1

Introduction to the Languages of Australia

1.1 Introduction and Overview

This is a book about the morphology and phonology – the *word structure* – of the Indigenous languages of Australia. While there have been previous general surveys of Australian languages in book form (in particular Bowern 2023; Dixon 1980, 2002; Evans 2003c; Koch & Nordlinger 2014) and of aspects of these languages such as discourse structure (Mushin & Baker 2008), to date there are no book-length overviews of the kind presented here. We provide up-to-date information on the segmental inventories, phonotactics, morpheme structure constraints, and morphological structures of Australian languages from across the continent, covering topics such as word minimality, metrical structure, vowel and consonant harmony, phonological alternations, markedness and universals, sound change, and opacity. Much of this material is not easily available to the broader linguistic community, even within Australia, and has not hitherto featured in theoretical discussions. We hope to rectify this lacuna in the literature and to encourage further research while there still remain speakers of this fascinating and unique group of languages. We especially hope to encourage students to work in this area. In this chapter we provide an overview of the genetic relationships of the Indigenous languages of Australia and discuss the possibility of links outside, and we briefly summarise the similarities among Australian languages. Section 1.5 presents a lengthy summary of aspects of the grammar of Australian languages, both as a reference for readers who are unfamiliar with them, to get an overview, and also to provide a way of finding information of interest within the book. Contributions from Australian languages to phonological and morphological theory are also touched on in this chapter, with more detailed expositions at the relevant sections in later chapters.

1.2 Coverage

We begin with some comments on what this volume aims to cover. Our primary focus in this volume is those phonological patterns that are

characteristic of Australian languages and their implications for phonological theory and speech science more generally. We are particularly keen to consider the implications for psycholinguistics and processing, two approaches from which researchers have barely started examining Australian languages but which we expect will come to dominate thinking in linguistics, precisely because they can reveal the behaviour and internal grammars of speakers. We aim to cover not just the descriptive contributions of fieldworkers but also the burgeoning literature on experimental, psycholinguistic, and instrumental studies of Australian languages that have appeared since Evans' (1995b) survey. This new work has profound ramifications for how we understand the sound patterns and word structure of Australian languages, and it represents data and analyses that were not available to previous overviews.

1.3 Previous Surveys of Australian Phonology and Morphology

In our view, the fundamental characteristics of Australian sound patterns were synthesised in three works: Dixon (1980), Busby (1980), and Hamilton (1996). Dixon (1970, 1972, 1980) provides overviews of the characteristic phonemic inventories, oppositions, and phonotactics (and see further Ch. 2). Busby's (1979) MA thesis, published as Busby (1980), is a survey of the segmental inventories and acoustic characteristics of Australian languages, the first of its kind and the most extensive survey of Australian phoneme inventories until Hamilton (1996) and Round (2023e). Butcher (1993) is an unpublished book manuscript with a wealth of original acoustic and articulatory data and discussion of several Australian languages, mostly of the Northern Territory. Evans (1995b) is an article-length survey of phonological patterns and phonotactics of Australian languages, with some discussion of issues relevant to theory – for example the nature of glottal stop and retroflexion as autosegments (see §2.4.3, §2.7.1), the question of whether some languages are best analysed in terms of a vowel–consonant (VC) syllable (see §3.5.1), the contribution of onsets to weight (see §4.5), evidence for constituents of the Prosodic Hierarchy (see §4.2, §5.1, §5.4.2), metrical structure and the importance of morphological structure (see §4.2.3, §4.3, §6.3.2), and reduplication (see §5.5). Hamilton (1996) is a comprehensive survey of consonant inventories and phonotactics in Australian languages, together with an extended analysis that uses Steriade's (2001) model of 'Licensing by Cue'; (see §3.3). Dineen (1990, published as Fabricius 1998) and Parncutt (2015) are, both, surveys of reduplication (reduplication is discussed in §5.5). Dixon (2002), a monograph on the languages of Australia with particular attention to comparative issues and reconstruction, includes a survey of phonemic inventories, phonotactics, and phonological behaviour (the latter very brief). Recent overviews of the inventories and phonetic characteristics of Australian languages can be found in Fletcher and Butcher (2014), Tabain (2023), and Round (2023e).

A chapter-length overview of phonological patterns (inventories, phonotactics, prosodic morphology, stress, alternations) is in Baker (2014), which updates and extends Evans (1995b) and was the basis for the current volume. A number of recent chapters by Round extend these previous works: Round (2023b) surveys lenition (see §3.6.1, §7.4) and assimilation; Round (2023c), nasal cluster dissimilation (see §5.4.3); and Round (2023d), phonotactics (Ch. 3).[1]

1.4 Structure of the Volume

Following the overview in §1.5, which follows, Chapter 2 discusses segmental inventories, segmental alternations, and theoretical proposals related to place, laryngeal, and manner contrasts in Australian languages. Chapter 3 focuses on phonotactics, the relationship between phonotactic alternations and the lexicon, loanwords, and contact phonology. Chapter 4 looks at metrical structure: stress in roots and complex words. Chapter 5 examines prosodic words, prosodically determined allomorphy, long-distance alternations, and reduplication. The topic of Chapter 6 is phonological domains: the characteristic features of utterance, phrase and word boundaries, the status of morphemes as domains, and the issue of how to model the phonology–morphology interface. Chapter 7 is a discussion of the evidence for sound change in Australian languages, the vexed problem of irregular sound change, and the relationship between sound change, the lexicon, and alternations, focusing on a few well-studied languages. There is no chapter on alternations per se; rather alternations are discussed throughout, in the relevant chapters.

1.4.1 What Counts as Phonology?

What counts as phonology very much depends on the standpoint of the analyst and the theory that the analyst assumes. As Kiparsky (2012 p. 119) puts it,

> What is the form of morphophonemic representations? How far removed are they from the phonetic or phonemic surface? The whole gamut of possible answers to this question has been given in modern linguistics.

[1] After submission of this manuscript for review, Bowern (2023) appeared, which contains several chapters on morphology, phonology, phonetics, and sound change in Australian languages, including those mentioned in this section. Round's chapters are quantitative in approach, are based on a database of nearly 400 language varieties, and to some extent update the quantitative surveys of Busby and Hamilton. They are valuable as a guide to the extent of certain kinds of phonological contrasts and alternations. One of Round's points of focus is the relationship between typological variation and genetic groupings; this relationship is not a focus of the current volume. We think it is fair to point out that, since Round's chapters are short, there is limited scope for an exemplification of the data or a discussion of theoretical issues, and thus his work is complementary to ours. We have added references to this work where appropriate throughout.

For some phonologists, 'phonology' is everything that an analyst can write a rule for; Chomsky and Halle (1968) is a well-known example of this approach. Chomsky and Halle were criticised for this, however, because the underlying forms they proposed in some cases were felt to deviate too far from what a child could reasonably be expected to infer about the lexicon (Kiparsky 1982a). Ultimately, we must base our understandings of what counts as phonology on the behaviour of speakers as revealed not only through their speech patterns, but also through their responses on experimental tasks, in first language acquisition, language change, neurolinguistic investigations, and so on. Although we have our own views on how much of the lexicon and sound patterns is accessible to speakers (see §6.3), in this book we take a fairly broad view of what is or might be relevant.

One recurrent factor that has played a role in discussions of this issue is productivity: to what extent can a phonological process be described just with reference to the phonological context, without reference to (for instance) exceptions, lists of morphemes to which it applies, or word classes? In Lexical Phonology, this distinction was captured by reference to 'levels' or 'strata': at the 'stem' level, phonological rules were subject to exceptions, could apply to classes of morphemes, produce phonemes as outputs, and be conditioned morphologically as well as phonologically. At the 'post-lexical' level, phonological rules were held to be exceptionless, produce allophones, and be conditioned only by phonological elements (natural classes and boundaries) (see, among many others, Borowsky 1993; Kaisse & Shaw 1985; Kiparsky 1982a; Rubach 1984, 2008; Rubach & Booij 1990). In between these levels, Lexical Phonology proposed a further, lexical level: the 'word' level, in which surface constraints (in OT terms) applied in order to produce primarily syllabification and juncture phenomena (Borowsky 1993) – although Kiparsky (2018) argues that this is the appropriate level for mental representations of lexemes. We have argued in previous work for the relevance of this kind of model to Australian languages (Baker 2008b; Baker & Harvey 2003; Borowsky & Harvey 1997; and see §6.3 for further discussion of this approach and its broader relevance).

Related to the question of a division between lexical and post-lexical phonology, there is also the long-standing problem of where to draw the line between historical and synchronic phonology, a line that has been conventional since de Saussure (1916). Writers such as Blevins (2004), however, have brought the issue back into focus (see also Kiparsky 2016; Labov 1994; Weinreich *et al.* 1968). Since variability and sound changes are a constant factor in languages and the researcher can decide only after the fact when a sound change has finished, it is impossible in principle to disentangle the domain of historical linguistics from that of phonology. These issues are touched on in §3.6.1, §4.4.3, §6.3, and more extensively in §7.4. Since Weinreich, Labov, and Herzog, the use of large corpora to investigate language variation and change has become standard. This kind of method is, however, limited by the lack of large corpora for most Australian languages, but Marley (2020) is an excellent recent example of the use of a diachronic corpus to

investigate a number of changes in the morphology and phonology of Bininj Gun-wok (one of these changes is discussed in §6.2.1.2). One of the characteristic patterns investigated in variationist work – vowel differences between speech communities – is largely absent from work on Australian languages. This may reflect a lack of attention or, more likely, a difference between the kinds of sound changes documented for Australian languages and those that appear to be prevalent elsewhere in the world (see §7.2).

1.5 Overview of Australian Languages

The following section provides an overview of the morphology, phonology, and word structure of Australian languages; in this we follow a model established by Dixon (1972). Readers familiar with these languages can safely pass over this section. For readers unfamiliar with Australian languages, this section will provide a handy summary of their main features and help to provide context for the chapters that follow; in particular, it describes many of the morphological structures of words that are not otherwise discussed in the book, but that are essential for understanding some of the phonological patterns examined throughout. It also provides a convenient reference point for later sections, which expand on each of these aspects of language. Some areas of grammar that are not extensively discussed elsewhere in the volume receive fuller treatment here.

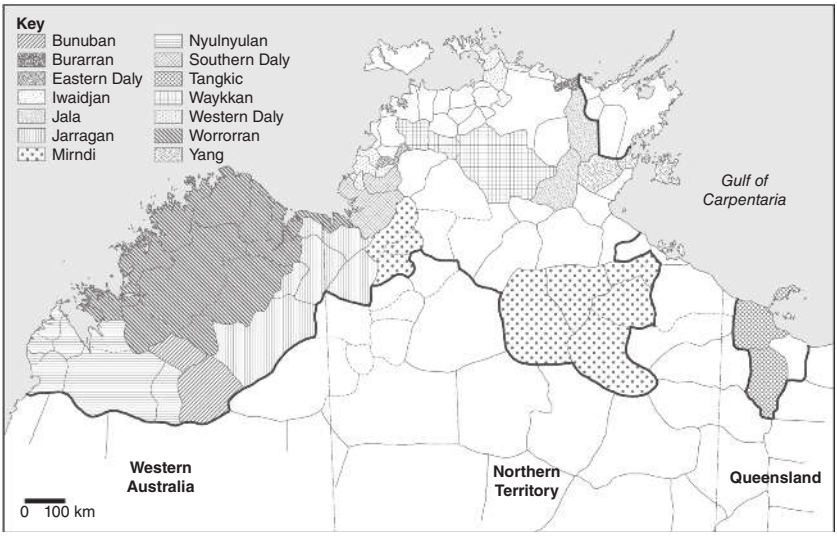
The consensus within Australianist circles is that all Australian languages are ultimately genetically related, with the possible exception of Tasmanian languages (Bowern 2012b; Crowley & Dixon 1981; Dixon 1980 p. 228) and the definite exception of the Eastern Torres Strait language Meriam (which is Papuan; see Evans 2005 p. 255).[2] Similarly, most Australianists are satisfied that there is a large and reasonably uniform family, the Pama–Nyungan family, which occupies 90 per cent of the continent (see Map 1.1). The languages that occupy the remaining 10 per cent of the continent are grouped together under the label 'Non-Pama–Nyungan'.

The Non-Pama–Nyungan languages do not constitute a family, but are rather a disparate collection of various families and isolates (Evans 2003b, Harvey & Mailhammer 2024; see Maps 1.2 and 1.3), hence are in fact literally non-Pama–Nyungan. But the term has been conventionalised to a group of Australian languages and so we refer to this group as 'Non-Pama–Nyungan'.

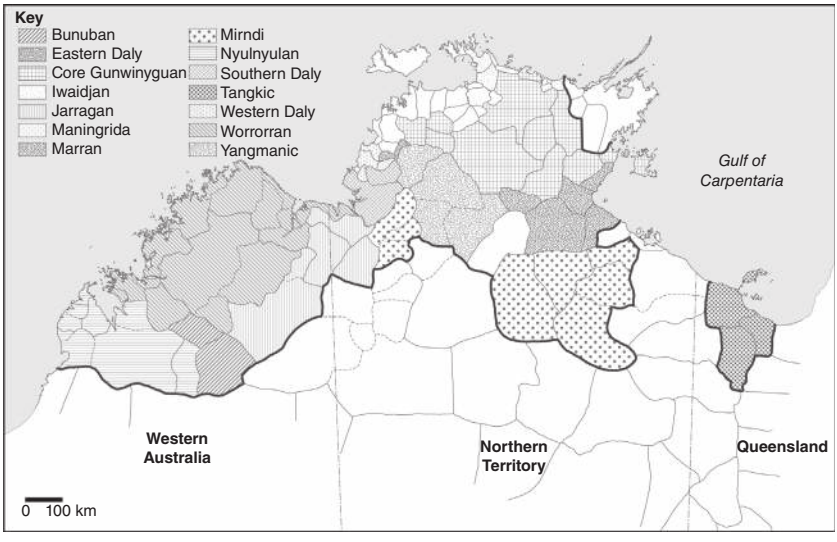
There are 'disconnected' languages from both groups, however: the Yolngu group of languages in north-east Arnhem Land is an enclave of Pama–Nyungan languages surrounded by Non-Pama–Nyungan ones. Similarly, the Tangkic group – which includes Lardil (made famous by Ken Hale), Kayardild, and Yukulta or Ganggalida in the southern Gulf of Carpentaria, around Bentinck Island and the adjacent coast – is a Non-Pama–Nyungan

[2] But Bowern (2023 p. lxviii) sounds a cautious note about whether the case for all Australian languages being genetically related has been proven.

Map 1.1 Pama–Nyungan vs Non-Pama–Nyungan.



Map 1.2 Established Non-Pama–Nyungan families and isolates.



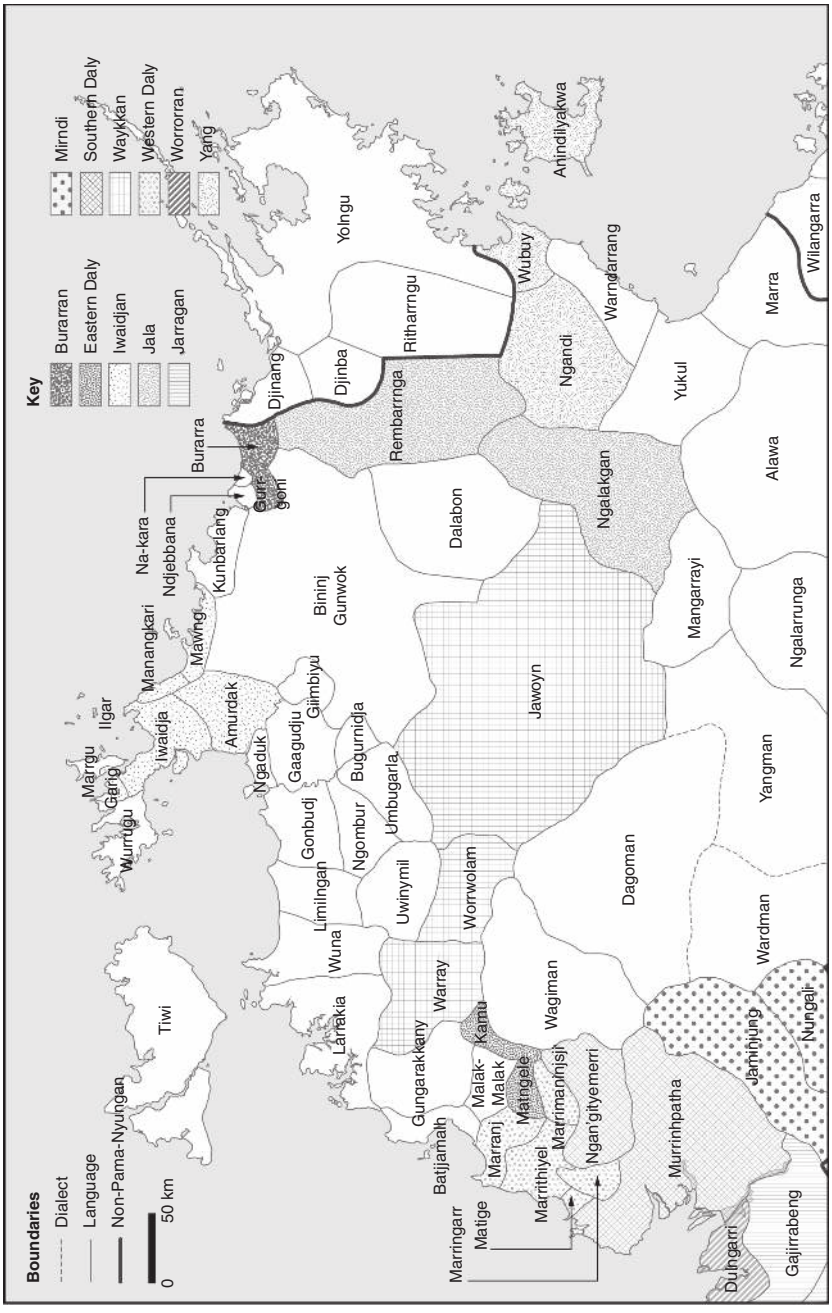
Map 1.3 Proposed Non-Pama-Nyungan families.

group (Blake 1988) surrounded by Pama–Nyungan languages and, like them, is exclusively suffixing. There are proposals that the Pama–Nyungan language family and all or most of the Non-Pama–Nyungan languages are daughters of a higher-level Proto-Australian (PA) family (Evans & Jones 1997; Harvey & Mailhammer 2017; Heath 1978a; O'Grady *et al.* 1966).[3]

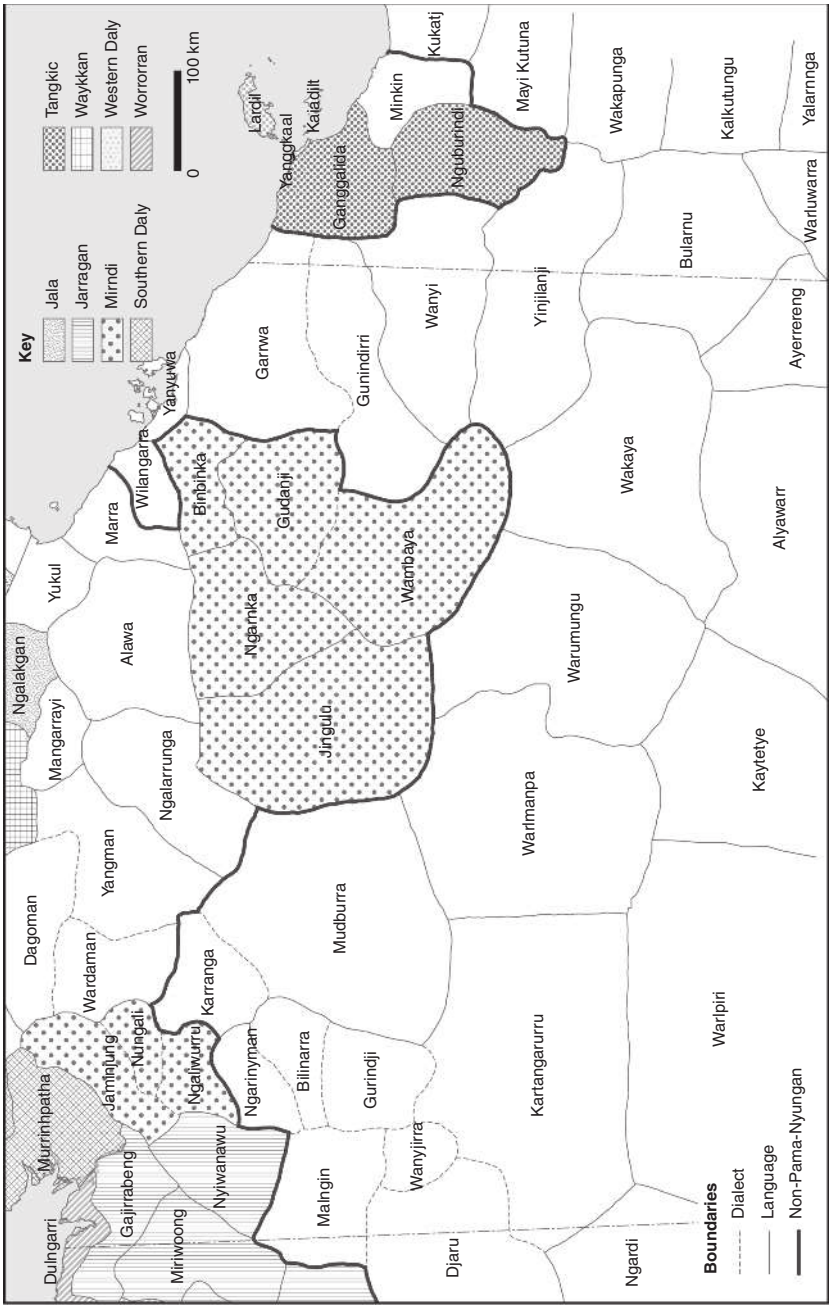
The genetic distinction between Pama–Nyungan and Non-Pama–Nyungan almost exactly matches a significant typological distinction. Nearly all Pama–Nyungan languages are exclusively or overwhelmingly suffixing, while Non-Pama–Nyungan languages have both suffixes and (typically) prefixes, as well as word structures and phonotactics that, quite often, are more complex than in Pama–Nyungan languages (Maps 1.4–1.6).[4]

[3] Since this manuscript was submitted for review, Harvey and Mailhammer (2024), which is an evaluation of the Proto-Australian (PA) hypothesis, has appeared. Harvey and Mailhammer conclude that the PA hypothesis is supported and that all the languages of Australia are related, apart from the languages of Tasmania and the Tiwi Islands. (They do not discuss the languages of Torres Strait.)

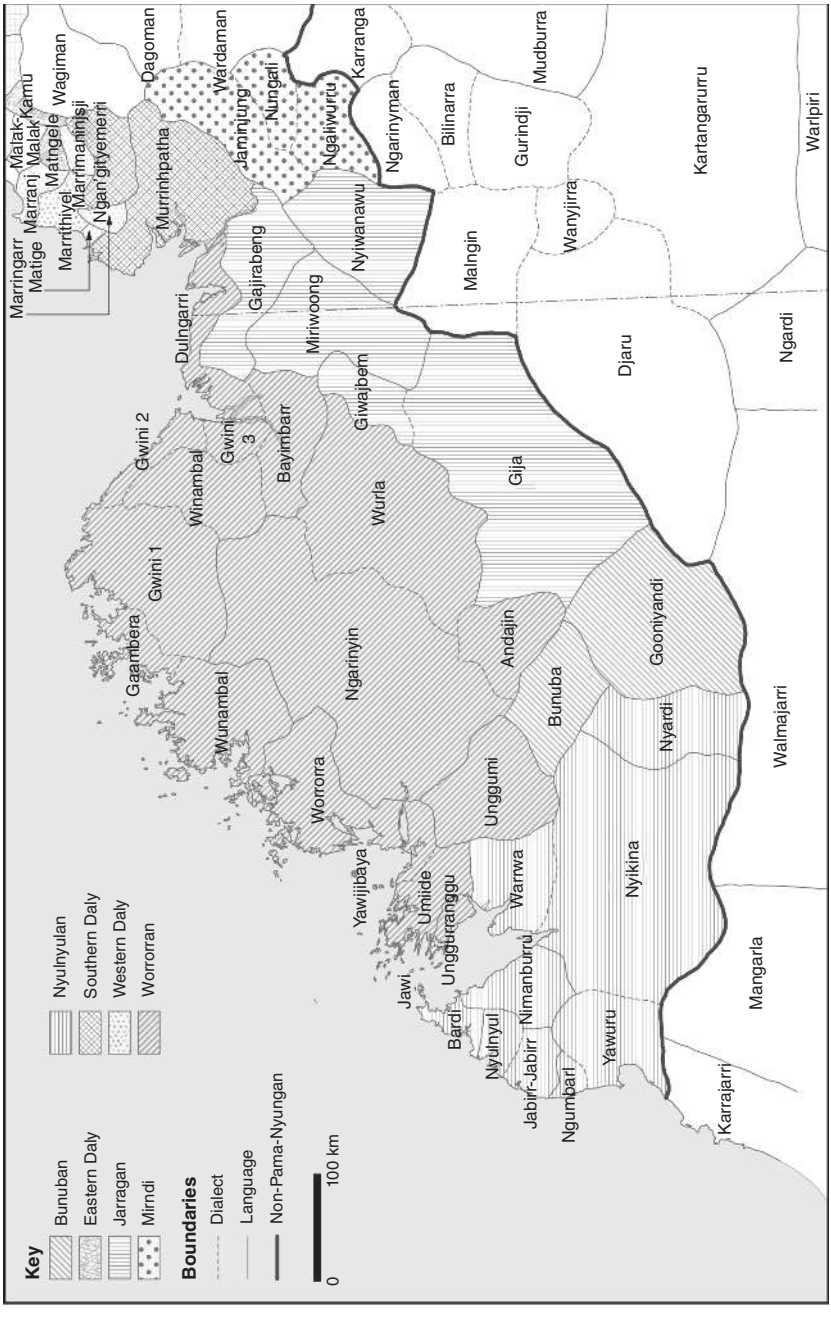
[4] The fact that all Australian languages appear to be related and the apparent genetic distinction between Pama–Nyungan languages and Non-Pama–Nyungan languages have engendered a great deal of speculation and debate, to which we can't hope to do justice here (see e.g. Alpher 2004; Bowern 2023; Bowern & Koch 2004; Dixon 1980, 2002; Evans 2005; Fitzgerald 1997; Miceli 2019; O'Grady *et al.* 1966; for some of the issues, and for the archaeological background and its potential relevance for language distribution, see Evans & Jones 1997; Hiscock 2008; McConvell 1990, 1996, 2001).



Map 1.4 Languages of the Top End.



Map 1.5 Languages of the Middle Northern Territory.



Map 1.6 Languages of the Kimberleys.