Aboriginal people have been in Australia for at least 40,000 years, speaking about 250 languages. Through examination of published and unpublished materials on each of the individual languages, Professor Dixon, a pioneering scholar in this field, surveys the ways in which the languages vary typologically and presents a profile of this long-established linguistic area. The parameters examined include phonological contrasts, types of nominal case marking, patterns of verb organisation, varieties of pronoun systems, the development and loss of bound pronouns and a prefixing profile, generic nouns, nominal classifiers and noun classes, and ergative/accusative characteristics. The areal distribution of most features is illustrated with more than thirty maps, showing that the languages tend to move in cyclic fashion with respect to many of the parameters. There is also an index of languages and language groups. Professor Dixon brings a unique perspective to this diverse and complex material which will appeal to researchers and students in linguistics as well as to anthropologists with linguistic interests.

R.M.W. Dixon is Director of the Research Centre for Linguistic Typology at La Trobe University. Professor Dixon’s book publications include grammatical studies of five Australian languages, of a dialect of Fijian, and of English, as well as Where have all the adjectives gone? and other essays in semantics and syntax (1982), Ergativity (Cambridge, 1994), and The rise and fall of languages (Cambridge, 1997).
CAMBRIDGE LANGUAGE SURVEYS

General editors
J. Bresnan (Stanford University)
B. Comrie (Max Planck Institute for Evolutionary Anthropology, Leipzig)
W. Dressler (University of Vienna)
C. Ewen (University of Leiden)
R. Lass (University of Cape Town)
D. Lightfoot (University of Maryland)
S. Romaine (University of Oxford)
N.V. Smith (University College London)

This series offers general accounts of the major language families of the world, with volumes organised either on a purely genetic basis or on a geographical basis, whichever yields the most convenient and intelligible grouping in each case. Each volume compares and contrasts the typological features of the languages it deals with. It also treats the relevant genetic relationships, historical development and sociolinguistic issues arising from their role and use in the world today. The books are intended for linguists from undergraduate level upwards, but no special knowledge of the languages under consideration is assumed. Volumes such as those on Australia and the Amazon Basin are also of wider relevance, as the future of the languages and their speakers raises important social and political issues.

Volumes already published include
Chinese Jerry Norman
The languages of Japan Masayoshi Shibatani
Pidgins and Creoles (volume I: Theory and structure; volume II: Reference survey) John H. Holm
The Indo-Aryan languages Colin Masica
The Celtic languages edited by Donald Macaulay
The Romance languages Rebecca Posner
The Amazonian languages edited by R.M.W. Dixon and Alexandra Y. Aikhenvald
The languages of Native North America Marianne Mithun
The Korean language Ho-Min Sohn
AUSTRALIAN LANGUAGES
Their Nature and Development

R. M. W. DIXON
Research Centre for Linguistic Typology, La Trobe University
for Michael Osborne
Vice-Chancellor with vision
CONTENTS

List of maps xiii
List of abbreviations and conventions xv
Preface xvii
Acknowledgements xxii
Conventions followed xxiv
List of languages and language groups xxx

1 The language situation in Australia 1
  1.1 A partial picture 1
  1.2 Social organisation and lifestyle 3
  1.3 The languages 4
  1.4 Prehistory 7
  1.5 Diffusion of non-linguistic traits 12

2 Modelling the language situation 20
  2.1 Preliminaries 20
      2.1.1 Assumptions 20
      2.1.2 Types of similarity 21
      2.1.3 Family trees 22
      2.1.4 Diffusion 24
      2.1.5 The 50 per cent equilibrium level 27
  2.2 The Punctuated Equilibrium model 31
      2.2.1 Linguistic equilibrium 32
      2.2.2 Punctuation 33
  2.3 The Australian scene 35
  2.4 Split and merger of languages 40
      2.4.1 Language split 40
      2.4.2 Language merger? 41
Appendix The ‘Pama-Nyungan’ idea 44

3 Overview 55
  3.1 Semantics 56
      3.1.1 Actual/potential 56
      3.1.2 Volitional/non-volitional 57
      3.1.3 Primacy of generic terms 57
Contents

5.4.4 Purposive, dative, genitive and allative 166
5.4.5 Ablative and causal 168
5.4.6 Comitative and privative 170
5.4.7 Aversive 171
5.4.8 Summary of relations between forms 171
5.5 Conclusion 173

6 Verbs 176
6.1 Transitivity 176
6.2 Manner adverbs 181
6.3 Simple and complex verbs 183
6.3.1 Types of verbal organisation 187
6.3.2 A cyclic pattern of change 197
6.4 Verbal derivations 201
6.4.1 Semantic derivations 201
6.4.2 Syntactic derivations 202
6.4.3 Deriving verbs from nominals 207
6.5 Verb forms and inflections 209
6.5.1 Forms of inflections 213
6.5.2 Forms of verbs and development of conjugations 215
6.5.3 Loss of conjugations 224
6.5.4 Extended fusion 234
6.6 Nominal suffixes onto verbs 237
6.7 Copula and verbless clauses 239

7 Pronouns 243
7.1 Pronoun systems 243
7.2 Number-segmentable pronoun systems 246
7.2.1 Forms 253
7.3 Non-number-segmentable pronoun systems 262
7.3.1 Forms 266
7.4 The evolution of pronoun systems 285
7.4.1 Diffusion 292
7.4.2 Recurrent features of change and reanalysis 294
7.5 Pronominal case forms 299
7.5.1 Stage A 300
7.5.2 Stage B 307
7.5.3 Stage C-i 310
7.5.4 Stage C-ii 312
7.5.5 Stage C-iii 314
7.5.6 Summary 314
7.5.7 Non-core functions 315
7.6 Reflexives and reciprocals 319
7.7 Interrogatives/indefinites 327
7.8 Demonstratives 335
8 Bound pronouns 337
  8.1 What are bound pronouns? 341
  8.2 The predicate arguments involved 344
    8.2.1 Which arguments? 344
    8.2.2 How many in a clause? 345
    8.2.3 Case systems 347
  8.3 Choices 351
  8.4 Forms 353
    8.4.1 Free and bound pronominal forms 354
    8.4.2 Zero realisation 363
  8.5 Categories 365
    8.5.1 Person 365
    8.5.2 Number 367
  8.6 Position 370
    8.6.1 Number of positions 370
    8.6.2 Order of transitive arguments 371
    8.6.3 Location in the clause 374
  8.7 Links with other categories 377
  8.8 Patterns of development 379
  8.9 Possessive bound pronouns 394

9 Prefixing and fusion 402
  9.1 Verbs, coverbs and pronominal placement 409
  9.2 Structure of the verb in prefixing languages 416
    9.2.1 Valency-changing affixes 418
    9.2.2 Directional markers 419
    9.2.3 Negation and number 420
  9.3 Nominal incorporation 423
  9.4 Pronominal prefixes and TAM 429
  9.5 Pronominal prefixes to transitive verbs 437
  9.6 Implications 447

10 Generic nouns, classifiers, genders and noun classes 449
  10.1 Generic nouns and classifiers 454
    10.1.1 Semantics 456
    10.1.2 Grammar 458
  10.2 Feminine suffix -gan 460
  10.3 Gender in free pronouns 461
  10.4 Noun classes in non-prefixing languages 463
  10.5 Nominal prefixes in prefixing languages 468
  10.6 Noun classes in prefixing languages 470
    10.6.1 Noun classes and number 473
    10.6.2 Where noun classes are marked 474
    10.6.3 Semantics 485
    10.6.4 Markedness 488
    10.6.5 Forms 492
## Contents

10.6.6 Development  497  
10.6.7 Loss  506  
10.7 Noun classes and case marking  508  
10.7.1 The loss of case marking  509  
10.8 Summary  513  

11 Ergative/accusative morphological and syntactic profiles  515  
11.1 Development of morphological marking  515  
11.2 Syntactic pivots  520  
11.3 Antipassive and passive  530  
11.3.1 *-dharri and other suffixes that can mark antipassive and/or passive function  531  
11.4 Shifts in profile  536  
11.5 Summary  545  

12 Phonology  547  
12.1 Canonical systems  548  
12.1.1 Consonants  550  
12.1.2 Vowels  552  
12.1.3 Phonotactics  553  
12.1.4 Stress  557  
12.2 Laminals  558  
12.3 Apicals, including rhotics  565  
12.3.1 Stops and nasals (and laterals)  567  
12.3.2 Rhotics  573  
12.3.3 The apical problem  581  
12.4 Initial dropping and medial strengthening  589  
12.4.1 Loss and lenition of initial consonant  593  
12.4.2 Loss or shortening of first vowel  595  
12.4.3 Changes affecting \( C_2 \)  597  
12.4.4 Vowel copying and metathesis  598  
12.4.5 Changes at \( V_2 \)  599  
12.4.6 An overall perspective  600  
12.5 Stop contrasts, and fricatives  602  
12.5.1 Historical development, and loss  603  
12.5.2 Occurrence  605  
12.6 Glottals  615  
12.7 Other types of change  619  
12.7.1 Assimilation  619  
12.7.2 Dissimilation  625  
12.7.3 Further changes  627  
12.8 Vowel systems  628  
12.8.1 Vowel quality  628  
12.8.2 Evolution of additional vowels  631  
12.8.3 Occurrence  634  
12.8.4 Vowel length  638  

© Cambridge University Press  
www.cambridge.org
xii Contents

12.9 On the margin of a word 643
  12.9.1 Vowel-final languages 644
  12.9.2 Consonant-final languages 648
  12.9.3 Non-prototypical consonant clusters 653

13 Genetic subgroups and small linguistic areas 659
  13.1 Some genetic subgroups 659
  13.2 Small linguistic areas 668
  13.3 Origin places and directions of expansion 680
  13.4 Shifting isoglosses 686

14 Summary and conclusion 690
  14.1 Outline of development 691
  14.2 Diffusional patterns and cyclic change 695

References 700
Index of languages, dialects and language groups 719
Subject index 731
**MAPS**

<table>
<thead>
<tr>
<th>Map Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>Master map of language groups and languages xxviii</td>
</tr>
<tr>
<td>0.2</td>
<td>Geographical regions, state boundaries and major cities xxix</td>
</tr>
<tr>
<td>1.1</td>
<td>Likely coastline (and maximal extent of inland lakes) for the Australia/New Guinea/Tasmania land mass at about 25,000 BP (with the modern coastline superimposed) 8</td>
</tr>
<tr>
<td>1.2</td>
<td>Places known to have lacked the curved boomerang as a hunting and fighting weapon at the time of European invasion 14</td>
</tr>
<tr>
<td>1.3</td>
<td>Areas in which circumcision and subincision were practised 15</td>
</tr>
<tr>
<td>1.4</td>
<td>Approximate distribution of moiety, section and subsection systems 17</td>
</tr>
<tr>
<td>5.1</td>
<td>Distribution of the ergative allomorph -ngu 160</td>
</tr>
<tr>
<td>6.1</td>
<td>Types of verbal organisation in the west and central north 189</td>
</tr>
<tr>
<td>7.1</td>
<td>Types of pronominal system 245</td>
</tr>
<tr>
<td>7.2</td>
<td>2n-sg and 2n-min forms relating to nu-, gu- and nugu- 257</td>
</tr>
<tr>
<td>7.3</td>
<td>Occurrence of 2pl form nhurre 269</td>
</tr>
<tr>
<td>7.4</td>
<td>Occurrence of 2du form nth(m)balV1/V2 270</td>
</tr>
<tr>
<td>7.5</td>
<td>Forms for 1pl(inc) 274</td>
</tr>
<tr>
<td>7.6</td>
<td>Forms for 1pl(exc) and gana 275</td>
</tr>
<tr>
<td>7.7</td>
<td>Languages in groups B–Y, WA–WM lacking goli 278</td>
</tr>
<tr>
<td>8.1</td>
<td>Distribution of bound pronouns 340</td>
</tr>
<tr>
<td>8.2</td>
<td>Languages of the Wik subgroup, Bc 387</td>
</tr>
<tr>
<td>8.3</td>
<td>Baagandji (V), its dialects and neighbours 391</td>
</tr>
<tr>
<td>9.1</td>
<td>Languages showing nominal incorporation 424</td>
</tr>
<tr>
<td>10.1</td>
<td>Languages with noun classes or genders 453</td>
</tr>
<tr>
<td>10.2</td>
<td>Languages with prefixes to nominals 468</td>
</tr>
</tbody>
</table>
List of maps

11.1 Languages with switch-reference marking 529
12.1 Languages with a laminal contrast 560
12.2 Languages lacking an apical contrast in stops and nasals 566
12.3 Languages with one, or with three, rhotic phonemes 577
12.4 Languages with initial dropping 592
12.5 Occurrence of stop contrasts and fricatives 607
12.6 Languages with more or less than three vowels 630
12.7 Languages with a length contrast for (some or all) vowels 642
12.8 Languages where the phonological form of every (or almost every) word ends in a vowel 646
12.9 Languages where the phonological form of every (or almost every) lexical root ends in a consonant 649
13.1 Daly River group, NH, and neighbours 677
13.2 Subgroup Ja, with approximate lexical percentage scores with neighbours, and possible directions of expansion 682
13.3 Subgroup Nc, with approximate lexical percentage scores with neighbours, and possible directions of expansion 685
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>transitive subject function</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative inflection</td>
</tr>
<tr>
<td>ABL</td>
<td>ablative case</td>
</tr>
<tr>
<td>inc</td>
<td>inclusive (addressee included)</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive case</td>
</tr>
<tr>
<td>INCH</td>
<td>inchoative</td>
</tr>
<tr>
<td>ALL</td>
<td>allative case</td>
</tr>
<tr>
<td>INST</td>
<td>instrumental case</td>
</tr>
<tr>
<td>ANTIPASS</td>
<td>antipassive verbal derivational suffix</td>
</tr>
<tr>
<td>INTERROG</td>
<td>interrogative</td>
</tr>
<tr>
<td>APPLIC</td>
<td>applicative derivational affix</td>
</tr>
<tr>
<td>LOC</td>
<td>locative case</td>
</tr>
<tr>
<td>aug</td>
<td>augmented</td>
</tr>
<tr>
<td>M, m</td>
<td>masculine</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary verb</td>
</tr>
<tr>
<td>min</td>
<td>minimal</td>
</tr>
<tr>
<td>AVERS</td>
<td>aversive case</td>
</tr>
<tr>
<td>NEUT</td>
<td>neuter</td>
</tr>
<tr>
<td>BENEF</td>
<td>benefactive</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative case</td>
</tr>
<tr>
<td>BP</td>
<td>before present</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>CAT</td>
<td>catalyst</td>
</tr>
<tr>
<td>n-f</td>
<td>non-feminine</td>
</tr>
<tr>
<td>CAUS</td>
<td>causal case</td>
</tr>
<tr>
<td>n-min</td>
<td>non-minimal</td>
</tr>
<tr>
<td>CC</td>
<td>copula complement</td>
</tr>
<tr>
<td>n-sg, n.sg</td>
<td>non-singular</td>
</tr>
<tr>
<td>COMIT</td>
<td>comitative affix</td>
</tr>
<tr>
<td>O</td>
<td>transitive object function</td>
</tr>
<tr>
<td>CONTIN</td>
<td>continuative affix</td>
</tr>
<tr>
<td>CS</td>
<td>copula subject</td>
</tr>
<tr>
<td>OBJ</td>
<td>object (O function)</td>
</tr>
<tr>
<td>DAT</td>
<td>dative case</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>du</td>
<td>dual</td>
</tr>
<tr>
<td>pX</td>
<td>proto-language for subgroup X</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative case</td>
</tr>
<tr>
<td>PERF</td>
<td>perfect aspect</td>
</tr>
<tr>
<td>exc</td>
<td>exclusive (addressee excluded)</td>
</tr>
<tr>
<td>PL, pl</td>
<td>plural</td>
</tr>
<tr>
<td>F, f</td>
<td>feminine</td>
</tr>
<tr>
<td>POS</td>
<td>positive</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive affix</td>
</tr>
<tr>
<td>POT</td>
<td>potential</td>
</tr>
<tr>
<td>H</td>
<td>stop homorganic with preceding segment</td>
</tr>
<tr>
<td>PRES</td>
<td>present tense</td>
</tr>
<tr>
<td>PRIV</td>
<td>privative</td>
</tr>
</tbody>
</table>
### List of abbreviations and conventions

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PURP</td>
<td>purposive inflection (on nouns and verbs)</td>
</tr>
<tr>
<td>RECIP</td>
<td>reciprocal</td>
</tr>
<tr>
<td>REDUP</td>
<td>reduplicated</td>
</tr>
<tr>
<td>REL</td>
<td>relative clause marking</td>
</tr>
<tr>
<td>S</td>
<td>intransitive subject function; stressed syllable</td>
</tr>
<tr>
<td>sg</td>
<td>singular</td>
</tr>
<tr>
<td>SUBJ</td>
<td>subject (S and A functions)</td>
</tr>
<tr>
<td>SUBORD</td>
<td>subordinate marking</td>
</tr>
<tr>
<td>TAM</td>
<td>tense and/or aspect</td>
</tr>
<tr>
<td>TR, tr</td>
<td>transitive</td>
</tr>
<tr>
<td>U</td>
<td>unstressed syllable</td>
</tr>
<tr>
<td>TR, tr</td>
<td>transitive</td>
</tr>
<tr>
<td>UA</td>
<td>unit augmented</td>
</tr>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>-</td>
<td>affix boundary</td>
</tr>
<tr>
<td>=</td>
<td>clitic boundary</td>
</tr>
</tbody>
</table>
I began the preface to *The languages of Australia* (*LoA*, published 1980) by stating that it was, in several ways, premature. By this I meant that more descriptions of languages would be forthcoming during the 1980s and 1990s (as, indeed, they have been), which would provide a surer basis for generalisation. I now realise that *LoA* was most importantly of all, conceptually premature.

I had learnt the principles of historical linguistics from my teachers Warren Cowgill and Calvert Watkins, and from reading Meillet, Benveniste and others. And I had assumed that the methodology which applies so well for the languages of Europe and North America and Oceania would also be appropriate for the linguistic situation in Australia. It is not, but it took me a long time to realise this.

I sometimes wondered whether my lack of success in applying the established methodology of historical linguistics to the Australian linguistic situation was a feature of that situation, or a reflection on my abilities. Then, in the 1990s, I did intensive field work on Jarawara, spoken in southern Amazonia, and undertook a comparative study of the six languages of the Arawá family, to which it belongs. I found that here the established methodology worked perfectly (it was like a dream, after my struggles with the Australian situation). I was able to establish correspondence sets, compare their distributions, and then to reconstruct the phoneme system, more than four hundred lexemes, quite a bit of morphology, and some of the syntax for proto-Arawá. This easy success with Arawá emphasised to me the unusual – and probably unique – nature of the language situation in Australia.

The languages of Australia show recurrent similarities, such that almost everyone who has studied several of them (beginning with Grey 1841) has inclined towards the opinion that they must all be related. Related how? Well, presumably in the way languages in other parts of the world are related, as one language family. I belonged to this band. There was, we assumed, likely to have been an ancestor language, proto-Australian. *LoA* was the first serious attempt to put forward a hypothesis concerning proto-Australian. But the procedure followed was flawed. I used a selection of data from the clearest and most accessible descriptions available, most of these being of non-prefixing languages. (In the late 1970s, when the book was completed, there were only a handful of descriptions available for prefixing languages; these were all made...
The method was selective; by comparing similar paradigms in a number of languages, I reconstructed proto-paradigms, which were certainly sound and valid with respect to the data employed. However, they did not justify the label ‘proto-Australian’.

In preparing the present volume I have made use of all the available material on each of the 240–50 autochthonous languages of mainland Australia, taking account not only of the sixty or so good grammars produced during the past twenty years, but also examining and analysing the old (and often far from satisfactory) materials on languages from southern regions, which fell out of use many years ago. It will be seen – from the surveys of phonological and grammatical features presented throughout this book – that no clear picture emerges of what the full inventory of lexical and grammatical forms could have been for a putative proto-Australian.

It is natural to work in terms of the prevailing body of opinion in any discipline. I began (in the 1970s) by hypothesising that the Australian languages were likely to constitute a genetic family (like Indo-European and Austronesian) and that ‘Pama-Nyungan’ was a high-level subgroup within it. But a subgroup may only be established on the basis of significant distinctive innovations. Although at the time I wanted ‘Pama-Nyungan’ to be a subgroup, it proved impossible to uncover sufficient distinctive innovations to justify this. On pages 255–6 of *LoA*, I stated: ‘Pama-Nyungan – although a useful label to cover the large class of Australian languages which have not undergone radical changes that involve the development of pronominal and other prefixes to the verb, and a generally polysynthetic structure – has not yet been shown to have any genetic significance.’

Some of the reviewers of *LoA* suggested that what I had reconstructed to be ‘proto-Australian’ was in fact better labelled ‘proto-Pama-Nyungan’. This is not so. For example, proponents of ‘Pama-Nyungan’ as a genetic group have typically taken ergative -gu to be a diagnostic feature of ‘proto-Pama-Nyungan’. But forms relating to -gu are found in only about one-third of the languages regarded as ‘Pama-Nyungan’ and this form cannot really be imputed to a ‘proto-Pama-Nyungan’. In a major contribution to comparative Australian studies, Sands (1996) showed that an original ergative form *-dhu explains the great majority of modern ergative forms right across the continent (in both prefixing and non-prefixing languages, in both ‘Pama-Nyungan’ and ‘non-Pama-Nyungan’ languages).

As explained in the appendix to chapter 2, ‘Pama-Nyungan’ was first introduced within a lexicostatistic classification of Australian languages, in terms of a view that every language should be able to be placed on a fully articulated family tree. The validity of such a family tree was assumed; it was then just a matter of discovering the place on the tree appropriate for each language. Later, the ‘Pama-Nyungan’/‘non-Pama-Nyungan’ division was redefined in terms of languages without and with number-segmentable non-singular pronouns (see §7.2). Lexicostatistics has been shown to be based on non-sustainable premisses and to have limited usefulness anywhere in the world. It has the lowest applicability in Australia, where there is no distinction between ‘core’
and ‘non-core’ vocabulary in terms of borrowability. Yet this is one of the few parts of the world where people still have recourse to lexicostatistics, as an indicator of genetic relationships.

I have – over the past thirty and more years – experimented with many ways of accounting for the relationship between Australian languages. In this volume I work in terms of what appears to me to be the only appropriate model for this quite special language situation. My essay *The rise and fall of languages* (1997) was conceived as a prolegomenon for the present work. It suggests that – during the hundred thousand years or more during which humankind has had language – there have been, in each geographical region, long periods of equilibrium broken by short periods of punctuation. During punctuation there is expansion and split of peoples and of languages; here a family tree diagram will model what happens. During equilibrium periods the number of languages spoken within a given region will remain roughly constant; there will be diffusion of cultural and linguistic traits so that the cultures will become more similar and the languages will tend to converge towards a common prototype.

It is generally accepted that the first people settled in the Australia/New Guinea land mass at least forty thousand (and probably fifty thousand) years ago. The spread of people and languages around the continent would have been a period of punctuation, which a family tree diagram would have modelled. But this is likely to have been completed within a few thousand years. It is likely that for tens of millennia the non-mountainous/non-forested part of the Australia/New Guinea land mass has constituted a linguistic equilibrium area. It is this which has to be investigated and described. Interestingly, there are a number of putative low-level genetic subgroups, pointing to minor punctuations in quite recent times (some probably due to expansion into previously unoccupied territory, as water resources became more abundant). A number of these genetic groups have been established, by reconstruction of the proto-language and the systematic changes through which the modern languages have developed; for others this remains to be done. On the evidence available, it seems most unlikely that the low-level genetic groups will be relatable together in terms of higher-level genetic groups. The question of whether all Australian languages go back to a single ancestor is not answerable, because of the great time-depth involved. All that we can perceive is a well-established equilibrium situation, across the continent; this is what must be studied. The most notable feature about the languages of Australia is that they do, without doubt, constitute the longest-established linguistic area in the world.

A major finding of the work reported in this volume is that Australian languages tend to vary in terms of a number of typological parameters, and to change with respect to them in a cyclic fashion – moving from type A to B to C and then back to A (some parameters shift in only one direction while others may be bidirectional). These cyclic changes are discussed through the volume and summarised in the final chapter, 14.

The only definite dates in this book are those provided by geographers for things like the rise and fall of sea level and by archaeologists for things like the earliest trace
of humans, and of dingos. I will often comment that a certain feature appears to be of relatively recent origin, and that something else appears to be relatively ancient. What dates should be attached to recent versus ancient? Should it be a few hundred years versus a few thousand? Or a few thousand versus a few tens of thousand? I don’t know.

This volume includes a number of maps, most of which show the isogloss for a linguistic feature. It will be noted that the isoglosses do not bunch. For ease of reference, the languages have been arranged in fifty groups (some genetic, some small linguistic areas, some simply geographical). The twelve groups relating to languages classified as ‘non-Pama-Nyungan’ are given a first letter N (NA–NL). This is to enable the reader to see that only one isogloss (that for number-segmentable non-singular pronouns) runs along the division between groups NA–NL, on the one hand, and the remaining groups (labelled A–Y and WA–WM), on the other hand. The ‘Pama-Nyungan’ idea is a pervasive one, used both by Australianists and by linguists at large. However, it is totally without foundation and must be discarded if any progress is to be made in studying the nature of the linguistic situation in Australia.

It should also be pointed out that comparison between work on Australian languages and work on languages from other parts of the world is otiose. For example, Crowley (1997: 275) offers the comment that ‘an Australian Dempwolff has yet to arrive on the scene’. Now Dempwolff (1934–8) compared the vocabularies of a number of Austronesian languages and reconstructed the forms of more than 2,200 lexemes in proto-Austronesian. There is no way that anyone – ‘a Dempwolff’ or anyone else – could do this for Australia. Capell (1956: 85–93) recognised thirty-six recurrent lexical forms which he labelled ‘common Australian’. In §4.2 below the inventory is expanded to about 120 recurrent lexemes, each of which is found beyond a single geographical region. Doubtless a number more could be added after intensive comparison of vocabularies from languages in different parts of the continent. But this is a long, long way short of 2,200. The language situation in Australia is simply unlike that of Austronesian; or of Indo-European or Uralic or Uto-Aztecan. It is unique.

The materials available on Australian languages are not of uniform quality; this applies to early materials from the nineteenth century and also to recent descriptions from the end of the twentieth century. Grammars vary in terms of their accessibility (how clearly organised they are, and how easy it is to find things in them) and – most importantly – in terms of their quality. When one examines some recent descriptions, one does not have full confidence that, for instance, the appropriate inventory of phonemes has been recognised, or that the morphological analysis has maximal explanatory power. It is a convention in academic society today (and perhaps slightly more in Australia than anywhere else in the world) that one should hesitate to criticise the work of colleagues. Indeed, there is a tendency on the part of many people to assume that EVERYTHING which has appeared in print is equally valid and correct. A close investigation shows that this is not the case. In some instances linguist A works on language X and publishes a slim grammar; then linguist B works on the same language and
Preface 21

produces a rich and exquisite description, revealing and explaining complications that had passed A by. In many cases only one linguist has worked on a language but it is not hard to assess – by the lack of internal consistency, and by the general way in which a grammar is written – that this work is of less than adequate quality.

In this volume I take account of all published and unpublished materials, but I have placed most reliance on those that are written in an accessible manner, and especially on those that I consider to be of good quality. For example, I look to see whether a writer provides explicit criteria for the analytic decisions they make rather than, say, assuming that the categories of traditional Latinate grammar (or those of some current formal theory) will apply.

The reader must be alerted to the fact that I have not attempted to provide an exhaustive survey of the literature. To have included a summary of every idea that has been put forward concerning the relationship between Australian languages would have added considerably to the length of this book, and would have made it less coherent and less readable. There are right indented and unjustified passages discussing a number of alternative analyses. In addition, references are provided within the text to good-quality discussions of points that I deal with. But I must ask the reader’s indulgence for only including bibliographical references that are strictly relevant to the overall thesis which is developed in the volume.

A companion volume is in preparation (but still some way from completion): Australian languages: a complete catalogue. This will consist of a short account of each of the 240–50 languages, giving tribal and dialect names, traditional territory and current situation, plus a summary of the main phonological, morphological and syntactic features, and an annotated list of published and unpublished source materials.

The older one gets the more one learns; and, at the same time, the greater the realisation one has of the vast amount one does not know. A young scholar is likely to have confidence that relevant problems can be neatly stated and satisfactorily answered. As the years advance, one tries rather to clarify the nature of the problems, and to formulate some ideas towards their solution.

All scientific progress is cumulative but sometimes a discipline becomes enmeshed in a cul-de-sac of its own making. One needs to make a sidestep in order then to continue to move forward. I have tried, in this book, to provide something of the foundation for further work on the indigenous languages of Australia which, hopefully, will enable future generations of scholars more fully to understand the nature of these languages, and of their development and interrelations.

July 1996 (Canberra) – April 2001 (Melbourne)
ACKNOWLEDGEMENTS

Over the past decades I have been helped by many people – by virtually everyone who has worked on Australian languages. The students whose theses I have supervised have, in their turn, educated me. My linguistic friends across the continent have become used to receiving questions (by phone or fax or letter); they have been unfailingly generous and courteous in providing answers. Over the years I have employed a variety of research assistants; the quality and quantity of their work has, of course, varied, but all have made useful contributions. The preparation of this volume has been greatly facilitated through the preparation of preliminary materials by Kristina Sands (concerning the morphology of nouns, verbs, pronouns, verbal forms, subordinate clauses and phonology); by Angela Terrill (on valency-changing derivations, moiety/section/sub-section systems, and in compiling the forms of twenty-two nominal lexemes across all languages of the continent); and by Lys Ford on the Daly River region as a small linguistic area.

The arrangement of languages into groups was done in collaboration with Kristina Sands and Angela Terrill for groups A–Y and WA–WM, and in collaboration with Rebecca Green, Ian Green and Kristina Sands for groups NA–NL.

Jennifer Elliott keyboarded a large part of the manuscript, Anya Woods checked the final draft and assisted with the indexes and Andrew Hardie drew the maps. I thank them for the skilful and intelligent way in which they performed these tasks. Kurt Lambeck and Tony Purcell were generous in supplying map 1.1, and the most up-to-date information concerning sea levels in times past.

The following linguists read all or part of a draft of the book and made most welcome suggestions for improvement: Juliette Blevins, Andrew Butcher, Alan Dench, Nicholas Evans, Lys Ford, John Hajek, Harold Koch and Tasaku Tsunoda. I owe a huge debt to those colleagues and friends who read patiently through every page, offering more data, cogent commentary and productive criticism – Alexandra Aikhenvald, Barry Alpher, Barry Blake, Gavan Breen and Peter Sutton.

The work on which this book is based has been supported, since 1972, by a succession of generous grants from the Australian Research Grants Committee and its successor the Australian Research Council. The Australian Research Council awarded me
Acknowledgements

a Senior Research Fellowship from 1991 until 1996, and then two further Fellowships (1996–2001 and 2001–6); in addition, they allocated me a Special Investigator Award from 1997 until 1999. These Fellowships and the Special Investigator Award have helped immeasurably in allowing me time to devote to the project, and in providing appropriate research assistance.

A great deal of the work for this book was done while I was at the Australian National University, from 1970 until 1999. During most of this period the ANU provided a fine work environment. From 2000 La Trobe University provided a new base for me, and for the Research Centre for Linguistic Typology. It is within this convivial and academically exciting environment that the volume has been brought to completion. My warmest thanks go to Professor Michael Osborne, Vice-Chancellor of La Trobe University, for inviting us to be a foundation unit within his Institute for Advanced Study, and for providing ideal working conditions within which the work of language description and comparison – and typological generalisation – can proceed apace.
Different scholars use different criteria and terminology. Plainly, in an integrated account such as the present volume, a consistent set of conventions must be adhered to. Inevitably I do, on occasion, describe and interpret the structures of an individual language in a way different from that of the linguist who wrote the grammar. I trust that my colleagues will appreciate that this is done not in a spirit of disagreement, but simply for pedagogic expediency and for consistency.

(1) The term language is used through this book in the technical sense of linguists: mutually intelligible forms of speech are regarded as dialects of a single language. It is feasible to produce an overall grammar of a language (in the linguistic sense) with notes on dialect variation. The other sense of ‘language’ treats the speech of each political group (in Australia: each tribe) as a distinct language; I refer to these as tribal dialects of languages (in the linguists’ sense). Note that in the linguists’ sense there are (or were) about 240–50 distinct languages in Australia (using ‘language’ in the political sense there would be at least seven hundred languages, possibly a thousand or more).

(2) The term subgroup is here used in a special way. Generally, a subgroup is a lower-level genetic grouping within a genetically established larger group, a language family. No large genetic groups are recognisable within the Australian linguistic area, but there are a number of low-level genetic groups, mostly consisting of just two or three languages. Rather than describing each of these as a small language family, I refer to them as ‘low-level genetic subgroups’. This leaves open the possibility that some of these subgroups may eventually be shown to be linked together in higher-level genetic groupings.

(3) I regard Australian languages as making up a large linguistic area, with the purpose of the present book being to try to characterise the Australian linguistic area. Within the larger area we can recognise a number of small linguistic areas; the languages in each area have much greater similarity to other languages in the area than...
to languages outside the area; but these are not sufficient to justify suggesting that they make up a genetic subgroup. I use the terms small linguistic area and areal group interchangeably.

(4) There are so many indigenous languages in Australia that referring to them in a book of this sort poses problems. It would be mind-numbing to have to refer to each language as an individual entity in describing areal patterns. For ease of reference I have organised them into fifty groups, labelled A–Y; WA–WM (where W stands for west) and NA–NL (where N stands for north). Each group includes between one and twenty-three languages. These are distinguished by the use of lower case letters (for groups within groups) and then numbers. For only two groups are their labels mnemonic – Y for the Yolngu subgroup; and WD for the Western Desert language.

- Some of the groups are tentatively identified as low-level genetic subgroups and labelled ‘subgroup’. A few of them include further subgroups as branches. For example, subgroup B, North Cape York, consists of further subgroups Ba, Northern Paman, and Bc, Wik; and also Bb, which is a single language, Umpila. There are six languages in Bc – Bc1, Wik-Ngathan, Bc2, Wik-Me’nh, etc. An asterisk, *, after a letter indicates the likelihood that all the languages in this group can be shown to make up a low-level subgroup, e.g. B* shows that B is probably a subgroup and Ba* that Ba is probably a subgroup within B. If two languages within a group are probably genetically related then * is included after each of their numbers, e.g. 1* and 2* within De (there is insufficient information on De3 to be able to decide whether this belongs in the subgroup with De1 and De2, although it is possible that it does).

I sometimes refer to a form reconstructed for the proto-language of a subgroup. The abbreviations ‘pZ’ is then used for ‘proto-Z’, where Z is the identificatory code for the subgroup.

- Other groups are tentatively identified as small linguistic areas and labelled ‘areal group’. For example, areal group U, Lower Murray, consists of five languages – U1, Yaralde, U2, Ngayawang, etc.

- The remaining groups (labelled just as ‘group’) simply consist of languages grouped together on a geographical basis; for example D, the South-east Cape York Peninsula group.

It must be stressed that the identification of a group of languages as a subgroup or as an areal group is in almost all instances tentative, and may need to be rethought when more descriptive material is available, and when more comparative work has been completed. For some languages that are no longer spoken the material available is poor, and it is likely that it will never be possible to arrive at definitive judgements concerning their affiliation.

The system of using code letters and numbers to refer to groups and to languages is intended to assist in describing areal patterns, and to help the reader identify where a given language is (or was) spoken. But its use does require some persistence. A course of action recommended to assiduous readers is to photocopy the master map
Conventions followed

(map 0.1) – and perhaps also the List of languages and language groups – and to keep these on the side when studying the volume. In this way, statements such as ‘feature so-and-so is found in languages from groups B–G, WH–WJ and NE–NG’ can easily be provided with a geographical reference.

(5) Through the book I shall often state that a certain category or form is found in a number of groups. For example, §4.2.6 states that a form *bula-* ‘two’ is found in groups H–R, T, V, E, WA–WB and WG. This indicates that the form is found in one or more languages from each of these groups, not (unless explicitly stated) in every language from each group.

(6) There is wide variation in orthographic conventions for writing Australian languages. For those sounds occurring across a fair range of languages I have used a practical orthography, which employs digraphs consisting just of letters of the roman alphabet (except that the velar nasal is always written as ʔ). This is set out in table 3.2 and in table 12.1. Other sounds which occur in just a small number of languages, are generally represented by IPA symbols (for example, ɓ, ʄ, ю).

For stop series I generally use voiced (b, g, d̪, d̪h, d̪, r̂d) or voiceless (p, k, tj, th, t, r̂t) symbols according to the convention normally followed for the language in question. In general and comparative discussion the voiced series is generally employed.

Vowels are written as i, a, u (plus e, o, where applicable). That is, I never write u as ‘oo’, or i as ‘ee’.

Material in Australian languages included in the text is generally in italics. Where italics are used this indicates that I believe the material is being given in phonemic form. Sometimes, material is quoted from an old source in the form in which it appears there (without any attempt to phonemicise it); it is then given in roman font within quotes, e.g. ‘possum’ is ‘pilla’ in WBb1, Parnkalla.

(7) In the introductory discussion (especially in parts of chapter 3) I have not hesitated to illustrate general points with data from languages I have worked on myself, since I am then certain of their appropriateness. In later chapters little data come from my own field work. If an example is taken from a published source (or a thesis), this is generally given (including the page on which it occurs in the source).

(8) My previous study, *The languages of Australia* (1980), began with a number of introductory chapters on the history of the study of Australian languages; tribe and language; speech and song styles (here briefly summarised in §3.4); and the role of language in Australian Aboriginal society today. These have dated very little and can still be read as a general orientation to the study of Australian languages. It did not seem appropriate to repeat them (in slightly revised form) here.

There is today great interest in assisting endangered languages to survive. And there is interest by communities whose language has ceased to be spoken in trying to start
speaking it all over again (even though, in most cases, the information recorded on the language – when it was still spoken – is limited). It did not seem appropriate, in a book of this nature, to comment on such projects.

The present volume is a reworking and extension of the typological survey in chapters 5–13 of the 1980 book. The great majority of points are new, but a few are the same as in the earlier survey. In such instances I have tried to use different examples to illustrate a point, rather than repeating an example given before. Only occasionally have I retained an earlier example.

(9) Very little was recorded on the languages of Tasmania before they passed into disuse, and there is no reliable information that could be included in the present survey. A summary of the available information is in Crowley and Dixon (1981).
Map 0.2 Geographical regions, state boundaries and major cities
LANGUAGES AND LANGUAGE GROUPS

The fifty groups into which I have divided the languages are listed here, together with some of the dialects within languages. Note that I have not attempted to provide an exhaustive list of all dialects of all languages.

There is generally a name for each tribal dialect but often no name for a language (in the linguists’ sense). In some such cases I have had to manufacture names – see §1.3. I have in most cases avoided names which appear to have been invented by White observers but had no validity for Aborigines in traditional times. For example, ‘Kurnay’ or ‘Gaanay’ (from the lexeme ‘man’) for the Gippsland language; I have preferred to label the language ‘Muk-thang’, the name of one of its dialects.

Language and dialect names are spelled in a variety of ways in the literature. I have generally spelled names according to the orthographic conventions followed in this book (see table 3.2 and table 12.1), but in some cases I have instead used a spelling that is well established in the literature. (Whichever spelling is used for a given language, it is likely that some people will applaud and others criticise.)

Some of the names for groups are taken over from the recent literature; for example, WL, Arandic, and WJb, Yapa. Others have been made up. The lexicostatistic classification often uses the term for ‘man’ but in many cases the form used is found in only some of the languages of the group to which it is applied. For instance, subgroup WJa was labelled ‘Ngumpin’, although this form is only found in one of the four languages in the subgroup (WJa3, Gurindji); I have avoided such inappropriate labels.

In writing this volume I have attempted to make use of all available sources on each language. For a number of languages there is a full or fullish grammar of good or quite good quality, or else one or more important papers. These are listed below, for the languages or language groups to which they apply. Where no reference is given here, there are simply a number of minor sources, which will be listed in the relevant entry in the planned companion volume.

A TORRES STRAIT GROUP (These are Papuan languages, not closely related to each other. A1 has a significant Australian substratum. A2 has been
said to be genetically linked to the languages of the Fly River region of New Guinea but this is not proven.)

A1 West Torres (also: Kalaw Kawaw Ya, Kala Lagaw Langgus, Yagar Yagar) – Ford and Ober (1986)
dialects on each island: Mahniag, Badu, Dauan, Saibai, Boigu, Moa, Yam, Sue, Yorke and Coconut

A2 East Torres (or Miriam or Meryam Mir) – Piper (1989)

R* NORTH CAPE YORK SUBGROUP

Ba* Northern Paman subgroup – Hale (1964, 1976b)
Ba1 Gudang
dialects include: Djabraga
Ba2 Uradhi – Hale (1976c), Crowley (1983)
dialects: Angkamuthi, Yadhaykenu, Atampaya
Ba3 Wuthati
Ba4 Luthigh
further dialect: Mpalitjanh
Ba5 Yinwum
probable further dialect: Njwadhai
Ba6 Anguthimiri – Hale (1966b), Crowley (1981)
dialects: Nggerikudi (or Yupungati), Tjungundji, Mpkwithi, Awnthim (with subdialects Mannngayt, Ntrwa’ngayth, Thyangngayth), Nta’angith, Alngith, Linnngithigh
Ba7 Ngkoth
dialects: Tootj (or Kauwala); Ngaawangati (or Ungauwangati)
Ba8 Arimatngithigh (or Ariminngayth)
further dialects: Latamngit
Ba9 Mbwyween
Ba10 Andjngith

Bb Umpila – Thompson (1988)
further dialects: Kunku Yani, Uutaalnganu, Kunku Ya’u, Kunku Iyu, Kaantju

Be* Wik subgroup
Be1 Wik-Ngathan (= Wik-lintjen) – Sutton (1978)
further dialect: Wik-Ngatharr (= Wik-Alken, = Wik-Elken)
Be2 Wik-Me’n
further dialects: Wik-Ep (= Wik-Iit), Wik-Keyenganh
Be3 Wik-Mungkhu (Wik-Munkan)
further dialect: Wik-Iyanj (= Wik-Iyjenj, = Wik-Iyjenji, = Mungkkanh)
Be4 Kugu-Muminh (Wik-Muminh) (or Kugu/Wik-Nganhcar) – Smith and Johnson (2000)
further dialects (all preceded by Kugu/Wik-): Mu’in, Uwanh, Ugbanh, Yi’an, Mangik, Iyanh
Be5 Bikanha (or Ayahakatiu)
Be6 Ayabadjhu

C Umbindhamu (or Umpithamu)

D SOUTH-EAST CAPE YORK PENINSULA GROUP

Da* Lama subgroup
Da1 Morroba-Lama (or Umbuyganu)
Da2 Lama-Lama (or Mba Rumbathama or Bakanambia or Wanbara)