This textbook investigates definiteness both from a comparative and a theoretical point of view, showing how languages express definiteness and what definiteness is. It surveys a large number of languages to discover the range of variation in relation to definiteness and related grammatical phenomena: demonstratives, possessives, personal pronouns. It outlines work done on the nature of definiteness in semantics, pragmatics and syntax, and develops an account on which definiteness is a grammatical category represented in syntax as a functional head (the widely discussed D). Consideration is also given to the origins and evolution of definite articles in the light of the comparative and theoretical findings. Among the claims advanced are that definiteness does not occur in all languages though the pragmatic concept which it grammaticalizes probably does, that many languages have definiteness in their pronoun system but not elsewhere, that definiteness is not inherent in possessives, and that definiteness is to be assimilated to the grammatical category of person.
DEFINITENESS
DEFINITENESS

CHRISTOPHER LYONS

LECTURER IN LINGUISTICS
UNIVERSITY OF CAMBRIDGE
To the memory of my parents,
Edith and Patrick Lyons
CONTENTS

Preface xv
Abbreviations xviii

1 Basic observations 1
1.1 What is definiteness? 1
  1.1.1 Simple definites and indefinites 1
  1.1.2 Familiarity and identifiability 2
  1.1.3 Uniqueness and inclusiveness 7
  1.1.4 Identifiability, inclusiveness and indefinites 12
  1.1.5 A unified account? 13
1.2 Types of definite noun phrase 15
  1.2.1 Testing for definiteness 15
  1.2.2 Demonstratives 17
  1.2.3 Proper nouns 21
  1.2.4 Possessives 22
  1.2.5 Personal pronouns 26
  1.2.6 Universal quantifiers 32
1.3 Simple and complex indefinites 33
  1.3.1 Indefiniteness and cardinality 33
  1.3.2 Complex indefinites 36
1.4 The noun phrase 41
1.5 Definiteness beyond the noun phrase 45

2 Simple definites and indefinites 47
2.1 Definiteness and indefiniteness marking 47
  2.1.1 The occurrence of articles in languages 48
  2.1.2 The distribution of definites and indefinites 51
2.2 Definite: semantic content and behaviour 52
  2.2.1 Optional definite articles 52
  2.2.2 Anaphoric articles 53

ix
## Contents

2.2.3 Deictic distinctions 55  
2.2.4 Definite and specific 57  
2.2.5 Nominalizing and other functions 60  

2.3 Definite article types 62  
2.3.1 Free-form articles 63  
2.3.2 Bound articles 68  
2.3.3 Phrasal clitics 72  
2.3.4 Mixed systems and double determination 77  
2.3.5 Definite adjectives 82  

2.4 Other ways of expressing definiteness 86  
2.4.1 Adpositional marking 86  
2.4.2 Agreement 86  
2.4.3 Pronominal marking 88  
2.4.4 Word order 88  

2.5 Indefinites 89  
2.5.1 Indefinite article 89  
2.5.2 Quasi-indefinite article 95  
2.5.3 Types of cardinal article 100  
2.5.4 Paritive indefinites 100  
2.5.5 Bare indefinites 103  

2.6 General remarks 105  

3 Complex definites and indefinites 107  
3.1 Demonstratives 107  
3.1.1 Deictic distinctions 107  
3.1.2 Non-deictic distinctions 113  
3.1.3 Forms and positions 116  
3.2 Proper nouns 121  
3.3 Possessives 124  
3.3.1 Pronoun and full noun phrase possessives 124  
3.3.2 Affixal possessives 125  
3.3.3 Alienable and inalienable 128  
3.3.4 Possessives and definiteness 130  
3.4 Personal pronouns 134  
3.4.1 Semantic and grammatical content 134  
3.4.2 Strong, weak and null pronouns 137  
3.4.3 Personal determiners 141  
3.4.4 Definite and demonstrative personal forms 145  
3.5 Universals and other quantifiers 148
## Contents

3.6 Indefinites 148
   3.6.1 Full noun phrases 149
   3.6.2 Pronouns 150
3.7 Indefinite demonstratives 151
3.8 Vocatives 152
3.9 Non-configurational languages 153

4 Some semantic and pragmatic distinctions 157
4.1 One definiteness or several? 157
   4.1.1 Identifiability and inclusiveness 158
   4.1.2 Anaphoric and non-anaphoric definites 158
   4.1.3 “Deictic” and non-“deictic” definites 160
4.2 Specificity and referentiality 165
   4.2.1 Opacity and scope ambiguities 166
   4.2.2 Transparent contexts 170
   4.2.3 A unified account of specificity? 172
4.3 Generics 179
   4.3.1 Generics in English 181
   4.3.2 Singular generics 184
   4.3.3 Non-singular generics 189
   4.3.4 Generics and proper nouns 193
4.4 Concluding comments 198

5 Interaction with other grammatical phenomena 199
5.1 Direct object marking 199
   5.1.1 Differential case marking 200
   5.1.2 Definite object marking 202
   5.1.3 Oblique direct objects 205
5.2 Verb agreement 207
   5.2.1 Direct object agreement 207
   5.2.2 Indirect object agreement 211
   5.2.3 Subject agreement 212
5.3 Definiteness and animacy 213
5.4 Articles and nominal feature marking 215
5.5 Null and implicit noun phrases 219
   5.5.1 Pro 220
   5.5.2 Traces 221
   5.5.3 PRO 223
   5.5.4 Implicit arguments 223
## Contents

5.6 Classifiers 225  
5.7 Concluding point 226  

6 Definiteness effects 227  

6.1 Discourse structure 227  
6.1.1 The organization of information 227  
6.1.2 Information structure and definiteness 232  

6.2 Existential sentences 236  
6.2.1 The phenomenon 237  
6.2.2 Explanatory accounts 241  

6.3 Other definiteness effects 246  
6.3.1 Superlatives 246  
6.3.2 PP-extraposition and *any* opacity 248  
6.3.3 Property predication 249  
6.3.4 Internal-head relatives 250  
6.3.5 Concluding remarks 251  

7 Defining definiteness 253  
7.1 The grammatical, logical and pragmatic traditions 253  
7.2 Hawkins: inclusiveness and location 260  
7.2.1 Shared sets, inclusiveness and exclusiveness 260  
7.2.2 Subsequent developments 262  
7.3 Definiteness and quantification 265  
7.3.1 Milsark and the quantification restriction 265  
7.3.2 Generalized quantifiers 266  
7.3.3 Discourse semantics: Heim 268  
7.4 Relevance theory 271  
7.5 Well, what is definiteness then? 274  
7.5.1 Grammaticalization 275  
7.5.2 Grammaticalization and definiteness 277  
7.5.3 Simple and complex definiteness 279  

8 Definiteness and noun phrase structure 282  
8.1 The noun phrase as NP 282  
8.1.1 The specifier position and definite Dets 283  
8.1.2 Affixal articles 287  
8.1.3 The definite constraint 288  

8.2 The DP hypothesis 290  
8.2.1 Movement processes in DP 291  
8.2.2 Other functional categories in the noun phrase 294  
8.2.3 Weaknesses in the DP analysis 296
Contents

8.3 A modified DP analysis 298
  8.3.1 D as definiteness 298
  8.3.2 The content of determiners 301
  8.3.3 The analysis of personal pronouns 303
8.4 Configurationality and definiteness 305
8.5 Definiteness and person 310
  8.5.1 Personal determiners and pronouns 310
  8.5.2 Person and definite determiners 313
  8.5.3 The nature of person-definiteness 318
8.6 Summary of proposals 320

9 Diachronic aspects 322
9.1 The emergence of functional structure 322
  9.1.1 The development of DP 323
  9.1.2 A definite article cycle 326
  9.1.3 Indefinites and CardP 330
9.2 The origin of articles 331
  9.2.1 Demonstrative to definite article 331
  9.2.2 Articles in competition 334
  9.2.3 Numeral to cardinal article 335
9.3 The longer perspective 336
  9.3.1 Expansion in article use 336
  9.3.2 The life cycle of definite articles 337
9.4 In conclusion 340

References 341
Index 359
This book is primarily a survey, but, unlike some other topic-based books in this series, it surveys two areas. First, it offers an account of the range of variation displayed by languages in relation to definiteness and related grammatical concepts. Most languages do not have “articles”, and in those that do they vary strikingly in both their form and their range of use. All languages have demonstratives, personal pronouns, possessives, and other expressions which either seem to be inherently definite or to interact in interesting ways with definiteness; but again, there is considerable variation in the ways in which these expressions relate to definiteness. Second, the book gives a (very selective) outline of the theoretical literature on definiteness. This literature is vast, consisting both of direct accounts of definiteness and of work mainly concerned with other phenomena on which definiteness impinges. Both the cross-linguistic survey and the theoretical survey are introductory and far from complete, and many of the choices I have made in reducing the material to manageable proportions are no doubt arbitrary. This is true particularly as regards the literature, where I have had to omit much which I see as important, and it is essential that the reader follow up further references given in the works I do refer to.

This is not just a survey, however. I am much too interested in the topic not to want to present my own view of what definiteness is, and I believe the work gains in coherence from the aim of reaching and defending (if in outline) a preferred account. Chapters 7 to 9, in particular, contain much discussion of the approach I believe to be the most promising. But in the earlier chapters too, I have not hesitated to advance far-reaching claims anticipating this approach. The view of definiteness I propose may be wrong, of course, but it will have achieved its purpose if a student reading my proposals is spurred to investigate further and show their inadequacy. My aim in this book is not to present a set of facts and analyses to be learned, but to offer a body of ideas to be thought about and improved upon.

The investigation of definiteness necessarily takes the reader into several domains of inquiry, some of which (like semantics and syntax) are highly technical. While I assume some familiarity on the part of the reader with the principles and
methods of linguistics, I do not assume advanced competence in these domains, and I have given at appropriate points brief outlines of essentials and references to further reading, where possible at an elementary level. But it must be stressed that the interested reader would need to follow up these references, sometimes to a fairly advanced point, in order to come fully to grips with the issues in question. I have in general maintained neutrality between different theoretical frameworks, except as regards syntax, where I assume the principles-and-parameters approach which is the most highly developed and best known. Most of the text of this book was written at a time when the current “minimalist” version of this approach was in its infancy and there were few accessible accounts available of this framework to refer the reader to, so I have taken little account of minimalism. But there is little in the syntax discussed here which cannot be easily recast in this paradigm.

There has been much debate over the years on the relative merits of, on the one hand, the wide-ranging descriptive work of typological studies, and, on the other, the deep analysis of a smaller range of languages done in theoretical work. I firmly believe that descriptive breadth and analytical depth benefit one another, but that the latter must be the ultimate goal, and I hope that the gulf between these two approaches to language is narrower now than it was. The “new comparative linguistics” in generative work indicates a recognition among theorists of the value of cross-linguistic investigation, though some of it can be criticized as too selective in scope. But, to repeat a familiar point, our understanding of the way language works is deepened by bringing to bear serious analyses of languages, not mere observational facts. And even the best descriptive grammars are rarely adequate by themselves to provide the basis for an analysis of any depth of a specific aspect of linguistic structure. Indeed, even the descriptive observations and generalizations made in typological work must be treated with great caution, partly because the descriptive grammars on which they are based are often unclear on crucial points or analytically unsophisticated, partly because the typologist looking at unfamiliar languages in pursuit of a generalization is prone to the same inaccuracy as the theorist aiming to prove a point of theory. In my own cross-linguistic survey here, I too will certainly have included inaccuracies, and I urge the reader to treat it as a guide and starting point, not as fully reliable data.

Many people have helped me in various ways in the course of my writing this book, and I wish to thank in particular the following friends and colleagues who have read and commented on the manuscript or sections of it, or discussed particular points with me: Nigel Vincent, Deirdre Wilson, Noel Burton-Roberts, Kasia Jaszczolt. Most special thanks to Ricarda Schmidt for constant intellectual and moral support.

xvi
Preface

A couple of points concerning the presentation of the material should be noted. Where items of literature discussed exist in different versions, I have tried to refer to the most easily accessible version. In the case of doctoral theses subsequently published this means the formal publication. The effect is sometimes that my reference is to a version dated several years later than the version most commonly cited. Finally, a note on my use of gender-marked personal pronouns in describing conversational exchanges: I follow the convention that the speaker is, unless otherwise stated, female, and the hearer or addressee male.
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>1EXC</td>
<td>first person exclusive</td>
</tr>
<tr>
<td>1INC</td>
<td>first person inclusive</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>ABL</td>
<td>ablative case</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive or absolute case</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative case</td>
</tr>
<tr>
<td>ADESS</td>
<td>adessive case</td>
</tr>
<tr>
<td>ANA,</td>
<td>anaphoric</td>
</tr>
<tr>
<td>ana</td>
<td></td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASP</td>
<td>aspect</td>
</tr>
<tr>
<td>ASS1,</td>
<td>associated with first person</td>
</tr>
<tr>
<td>Ass1</td>
<td></td>
</tr>
<tr>
<td>ASS1EXC</td>
<td>associated with first person</td>
</tr>
<tr>
<td>ASS1INC</td>
<td>exclusive</td>
</tr>
<tr>
<td>ASS2,</td>
<td>associated with second person</td>
</tr>
<tr>
<td>Ass2</td>
<td></td>
</tr>
<tr>
<td>ASS3</td>
<td>associated with third person</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
</tr>
<tr>
<td>C</td>
<td>common gender</td>
</tr>
<tr>
<td>CL</td>
<td>clitic</td>
</tr>
<tr>
<td>CLASS</td>
<td>classifier, class marker</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous aspect</td>
</tr>
<tr>
<td>DAT</td>
<td>dative case</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative</td>
</tr>
<tr>
<td>DEF,</td>
<td>definite</td>
</tr>
<tr>
<td>Def</td>
<td></td>
</tr>
<tr>
<td>DEM,</td>
<td>demonstrative</td>
</tr>
<tr>
<td>Dem</td>
<td></td>
</tr>
<tr>
<td>DIR</td>
<td>direction</td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
</tr>
<tr>
<td>DU</td>
<td>dual number</td>
</tr>
<tr>
<td>ELAT</td>
<td>elative case</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERG</td>
<td>ergative case</td>
</tr>
<tr>
<td>EXP</td>
<td>experiential aspect</td>
</tr>
<tr>
<td>F</td>
<td>feminine gender</td>
</tr>
<tr>
<td>FUT</td>
<td>future tense</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive case</td>
</tr>
<tr>
<td>GENR</td>
<td>generic aspect</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual aspect</td>
</tr>
<tr>
<td>HON</td>
<td>honorific</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>IMPF</td>
<td>imperfect, imperfective aspect</td>
</tr>
<tr>
<td>IMPRS</td>
<td>impersonal</td>
</tr>
<tr>
<td>INAN</td>
<td>inanimate</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INESS</td>
<td>inessive case</td>
</tr>
<tr>
<td>INST</td>
<td>instrumental case</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>LINK</td>
<td>linker</td>
</tr>
<tr>
<td>LOC</td>
<td>locative case</td>
</tr>
<tr>
<td>M</td>
<td>masculine gender</td>
</tr>
<tr>
<td>N</td>
<td>neuter gender</td>
</tr>
<tr>
<td>NEG</td>
<td>negative</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative case</td>
</tr>
<tr>
<td>NONPAST</td>
<td>non-past tense</td>
</tr>
<tr>
<td>NUN</td>
<td>nunation</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique case</td>
</tr>
<tr>
<td>PART</td>
<td>partitive case</td>
</tr>
<tr>
<td>PASS</td>
<td>passive voice</td>
</tr>
<tr>
<td>PAST</td>
<td>past tense</td>
</tr>
<tr>
<td>PERS</td>
<td>person</td>
</tr>
<tr>
<td>PL</td>
<td>plural number</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>PRES</td>
<td>present tense</td>
</tr>
<tr>
<td>PRF</td>
<td>perfective aspect</td>
</tr>
<tr>
<td>PRI</td>
<td>primary case</td>
</tr>
<tr>
<td>PROX, Prox</td>
<td>proximal</td>
</tr>
<tr>
<td>PRT</td>
<td>particle</td>
</tr>
<tr>
<td>REFL</td>
<td>reflexive</td>
</tr>
<tr>
<td>REL</td>
<td>relative marker</td>
</tr>
</tbody>
</table>
Abbreviations

SG, Sg     singular number
SUBJ       subject
TNS        tense
TOP        topic
WH         interrogative

The standard labels are used for syntactic categories (N, V, D and Det, Agr etc.). Any idiosyncratic or non-standard labels used are explained at the appropriate point in the text.