategy 2: 399–415
 relative clause strategy 2: 399–414
 serial verb construction strategy 2: 404–14
complementary distribution 1: 285
complementizer 1: 333
completion 1: 153
composition 1: 153–4
compounding 1: 138–9, 304–5; 2: 23, 26, 56, 155
concord 1: 230
condensed relative clause 2: 359–60
conditional 1: 135–6
configurational language 1: 72
congruent/non-congruent, *see* conjunct/disjunct contrast
conjugation 1: 207–12, 239
conjunct/disjunct contrast 1: 334; 2: 222–3, 259
conjunction 1: 134–7
conjunctive writing system 2: 6
consanguineal relation 1: 6–7
consequence type of clause linking 1: 134–6
consonant system 1: 7–8, 250, 266–73
constituent 1: 232
constituent order 1: 37–8, 71–5, 126, 233–4, 254–5; 2: 164–5
construct state 2: 310
construction 1: 228–9
content interrogative/question 1: 95–6, 159; 2: 233–4, 346, 364, 368–9
 also see interrogative/indefinite word
continuous 1: 153
contrast type of clause linking 1: 134–6
contrastive focus 1: 174–5
convenience sample 1: 263
copula clause 1: 100–1; 2: 66–114, 159, 188, 300–3
copula complement 1: 100–1; 2: 66–114, 159, 188
copula subject 1: 100–1; 2: 66–114, 159–168
core argument 1: 97–101, 122–8, 246–7; 2: 116–59
 also see A, S and O
co-relative construction 2: 356–8
CORPOREAL semantic type 1: 54, 300; 2: 147, 394
correlative construction, *see* co-relative construction
coverb 1: 52, 103, 305
creole 1: 21–2; 2: 176, 343
- dative 1: 128; 2: 148–52, 290–1
DECIDING semantic type 2: 397
declarative mood 1: 95–7

- definiteness 1: 49, 160–2; 2: 55
 deictic reference, *see* demonstrative
 demonstrative 1: 71, 108, 117, 159; 2: 183, 188,
 227–47, 346, 364, 369
 local adverbial demonstrative 2: 224–51
 manner adverbial demonstrative 2: 224,
 229, 233
 nominal demonstrative 2: 224–51
 verbal demonstrative 2: 224, 229–31, 242,
 251
 dependencies between grammatical
 systems 1: 162–5, 181, 255
 derivation 1: 142–52, 180, 218–21;
 2: 15–16, 46–50, 56, 61, 385–6
 determiner 1: 27–9, 128, 180
 detransitivizing derivations 1: 165–8, 175–7
 dictionary 1: 48, 215
 DIFFICULTY semantic type 2: 74, 76, 95
 DIMENSION semantic type 1: 53, 114, 194–6,
 304; 2: 73–6, 79, 92–5, 104, 114
 diphthong 1: 198–9
 direct speech 1: 307–8; 2: 171, 397–8, 419
 disjunct/conjunct contrast 1: 334; 2: 222–3,
 259
 disjunction type of clause linkage 1: 136–7
 disjunctive writing system 2: 6
 dissimilation 1: 270
 ditransitive, *see* extended transitive
 Donor semantic role 1: 63–4, 115, 229;
 2: 127–8, 134–7
 double case 1: 45, 56
 dual number 1: 9–10, 158; 2: 191–217
 durative 1: 153

 E syntactic function 2: 116–19, 161–2
 ejective 1: 271, 313
 elaboration, *see* pronoun elaboration
 enclitic, *see* clitic
 environment affecting language 1: 15–17
 equipollent opposition 1: 236, 272
 ergativity 1: 76, 82, 86, 89, 91, 123–8, 188–9,
 246, 261; 2: 116–23, 147–56, 162, 165, 167,
 169, 291
 essive 2: 170
 ethics of fieldwork 1: 311
Ethnologue 1: xiii, 73, 91; 2: xiv
 evaluation 1: 4
 evidentiality 1: 13, 18, 56, 87, 153, 162–4,
 260–1; 2: 260
 exclusive 1: 335; 2: 194–6, 258
 Existence semantic relation 2: 160, 174
 Experiencer semantic role 1: 53–4, 104–5;
 2: 127–30, 150–1

 explanation 1: 205–13
 extended intransitive 1: 99–100, 229;
 2: 116–24, 144, 150
 extended transitive (or ditransitive)
 1: 99–100; 2: 116–18, 134
 extra-language typology 1: 247–8
 extent (non-spatial setting) 1: 153

 Fact type of complement clause 2: 380–421
 feminine
 as unmarked gender 1: 240
 also see noun classes, gender
 fieldwork 1: 297–9, 209–30
 finite 1: 80, 91
 fluid-S marking 1: 77–8, 124–5; 2: 121, 126,
 141
 focal clause 1: 133–6
 focus 1: 174–5
 focus system 2: 52
 foot 1: 148, 206
 formal markedness 1: 237–40
 formal theories 1: 3–4, 183–4
 fourth person pronoun 1: 260–1; 2: 203–5
 free form 1: 145, 217
 fricative 1: 269
 functional markedness 1: 237–40
 fused relative clause 2: 356–60
 fusion of morphemes 2: 215–6
 fusional language 1: 226–7
 future time marking 1: 154

 gender 1: 12–13, 43, 87, 155–8, 180, 290, 335;
 2: 54–5, 86–7, 200–222, 246, 259, 297
 also see noun classes
 generic noun 1: 300–2
 genitive 1: 44–5, 73; 2: 123, 148–50, 167,
 268–312
 Gift semantic role 1: 53, 104–6, 229; 2: 127,
 134–7
 GIVING semantic type 1: 104–6; 2: 127–8,
 134–7, 145–6, 157, 394
 glossing, conventions for 1: 61, 216
 glottalic airstream mechanism 1: 271
 glottalization 1: 280
 government 1: 231
 gradual opposition 1: 236, 272
 grammatical word 1: 93, 108, 116, 138,
 221–4; 2: 1–36
 criteria for 2: 12–19

 half-conjunction 1: 134–6
 harmonic pronoun 1: 17
 ‘have’ 2: 290, 298–302

- head (of noun phrase) 1: 229–32; 2: 296–8
 heterorganic 1: 276
 historical explanation 1: 62–3, 205, 209–13
 homonymy, multiple 1: 290
 homorganic 1: 198, 276
 honorific pronoun 1: 17–18
 HUMAN PROPENSITY semantic type 1: 53,
 114, 304; 2: 73–6, 79, 81, 92–5, 104, 114,
 146
 iconicity 2: 289–90
 ideophone 1: 302; 2: 30
 Identity semantic relation 1: 101; 2: 159,
 170–83, 187
 immersion fieldwork 1: 317–25
 imperative 1: 13, 77, 95–7; 2: 78, 154, 181,
 185–6
 imperfect 1: 153
 imperfective 1: 154
 impersonal form 2: 204–5
 Impression semantic role 1: 104–5;
 2: 127–30, 150–1
 inalienable possession 1: 5–7, 11–12, 230;
 2: 277–312
 inclusive 2: 194–6, 258
 inclusory construction, *see* pronoun
 elaboration
 indefinite form 2: 205
 index of comparison 1: 177–8
 indicative mood, *see* declarative mood
 indirect speech 2: 397–8, 419
 inductive basis for generalizations 1: 2,
 184h
 infinitive 1: 79–80
 infix 1: 141, 146–7, 180, 217; 2: 52, 215, 259,
 305
 inflection 1: 96–7, 142–7, 203, 213–21;
 2: 15–16
 ingressive airstream mechanism 1: 271
 inner locative 1: 119
 instrument(al) 1: 98–9, 126–8, 188–9;
 2: 121–3, 170
 interjection 1: 283; 2: 27–30, 36
 internal change 1: 140
 interrogative mood 1: 95–7; 2: 186
 interrogative/indefinite word 1: 17–18, 216;
 2: 233–4
 also see content interrogative/question
 intonation 1: 24, 67–8, 75, 95–6, 133, 137,
 187–8, 283
 intra-language typology 1: 247–8
 intransitive, *see* transitivity
 irrealis 1: 153
 isolating language 1: 226–7
 item-and-arrangement model 1: 146
 item-and-process model 1: 146
 Judgment to complement clause 1: 31–33,
 42; 2: 395
 karaka 1: 45
 kin term 1: 5–7, 26, 256, 262, 300; 2: 40, 47,
 262–6, 278, 298
 labile, *see* ambitransitive
 laminal place of articulation 1: 267, 276
 language contact 1: 15, 83, 209, 283; 2: 205,
 220
 language planning 1: 20
 lateral 1: 269, 275–6
 lexicon (and lexeme) 1: 47–54, 214–7, 253–5,
 289–308; 2: 4–5
 lexicostatistics 1: 215, 240
 LIKING semantic type 1: 32–4, 104–5;
 2: 127–30, 148–50, 157, 397–413
 liquid 1: 269
 Location semantic relation 1: 101; 2: 159–61,
 169, 172, 176, 179, 180–8
 locative 1: 128; 2: 151
 locutor/non-locutor, *see* conjunct/disjunct
 contrast
 logophoric pronoun 2: 252–4, 419
 long vowel 1: 196–9, 209–13, 281–2
 Manip semantic role 1: 99, 104–5; 2: 127–33
 manner of articulation 1: 269–71, 276
 mark of comparison 1: 177–8
 marked nominative 2: 167–8, 172
 markedness 1: 235–41, 272–3; 2: 120, 137–40,
 166–8, 181, 246–7
 masculine
 as unmarked gender 1: 237
 also see noun classes, genders
 medial clause 2: 410
 Medium semantic role 1: 53, 104–6; 2:
 127–31
 Message semantic role 1: 53, 104–6; 2:
 127–31, 136
 middle 1: 235
 minimal pair 1: 273
 minimal term in pronoun system 1: 115;
 2: 196–9, 258
 modal verb 1: 109; 2: 402
 modality 1: 96, 153; 2: 52–4, 301
 modifier 1: 106–8, 230
 mood 1: 95–6

- mora 1: 197–8, 281–2
 morpheme 1: 145, 180
 morphological processes 1: 83–4, 138–44,
 217–8; 2: 53
 morphology 1: 85–6, 89, 138–52, 180, 217–21
 ‘mother-in-law language’, *see* avoidance
 style
 MOTION semantic type 1: 119; 2: 145–6, 394
 multilateral opposition 1: 70–1, 235–6
 mutation 2: 12

 naming 2: 173–4
 narrative case 2: 168
 nasal 1: 269, 276
 nasalization 2: 9–10
 negation 1: 83, 137–8; 2: 57, 186, 385, 401–2
 negative copula 2: 178–9
 neutralization 1: 164, 272; 2: 199–200, 217
 nominal hierarchy 1: 123, 179; 2: 137–41
 nominalization 1: 129–31, 150–1; 2: 263–7,
 292, 326, 363, 403–14
 nominative 1: 76, 122–3; 2: 116–23, 147–52,
 167–72
 non-canonical marking of core
 arguments 1: 127; 2: 147–52
 non-canonical relative clauses 2: 356–61
 non-configurational languages 1: 72
 non-nuclear verb 1: 293–5
 non-restrictive relative clause 2: 314–5,
 352–4
 non-spatial setting 1: 152–5
 noun 1: 51–3, 102–3, 302–4; 2: 37–61
 criteria for 2: 38–41
 grammatical categories associated
 with 2: 54–5
 noun classes (and genders) 1: 12–13, 27–31,
 87, 155–8, 163–5; 2: 54–5, 86–9, 239,
 245–6
 noun incorporation 1: 68–9, 89; 2: 155,
 310–11
 noun phrase (NP) 1: 106–8, 229–30;
 2: 50–51, 68–9, 85
 nuclear verb 1: 293–5
 number system 1: 9–11, 43, 87, 143, 153–9,
 163–5, 184–5, 237; 2: 55, 87, 154, 191–2,
 246
 numeral (number) 1: 175, 207, 301; 2: 72, 93

 O, S and A core arguments 1: 76–7, 98–100;
 122–5, 228–9; 2: 116–23, 129–33, 139–40,
 147–55, 161–2, 166–70, 292–4
 object incorporation, *see* noun
 incorporation

 obstruent 1: 269
 omission of copula verb 2: 180–2
 onomatopoeia 1: 68–9; 2: 30
 oppositions 1: 70–1, 236, 272–3
 orthography 1: 66–7, 90, 286
 also see alphabet
 outer locative 1: 119
 overlapping between word classes,
 semantic 2: 99–103

 parameter of comparison 1: 177–9
 particle 2: 37
 partitive 2: 148–50, 158, 167
 passive articulator 1: 268
 passive voice 1: 165–8, 172–4, 240–1, 245–6;
 2: 136–7, 183, 373–4, 377–8
 paucal number 1: 9–10, 158; 2: 101–2, 199,
 202
 pause 2: 9–10, 18–26, 35
 Perceiver semantic role 1: 53–4, 104–5;
 2: 117, 127–30, 150–3
 perfect 1: 153
 perfective 1: 153–4
 peripheral argument 1: 97–101, 126–8,
 189–91; 2: 116
 peripheral place of articulation 1: 270, 276
 person system 1: 115–16, 163–5; 2: 191–2,
 245–6
 also see pronoun
 pertensive 2: 268–312
 phase of activity 1: 153, 155
 phoneme 1: 88, 145, 264–6, 287
 phonoaesthesia, *see* sound symbolism
 phonological rules 2: 11–12
 phonological word 1: 93, 197–200, 206–7,
 221–5, 249–51, 267; 2: 1–36
 criteria for 2: 7–12
 phonology 1: 200–1, 216–7, 249–51, 264–88
 phonotactics 1: 273–9, 290
 phrasal verbs in English 1: 35–7, 67, 290;
 2: 5, 395
 phrase 1: 106–10, 254–5
 PHYSICAL PROPERTY semantic type 1: 53,
 114, 202, 205, 314; 2: 73–6, 79, 92–5, 104,
 114
 pidgin 1: 21
 pivot 1: 168, 172–5, 235; 2: 237, 321, 411–12
 place of articulation 1: 268–9, 276
 plural number 1: 9–10, 158–9; 2: 191–261
 polar question/interrogative 1: 95–6; 2: 166
 polarity 1: 138, 162–5
 politeness strategies 1: 17–18; 2: 201–3
 polysynthetic 1: 226–8, 241

- portmanteau form 2: 215–6
 Possessed, parameters relating to 2: 277–90
 possession 1: 5–7, 11–12; 2: 51, 55, 219–22, 262–312
 Possession semantic relation 2: 171, 174–81
 possessive phrase 1: 107, 229–30
 possessive relation, nature of 2: 274–7
 Possessor, parameters relating to 2: 271–4
 possible consequence type of clause
 linking 1: 152
 postposition, *see* adposition
 posture verb 1: 257, 262, 307
 Potential type of complement clause
 2: 392–421
 pragmatics 1: 68, 95–7, 246, 325, 327; 2: 61, 109, 119, 161, 172–3, 179, 332, 334, 352, 356, 398, 403
 predicate 1: 78–9, 98–101; 2: 40–1, 51–2, 77–8, 162–4
 prediction 1: 4
 prefix, *see* affix(ation)
 preposition, *see* adposition
 preverb 2: 5
 primary sources, need to refer to
 1: 64–6
 primary verbs 1: 54, 131; 2: 394
 primitive languages, lack of 1: 21
 privative marking 2: 11
 privative opposition 1: 236, 272
 proclitic, *see* clitic
 progressive 1: 153
 pronoun (personal) 1: 115–17; 2: 189–223, 247–61
 harmonic 1: 17
 honorific 1: 17–18
 also see bound pronoun
 pronoun elaboration 2: 207–10, 259
 proper noun/name 1: 102, 108
 prosody 1: 251, 279–83; 2: 10–11
 prototypical pattern 1: 4–9
 psychological reality of phonological
 units 1: 26; 2: 26, 31
 pulmonic airstream mechanism 1: 271
 punctual 1: 153
 purposive construction 1: 96–7, 124, 136;
 2: 142, 291, 399, 415

 QUALIFICATION semantic type 2: 74
 QUANTIFICATION semantic type 2: 74, 76, 93, 95
 quantifier 1: 107, 301; 2: 50–1, 59, 81, 101, 189, 230
 question 1: 95–6

 realis 1: 153
 reality 1: 153
 Recipient semantic role 1: 53, 104–6, 126–8, 229; 2: 127, 134–7
 reciprocal 1: 176–7
 recursion 2: 16–17
 reduplication 1: 139–40, 180, 253–4, 262, 294; 2: 13–16, 24–7, 55–6, 207
 for distinguishing between word
 classes 2: 65–6, 69, 78–9, 87, 114
 reflexive 1: 77, 176–7; 2: 154, 185, 347
 relative clause 1: 23–4, 89, 93–4, 107, 120, 186–8; 2: 81, 142, 185, 313–69, 374–5, 399–414
 function of 2: 350–2
 marking of 2: 338–48
 meaning of 2: 352–4
 non-canonical 2: 356–61
 restrictive and non-restrictive 2: 314–6, 352–4
 structure of 2: 348–50
 relative pronoun 2: 344–8, 363–4
 REST semantic type 1: 119; 2: 145–6, 394
 restrictive relative clause 2: 314–6, 352–4
 retroflex 1: 268, 279–80; 2: 10
 rhotic 1: 248, 265–6, 269–71, 284–6, 288
 root 1: 138–48, 199–200, 217–8
 rounding 1: 273–4

 S, A and O core arguments 1: 76–7, 98–100; 122–5, 228–9; 2: 116–23, 129–33, 139–40, 147–55, 161–2, 166–70, 292–4
 S = A type ambitransitive, *see* ambitransitive
 S = O type ambitransitive, *see* ambitransitive
 sampling 1: 257–63
 sandhi rules 2: 12
 science, linguistics as a branch of
 1: 1–4
 scrambling rule 1: 72
 secondary concept 1: 50, 131; 2: 394–5, 399–400
 secondary predicate 1: 41–3
 secondary verb 1: 54, 131; 2: 394–414
 segmental features as criteria for
 phonological word 2: 7–10
 self repair 2: 19, 30
 semantic overlap between word classes
 2: 99–103
 semantic roles 1: 53–4, 104–6; 2: 127
 semantic types 1: 31–3, 53–4, 102–6, 300; 2: 127, 394–421

- semi-vowel 1: 269–71, 275, 284
 sentence 1: 75–6, 91, 132–7
 serial verb construction 1: 83, 109, 132, 139, 178, 290, 305; 2: 52, 58–60, 69, 82, 136, 404–21
 shifter 1: 114; 2: 189
 also see pronoun, demonstrative, temporal item
 sign languages 1: 90; 2: 12
 SIMILARITY semantic type 2: 74, 76
 singular number 1: 9–10, 158–9
 social niceties 2: 201–2
 sonorant 1: 269
 sound symbolism 1: 68–70, 90; 2: 242
 spatial item 1: 107, 118–22, 152
 spatial shifter 1: 114
 Speaker semantic role 1: 53–4, 104–6; 2: 127–31
 SPEAKING semantic type 1: 104–6; 2: 127–31, 146, 397–413
 SPEED semantic type 2: 73–6, 79, 93–5, 104
 split ergativity 1: 122–6; 2: 137–42, 206, 208, 212, 234
 split-S marking 1: 77–8, 82, 124–5; 2: 120, 126, 140–1, 155–6, 221
 stance verb 2: 161, 182–8
 standard error in sampling 1: 258–9
 standard of comparison 1: 177–9
 statement 1: 95–6
 stative/active, *see* split-S marking, fluid-S marking
 stem 1: 138–42
 Stimulus semantic role 1: 53, 104–5; 2: 127–30, 150–1
 stop 1: 269, 276
 stress (or accent) 1: 251, 280–3; 2: 10–12
 stress shift 1: 14, 160–1, 180
 subject 1: 76–7, 98, 229
 subjunctive 1: 97; 2: 183, 391
 substitution anaphora/cataphora 2: 248–50
 subtraction 1: 140–1
 suffix, *see* affix(ation)
 supporting clause 1: 133–6
 suprasegmental 1: 279, 297; 2: 10
 switch-reference marking 1: 82, 174; 2: 185
 syllable 1: 249, 277–9
 syllable structure 1: 9, 249–50, 275–9
 symbolic type of morphological process 1: 226
 synharmonism 2: 11
 synthetic language 1: 226–8
 systems, grammatical 1: 247, 252
 tabooing 2: 31
 Target semantic role 1: 99, 104–5; 2: 127–33
 telic 1: 153
 temporal item 1: 107, 114, 118–22, 248; 2: 107, 222, 242
 temporal type of clause linking 1: 134–5
 tense 1: 12–14, 154–5, 162–5, 239; 2: 52–4, 181–4, 301
 motivating split marking 2: 141
 textual anaphora/cataphora 2: 248–50
 thesaurus 1: 296–9, 308, 319
 THINKING semantic type 1: 32–4, 104; 2: 127–30, 146–7, 396–413
 third person 2: 189–90, 203
 Thought semantic role 1: 104; 2: 127–30
 time word, *see* temporal item
 tones 1: 140, 251, 279; 2: 10–12, 178
 topic 1: 171–5, 234–5; 2: 172
 topicalization 1: 235
 traditional grammar 1: 7, 114; 2: 123, 225, 369
 transitivity 1: 89, 103–5, 165, 168–71; 2: 115–58
 tree structure 1: 48, 292
 trial number 1: 9–10, 158
 tripartite marking of S, A and O 1: 123–4; 2: 118–9, 139, 154–7
 typology, linguistic 1: 242–63

 unaccusative 2: 155–6
 unergative 2: 155–6
 unit augmented term in pronoun system 2: 180–5, 294
 unmarked, *see* formal markedness, functional markedness

 valency-changing derivation 1: 165–71, 227, 235–6
 VALUE semantic type 1: 114; 2: 73–6, 79, 92–5, 104, 114
 verb 1: 52–4, 103, 305–6; 2: 37–61
 criteria for 2: 38–41
 grammatical categories associated with 2: 52–4
 verb phrase 1: 108–10
 verbalization 1: 150–1
 verbless clause 1: 341; 2: 160–188
 verbless clause complement 1: 341; 2: 160–188
 verbless clause subject 1: 341; 2: 160–188
 vicious circles in definitions 1: 292

- visibility 2: 243–4
vocabulary, *see* lexicon
voice system 1: 167, 273–5
voicing 1: 271–2
vowel harmony 1: 142, 251, 274, 279;
2: 7, 10, 12, 20, 24, 32–3
vowel system 1: 7–8, 249–50
- WANTING semantic type 2: 403–13
word 1: 92–3, 223–4; 2: 1–36
definitions of 2: 1–5, 34
word classes, recognition of 1: 25–7, 102,
191–3; 2: 38–41
- word classes and clause structure 1: 110–12;
2: 41–50
word order 1: 37–8, 71–5, 233–4; 2: 15
also see constituent order
word-and-paradigm model 1: 144
writing a grammar 1: 57–63
writing systems 2: 5–6
- yes/no question, *see* polar
question/interrogative
- zero 1: 143–4
zero derivation 2: 46–50

Books by R. M. W. Dixon

ON LINGUISTICS

Linguistic science and logic

What *is* language? A new approach to linguistic description

The Dyirbal language of North Queensland

A grammar of Yidij

The languages of Australia

Where have all the adjectives gone? and other essays in semantics and
syntax

Searching for Aboriginal languages, memoirs of a field worker

A grammar of Boumaa Fijian

Words of our country: stories, place names and vocabulary in Yidiny,
the Aboriginal language of the Cairns-Yarrabah region

Ergativity

The rise and fall of languages

Australian languages: their nature and development

The Jarawara language of southern Amazonia

A semantic approach to English grammar

with Bruce Moore, W. S. Ramson, and Mandy Thomas

Australian Aboriginal words in English, their origin and meaning

*with Grace Koch*Dyirbal song poetry, the oral literature of an Australian rainforest
people

ON MUSIC

with John Godrich

Recording the blues

with John Godrich and Howard Rye

Blues and gospel records, 1890–1943

NOVELS (*under the name Hosanna Brown*)

I spy, you die

Death upon a spear

EDITOR OF BOOKS ON LINGUISTICS

Grammatical categories in Australian languages
Studies in ergativity

with Barry J. Blake

Handbook of Australian languages, Vols 1–5

with Martin Duwell

The honey ant men's long song, and other Aboriginal song poems

Little Eva at Moonlight Creek: further Aboriginal song poems

with Alexandra Y. Aikhenvald

The Amazonian languages

Changing valency: case studies in transitivity

Areal diffusion and genetic inheritance: problems in comparative linguistics

Word, a cross-linguistic typology

Studies in evidentiality

Adjective classes: a cross-linguistic typology

Serial verb constructions: a cross-linguistic typology

Complementation: a cross-linguistic typology

Grammars in contact: a cross-linguistic typology

The semantics of clause linking: a cross-linguistic typology

with Alexandra Y. Aikhenvald and Masayuki Onishi

Non-canonical marking of subjects and objects