

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

---

## Endangered Languages and New Technologies

At a time when many of the world's languages are at risk of extinction, the imperative to document, analyse and teach them before time runs out is very great. At this critical time new technologies such as visual and aural archiving, digitization of textual resources, electronic mapping and social media have the potential to play an integral role in language maintenance and revitalization.

Drawing on studies of endangered languages from around the world – Europe, Asia, Africa and North and South America – this volume considers how these new resources might best be applied, and the problems that they can bring. It also reassesses more traditional techniques of documentation in light of new technologies and works towards achieving a practicable synthesis of old and new methodologies.

This accessible volume will be of interest to researchers in language endangerment, language typology and linguistic anthropology, and to community members working in native language maintenance.

MARI C. JONES is Reader in French Linguistics and Language Change at the University of Cambridge and Fellow in Modern and Medieval Languages at Peterhouse, Cambridge. A highly experienced fieldworker, she has published extensively on language obsolescence and revitalization in relation to Insular and Continental Norman, Welsh and Breton. Her recent publications include *Keeping Languages Alive* (2013), edited with Sarah Ogilvie.

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

---

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

---

# Endangered Languages and New Technologies

---

*Edited by*

Mari C. Jones



**CAMBRIDGE**  
UNIVERSITY PRESS

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

**CAMBRIDGE**  
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781107627840](http://www.cambridge.org/9781107627840)

© Cambridge University Press 2015

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 2015

Printed in the United Kingdom by Clays, St Ives plc

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

*Library of Congress Cataloguing in Publication data*

Endangered languages and new technologies / Edited by Mari C. Jones.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-107-04959-8 (hardback : alk. paper) – ISBN 978-1-107-62784-0

(paperback : alk. paper) 1. Language obsolescence – Technological innovations.

2. Language and languages – Technological innovations. I. Jones, Mari C., editor.

P40.5.L33E533 2015

417'.7 – dc23 2014032212

ISBN 978-1-107-04959-8 Hardback

ISBN 978-1-107-62784-0 Paperback

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.

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

## Contents

---

<i>List of figures</i>	page vii
<i>List of tables</i>	viii
<i>List of maps</i>	ix
<i>List of contributors</i>	x
<i>Preface</i>	xiii
MARI C. JONES	
Introduction: Endangered languages in the New Multilingual Order <i>per genus et differentiam</i>	1
NICHOLAS OSTLER	
<b>Part 1 Creating new technologies for endangered languages</b>	
1 The Kiranti comparable corpus: A prototype corpus for the comparison of Kiranti languages and mythology	17
AIMÉE LAHAUSSOIS	
2 European Dialect Syntax: Towards an infrastructure for documentation and research of endangered dialects	35
SJEF BARBIERS	
3 Keyboard layouts: Lessons from the Me'phaa and Sochiapam Chinantec designs	49
HUGH PATERSON III	
4 Rule-based machine translation for Aymara	67
MATT COLER AND PETR HOMOLA	
5 Data management and analysis for endangered languages	81
DOROTHEE BEERMANN	
6 Endangered languages, technology and learning: Immediate applications and long-term considerations	95
RUSSELL HUGO	

v

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

vi Contents

**Part 2 Applying new technologies to endangered languages**

- |    |  |     |
|----|--|-----|
| 7  | Digital curation and event-driven methods at the service of endangered languages<br>BERNARD BEL AND MÉDÉRIC GASQUET-CYRUS  | 113 |
| 8  | ‘Allant contre vent et mathée’: Jèrriais in the twenty-first century<br>ANTHONY SCOTT WARREN AND GERAINT JENNINGS  | 127 |
| 9  | The use of new technologies in the preservation of an endangered language: The case of Frisian<br>TJEERD DE GRAAF, COR VAN DER MEER AND<br>LYSBETH JONGBLOED-FABER | 141 |
| 10 | Language description and documentation from the native speaker’s point of view: The case of the Tundra Yukaghir<br>CECILIA ODÉ                                     | 150 |
| 11 | American Indian Sign Language: Documentary linguistic methodologies and technologies<br>JEFFREY E. DAVIS   | 161 |
|    | <i>Bibliography</i>  | 179 |
|    | <i>Index</i>   | 202 |

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

## Figures

---

0.1	Content languages for websites	page 8
0.2	Percentage of internet users by language	9
1.1	The Integral Text View	28
1.2	The Similarity View	29
1.3	Concordance results for the English term ‘to die’	30
3.1	First stroke of a two-keystroke character, highlighting the diacritic in OSX	56
3.2	Visual feedback loop	57
3.3	Conceptualization of a graphical element	59
4.1	A comparison of tree structures in English and Aymara	70
4.2	F-structure and dependency tree for <i>naja-w aka ut uta-ča-ya:t-xa</i>	75
4.3	F-structure and dependency tree for <i>huma-n-x hiw-i-w kimsa č'ijar p<sup>h</sup>isi-ma-xa</i>	76
4.4	Constituent structure for <i>naja-x kułaka-ma-r uñh-t-wa</i>	77
4.5	F-structure for <i>naja-x kułaka-ma-r uñh-t-wa</i>	77
5.1	Interlinear Glossed Text seen from within the TypeCraft Editor	85
5.2	The TypeCraft data model	87
5.3	External and internal convertibility of linguistic data	89
5.4	A XLFG5 c-structure parse of the Akan sentence: <i>Papa no kyereε mmɔfra no kwa</i> ‘the man showed the children the way’	92
5.5	HPSG syntactic and semantic parsing from IGT data	93
6.1	Example work flow diagram with a central coordinator	108
9.1	The Dictionary Room at the Frisian Academy	146
9.2	The original word cards	147
9.3	The Dictionaries of the Frisian Language	147
9.4	The online version of the Dictionary of the Frisian Language	147
10.1	A TY classroom	153
10.2	Alayi	158
11.1	Sustainability model for AISL documentation and revitalization	172

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

## Tables

---

0.1 An overview of language technologies	<i>page</i> 10
3.1 Options available to encode ñ in the Me'phaa and Chinantec layouts	60
3.2 Options available to encode á in the Me'phaa and Chinantec layouts	61
3.3 The conceptual construction of ⟨ 𐄀 ⟩ in Unicode	61
4.1 Information used in DTs	72
4.2 C-structures, f-structures and DTs in a parallel Aymara-English corpus of around 200 sentences	73
5.1 Tool-supported linguistic work flow, output, and impact	88
6.1 A summary of the fundamental questions for endangered language learning technology projects	104



Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

## Maps

---

0.1 GDP – growth or recession 2007	<i>page</i> 4
0.2 Official zones for the world’s major languages in 2010	7
1.1 Map of the Kiranti area	20
6.1 Map of Washington State as a distribution model example	107
10.1 The Tundra Yukaghir speech community	152
11.1 Great Plains cultural and geographic area	163

Cambridge University Press

978-1-107-04959-8 - Endangered Languages and New Technologies

Edited by Mari C. Jones

Excerpt

[More information](#)

---

## Contributors

---

MARI C. JONES

Reader in French Linguistics and Language Change, Peterhouse and  
Department of French, University of Cambridge, United Kingdom

SJEF BARBIERS

Senior researcher, Meertens Institute and Professor of Dutch Variation  
Linguistics, the University of Utrecht, the Netherlands

DOROTHEE BEERMANN

Professor, Norwegian University of Science and  
Technology, Norway

BERNARD BEL

Ingénieur de recherche, University of Aix-Marseille, France

MATT COLER

Researcher, INCAS<sup>3</sup>, the Netherlands

JEFFREY E. DAVIS

Professor, Educational Interpreter Programme, University of  
Tennessee, USA

MÉDÉRIC GASQUET-CYRUS

Lecturer, University of Aix-Marseille, France

TJEERD DE GRAAF

Research Associate, Fryske Akademy, the Netherlands

PETR HOMOLA

Researcher, Codesign, s.r.o, Czech Republic

RUSSELL HUGO

Researcher, University of Washington, USA

GERAINT JENNINGS

Assistant Language Support Officer, L'Office du Jèrriais, Jersey,  
Channel Islands

x