

The Cambridge Handbook of

# Endangered Languages

edited by **Peter K. Austin**  
and **Julia Sallabank**

# The Cambridge Handbook of Endangered Languages

Edited by  
**Peter K. Austin**  
and  
**Julia Sallabank**



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town,  
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press  
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)  
Information on this title: [www.cambridge.org/9780521882156](http://www.cambridge.org/9780521882156)

© Cambridge University Press 2011

This publication is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without the written  
permission of Cambridge University Press.

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this publication is available from the British Library*

*Library of Congress Cataloguing in Publication data*

The Cambridge handbook of endangered languages / [edited by] Peter K. Austin,  
Julia Sallabank.

p. cm. – (Cambridge handbooks in language and linguistics)

ISBN 978-0-521-88215-6 (hardback)

1. Language obsolescence—Handbooks, manuals, etc. 2. Language and  
languages—Handbooks, manuals, etc. I. Austin, Peter. II. Sallabank, Julia.

III. Title. IV. Series.

P40.5.L33C36 2011

409—dc22

2010051874

ISBN 978-0-521-88215-6 Hardback

Cambridge University Press has no responsibility for the persistence or  
accuracy of URLs for external or third-party internet websites referred to in  
this publication, and does not guarantee that any content on such websites is,  
or will remain, accurate or appropriate.

# Contributors

**Peter K. Austin**, Department of Linguistics, School of Oriental and African Studies

**Josh Berson**, Max Planck Institute for the History of Science, Berlin

**Michel Bert**, Laboratoire DDL, Université Lumière Lyon 2

**Claire Bowern**, Department of Linguistics, Yale University

**David Bradley**, Department of Linguistics, La Trobe University

**Lyle Campbell**, Department of Linguistics, University of Utah

**Lisa Conathan**, Beinecke Rare Book and Manuscript Library, Yale University

**Serafin M. Coronel-Molina**, Department of Education, Indiana University, Bloomington

**Lise M. Dobrin**, Department of Anthropology, University of Virginia

**Jeff Good**, Department of Linguistics, University of Buffalo

**Lenore A. Grenoble**, Department of Slavic Languages and Literatures, University of Chicago

**Colette Grinevald**, Laboratoire DDL, Université Lumière Lyon 2

**Wayne Harbert**, Department of Linguistics, Cornell University

**Leanne Hinton**, Department of Linguistics, University of California, Berkeley

**Gary Holton**, Department of Linguistics, University of Fairbanks

**Anthony Jukes**, Department of Linguistics, La Trobe University

**Friederike Lüpke**, Department of Linguistics, School of Oriental and African Studies

**Teresa L. McCarty**, Division of Educational Leadership and Policy Studies, Arizona State University

**Lev Michael**, Department of Linguistics, University of California, Berkeley

**Máiréad Moriarty**, Department of Languages and Cultural Studies, University of Limerick

**Ulrike Mosel**, Department of Linguistics, University of Kiel

**David Nathan**, Department of Linguistics, School of Oriental and African Studies

**Carmel O'Shannessy**, Department of Linguistics, University of Michigan

**Naomi Palosaari**, Department of Linguistics, University of Utah

**Julia Sallabank**, Department of Linguistics, School of Oriental and African Studies

**Bernard Spolsky**, Department of English, Bar-Ilan University

**Anthony C. Woodbury**, Department of Linguistics, University of Texas at Austin

# Acknowledgements

The idea for a *Handbook of Endangered Languages* was first proposed to Peter Austin by Andrew Winnard of Cambridge University Press back in 2006. Peter prepared a book outline, contacted potential authors, and submitted a detailed proposal describing the overall conception of the book and the scope of the individual chapters to CUP, which was duly refereed and accepted with some minor changes. The project stumbled along for a year and a half until Julia Sallabank joined SOAS in 2008 and agreed to co-edit the *Handbook*, bringing with her seventeen years of editorial experience at Oxford University Press, and a different and complementary perspective on language endangerment (as well as a different network of colleagues). The individual authors were then contracted to write their chapters, and the serious business of reading, editing and correcting drafts, cajoling and encouraging authors, and knocking the book into shape began in earnest.

Preparing a handbook with twenty-three chapters by authors located around the world has been a major undertaking and its completion has only been possible as a result of generous contributions from a number of sources. We would like to thank the following:

- the authors, who willingly took on the task of preparing their chapters and patiently responded to our suggestions for changes;
- Stuart McGill, who joined us during 2009 as a research assistant, reading all of the chapters and making numerous detailed and insightful comments on them;
- Jennifer Marshall, who worked tirelessly for many months checking and correcting bibliographical references;
- Andrew Winnard, for originally sponsoring the project, and for cajoling and nagging us at appropriate times to ensure that we were on track for completion.

- SOAS Department of Linguistics (and especially the Head of Department, Peter Sells), for ongoing support throughout the project, and for granting Peter Austin sabbatical leave, during which the final compilation and editing on the volume was done. Thanks are due to the Research Centre for Linguistic Typology, La Trobe University (and its Director, Randy La Polla), for hosting Peter during these final and vital stages.

# Contents

<i>List of figures</i>	page vii
<i>List of tables</i>	viii
<i>List of contributors</i>	ix
<i>Acknowledgements</i>	xi
 1 Introduction <i>Peter K. Austin and Julia Sallabank</i>	 1
<b>Part I</b> Endangered languages	25
2 Language ecology and endangerment <i>Lenore A. Grenoble</i>	27
3 Speakers and communities <i>Colette Grinevald and Michel Bert</i>	45
4 A survey of language endangerment <i>David Bradley</i>	66
5 Language contact and change in endangered languages <i>Carmel O'Shannessy</i>	78
6 Structural aspects of language endangerment <i>Naomi Palosaari and Lyle Campbell</i>	100
7 Language and culture <i>Lev Michael</i>	120
8 Language and society <i>Bernard Spolsky</i>	141
<b>Part II</b> Language documentation	157
9 Language documentation <i>Anthony C. Woodbury</i>	159
10 Speakers and language documentation <i>Lise M. Dobrin and Josh Berson</i>	187
11 Data and language documentation <i>Jeff Good</i>	212
12 Archiving and language documentation <i>Lisa Conathan</i>	235
13 Digital archiving <i>David Nathan</i>	255
<b>Part III</b> Responses	275
14 Language policy for endangered languages <i>Julia Sallabank</i>	277
15 Revitalization of endangered languages <i>Leanne Hinton</i>	291
16 Orthography development <i>Friederike Lüpke</i>	312



17	Lexicography in endangered language communities <i>Ulrike Mosel</i>	337
18	Language curriculum design and evaluation for endangered languages <i>Serafin M. Coronel-Molina and Teresa L. McCarty</i>	354
19	The role of information technology in supporting minority and endangered languages <i>Gary Holton</i>	371
	<b>Part IV</b> Challenges	401
20	Endangered languages and economic development <i>Wayne Harbert</i>	403
21	Researcher training and capacity development in language documentation <i>Anthony Jukes</i>	423
22	New roles for endangered languages <i>Máiréad Moriarty</i>	446
23	Planning a language-documentation project <i>Claire Bower</i>	459
	References	483
	<i>Index of language names</i>	554
	<i>Index</i>	558

# Figures

1.1	The proportion of languages in each continent of the world	page 5
16.1	Sample pages of the Jalonke primer using a Latin-based orthography (Lüpke <i>et al.</i> , 2000)	319
16.2	Berber flag adopted in 1998 by the Amazigh World Congress, featuring the Tifinagh grapheme <i>yaz</i>	324
16.3	Sample pages from the Manding-English dictionary: bigraphic in Latin and Nk'o script (Vydrine 1999)	325
16.4	Palestinian cartoon by Naji al-Ali (2009: 32)	328
19.1	Southern Tutchone text image. After Yukon Native Language Centre (1996)	379
19.2	Story excerpt from Dane Wajich (Doig River First Nation 2007)	380
19.3	Screenshot of RosettaStone Iñupiaq	382
19.4	Screenshot of Lexique Pro viewing the Western Pantar lexicon	385
19.5	Screenshot of <i>Kirrkirr</i> program showing network of semantic relationships	386
19.6	An all-too-common response from an endangered language website	397

# Tables

1.1	UNESCO's Language Vitality and Endangerment framework	page 3
2.1	Language endangerment scale proposed by Krauss (1997)	41
4.1	Degrees of endangerment by continent	68
4.2	Degrees of endangerment by country, mainland South-East Asia	69
5.1	Thomason and Kaufman's (1988) borrowing scale, as in Winford (2003: 30)	81
6.1	Nivacle consonant inventory	102
6.2	Nivacle demonstratives	108
11.1	Tabular implementation of word list entries	221
13.1	Binary symbols in files are mapped onto orthographic characters through standardized character sets or encodings	259
14.1	Definitions of language policy and planning	278
16.1	Differences between NTM alphabet and national alphabet for Baïnouk with IPA correspondences	317
16.2	Hausa Ajami chart after Philips (2000: 21f.)	327
21.1	DOBES training course June 2009	433
21.2	ELDP grantee training course timetable, 27 August–2 September 2008. 90-minute sessions	434
21.3	The InField 2008 workshop tracks	435

# Introduction

Peter K. Austin and Julia Sallabank

## 1.1 Language endangerment

It is generally agreed by linguists that today there are about 7,000 languages spoken across the world; and that at least half of these may no longer continue to exist after a few more generations as they are not being learnt by children as first languages. Such languages are said to be **ENDANGERED LANGUAGES**.<sup>1</sup>

Current language and population distributions across the world are heavily skewed: there is a small number of very large languages (the top twenty languages, like Chinese, English, Hindi/Urdu, Spanish have over 50 million speakers each and are together spoken by 50 per cent of the world's population), and a very large number of small languages with speaker communities in their thousands or hundreds. Economic, political, social and cultural power tends to be held by speakers of the majority languages, while the many thousands of minority languages are marginalized and their speakers are under pressure to shift to the dominant tongues. In the past sixty years, since around the end of World War II, there have been radical reductions in speaker numbers of minority **AUTOCHTHONOUS** languages, especially in Australia, Siberia, Asia and the Americas. In addition, the languages under pressure show shifting age profiles where it is only older people who continue to speak the threatened languages and younger people typically show **LANGUAGE SHIFT**, meaning they move to using more powerful regional, national or global languages. Language shift can take place rapidly, over a generation or two, or it can take place gradually, but continuously, over several generations. Language shift often takes place through a period of **UNSTABLE BILINGUALISM OR MULTILINGUALISM**, that is, speakers use two or more languages but one (or more) of them is more dominant and used increasingly widely until finally it (or they) take over the roles previously carried by the endangered language(s).

Linguists are becoming increasingly alarmed at the rate at which languages are going out of use. A special issue of the journal *Language* (Hale *et al.* 1992), based on a colloquium held at the 1991 annual meeting of the Linguistic Society of America, drew the attention of the linguistics profession to the scale of language endangerment, and called for a concerted effort by linguists to record the remaining speakers, and to create linguistic archives for future reference. In this issue of *Language*, Krauss (1992) estimated that 90 per cent of the world's languages would be severely endangered or gone by 2100. According to more optimistic estimates such as Nettle and Romaine (2000) and Crystal (2000), 'only' 50 per cent will be lost.

This 'call to action' reinvigorated fieldwork and documentation of languages, which had characterized an earlier era of linguistics (associated with the work of Franz Boas and his students). In the past ten years a number of initiatives responding to the call of Hale, Krauss, Grinevald and Yamamoto (and others) have been launched, including:

- the Hans Rausing Endangered Languages Project,<sup>2</sup> funded by Arcadia, which gives research grants for language documentation projects, maintains a digital archive of recordings, transcriptions and meta-data, and runs an academic programme with newly introduced MA and PhD degrees to train linguists and researchers;
- the Volkswagen Foundation's sponsorship of the DoBeS (Dokumentation Bedrohter Sprachen)<sup>3</sup> project;
- the US National Science Foundation (NSF) and National Endowment for the Humanities (NEH) Documenting Endangered Languages initiative (DEL), 'a new, multi-year effort to preserve records of key languages before they become extinct';<sup>4</sup>
- the European Science Foundation Better Analyses Based on Endangered Languages programme (EuroBABEL) whose main purpose is 'to promote empirical research on under-described endangered languages, both spoken and signed';<sup>5</sup>
- The Chirac Foundation for Sustainable Development and Cultural Dialogue Sorosoro programme 'so the languages of the world may prosper';<sup>6</sup>
- The World Oral Literature Project based at Cambridge University, 'to record the voices of vanishing worlds';<sup>7</sup>
- smaller non-profit initiatives, notably the Foundation for Endangered Languages,<sup>8</sup> the Endangered Language Fund,<sup>9</sup> and the Gesellschaft für bedrohte Sprachen.<sup>10</sup>

Intergovernmental agencies have taken on board the problem of the loss of linguistic diversity. The United Nations has a number of policy papers and guidelines for governmental action plans on the UNESCO website under the heading of safeguarding 'intangible cultural heritage' (UNESCO 2003a; 2003b; see Section 1.5.4 below for further discussion).

Table 1.1. UNESCO's Language Vitality and Endangerment framework

Degree of endangerment	Intergenerational language transmission
Safe	language is spoken by all generations; intergenerational transmission is uninterrupted
Vulnerable	most children speak the language, but it may be restricted to certain domains (e.g., home)
Definitely endangered	children no longer learn the language as mother tongue in the home
Severely endangered	language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves
Critically endangered	the youngest speakers are grandparents and older, and they speak the language partially and infrequently
Extinct	there are no speakers left

One of the tasks that UNESCO has tried to tackle is how to categorize levels of endangerment. Assessing levels of language knowledge and use is an important element of language documentation and planning because 'a language spoken by several thousand people on a daily basis presents a much different set of options for revitalization than a language that has a dozen native speakers who rarely use it' (Grenoble and Whaley 2006: 3). Although numerous schemes have been proposed, the most comprehensive is UNESCO's Language Vitality and Endangerment framework,<sup>11</sup> which is shown in Table 1.1. It establishes six degrees of vitality/endangerment based on nine factors. Of these factors, the most salient is that of intergenerational transmission: whether or not a language is used in the family and passed from an older generation to children. This factor is generally accepted as the 'gold standard' of language vitality (Fishman 1991). (For more on measuring language vitality, see Grenoble, Chapter 2, and Grinevald and Bert, Chapter 3.)

## 1.2 Counting languages

Overviews of the study of language endangerment usually start with a list of statistics about the number of languages in the world, the proportion considered endangered, and so on. The usual source of statistics concerning the number of languages and their users is *Ethnologue* (Lewis 2009), which listed 6,909 living languages at the time of going to press.

However, this headline figure masks inherent problems in the counting of languages, as the Introduction to *Ethnologue* itself recognizes. Many linguists use the criterion of MUTUAL INTELLIGIBILITY to distinguish languages: if users of two language varieties cannot understand each other, the varieties are considered to be different languages. If they can

understand each other, the varieties are considered mutually comprehensible dialects of the same language. However, mutual intelligibility is to a certain extent a function of attitudes and politics; that is, whether or not people want to understand each other. Such attitudes are, in part, linked to whether a community considers itself to have a distinct ethno-linguistic identity, but members of a community may not agree about this. Because of such issues, some linguists (especially sociolinguists and anthropological linguists influenced by postmodern theories) now question whether language boundaries can be identified at all.

Politics also plays an important part in language differentiation. Following nineteenth-century philosophers such as Herder, language has been considered a crucial element of national identity, with 'one state, one people, one language' being seen as the ideal. But languages do not necessarily follow political boundaries. For example, Quechua is often thought of as one language, but in fact this is an overarching name which denotes a group of related language varieties (Coronel-Molina and McCarty, Chapter 18). Linguists distinguish between twenty-seven Quechuan indigenous languages in Peru, but the Peruvian government only recognizes six of these as languages (the official national language is the colonial language, Spanish). Minority groups may claim full 'language' status for their variety, especially if it has been disregarded as a 'substandard' dialect in the past (e.g. Aragonese in Spain). Separatist groups may highlight linguistic differences to support their cause, while national governments may play these down. Paradoxes such as the mutual incomprehensibility of Chinese 'dialects' compared to the mutual comprehensibility of mainland Scandinavian languages are clearly motivated by political and nationalistic considerations rather than linguistic ones. (See Bradley, Chapter 4, on the many complex issues connected to delineating languages, with other examples from South-East Asia.)

In addition, complete information on all of the world's languages is not available: the majority have not been recorded or analysed by linguists, have no dictionaries or even written form, and are not recognized officially in the countries in which they are spoken. What information there is available, is often out of date: for example, for Guernesiais (Channel Islands, Europe) the information in *Ethnologue* is based on a 1976 estimate and ignores more recent data such as the 2001 census.

The Introduction to *Ethnologue* admits that: 'Because languages are dynamic and variable and undergo constant change, the total number of living languages in the world cannot be known precisely.' Nevertheless, the traditional approach to counting languages is still followed by most researchers, and also by the UNESCO *Atlas of Languages in Danger of Disappearing* (Moseley 2009). Despite their shortcomings however, at the very least these compendia provide a useful guide to relative levels of linguistic diversity around the world. Figure 1.1 shows the proportion of languages in each continent. It can be seen that Europe is by far the least

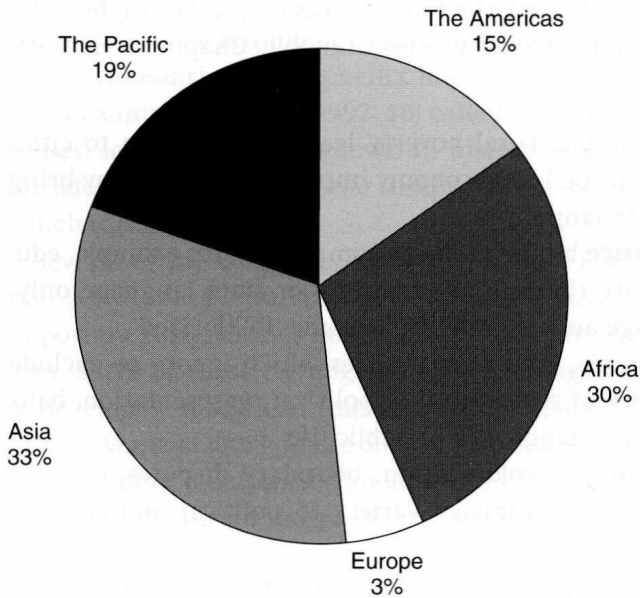


Figure 1.1. *The proportion of languages in each continent of the world*

linguistically diverse continent, which is worrying if other parts of the world continue to follow European trends.

### 1.3 Causes of language endangerment

The causes of language endangerment can be divided into four main categories (synthesized from Nettle and Romaine 2000; Crystal 2000; see also Grenoble, Chapter 2):

- natural catastrophes, famine, disease: for example, Malol, Papua New Guinea (earthquake); Andaman Islands (tsunami)
- war and genocide, for example, Tasmania (genocide by colonists); Brazilian indigenous peoples (disputes over land and resources); El Salvador (civil war)
- overt repression, often in the name of 'national unity' or ASSIMILATION (including forcible resettlement): for example, Kurdish, Welsh, Native American languages
- cultural/political/economic dominance, for example, Ainu, Manx, Sorbian, Quechua and many others.

Factors often overlap or occur together. The dividing lines can be difficult to distinguish. For example, in the Americas and Australia disease and suppression of indigenous cultures spread after colonization, and in Ireland many Irish speakers died or emigrated due to colonial government inaction which compounded the effects of the potato blight famine in the nineteenth century.



The fourth category, which is the most common, can be further subdivided into five common factors (see also Grenoble, Chapter 2; Harbert, Chapter 20):

- **economic:** for example, rural poverty leads to migration to cities and further afield. If the local economy improves, tourism may bring speakers of majority languages
- **cultural dominance** by the majority community, for example, education and literature through the majority or state language only; indigenous language and culture may become 'folklorized'
- **political:** for example, education policies which ignore or exclude local languages, lack of recognition or political representation, bans on the use of minority languages in public life
- **historical:** for example, colonization, boundary disputes, the rise of one group and their language variety to political and cultural dominance
- **attitudinal:** for example, minority languages become associated with poverty, illiteracy and hardship, while the dominant language is associated with progress/escape.

More recently, there have been many community initiatives for LANGUAGE REVIVAL OR LANGUAGE REVITALIZATION of endangered languages to expand the contexts in which they are used and to ensure they continue to be passed on to new generations (for examples see Grenoble and Whaley 2006; Hinton and Hale 2001; Hinton, Chapter 15; Moriarty, Chapter 22).

## 1.4 Why worry about language endangerment?

### 1.4.1 Value to linguistic science

Throughout history languages have died out and been replaced by others through LANGUAGE CONTACT; that is, contact between groups of people speaking different languages, or through DIVERGENCE due to lack of communication over distances (Dalby 2002). Until recently this was seen as a natural cycle of change. But the growing number of linguistic varieties no longer being learnt by children, coupled with a tendency for language shift, where speakers move to languages of wider communication (especially major languages like English or Spanish), means that unless the myriad inventive ways in which humans express themselves are documented now, future generations may have no knowledge of them. For example, Ubykh, a Caucasian language whose last fully competent speaker (Tevfik Esenç) died in 1992, has eighty-four distinct consonants and, according to some analyses, only two phonologically distinct vowels. This is the smallest proportion of vowels to consonants known, and the possibility that such languages could exist would have been unknown