

Trends in Linguistics
Studies and Monographs 181

Editors

Walter Bisang
Hans Henrich Hock
Werner Winter

Mouton de Gruyter
Berlin · New York

Language Diversity
Endangered

edited by

Matthias Brenzinger

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

⊗ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Language diversity endangered / edited by Matthias Brenzinger.
p. cm. — (Trends in linguistics. Studies and monographs ; 181)
Includes bibliographical references and indexes.
ISBN 978-3-11-017049-8 (hc. : alk. paper)
ISBN 978-3-11-017050-4 (pbk. : alk. paper)
1. Language attrition. I. Brenzinger, Matthias, 1957—
P40.5.L28L36 2007
306.44—dc22

2006035522

ISBN 978-3-11-017049-8
ISSN 1861-4302

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© Copyright 2007 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin
All rights reserved, including those of translation into foreign languages. No part of this
book may be reproduced or transmitted in any form or by any means, electronic or mechan-
ical, including photocopy, recording or any information storage and retrieval system, with-
out permission in writing from the publisher.
Cover design: Christopher Schneider, Berlin.
Printed in Germany.

Preface

The present volume aims to familiarize interested readers with the extent and variation of the accelerating phenomena of language endangerment. They will find global overviews on endangered languages in chapters dealing with all major geographic regions of the world. These contributions provide insights into the specific areal dynamics of language endangerment, past and present. In addition, the authors discuss numerous key issues concerning the documentation of endangered languages. This book is aimed not only at scholars and students from the various sub-disciplines of linguistics, but also addresses issues that are relevant to educators, language planners, policy makers, language activists, historians and other researchers in human science.

The volume comprises updated versions of presentations from the Colloquium *Language Endangerment, Research and Documentation – Setting Priorities for the 21st Century* held in Bad Godesberg from February 12th–17th, 2000 and sponsored by the *Volkswagen Foundation*. Besides the present publication, the colloquium had a substantial impact on the genesis of the UNESCO report *Language Vitality and Endangerment*, as well as the *Recommendations for Action Plans*. Between 2001 and 2003, a UNESCO ad-hoc expert group on endangered languages (co-chaired by Akira Yamamoto and Matthias Brenzinger) collaborated with the UNESCO Intangible Cultural Heritage Section in Paris to draft a preliminary version.

I wish to thank the *Volkswagen Foundation* for their financial support in organizing the colloquium. Thanks also to Monika Feinen, cartographer at the Institut für Afrikanistik, University of Cologne, for her professional contribution. I would like to thank Anke Beck (Mouton de Gruyter) for her sustained support in this enduring publishing project. Thanks are also due, of course, to the authors of the volume and to all colleagues who contributed by sharing their experience in the study of endangered languages. I am particularly grateful to the *Research Institute for Languages and Cultures of Asia and Africa* at the *Tokyo University of Foreign Studies, Tokyo* for enabling me to finish the book manuscript during my stay there as a visiting professor in 2005–06.

Tokyo, April 2006

Matthias Brenzinger

Table of Contents

Preface	v
<i>Matthias Brenzinger</i>	
Language Endangerment Throughout the World	ix
<i>Matthias Brenzinger</i>	
Chapter 1	1
Classification and Terminology for Degrees of Language Endangerment	
<i>Michael Krauss</i>	
Chapter 2	9
Threatened Languages in Hispanic South America	
<i>Willem F. H. Adelaar</i>	
Chapter 3	29
Endangered Languages of Lowland Tropical South America	
<i>Denny Moore</i>	
Chapter 4	59
Endangered Languages of Mexico and Central America	
<i>Colette Grinevald</i>	
Chapter 5	87
Endangered Languages in USA and Canada	
<i>Akira Yamamoto</i>	
Chapter 6	123
Language Endangerment in Northern Africa	
<i>Matthias Brenzinger</i>	
Chapter 7	140
Endangered Languages in West Africa	
<i>Roger Blench</i>	
Chapter 8	163
Endangered Languages in Central Africa	
<i>Bruce Connell</i>	

Chapter 9	179
Language Endangerment in Southern and Eastern Africa	
<i>Matthias Brenzinger</i>	
Chapter 10	205
Endangered Languages in Europe	
<i>Tapani Salminen</i>	
Chapter 11	233
Language Endangerment in the CIS	
<i>Olga Kazakevich and Aleksandr Kibrik</i>	
Chapter 12	263
Endangered Languages of the Middle East	
<i>Jonathan Owens</i>	
Chapter 13	278
Language Endangerment in China and Mainland Southeast Asia	
<i>David Bradley</i>	
Chapter 14	303
Endangered Languages of South Asia	
<i>George van Driem</i>	
Chapter 15	342
Warramurrungunji Undone: Australian Languages in the 51st Millennium	
<i>Nicholas Evans</i>	
Chapter 16	374
Threatened Languages in the Western Pacific Area from Taiwan to, and Including, Papua New Guinea	
<i>Stephen A. Wurm</i>	
Chapter 17	391
The Languages of the Pacific Region: the Austronesian Languages of Oceania	
<i>Darrell Tryon</i>	
Indexes	411

Language Endangerment Throughout the World

Matthias Brenzinger

1. Introduction

Questions concerning the origin of human language have recently regained significant scholarly attention and it is expected that ongoing studies may produce important new insights into basic issues of language evolution. One of the most fundamental questions in this context is: Was there ever a *Homo sapiens* proto-language which existed some 100,000 or 200,000 years ago and then gave rise to large number of very distinct languages about 12,000 years ago, which for their part have been reduced to the approximately 6,000 of today? No matter what possible scenario for language genesis may be conjectured, it seems most likely that the large number of languages spoken on earth in some distant past dramatically dropped when hunter-gatherers changed to a pastoral lifestyle and even more so, when humans become sedentary farmers. The few thousand languages currently spoken are remaining relics of a once much richer pool of languages, and the shrinking of language diversity has accelerated during the last few thousand years.

The currently disappearing and endangered languages of the world, featured in the present volume, are essential sources for studying not only diachronic and synchronic aspects of human language. They are of eminent importance in attaining knowledge on human prehistory in general. Languages are formed by and reflect the most basic human experiences. Without proper scientific documentation, the decline of these languages will result in the irrecoverable loss of unique knowledge that is based on specific cultural and historical experience. Furthermore, the speech communities themselves will often suffer from the loss of their heritage language as a crucial setback of ethnic and cultural identity.

2. Indicators for assessing language vitality

The evaluation of the state of vitality of any language is a challenging task, as one has to consider different, intertwining factors. Speech communities are complex and patterns of language use within these commu-

Chapter 15

Warramurrungunji Undone: Australian Languages in the 51st Millennium

Nicholas Evans

1. Initial orientation to Australian languages

Australia has been occupied by humans for around ^{50,}5,000 years, and its initial colonization represents the first known major ocean crossing. Thereafter, though there were minor contacts with Indonesia, there is no clear archaeological evidence of subsequent major waves of immigrants until the British established the first European colony in Sydney in 1788. While agriculture and then state-based societies developed on all other continents the original inhabitants of Australia were able to continue living as hunters, fishers and gatherers throughout 50 millennia.

Virtually all scholars agree that *Australian* languages are all related at a deep level. This makes Australia unique in being the only continent entirely occupied by related languages, though the degree of time depth is such that the term 'phylic family' is more appropriate than 'family' to describe the relationship.¹ It also makes the *Australian* language family specially significant because of its likely great time depth. This phylic family comprises around 250 languages, or about 700 distinct linguistic varieties if one follows Aboriginal practice in how to measure language diversity; they are relatively uniform phonologically but exhibit enormous variation in grammatical organization.

Because Australia was the only continent entirely occupied by hunter-gatherers in historical times, its sociolinguistic situation holds great significance for our understanding of the processes of language change in a social universe entirely made up of small groups. Nettle (1999: 93) argues that 'there would have been more language diversity in the world, certainly in relative terms and perhaps absolutely, before the origin of agriculture', so *Australian* languages are vital to understanding the forces that have shaped linguistic diversification for the 99% of our evolutionary history during which all humans were hunter-gatherers (Livi-Bacci 1992). Certainly the pattern described by Nettle, according to which languages spoken by hunter-gatherers rarely have over 5,000 speakers, is played out in Australia where the average number of

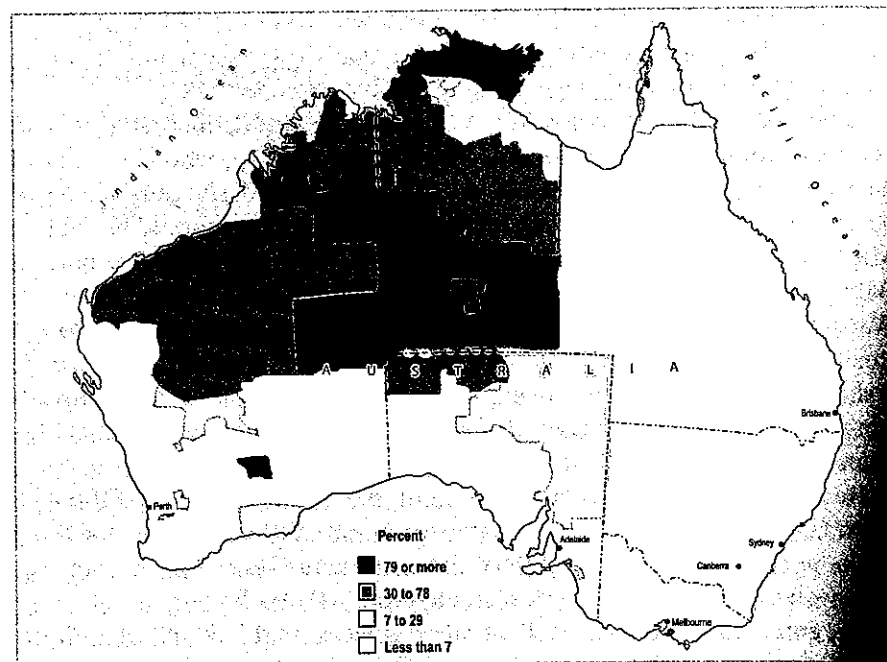
speakers per language would have been somewhere between 1,000 and 3,000, with some languages counting hardly more than 50 speakers and the maximum size for any one language unlikely to have exceeded 5,000.

In the oral traditions of north-western Arnhem Land, the first human to enter the continent was a woman, Warramurrungundji, who came out of the sea on the Cobourg Peninsula. As she heads inland, she puts groups of people into particular areas and decrees which languages should be spoken where. Unlike the Judaeo-Christian tradition which sees the aftermath of Babel as a negative outcome punishing humans for their presumption, the Warramurrungunji myth reflects a point of view much more common in small speech communities: that language diversity is a good thing because it shows where every² comes from and belongs.² Widespread multilingualism meant that even languages differing radically in their grammatical organization posed little obstacle to inter-group communication.

Genetic diversity is not evenly distributed over the continent. Of the 25 or so language families found on the continent, all but one are confined to the north-western one-eighth of the continent (*Map 23*). The remaining seven-eighths is covered by a single family, Pama-Nyungan.

The extraordinarily rich intellectual production of this 50,000 year flourishing, some of which I will touch on briefly in this chapter, is now collapsing over most of the continent, under the impact of the colonial language *English* and its associated culture of monolingualism. Indeed Australia is probably the continent currently experiencing the most rapid and drastic effects of language loss. It seems likely that by 2088 (three hundred years after the first European colonization), at most a dozen indigenous languages of the continent will have viable speech communities – i. e. a 95% extinction rate, over the three centuries of European rule that will then comprise one two-hundredth of the continent's history of occupation, of the accumulated linguistic legacy of the preceding fifty millennia.

Map 22 gives a preliminary impression of the fragility of the current situation. Of the 605 language varieties mapped (territories based on estimates of traditional locations), only 25 have over 1,000 speakers, and only 36 have over 500 speakers. The vast majority (526/605) have fewer than 100 speakers. The distribution of 'healthy' languages exhibits a clear geographical pattern – they are confined to those parts of the continent furthest from urban settlements, such as the Western Desert, Arnhem Land and the Torres Strait. Even having a 'remote' location, however, is no guarantee that a language is healthy – examples of extinct languages from quite remote locations are *Wurrugu* from the Cobourg Peninsula and *Linngithigh* from Cape York.



Map 22. Indigenous language speakers as proportion of indigenous population (map S1.5 from the ABS website; reproduced from McConvell and Thieberger 2000)

2. Special characteristics of Australian languages

In this section I give a taste of some of the ways *Australian* languages are important for our general understanding of linguistics, exhibiting characteristics that are totally or largely absent elsewhere in the world.

As an undergraduate student of psychology at the Australian National University in the 1970s, excited by reading Whorf's *Language, Thought and Reality* for the first time, I asked my lecturer why all the 'exotic' language examples were *Amerindian* and none *Australian*. He replied that 'we don't know much about *Aboriginal* languages, but they don't seem to be as different or interesting as *American Indian* ones'. Perhaps I wouldn't have said this if he had been more aware of the research going on at the time in the Linguistics department a few hundred metres away, but it was a characteristic response to the somnolent history of indifference to *Australian* languages that had characterised most of the centuries since white occupation of the continent.

Although there were some honourable exceptions, the overwhelming impression until the 1960s is one of neglectful, amateurish and uninspired efforts that rarely succeed in documenting the languages in a way remotely comparable to the work being carried out by Boas, Sapir and their students in North America, or by Dutch scholars in Indonesia. It was not until the 1970s – 200 years after the British claimed New South Wales, and at a time when a huge number of languages had already disappeared completely – that the first reasonable grammars and dictionaries of *Australian* languages began to appear (see Austin 1991 for a good survey).

We now have around thirty high-quality grammars and workmanlike grammars of as many again (see below for the situation with dictionaries and text collections). However, the concentration of typological and genetic diversity in the north-west (and the fact that most languages of the south-east became dead or moribund before modern research could be undertaken) means that our coverage is far from even.

In many cases, earlier descriptions simply failed to notice central and interesting facts about the language – if we had ticked these languages off as 'already described' on the basis of the short descriptions available, we would have missed the really crucial lessons they have to offer the science of linguistics. I will describe three examples of this in §2.1., concentrating on grammar and phonology, the heartland of the last three decades of research. In §2.2.–2.4. I look at some broader problems where our understanding is much further behind: lexicon and texts (§2.2.), special registers (§2.3.), and the area of multilingualism and the use of linguistic variation more generally to index differences in social group membership (§2.4.).

2.1. Grammar, phonology, and the unimaginable human language

One of the most exciting things about fieldwork is its potential to continuously broaden our preconceptions of what is possible for the human mind, by stumbling upon 'unimaginable' languages – those that would never have been thought possible.³

Work on *Australian* languages has had its fair share of such revelations since the 1960s, whether it be *Dyirbal* [1], the first language in which ergative syntax was comprehensively described (Dixon 1972), *Warlpiri* [2], in which the work of Ken Hale and others has uncovered the workings of non-configurational syntax in greater detail than any comparable language (e.g. Hale 1983) or, more recently, the way in which many north-western languages break their representation of event structure into two

parts – a small set of generic verbs that carves up the semantic space of event representations into one or two dozen main types, and a larger set of coverbs that combines with these to portray the event in detail (see e.g. Schultze-Berndt 2000, McGregor 2002).

In most cases it was far from obvious, at the stage of initial survey work that led to descriptions at the level of a short grammar, where the most interesting issues would arise. In each of above cases the interest of the language only became evident when 4 conditions came together – the opportunity for detailed and sustained research, an open-minded linguist with a flair for noticing new issues, outstanding native speakers with an interest in the work, and a favourable intellectual climate.

It is certainly not the case that either the prior existence of a grammatical sketch or of detailed knowledge of a closely related language would lead a researcher to suspect what treasures lay beneath the surface. I now consider two examples of how easy it is to overlook these even when relatively good survey-level documentation is available.

(a) *Kayardild* [3] is spoken in the South Wellesley Islands in north-west Queensland. It is closely related to *Yukulta* [4], with which it shares around 80% of its vocabulary, and a little less closely related to *Lardil* [5] on Mornington Island. Sandra Keen, who undertook research on *Yukulta* in 1969-70, described the situation as follows: 'the small amount of comparative work done indicates that *Yukulta*, *Gayardilt* [= *Kayardild*] and *Yanggal* [6] [= *Yangkaal*] belong as dialects of one language while *Lardil* is a different but closely related language. *Yukulta*, *Gayardilt* and *Yanggal* have considerable grammatical similarity as well as a high percentage of shared vocabulary and speakers of these dialects can hold quite detailed conversations with each other' (Keen 1983). Hale (1983), on the basis of fieldwork on Mornington Island that concentrated on *Lardil* but included brief recordings of *Yangkaal*, made similar observations. Any setting of research priorities that de-prioritized work if good descriptions of closely-related languages already existed would certainly not have identified *Kayardild* as needing special attention in the 1980s.

Luckily other developments led, quite fortuitously, to more detailed research. In 1981 Ken Hale visited Mornington Island to present the *Lardil-English* dictionary. The *Kayardild* community then concerned about the fact their language was not being passed on, asked Hale whether a similar project could be undertaken for their language and a chain of events led to the community approaching Bob Dixon, who, to my great good fortune, asked me if I would be interested in doing my Ph.D. on the language.

As it turns out, *Kayardild* is a superb example of an 'unimaginable language', and although it shares most of its vocabulary with *Yukulta* and the phonological differences are only slight, its grammar is organized in a quite different way. Perhaps its most outstanding feature is the way it keeps passing down all information from higher syntactic levels, so that each word projects up a host of information about all structural levels above it, rather than just its local environment. This can be illustrated by the phenomenon of multiple case marking. To say 'with a/the boomerang', the word for 'boomerang' is inflected with the instrumental case (1) – nothing unusual about this, from the perspective of *Russian* or *Latin*. If you go on to say 'with brother's boomerang', as in (2), you notice that the instrumental spreads over the whole phrase, being placed outside the genitive case on 'brother' as well as on 'boomerang'.

1. *wangal-nguni*
boomerang-INSTR
'with the boomerang'
2. *thabuju-karra-nguni* *wangal-nguni*
brother-GEN-INSTR boomerang-INSTR
'with brother's boomerang'

Although somewhat unusual typologically, double case-marking is far from unique – as well as being found in many other *Australian* languages, it is a trait found in *Central Cushitic* in Ethiopia, the ancient near East, and the Caucasus – see Plank (1995) for a collection of descriptions and a historical survey. But in *Kayardild* this is only the beginning. If you then embed (2) in a clause in the past tense, each noun phrase other than the subject takes an additional ablative suffix, as illustrated in (3).

- | | | |
|--|------------------------|--------------------|
| <i>dangka-a</i> | <i>burldi-jarra</i> | <i>yarbuth-ina</i> |
| man-NOM | hit-PAST | bird-ABL |
| <i>thabuju-karra-nguni-na</i> | <i>wangal-nguni-na</i> | |
| brother-GEN-INSTR-ABL | boomerang-INSTR-ABL | |
| the man hit the bird with brother's boomerang' | | |

an example of the 'modal case' system, which gives *Kayardild* a secondary way of giving information about tense/mood, linked with but partly independent of the main case system.

dependent of its verb inflections; it originated through the use of case to mark inter-clausal relations ('from/after spearing the fish, the man cooked it') and then became an independent marker of tense/mood when the main clause got ellipsed. In fact, it is precisely because *Kayardild* is so close to *Yukulta* (which has the interclausal use) that we have been able to figure out the historical origins of modal case. Among all the world's languages, only *Kayardild* and *Lardil* have modal case systems, and various changes that have taken place in *Lardil* (such as the specialization of one of the modal cases to the point where its basic use has been lost) makes it much harder to work out what is going in *Lardil* than in *Kayardild*.

To top things off, there is a fourth layer of case marking in *Kayardild*. When one clause is embedded in another and certain other conditions are met (e.g. the two clauses not having the same subject), every word in the subordinate clause takes a further case suffix (most commonly the 'oblique', historically a dative). An example is (4); note that while the nominative on 'man' in (3) is simply replaced (basically because the nominative only goes on words that have missed out on getting in any other case), all other words (including the past-inflected verb 'hit') just tack on the 'complementizing' oblique suffix after all other inflections; in the case of *thabuju* 'brother' this means it ends up with 4 case suffixes. No other language in the world takes case stacking to this extreme – comparative morphological evidence shows that *Lardil* once had it, but various radical phonological changes that have chopped off the ends of *Lardil* words have produced a somewhat different system in the modern language.

4. *ngada* *kurri-jarr*, *dangka-ntha* *burldi-jarra-ntha*
1sgNOM see-PAST man-OBL hit-PAST-OBL

yarbuth-inaa-ntha *thabuju-karra-nguni-naa-ntha*
bird-ABL-OBL brother-GEN-INSTR-ABL-OBL

wangal-nguni-naa-nth
boomerang-INSTR-ABL-OBL

'I saw that the man had hit the bird with brother's boomerang'

This does not exhaust the unique features of *Kayardild* grammar (see Evans (1995) for a detailed description). For example, in addition to 'modal' case inflections like those just given, which leave nouns and verbs unchanged, there is another whole set that turns nouns into morphologically

pable of taking the full range of verbal inflections (but which still function as nouns syntactically) – a counter-example to the often-made claim that inflections do not change word-class. The point is that here is a prime 'unimaginable' language whose importance was still masked when we had quite full information on two closely related languages – and also one where it would have been very difficult to work out the system had we not already had outstanding descriptions of *Yukulta* and *Lardil*. It would also have been impossible to figure out its structure if research had begun twenty years later – the number of fluent speakers has dwindled from around 40 in 1982 to under 10 in 2000. Most crucially, the fact that the generation who would now be in their 60s has been disproportionately affected by premature deaths mean that there is effectively only one bilingual speaker left, yet it was the bilingual generation (all born between 1935 and 1945) who were the key to understanding how the language functioned. The premature loss of this generation would certainly not have been predicted, though retrospectively it reflects where the major health stresses of the transition from traditional to mission to 'modern' living conditions have fallen most heavily.

(b) *Dalabon* [7] is spoken in central Arnhem Land. The pioneering Australianist Capell published a grammar of it in 1962 as one of 4 exemplified in his book *Some Linguistic Types in Australia*. Capell's grammar typifies the way a sketch grammar can overlook centrally interesting and unusual facts about a language – in this case, I believe, because most of the data was based on a questionnaire which only allowed the investigator to find what had already been foreseen as a possibility. Consider the following pair of sentences (adapted from Alpher 1982):

darruh-ko *barra-h-bo-n*
brother-DYAD 3dual(harmonic)-REALIS-go-PRES
'They two, brothers, go.'

be-ko *ke-h-bo-n*
man's son-DYAD 3dual(disharmonic)-REALIS-go-PRES
'They two, father and son, go.'

This pair illustrates, the pronominal prefixes on the verb, in addition to number, are sensitive to a social factor usually called 'harmonic'. The harmonic forms are used if members of the referent set are in the same generation (or even-numbered generations, such as grandpar-

ents and grandchildren), while the disharmonic forms are used if they are in odd-numbered generations (e. g. parent and child).

The existence of kinship-sensitive pronominal prefixes was completely overlooked by Capell's grammar. The most likely explanation for this is that it was largely based on a survey-type questionnaire, used widely by himself and others from the 1940s to the 1960s. Although this included a thorough examination of paradigmatic combinations like 'we two saw the dog', 'we three saw the dog', etc., it did not ever try to vary the kin relationship between the members of a set like 'we two' – presumably because no-one had reported this as being relevant at the time, the first published discovery of the phenomenon in any language being Hale (1966) for *Lardil*. It was only when Barry Alpher began work on *Dalabon* in 1974 that this material came to light, partly because of the involvement in this research of a native *Dalabon* speaker, David Jentian Nangan-golod, who was training in linguistics at the then School of Australian Linguistics.

Now it is perfectly understandable that Capell, as the first investigator to report on the language, should overlook many significant features, and his sketches have been put to good use by subsequent investigators (though we still lack a comprehensive description of *Dalabon*). My point is that survey work is inherently prone to overlooking the most interesting features of languages, and hence provides a poor guide to locating unusual grammatical and phonological characteristics.

2.2. Lexicography and texts

The trajectory of scientific interest since the 1960s has favoured work on grammars at the expense of dictionaries, which have remained 'the poor cousin of grammatical description' (Austin 1991). A recent survey of Australian lexicography by Goddard and Thieberger (1997) listed only 17 'reasonably large' dictionaries of *Australian* languages, even though their definition of 'reasonably large' is an unambitious '2,000+ entries'. Only 5 out of the *non-Pama-Nyungan* families are represented in this list. Yet for most linguists dictionaries are the most important type of documentation – certainly the speech communities themselves are more interested in dictionaries than in grammars; encyclopaedic dictionaries are an excellent source of cultural documentation relevant to community-based enterprises such as eco-tourism, gathering of wild plant seeds for native nurseries, and documentation of bark paintings. For historical linguistics the lack of good documentation remains the single greatest handicap to detailed work.

A successful initiative in the early 1990s, resulting from a one-off grant by the Federal Government to the Australian Institute of Aboriginal and Torres Strait Islander Studies, encouraged linguists to bring out provisional versions of dictionaries based on materials they had already gathered, leading to the publication of dozens of dictionaries and word-lists. However, the interaction of the much slower gestation time of dictionaries than grammars with the short funding cycle of most supporting institutions has made it hard to establish the sort of twenty-year project needed to produce proper dictionaries. Although there are isolated cases of comprehensive dictionaries produced by individual researchers (e. g. Alpher 1991), it has more often been the case that good dictionaries reflect long-term projects, based in the community or a local language centre, that have gathered lexical material and written definitions as part of a more wide-ranging program of language work. In some cases this has involved missionary organizations who over the years have built up a team comprising native speakers engaged in Bible translation (e. g. the *Wik-Mungkan* [8] dictionary by Kilham et al 1986).

The most successful dictionary-making program has been based in Alice Springs at the Institute for Aboriginal Development, which so far has produced three excellent dictionaries (Goddard's 1992 *Pitjantjatjara/Yankunytjatjara* [9] to *English* dictionary, Henderson and Dobson's 1994 *Eastern and Central Arrernte* [10] to *English* dictionary, and Green's 1992 *Altyawarr* [11] to *English* dictionary). Crucial to the quality of their dictionary-making programs has been (a) cooperation of trained linguists with native speakers who combine practical training in lexicography with an ongoing range of language-related jobs such as interpreting, language teaching, cultural documentation, and producing vernacular literacy materials (b) parallel production of several dictionaries with sharing of ideas and techniques between them (c) success in bringing together sufficient funding, over a relatively long term, to build stable lexicography teams (d) the capability of publishing their own materials.

Text material has lagged even further behind grammars than dictionaries; the seriousness of this gap is becoming more apparent as linguistics makes more and more use of extensive corpora. Publishing of bilingual texts has neither been commercially feasible nor a high academic priority, and many text collections, submitted to publishers as bilingual, had to excise much of the vernacular text as the price of publishability (see Laughren 2000 for a fuller listing) include Strehlow's 'Songs of Central Australia', the collection of *Warlpiri* [2] 'dreams and histories' in Napaljarri and Cataldi (1994), of *Kunwinjku* [12]

texts in Nganjmira (1997), and, more academically oriented in terms of glossing and linguistic commentary, Heath (1981) on *Nunggubuyu* [13] and Schebeck (1974) on *Adnyamathanha* [14]. In recent years it has been local *Aboriginal* language centres that have led the way in the publication of bilingual text materials, which tend to be directly written in the vernacular rather than transcribed from speakers. As with dictionaries, text collections are valued by the speech communities themselves because of the high content of cultural information; they are also useful 'insurance' against an overly-narrow focus by linguists on the questions of the day.

2.3. Multilingualism in speech communities

There is growing interest in the question of how linguistic diversity is engendered and maintained (e.g. Nettle 1999), and in the social mechanisms that favour its development. Australia provides a clear example of how small, culturally similar groups that are in regular contact and whose members normally know a number of neighbouring languages, consistently develop and maintain language diversity. The reigning social model over much of Australia posits a direct relationship between land and language (e.g. Merlan 1981), as well as between language and particular social groupings, such as clans. Individuals then derive the right to be recognized as speakers of particular languages indirectly, through their membership in clan or other groups:

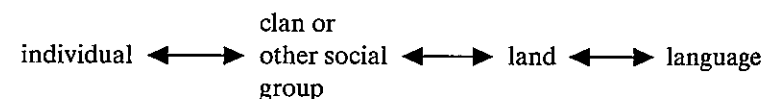


Figure 1. The indirect relationship between individuals and 'their' language.

This ideology leads to a widely-maintained distinction between speaking and owning a language: a statement that 'X speaks Y' may simply mean that 'X belongs to the group associated with the country where Y is spoken', and some other person who actually speaks Y fluently because of their life history, but does not belong to the requisite grouping, may not be volunteered as a speaker. I discuss elsewhere (Evans 2001) some of the difficulties this can create in discovering last speakers.

The mapping of language onto country creates an interesting set of sociolinguistic practices (see e.g. Brandl and Walsh 1982, Trigg 1991).

is a regular practice to change the language one is speaking when entering a new territory, to address particular locales (e.g. wells, or dangerous places) in the local language, for characters in myths to switch languages as they move about the country (sometimes this metonymic device may be the only indication that a character has moved from one place to another), and for large-scale song and myth cycles to incorporate a number of 'legs', each in a different language and typically told by different speakers in relay – as if the *Odyssey*, say, passed through half-a-dozen *Mediterranean* languages and story-tellers as the action shifted.

Typically there is no etiquette that all participants in a conversation should speak the same language, and it is quite normal to witness multilingual conversations in which each participant speaks their own language. This is one factor that may help speakers of critically endangered languages remain fluent the fact that their conversational partners don't speak their languages doesn't mean that they have to give it up. It is striking how often, compared to cases reported in some other parts of the world, the 'last speakers' of *Australian* languages have had an excellent command of their morphosyntactic intricacies (see e.g. Harvey's (1992) grammar of *Gaagudju* [15] and Dench's (1995) grammar of *Martuthunira* [16], both based on work with last speakers).

It is also a widespread practice for a number of languages in a region to be organized into an overarching system, such that speakers' metalinguistic awareness extends well beyond their 'own' speech community or territory. The clearest example is in North-Eastern Arnhem Land (Morphy 1997, Wilkinson 1991), where each language variety is associated with one of two patrimoieties; features distinguishing 'moiety lects' (e.g. dropping of final vowels) cross-cut features distinguishing dialect groupings set up on geographical grounds (e.g. loss of initial syllables of some pronouns). Social processes by which differentiation of new social groups (e.g. splits) lead to the formation of new 'patrilects' in very small-scale speech communities have been described for Western Cape York by Johnson and Johnson (1986).

So far little theorization in sociolinguistics has appreciated the importance of carrying out detailed work on super-speech-communities made up of interlocked and intermarrying small languages – and obvious this has to be achieved until basic grammatical and lexicographic work has been carried out (though see Sutton 1978 for an example of what is possible). Psycholinguistics, with its belief in a 'critical period' based on observations of language-learning in large, literate speech communities, may discover surprises if it were to tackle the problem of second- lan-

guage learning in speech communities where it is normal for adults to keep learning new languages throughout their lives. This is also relevant to the question of whether languages can still be passed on even when they are not learned as children. In short, full documentation of language use in such communities is one of the highest priorities if we are to understand the forces that have shaped linguistic diversity for the majority of our human past, in which speech communities throughout the world are likely to have been small.

2.4. Special language functions

A commonly used index of language endangerment is to look at the encroachment of a replacing language on the 'language functions' of the threatened language. Frequently cited functions include 'religious practice', 'education', 'inter-ethnic communication' and so forth. However, if we take seriously the emic premise of examining functions from within the perspective of each speech community, we need to allow for the possibility that there exist culture-specific functions that we would never have imagined before undertaking detailed field research. These are of particular scientific interest for the way they encode interpersonal relationships in language, and because in many cases they embody an indigenous tradition of analysis that gives great insight into semantic structure.⁴

2.4.1. Respect registers

Many *Australian* languages have special 'respect registers', sometimes known as 'brother-in-law languages' or 'mother-in-law languages', which are used as a polite way of talking between certain types of relative. They normally involve substitution of a proportion of the vocabulary, which may range from a few words to virtually every word in the language, by different lexical items, while leaving the grammar intact. For example in *Uw-Oykangand* [17], a language of Cape York, a man talking to a potential mother-in-law must use a special register known as *Olkel-Ilmbanhthi*; to speak this one leaves affixes and function words like 'I' intact but replaces the remaining vocabulary items (Alpher 1993: 98):

- 7a. *Alka-nhdh* *idu-rr* *ay*
 spear-INSTR spear-PAST I
 'I speared it with a spear.' (ordinary register)

- 7b. *Udnga-nhdh* *yanganyunyja-rr* *ay*
 spear-INSTR spear-PAST I
 'I speared it with a spear.' (respect register)

Because respect registers typically have fewer words than the everyday register, their meanings are correspondingly more general, and several linguists have used this as a way of investigating semantic structure by studying the many-to-one meaning relationships between everyday register words and their respect-register equivalents (Dixon 1971, Evans 1992).

There have been few comprehensive studies of respect registers, and most have been based on what a few old people remember rather than observations of actual use. This is partly because it takes time for fieldworkers to discover their existence, but also because respect registers are far more fragile than the everyday form of a language, not being mastered until well after adolescence. For example, at Oenpelli all children still learn *Kunwinjku* [12] as their mother tongue, but the youngest two generations are not mastering the respect register, known as *Kunbalak* or *Kunkurrng*, to the great concern of the speech community. This has prompted Andrew Manakgu, a *Kunwinjku* linguist working at the *Kunwinjku* Language Centre, to specially produce a book of stories with parallel *Kunwinjku* and *Kunbalak* versions as an attempt to restore knowledge of this register in the younger generations (Manakgu 1998).

2.4.2. Trirelational kin terms

Another way in which many *Australian* languages reflect a sensitivity to kinship relations between speaker and hearer is through systems of 'tri-relational' kin terms, which offer a number of ways of referring to kin based on simultaneously figuring out the relationship of the referent to speaker and hearer. Consider the following terms from the *Gun-dembui* or tri-relational register of *Gun-djeihmi* [18]; all are ways of referring to the mother of speaker and/or hearer in a range of circumstances:

8. *al-garrng* 'the one who is *your mother* and *my daughter*, given that I am your mother's mother'
 al-doingu 'the one who is *your daughter* and *my mother* (given that you are my mother's mother)'
 al-gakkak 'the one who is *your maternal grandmother* and *my mother*, given that I am your mother'

- arduk gakkak* 'the one who is my maternal grandmother and mother, given that I am your mother'
al-bolo 'the one who is mother of one of us and mother-in-law of the other, given that we are husband and wife'

These systems are important for cognitive psychology because of the way they require the speaker to take two perspectives at once – their own and that of their interlocutor. See Merlan (1989) and Laughren (1982) for discussions of two such systems.

Not surprisingly, given the fact that the ability to outgrow an egocentric perspective only comes with maturity, they are not mastered until mid-adulthood (though we have no proper developmental studies), and their correct use is regarded as polished and courteous. Like respect registers with which they may coexist in the same language,⁵ the fact that they are not mastered until adulthood renders them particularly fragile.

2.4.3. Initiation register

A further type of special register is that taught to ceremonial initiates in certain Australian communities as part of the process of formal religious education. I shall briefly discuss two of these here – *Jiliwirri* and *Damin* – to give some idea of how important they are as part of humankind's intellectual heritage; this point has been elaborated elsewhere by Hale (1999), who is responsible for most of the research on which this section is based.

Warlpiri [2] initiates learned, as part of the process of initiation, a special register known as *Jiliwirri* (Hale 1971), based on the principle of replacing all lexical items (though not grammatical affixes other than pronouns) with their opposites. As the following example shows, to convey the proposition 'I am sitting on the ground', one must use a *Jiliwirri* utterance which would translate literally into everyday *Warlpiri* as 'someone else is standing in the sky':

9. *ngaju ka-rna walya-ngka nyina-mi*
 I PRES-1sg ground-LOCATIVE sit-NONPAST
 'I am sitting on the ground.' (ordinary *Warlpiri*)
10. *kari ka-ø nguru-ngka karri-mi*
 other PRES-3sg sky-LOCATIVE stand-NONPAST
 'I am sitting on the ground.' (*Jiliwirri*)

Warlpiri has been used to investigate antonymy in *Warlpiri* lexical semantics, including such non-obvious issues as whether the perception verbs 'see', 'hear', etc. have antonyms, and on how one determines antonyms for natural species names like 'red kangaroo'. See Hale (1971) for full discussion.

Even more spectacular is a special initiation register known as *Damin*, which was taught to *Lardil* [5] men on Mornington Island as part of their initiation as *Warama* (second degree initiates) – see Hale and Nash (1997) and McKnight (1999). It is said to have been created by an ancestor known as *Kaltharr* (Yellow Trevally fish), and has a rich inventory of sounds, some echoing what 'fish talk' would sound like. In fact, its phoneme inventory is unique among the world's languages in employing 5 distinct airstream initiation types: pulmonic egressive (the normal type in all other Australian languages), pulmonic ingressive (the ingressive lateral fricative *l**, velaric ingressive (nasal clicks like *n!*), labiovelar lingual egressive (*p'*) and glottalic egressive (*k'*).⁶ Of these, the pulmonic ingressive and labiovelar lingual egressives are found nowhere else in the world's languages, at least as phonemic elements.

Because grammatical affixes are simply taken over from everyday *Lardil*, it is only the lexical roots that display these special phonemes, as can be illustrated by the following sentence equivalents from everyday *Lardil* (11a) and *Damin* (11b): *Damin* substitutes *n!aa* for *ngada*, *didi* for *ji-* and *l*i* for *yak-*, but leaves the grammatical suffixes intact (other than allomorphic differences between postconsonantal *-ur* and postvocalic *-ngkur*).⁷

- 11a. *Ngada ji-thur yak-ur*
 I eat-FUT fish-OBJ
 'I will eat fish.'

- 11b. *N!aa didi-thur l*i-ngkur*

Equally unique is the semantic structure of *Damin*, which maps the many thousand lexical items of everyday *Lardil* into around 200 words by a combination of highly abstract semantics, extended polysemic chains, paraphrase, and supplementation by hand signs. In the above example, *n!aa* does not simply correspond to *ngada* 'I' but can denote any group including ego, which in *Lardil* includes 9 possible pronouns, representing the three-dimension matrix formed by *ngada* 'I' plus $2 \times 2 \times 2$ combinations of the oppositions dual vs plural, inclusive vs exclusive, and harmonic vs disharmonic. Likewise *didi* does not simply correspond to *jitha* 'eat'

but also includes all actions of harmful affect, such as *barrki* 'chop', *betha* 'bite', *bunbe* 'shoot', and *kele* 'cut'. The net effect is to produce a totally indigenous analysis of the semantics of the entire vocabulary into a small number of elements, and Hale (1982: 32) justifiably refers to *Damin* as a 'monument to the human intellect'. Elsewhere he has drawn attention to the fact that its association with rituals outlawed by the missionaries in power on Mornington Island meant that its transmission was interrupted well before the transmission of everyday *Lardil*, as well as to the invisibility of this astounding achievement to the outside world:

The destruction of this intellectual treasure was carried out, for the most part, by people who were not aware of its existence, coming as they did from a culture in which wealth is physical and visible. *Damin* was not visible for them, and as far as they were concerned, the *Lardil* people had no wealth, apart from their land. (Hale 1998: 211–212)

Each of the special registers described in this section has been developed to fulfil particular 'language functions' that lie outside the usual typology of written vs. spoken, high vs. low that we are familiar with from the languages of literate societies. In addition to their sociolinguistic interest, all bring unusual and sometimes unparalleled semantic insights into the study of the 'everyday' forms of the corresponding languages. At the same time they are among the most fragile and easily-overlooked phenomena in need of documentation – in other words, they are normally much more sensitive to loss than the 'normal' language. Indeed, as we saw in the case of the *Kun-balak* and *Kun-debi* registers of *Kunwinjku*, it may happen that an apparently-healthy language is regarded by its speakers as under threat precisely because one or more of these special functions is no longer being carried out. It is essential that prioritizations of needs for language documentation not overlook these exceptionally interesting language varieties, whose existence often does not emerge until detailed work has been carried out on the everyday variety.

3. Assessing endangerment across the continent

I lack the space here to give full figures on the status of every language on the continent. Rather, I shall summarize the overall picture, give some examples illustrating some of the difficulties of assessment, and point the reader to more comprehensive sources.

Only since 1986 questions about language been asked in the Australian census, and only in the last census (1996) have indigenous languages been identified individually by name. The exact current form of the question – 'do you speak a language other than *English* at home?' – has various difficulties, such as disallowing the possibility that the individual speaks more than one language other than *English*, and allowing for self-reporting errors. The existence of multiple language names and spellings causes confusion in interpreting the results, although in some areas assistance given to the census-takers helped standardize this. Finally, confidentiality requirements mean that for languages with very low numbers of speakers the figures are not made available, obviously a grave shortcoming when dealing with endangered languages.

Before census figures began being compiled, a number of surveys had been published, beginning with Oates and Oates (1970), which attempted a complete listing of languages, giving estimates of speaker numbers, language viability (including the proportion of the population who still use the language), documentation of grammar/phonology, vocabulary, text/tape, and names of investigators involved with the language. At the stage when they compiled this the status of many varieties as distinct languages or dialects was still unclear (see *Notes to Figure 2* for two examples).

More recently, comprehensive 'handbooks' have appeared for several regions: Central Australia (Menning and Nash 1981), the Kimberleys (McGregor 1988), the Pilbara (Sharp and Thieberger 1992), and Western Australia south of the Kimberleys (Thieberger 1993); the latter two overlap in coverage. Each of these gives estimates of speaker numbers and language situation, detailed information on available sources, school programs and language learning materials, and basic word lists.

Significantly, the first three handbooks were commissioned by local Aboriginal organizations:⁸ an important development since the mid 1980s has been the setting up of Aboriginal-run regional 'Language Centres' acting as a local focus for language maintenance, documentation and archiving (Laughren 2000 has a full listing), and these have played an important role in involving community members in a wide range of language activities, as well as in matching researchers with speech communities and negotiating acceptable conditions of involvement by academic researchers.

A further set of figures, for the whole of Australia, was compiled by Schmidt (1990) as part of a nation-wide survey commissioned by the Australian Institute of Aboriginal and Torres Strait Island Studies (AIATSIS) in Canberra. Schmidt's figures on speaker numbers draw on a number of the above sources plus additional surveys in some areas. However,

they only extend to 90 languages: the 20 languages she identifies as 'healthy/strong', defined as 'one which is transmitted to children and actively spoken by all generations in a wide range of social contexts' (p. 2), and ranging from 200 to 3,000+ speakers, plus a further 70 'severely threatened' languages estimated to have at least 10 speakers (p. 5), and ranging upwards from 10 to 1,000 speakers. The author notes that this list is not exhaustive, and than languages with fewer than 10 speakers are not listed at all. Additional comprehensive figures are the constantly-updated source Ethnologue, and Wurm's 1999 listing of endangered languages, both compiled from a number of sources.

The difficulties in coming up with accurate figures are illustrated in *Table 1*, which compares sources for a number of languages that I have worked on sufficiently to come up with my own estimate, and to make a sober assessment of claimed knowledge. The discrepancies are representative of the mistakes that arise when figures are quoted without intensive investigation on the language having been undertaken, but also of the way in which speaker numbers can collapse very quickly in communities with life expectancies decades below those for non-indigenous Australians. For *Kayardild* [3], for example, Wurm's estimate of 50 speakers in 1981 would not have been far out; the decline to the current figure of 8 simply reflects the death of most *Kayardild* speakers in the last 20 years.

They also show how speakers of languages believed extinct can still come out of the woodwork, as happened with *Ilgar* [19]; another recent example is the Kimberley language *Andajin* [20] (Saunders 1999).⁹

Finally, it illustrates the fact that a small number of *Australian* languages are actually increasing in number, usually through a combination of demographic growth and language shift by speakers of neighbouring languages. *Kunwinjku* (*Bininj Gun-wok*) is an example of this (Evans 2003): the more than doubling since the 1980s estimates reflects both factors. Families at Korlobidahdah outstation in Central Arnhem Land, for example, have shifted in the last three generations from being bilingual in *Dalabon* and the *Kune* dialect of *Bininj Gun-wok*, to being monolingual *Kune* speakers (with some *English*); the youngest residents of this community who know *Dalabon* are in their late 40s. In this case, as in many other expanding indigenous languages, a major factor favouring *Bininj Gun-wok* growth is its role as a lingua franca in the staging of traditional religious ceremonies. In other cases one indigenous language has spread at the expense of others through its use as a lingua franca on missions – see Smith (1986) on the displacement of Kugu *Muminh* [24] by *Wik-Mungkan* at Aurukun.

Table 1. Examples of speaker estimates of selected languages from various sources

	Oates and Oates 1970	Ethnologue 1988	Black 1983 (NT only)	Schmidt 1990	Wurm 1999	Australian Census 1966	My assess- ment
<i>Kayardild</i> [3]	[no figures]	no mention	no mention	70	50 in 1981, lower today	no mention	6
<i>Lardil</i> [5]	'Entire popula- tion of Morn- ington Island'	50 possibly	no mention	50	possibly 50 in 1981	no mention	1
<i>Dalabon</i> [7] (<i>Ngalkbon</i>)	33*	100–200 (<i>< Black 1983</i>)	100–200	100–200	100–200 in 1983	no mention	20
<i>Kunwinjku</i> (<i>Bininj Gun- wok</i>) [12]	180	900 inc. L2 speakers	400–900	900	no mention	1403	1500–2000
<i>Iwaidja</i> [21]	120	180 (<i>< SIL</i>)	180	180	180 in 1983	no mention	c. 100
<i>Garig</i> [22]	**	no mention	no mention	no mention	no mention	no mention	0
<i>Ilgar</i> [19]	**	no mention	no mention	no mention	no mention	no mention	2 'hearers'
<i>Marrgu</i> [23]	12	4 possibly (<i>< Wurm and Hattori 1981</i>)	no mention	no mention	'recently extinct'	no mention	1 semi-fluent

* Oates and Oates (1970) discuss the language I am calling *Dalabon* here under three headings: *Gun-dangbon* (which they mistakenly list as a dialect of *Kunwinjku* ('*Gunwingu*')), *Dalabon* and *Boun*. The figure given here adds the figures for these three together.

** Oates and Oates do not mention *Ilgar*; they mistakenly subsume *Garig* as a variant of *Iwaidja*.

In the last two years an important initiative has led to a comprehensive compilation of data on *Australian* languages, the State of *Indigenous Languages* audit. This was commissioned by Environment Australia and treats indigenous languages as part of the nation's natural and cultural heritage. An important feature is that, since it has been introduced as part of a 5-yearly round of reporting, it will be possible to make time-based comparisons from subsequent surveys. The current report (McConvell and Thieberger 2000) includes a full data-base, remarks on factors affecting language maintenance, as well as recommendations for data-gathering methods that will produce a more accurate picture for future audits.

Although this section has focussed on the problems of accurate assessment, it is worthwhile concluding this with some overall approximate statistics revealing the extent of endangerment and extinction.

Of the 250 or so 'linguist's languages' (i. e. grouping together dialects as one language on structural grounds, in a way that speech community members may not acknowledge as valid) spoken in Australia prior to European colonization, fewer than 20 are being passed on to children today, and fewer than one hundred have more than ten fluent speakers. The number that are extinct is harder to quantify exactly, though Wurm 1999 lists 131 extinct languages.

Another way of taking stock is by language family. Of the 25 language families shown on Map 2, only 6 contain 'healthy' languages: *Pama-Nyungan* contains more than half the 'healthy' languages, including *Warlpiri* [2] (3,000+), *Arrernte* [10] (3,000+), *Western Desert* [25] (4,000+), *Yolngu-Matha* [26] (2,500+), *Kala Kawaw Ya* [27] (3,000+), *Wik Mungkan* [8] (1,000), *Nyangumarta* [28] (700–800), and *Thaayorre* [29] (500). Five other language families contain the remaining 'healthy' languages: *Tiwi* [30] (1,400) and *Anindilyakwa* [31] (1,000+) are family-level isolates, *Southern Daly* contains *Murrinhpatha* [32] (900+), *Gunwinyguan* contains *Bininj Gun-wok* [12, 18] (1,500), and the *Maningrida* family contains *Burarra* [67] (400–600) and *Ndjébbana* [33] (200–). Of the remaining 21 families, 6 or extinct or virtually so – these are *Gaagudju* [15] (family-level isolate), *Umpugulan*, *Geimbiyo*, *Larrakiyan*, *Limulngan* and *Kungarakany* (isolate). For a further 7, their 'largest' language numbers fewer than 20 speakers – these are *Tangkic* (largest language *Kayardild* [3], 8–10), *Anson Bay* (largest language *Batjamalh* [34], 20), *Eastern Daly* (largest language *Matngele* [35], perhaps 10), *Western Daly* (*Marrithiyel* [36] and *Marri-Ngarr* [37] have around 20 each), *North Daly* (*Malak-Malak* [38], perhaps 10), and *Wagiman* [39] (10?). In the 8 further families, the healthiest languages are severely threatened, with speaker numbers between 20

and 200: these are *Maran* (*Alawa* [40], perhaps 50), *Iwaidjan* (*Iwaidja* [21] and *Maung* [41], perhaps 150 each), the *Wardaman* group (*Wardaman* [42], 30), *Mindi* (*Jaminjung/Ngaliwurru* [43], 50–150), *Garrwan* (*Garrwa* [44], 200) *Jerragan* (*Kija* [45], perhaps 100), *Worrorran* (*Ungarinyin* [46], 100–), *Bunaban* (*Bunaba* [47], 100), *Nyulnyulan* (*Bardi* [48], 75). *Meryam Mir* [49], the only *Papuan* language spoken in Australia, has around 200 speakers, none fluent under 40.

There have been many positive turn-arounds in public policy and attitudes towards indigenous languages in the last decades – including public recognition, provision of (some) interpreting services, support for regional language centres, appearance of indigenous languages on the airwaves, return of some Aboriginal lands to their traditional owners, some introduction of *Aboriginal* languages into schools (though bilingual education has actually been scaled down in the last few years), and language revival programs such as the *Kaurina* [50] revival program in Adelaide (see Laughren 2000 for a survey of these developments). However, it is not clear that this has significantly retarded the rate of language loss, and other factors, such as the penetration of *English*-language television and videos into remote areas, appear to be at least as powerful in accelerating language loss. It is unlikely that Australia will lose the dubious distinction of being the continent in which language loss is proceeding most rapidly.

4. Priorities for documentation: a survey of the Australian situation

A situation such as that described in §3 obviously places a large number of languages on the high-priority list for documentation, and prioritising them is an invidious task. In this section I give a brief personal view of the 4 main factors I believe should be taken into account in formulating research in the decades to come.

Firstly, there are languages about which virtually nothing is known, and very low numbers of speakers remain. Examples are *Pakanha* [51] and *Ayupathu* [52] in Cape York, virtually undescribed, and with a couple of fluent old speakers each (Barry Alpher p. c.), or *Andajin* [20] in the Kimberleys, for which a speaker has recently been discovered. Probably fewer than ten languages are in this category. Clearly research with languages like this is high-risk, since the speakers could die at any time, and it depends very much on the personality and state of health of the last individuals, but there have been some very successful documentation projects in

similar situations (*Martuthunira* [16] and *Gaagudju* [15] being two, as mentioned above).

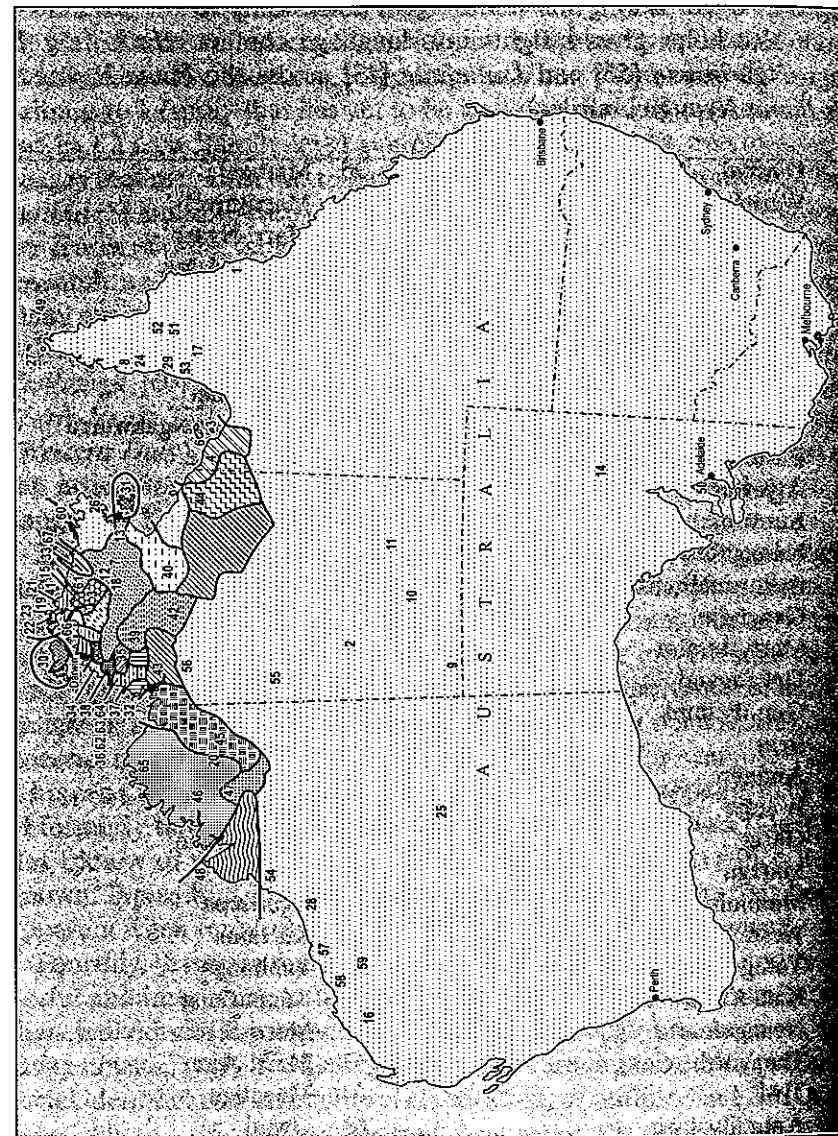
Secondly, there is a large number of languages with a number of speakers ranging from half-a-dozen to 200, and some sort of limited documentation. Effective research is less risky in such situations than when there is only one or two speakers, and particularly important when the language belongs to a family that has yet to be well-described. But, as also pointed out in §2.1., a 'family-level sampling' approach risks overlooking languages that may turn out to have something equivalent in interest to *Kayardild* [3] multiple-case or *Dalabon* [7] disharmonic pronouns, and there are scores of other languages where it is still feasible to carry out detailed research and the descriptions are relatively rudimentary: among the many examples are *Thaayorre* [29], *Koko-Bera* [53], *Karajarri* [54], *Ngar-di* [55], *Ngarinyman* [56], *Ngarla* [57], *Ngarluma* [58], *Kurrama* [59], and *Yanhangu* [60] (all *Pama-Nyungan*), *Kunbarlang* [61] (*Gunwinyguan*), *Marri-Ngarr* [37], *Marri-Sjabin* [62], *Marri-Amu* [63] and *Magatige* [64] (*Western Daly*), and *Gwini* [65] (*Worrorran*). (Though in the time since this article was originally written, major new projects on *Karajarri*, *Yanhangu* and *Kunbarlang* have commenced).

Thirdly, as indicated in §2.2., special registers such as mother-in-law varieties, initiation languages, and systems of trirelational kin terms have great sociolinguistic and semantic interest and are of major cultural importance to the communities that use them, but are far more fragile than the basic register of the language. The fragility and importance of such sub-systems is often overlooked in assessing priorities for documentation, and does not usually appear on lists of speaker numbers. For example, *Tiwi* [30] is known to have had a special 'spirit language', which appears to be a pool of formerly tabooed terms from which names can be drawn (Hart 1930). Yet I know of no recent source setting out the extent to which it is still known. Typically, too, the existence of such special registers only becomes apparent when research on the ordinary variety is well-advanced.

Fourthly, there are many speech communities where it would be more useful and interesting to take the multilingual speech community, rather than individual languages, as the unit of investigation, because of the many unusual ways in which multiple codes are organized for purposes of social identification and cultural elaboration, as outlined in §2.3. In many cases where critically endangered languages are encapsulated within slightly less endangered ones (e.g. *Amurdak* [66] in *Iwaidja*, or some of the *Wik* languages within *Wik-Mungkan*), taking the multilingual speech community as the domain for research reduces the dependence of a

project on the survival of one or two speakers, and is more efficient in terms of gathering cultural information since in situations like gathering plant vocabulary speakers normally volunteer material on several languages at once. Taking the multilingual community as the focus of research also helps give insights into language contact effects (e.g. between *Ngarinman* [56] and *Jaminjung* [43] across the *Pama-Nyungan* / *non-Pama-Nyungan* border).

1	Dyirbal	35	Matngele
2	Warlpiri	36	Marrithiyel
3	Kayardild	37	Marri-Ngarr
4	Yukulta	38	Malak-Malak
5	Lardil	39	Wagiman
6	Yanggal	40	Alawa
7	Dalabon	41	Maung
8	Wik-Mungkan	42	Wardaman
9	Pitjantjatjara/Yankunytjatjara	43	Jaminjung/Ngaliwurru
10	Eastern and Central Arrernte	44	Garrwa
11	Alyawarr	45	Kija
12	Kunwinjku	46	Ungarinyin
13	Nunggubuyu	47	Bunaba
14	Adnyamathanha	48	Bardi
15	Gaagudju	49	Meryam Mir
16	Martuthunira	50	Kaurua
17	Oykangand	51	Pakanha
18	Gun-djeihmi	52	Ayupathu
19	Ilgar	53	Koko-Bera
20	Andajin	54	Karajarri
21	Iwaidja	55	Ngaridi
22	Garig	56	Ngarinyman
23	Marrgu	57	Ngarla
24	Muminh	58	Ngarluma
25	Western Desert	59	Kurrama
26	Yolngu-Matha	60	Yanhangu
27	Kala Kawaw Ya	61	Kunbarlang
28	Nyangumarta	62	Marri-Sjabin
29	Thaayorre	63	Marri-Amu
30	Tiwi	64	Magatige
31	Anindilyakwa	65	Gwini
32	Murrinhpatha	66	Amurdak
33	Ndjébbana	67	Burarra
34	Batjamalh		



Map 23. Australia

5. Final remarks

In closing it is worth remarking on the rapidly evolving changes to the relationship between researcher, research institution, speaker, and local community. As indigenous groups become less disenfranchised, the relationships between all parties have become more complex. One form this has taken has been the drawing up of a code of ethics for linguists working in indigenous communities in Australia by the Australian Linguistic Society in the 1980s, more or less simultaneously with the drafting of a statement of linguistic rights by the *Aboriginal* Languages Association, an indigenous national association. There has been considerable debate about issues of intellectual property, responsibilities of the linguist to the language community, appropriate training of members of the language community as part of the research process, and a trade-off between outside researchers' goals and those of the community (e.g. an agreement to produce pedagogical or dictionary materials in return for community support for research on a more 'academic' topic, such as syntax or phonology) – see e.g. Wilkins (1992). It is now normal for regional language centres to expect full consultation on such matters as part of the process of planning, execution and 'bringing back' of research projects.

Although some academics have seen this as an unwelcome limitation on their intellectual freedom, and there have been rare cases where it has created difficulties in publishing their research, the net effect has been to make for a richer and more interesting understanding of *Australian* languages, better able to incorporate the cultural insights of native speaker researchers. Some examples of the fecundity of such 'two-way research' were given in §2.2. under the discussion of lexicography and text material. As we strive to document what we can of the extraordinary linguistic legacy that will fade away over the next century, a research process that combines the questions and insights of insiders with those of outsiders stands the best chance of describing languages, on their own terms and in ways that capture possibilities beyond our current imaginations.

Notes

1. There is a single exception to this within Australia's political borders, though not on the continent proper: *Meryam Mir*, spoken in the eastern Torres Strait between Australian and Papua New Guinea, belong to the *Trans-Fly* family whose other members are located in Papua New Guinea.

2. On similar attitudes in Melanesia see Laycock (1982: 33), who quotes a Sepik man saying 'it wouldn't be any good if we all talked the same; we like to know where people come from'.
3. It is a peculiar fact about linguistics that practically none of the most astonishing typological features eventually discovered empirically have been anticipated through prior speculation, whether by philosophers or armchair linguists. This is in marked contrast to the natural sciences where, for example, mathematicians had anticipated the possible existence of the Einsteinian universe by playing speculatively with non-Euclidean geometry, or chemistry, where the likely existence of a large number of elements was deduced by Mendeleev before they were discovered.
4. This is by no means exhaustive – e.g. I omit the special systems of sign language, of interest to our understanding of grammatical and semantic structure – see Kendon (1988), Wilkins (1997).
5. In fact it is possible to see both as special registers: mother-in-law registers index a simple two-value opposition in terms of kin relationships (respectful vs unmarked) but do so by distinct words over the whole vocabulary, whereas trirelational kin terms index a much larger set of relationships between speaker and hearer (as indicated by the translations for (8)), but only register this on that part of the vocabulary dealing with kinship.
6. These forms are cited in a practical orthography. See Hale and Nash 1997.
7. This is slightly oversimplified, since the allomorph *-ngkur* is restricted to Damin and may represent an archaic post-vocalic form – see Hale (1973).
8. The Institute for Aboriginal Development in Alice Springs commissioned the Central Australian handbook; this institution has a broader brief than just language but includes a very active language department (see discussion of lexicography in §2.2.). The Kimberley survey was commissioned by the Kimberley Language Resource Centre, and the Pilbara survey by the Pilbara Aboriginal Language Centre.
9. As far back as Tindale (1974: 225), the group inhabiting Ilgar country (though labeled Gaari by him) was described by him as 'a now extinct people'. Chir Wardaga, a knowledgeable *Ilgar* speaker, came forward as an *Ilgar* speaker during research for a Native Title hearing by anthropologist Jeannie Devitt in 1993.

References

Alpher, Barry

- 1982 Dalabon dual-subject prefixes, kinship categories, and gender skewing. In Heath, Merlan and Rumsey (eds.) 1–18.
- 1991 *Yir-Yoront lexicon. Sketch and dictionary of an Australian language*. Berlin: Mouton de Gruyter.
- 1993 Out-of-the-ordinary ways of using a language. In M. Walsh and J. Yallop (eds.) *Language and culture in Aboriginal Australia*.

Austin, Peter

- 1991 Australian Aboriginal languages. In Michael G. Clyne (ed.) *Linguistics in Australia: trends in research*. Canberra: Academy of the Social Sciences in Australia. 55–74.

Black, Paul

- 1983 *Aboriginal languages of the Northern Territory*. Darwin: School of Australian linguistics.

Brandl, Maria and Michael Walsh

- 1982 Speakers of many tongues: toward understanding multilingualism among Aboriginal Australians. In G. R. McKay (ed.) *Australian Aborigines: sociolinguistic studies*. (International journal of the sociology of language 36). Berlin: Mouton. 71–81.

Capell, Arthur C.

- 1962 *Some linguistic types in Australia*. Sydney: Oceania Linguistic Monograph 7.

Dench, Alan

- 1995 *Martuthunira, a language of the Pilbara region of Western Australia*. Canberra: Pacific Linguistics.

Dixon, R. M. W.

- 1971 A method of semantic description. In D. D. Steinberg and L. A. Jakobovits (eds.) 436–471.
- 1972 *The Dyirbal language of north Queensland*. London: Cambridge University Press.

Evans, Nicholas

- 1992 Multiple semiotic systems, hyperpolysemy and the reconstruction of semantic change in Australian languages. In Michael Morrissey and Günter Kellermann (eds.), *Diachrony within synchrony: language, history and cognition*. Bern: Peter Lang. 475–508.
- 1995 *A grammar of Kayardild, with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
- 2000 Iwaidjan, a very un-Australian language family. *Linguistic typology* 4.1: 91–112.
- 2001 The last speaker is dead – long live the last speaker! In Paul Newman and Martha Ratcliff (eds.) *Linguistic fieldwork*. Cambridge: Cambridge University Press. 250–281.
- 2003 *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.

Hard, Cliff

- 2002 *Pitjantjatjara / Yankunytjatjara to English dictionary*. (2nd edition). Alice Springs: Institute for Aboriginal Development.

Hard, Cliff and Nicholas Thieberger

- 2007 Lexicographic research on Australian Aboriginal languages 1968–1993. In Tryon and Walsh (eds.) 175–208.

- Green, Jenny
1992 *Alyawarr to English dictionary*. Alice Springs: Institute for Aboriginal Development.
- Grimes, Barbara F.
1988 *Ethnologue. Languages of the World* (11th edition). Dallas: Summer Institute of Linguistics.
- Hale, Kenneth L.
1966 Kinship reflections in syntax: some Australian examples. *Word* 22: 318–324.
1971 Notes on a Walpiri tradition of antonymy. In Steinberg and Jakobovits (eds.) 472–482.
1973 Deep-surface canonical disparities in relation to analysis and change: an Australian example. In T. A. Sebeok (ed.) *Current trends in linguistics 8: Linguistics in Oceania*. The Hague: Mouton. 401–458.
1982 The logic of Damin kinship terminology. In Heath, Merlan and Rumsey (eds.) 31–37.
1983 Warlpiri and the grammar of non-configurational languages. *Natural language and linguistic theory* 1,1: 5–47.
1998 On endangered languages and the importance of linguistic diversity. In Lenore A. Grenoble and Lindsay J. Whaley (eds.) *Endangered languages: Language loss and community response*. Cambridge: Cambridge University Press. 192–216.
- Hale, Kenneth L. and David Nash
1997 Damin and Lardil phonotactics. In Darrell Tryon and Michael Walsh (eds.) *Boundary rider: essays in honours of Geoffrey O'Grady*. Canberra: Pacific Linguistics. 247–259.
- Hart, Charles W. M.
1930 Personal names among the Tiwi. *Oceania* 1,3: 28–290.
- Harvey, Mark
1992 *The Gaugudju people and their language*. Unpublished Ph.D. Thesis, University of Sydney.
- Heath, Jeffrey
1981 *Nunggubuyu myths and ethnographic texts*. Canberra: Australian Institute of Aboriginal Studies.
- Heath, Jeffrey, Alan Rumsey and Francesca Merlan (eds.)
1982 *Languages of kinship in Aboriginal Australia*. Sydney: Oceania Linguistic Monographs No. 24.
- Henderson, John and Veronica Dobson
1994 *Eastern and Central Arrernte to English dictionary*. Alice Springs: Institute for Aboriginal Development.
- Keen, Sandra
1983 Yukulta. In R. M. W. Dixon and Barry J. Blake (eds.) *Handbook of Australian languages, Vol. 3*. Canberra: Australian National University Press. 190–304.

- Kendon, Adam
1988 *Sign languages of Aboriginal Australia*. Cambridge: Cambridge University Press.
- Kilham, Christine, Mabel Pamulkan, Jennifer Pootchemunka and Topsy Wolmby.
1986 *Dictionary and source book of the Wik-Mungkan language*. Darwin: Summer Institute of Linguistics.
- Laughren, Mary
1982 Warlpiri kinship structure. In Heath, Merlan and Rumsey (eds.). 72–85.
2000 Australian Aboriginal languages: their contemporary status and functions. In R. M. W. Dixon and Barry J. Blake (eds.) *Handbook of Australian languages, Vol. V*. Melbourne: Oxford University Press. 1–32.
- Laycock, Don
1982 Linguistic diversity in Melanesia: a tentative explanation. In Rainer Carle, Martina Henschke, Peter W. Pink, Christel Rost and Karen Stadtlender (eds.) *GAVA: studies in Austronesian languages and cultures, dedicated to Hans Kähler*. Berlin: Dietrich Reimer. 31–37.
- Livi-Bacci, Massimo
1992 *A concise history of world population* (transl. Carl Ipsen). Cambridge, Massachusetts: Blackwells.
- McConvell, Patrick and Nicholas Thieberger
2000 *The state of indigenous languages in Australia*. Canberra: Australian Institute of Aboriginal Studies.
- McGregor, William B.
1988 *Handbook of Kimberley languages. Volume 1. General information*. Canberra: Pacific Linguistics.
2002 Verb classification in Australian languages. MS.
- McKnight, David
1999 *People, countries and the Rainbow Serpent*. Oxford: Oxford University Press.
- Manakgu, Andrew
1998 *Kunbalak. Stories for Kunwinjku young people in mother-in-law language, ordinary Kunwinjku and English* (2nd edition). Kunbarllanjja (Oenpelli): Kunwinjku Language Centre.
- Merlan, Francesca
1981 Land, language and social identity in Aboriginal Australia. *Mankind* 13,2: 133–148.
1989 Jawoyn relationship terms: interactional dimensions of Australian kin classification. *Anthropological linguistics* 31,3/4: 227–263.
- Morphy, Frances
1977 Language and moiety: sociolectal variation in a Yu:ngu language of north-east Arnhem Land. *Canberra anthropology* 1,1: 51–60.
- Napaljarri, Peggy and Lee Cataldi
1994 *Yimikirli – Warlpiri dreamings and histories*. San Francisco: Harper Collins.

- Nash, David (ed.) and Kathy Menning (comp.)
1981 *Source book for Central Australian languages*. Alice Springs: Institute for Aboriginal Development.
- Nettle, Daniel
1999 *Linguistic diversity*. Cambridge: Cambridge University Press.
- Nganjmira, Nawakadj
1997 *Kunwinjku spirit*. Melbourne: Miegunyah Press.
- Oates, William J. and Lynette F. Oates
1970 *A revised linguistic survey of Australia*. Canberra: Australian Institute of Aboriginal Studies.
- Plank, Frans (ed.)
1995 *Double case: agreement by Suffixaufnahme*. Oxford: Oxford University Press.
- Pym, Noreen (with Bonnie Larrimore)
1979 *Papers on Iwaidja phonology and grammar*. Darwin: Summer Institute of Linguistics, Australian Aborigines Branch.
- Saunders, Thomas
1999 A first look at Andajin. Paper presented at Australian Linguistics Conference, University of Western Australia, Perth, September 30, 1999.
- Schebeck, Bernhard
1974 *Texts on the social system of the Atinyamatana people, with grammatical notes*. Canberra: Pacific Linguistics.
- Schmidt, Annette
1990 *The loss of Australia's Aboriginal language heritage*. Canberra: Aboriginal Studies Press.
- Schultze-Berndt, Eva
2000 *Simple and complex verbs in Jaminjung*. A study of event categorization in an Australian language. Proefschrift, Katholieke Universiteit Nijmegen.
- Sharp, Janet and Nicholas Thieberger
1992 *Bilybara. The Aboriginal languages of the Pilbara region of Western Australia*. Port Hedland: Wangka Maya, The Pilbara Aboriginal Language Centre.
- Smith, Ian
1986 Language contact and the life or death of Kugu Muminh. In Joshua A. Fishman, Andr  e Tabouret-Keller, Michael Clyne, Bh. Krishnamurti and Mohamed Abdulaziz (eds.) *The Fergusonian impact: in honor of Charles A. Ferguson on the occasion of his 65th birthday, volume 2: Sociolinguistics and the sociology of language*. Berlin: Mouton de Gruyter. 513-532.
- Smith, Ian and Steve Johnson
1986 Sociolinguistic patterns in an unstratified society: the patilects of Kugu Nganhcara. *Journal of the atlantic provinces linguistic association* 8: 29-43.

- Steinberg, D. D. and L. A. Jakobovits (eds.)
1971 *Semantics*. Cambridge: Cambridge University Press.
- Strehlow, Theodor G. H.
1971 *Songs of Central Australia*. Sydney: Angus and Robertson.
- Sutton, Peter
1978 Wik: Aboriginal society, territory and language at Cape Keerweer, Cape York Peninsula, Australia. University of Queensland: unpublished Ph.D. Dissertation.
- Thieberger, Nicholas
1993 *Handbook of WA Aboriginal languages south of the Kimberley Region*. Canberra: Pacific Linguistics.
- Thieberger, Nicholas and William B. McGregor (eds.)
1994 *Macquarie Aboriginal words*. Macquarie University: Macquarie Dictionary.
- Tindale, Norman B.
1974 *Tribal boundaries in Aboriginal Australia*. (2nd edition). Berkeley: University of California Press.
- Trigger, David
1987 Languages, linguistic groups and status relations at Doomadgee, an Aboriginal settlement in north-west Queensland, Australia. *Oceania* 57: 217-238.
- Tryon, Darrell and Michael Walsh (eds.)
1997 *Boundary rider: essays in honour of Geoffrey O'Grady*. Canberra: Pacific Linguistics.
- Wilkins, David
1992 Linguistic research under Aboriginal control: a personal account of fieldwork in Central Australia. *Australian journal of linguistics* 12,1: 171-200.
1997 Handsigns and hyperpolysemy: exploring the cultural foundations of semantic association. In Tryon and Walsh (eds.) 413-444.
- Wilkinson, Melanie
1991 Djamparrpuynu: a Yolngu variety of Northern Australia. Ph.D. Dissertation, University of Sydney.
- Wurm, Stephen
1996 *Atlas of the world's languages in danger of disappearing*. Paris/Canberra: Unesco Publishing/Pacific Linguistics.
forthcoming a. Alphabetical list of entries for threatened and extinct languages.
forthcoming b. Australian Aboriginal languages which are endangered in some form, and additional extinct ones.