

# Africa's Endangered Languages

Documentary and  
Theoretical Approaches



EDITED BY

JASON KANDYBOWICZ

HAROLD TORRENCE

# Africa's Endangered Languages



# Africa's Endangered Languages

DOCUMENTARY AND THEORETICAL APPROACHES

Edited by Jason Kandybowicz  
and  
Harold Torrence

**OXFORD**  
UNIVERSITY PRESS

**OXFORD**  
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America.

© Oxford University Press 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data

Names: Kandybowicz, Jason, editor. | Torrence, Harold, editor.

Title: Africa's endangered languages : documentary and theoretical approaches/  
edited by Jason Kandybowicz and Harold Torrence.

Description: New York, NY : Oxford University Press, [2017] |  
Includes bibliographical references and index.

Identifiers: LCCN 2016042230 (print) | LCCN 2017008543 (ebook) |  
ISBN 9780190256340 (hardback : alk. paper) | ISBN 9780190256357 (pdf) |  
ISBN 9780190256364 (online course)

Subjects: LCSH: Endangered languages—Africa. | Language  
obsolescence—Africa. | Linguistic change—Africa. | Africa—Languages  
Classification: LCC P40.5.E532 A37 2017 (print) | LCC P40.5.E532 (ebook) |  
DDC 496—dc23

LC record available at <https://lcn.loc.gov/2016042230>

1 3 5 7 9 8 6 4 2

Printed by Sheridan Books, Inc., United States of America

We dedicate this volume to the memory and legacy of Russ Schuh (1941–2016),  
a mentor, a colleague, and an inspiration.



## CONTENTS

|  |            |
|--|------------|
| List of Contributors   | ix         |
| <b>1. Africa's Endangered Languages: An Overview</b>   | <b>1</b>   |
| Jason Kandybowicz and Harold Torrence  |            |
| <b>2. The Challenge of Documenting Africa's Least-Known Languages</b>  | <b>11</b>  |
| Bonny Sands  |            |
| <b>3. The Nata Documentation Project: An Overview</b>  | <b>39</b>  |
| Joash J. Gambarage, Andrei Anghelescu, Strang Burton, Joel Dunham,<br>Erin Guntly, Hermann Keupdjio, Zoe Wai-Man Lam, Adriana Osa-Gomez,<br>Douglas Pulleyblank, Dayanqi Si, Yoshiko Yoshino,<br>and Rose-Marie Déchaine |            |
| <b>4. Tongue Root Harmony in Nata: An Allomorphy-based Account</b>   | <b>57</b>  |
| Joash J. Gambarage and Douglas Pulleyblank   |            |
| <b>5. Nominal and Verbal Tone in Nata: An Allomorphy-based Account</b>   | <b>87</b>  |
| Andrei Anghelescu, Joash J. Gambarage, Zoe Wai-Man Lam,<br>and Douglas Pulleyblank   |            |
| <b>6. Nata Deverbal Nominalizations</b>  | <b>125</b> |
| Rose-Marie Déchaine, Dayanqi Si, and Joash J. Gambarage  |            |
| <b>7. Busy Intersections: A Framework for Revitalization</b>   | <b>145</b> |
| G. Tucker Childs   |            |
| <b>8. Documenting Ekegusii: How Empowering Research Fulfills Community<br/>and Academic Goals</b>  | <b>165</b> |
| Carlos M Nash  |            |
| <b>9. The Role of Theory in Documentation: Intervention Effects and<br/>Missing Gaps in the Krachi Documentary Record</b>  | <b>187</b> |
| Jason Kandybowicz and Harold Torrence  |            |
| <b>10. Documenting Raising and Control in Moro</b>   | <b>207</b> |
| Peter Jenks and Sharon Rose  |            |
| <b>11. The Linker in the Khoisan Languages</b>   | <b>237</b> |
| Chris Collins  |            |

|  |            |
|--|------------|
| <b>12. Theory and Description: Understanding the Syntax of Eegimaa Verb Stem Morphology</b>                  | <b>267</b> |
| Mamadou Bassene and Ken Safir  |            |
| <b>13. On (Ir)realis in Seenku (Mande, Burkina Faso)</b>   | <b>297</b> |
| Laura McPherson  |            |
| <b>14. Contributions of Micro-comparative Research to Language Documentation: Two Bantu Case Studies</b>     | <b>321</b> |
| Michael R. Marlo   |            |
| <b>15. Sebirwa in Contact with Setswana: A Natural Experiment in Learning an Unnatural Alternation</b>       | <b>343</b> |
| Elizabeth C. Zsiga and One Tlale Boyer   |            |
| <b>16. Three Analyses of Underlying Plosives in Caning, a Nilo-Saharan Language of Sudan</b>                 | <b>367</b> |
| Timothy M. Stirtz  |            |
| <b>17. Exceptions to Hiatus Resolution in Mushunguli (Somali Chizigula)</b>                                  | <b>411</b> |
| Katherine Hout   |            |
| <b>18. Acoustic and Aerodynamic Data on Somali Chizigula Stops</b>   | <b>427</b> |
| Michal Temkin Martinez and Vanessa Rosenbaum   |            |
| <b>19. Unmasking the Bantu Orthographic Vowels: The Challenge for Language Documentation and Description</b> | <b>449</b> |
| Joash J. Gambarage   |            |
| General Index  | 485        |
| Author Index   | 493        |
| Language Index   | 499        |

## LIST OF CONTRIBUTORS

**Andrei Anghelescu** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Mamadou Bassene** is an independent researcher.

**One Tiale Boyer** is an independent researcher.

**Strang Burton** is Instructor in the Department of Linguistics at the University of British Columbia.

**G. Tucker Childs** is Professor of Linguistics in the Department of Applied Linguistics at Portland State University.

**Chris Collins** is Professor of Linguistics in the Department of Linguistics at New York University.

**Rose-Marie Déchaine** is Associate Professor of Linguistics in the Department of Linguistics at the University of British Columbia.

**Joel Dunham** is a postdoctoral fellow in the Department of Classics, Modern Languages and Linguistics at Concordia University.

**Joash J. Gambarage** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Erin Guntly** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Katherine Hout** is a Ph.D. student in the Department of Linguistics at the University of California, San Diego.

**Peter Jenks** is Assistant Professor of Linguistics in the Department of Linguistics at the University of California, Berkeley.

**Jason Kandybowicz** is Associate Professor of Linguistics in the Linguistics Program at The Graduate Center, City University of New York.

**Hermann Keupdjio** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Zoe Wai-Man Lam** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Michael R. Marlo** is Associate Professor of Linguistics in the Department of English at the University of Missouri.

**Michal Temkin Martinez** is Associate Professor of Linguistics in the Department of English at Boise State University.

**Laura McPherson** is Assistant Professor of Linguistics in the Program in Linguistics at Dartmouth College.

**Carlos M Nash** is Lecturer of Linguistics in the Department of Linguistics at the University of California, Santa Barbara.

**Adriana Osa-Gomez** is a Ph.D. student in the Department of Linguistics at the University of British Columbia.

**Douglas Pulleyblank** is Professor of Linguistics in the Department of Linguistics at the University of British Columbia.

**Sharon Rose** is Professor of Linguistics in the Department of Linguistics at the University of California, San Diego.

**Vanessa Rosenbaum** was an undergraduate linguistics student at Boise State University and now works in the private sector.

**Ken Safir** is Professor of Linguistics in the Department of Linguistics at Rutgers University.

**Bonny Sands** is Adjunct Professor in the Department of English at Northern Arizona University.

**Dayanqi Si** is a member of the Faculty of Information and Media Studies at the University of Western Ontario.

**Timothy M. Stirtz** is Linguistics Coordinator and Linguistics Consultant at SIL South Sudan.

**Harold Torrence** is Associate Professor of Linguistics in the Department of Linguistics at the University of California, Los Angeles.

**Yoshiko Yoshino** was an undergraduate linguistics student at the University of British Columbia and now works in the private sector.

**Elizabeth C. Zsiga** is Professor of Linguistics in the Department of Linguistics at Georgetown University.

# Africa's Endangered Languages



# Africa's Endangered Languages

## AN OVERVIEW

Jason Kandybowicz and Harold Torrence

### 1.1. Language endangerment in Africa

Perhaps one of the most disturbing trends of our time is the accelerating rate of language extinction and endangerment.\* Most of today's languages are struggling to survive, clinging to life in a world of diminishing linguistic diversity. The phenomenon is not relegated to the planet's most remote linguistic outposts. Wherever we find languages, we find language endangerment.

The African continent hosts roughly one-third of the world's approximately 7,000 living languages. We might expect, therefore, to find a rich deposit of endangered languages within its borders. But we would be wrong, according to some. Ever since the late 1980s and early 1990s, the overall threat level of language endangerment in sub-Saharan Africa has been characterized and widely accepted as "low" (Sommer 1992; Brenzinger 1992, 1998; Wurm 1996; Anderson & Harrison 2006), owing in all likelihood to misleadingly high population counts. In addition, it has been claimed that the rate of language endangerment is significantly lower in sub-Saharan Africa than in other parts of the world (Simons & Lewis 2013) owing to diverse factors such as extensive multilingualism (Childs et al. 2014), urbanization (Simons & Lewis 2013), and the effects of settlement colonization versus exploitation colonization on language ecologies (Mufwene 2002). Consequently, research on Africa's endangered languages has lagged behind endangered language research in other parts of the world.

Nonetheless, there have been some attempts to document the endangered languages of Africa, to ascertain their relative threat levels, and to catalogue the continent's extinct languages. Among them are the Summer Institute of Linguistics' (SIL)

*Ethnologue* project (Lewis et al. 2015), the School of Oriental and African Studies' (SOAS) Endangered Languages Documentation Program, Sommer's (1992) survey, the collection of articles in Brenzinger (1992, 1998), works such as Wurm (1996), Haarmann (2001), and Batibo (2005), and the collection of Africa-specific articles in Brenzinger (2007a), to name a few.<sup>1</sup> Although they vary in the details, all such projects paint a consistent picture with respect to the state of language endangerment in sub-Saharan Africa. They note that threatened languages or families are distributed across several geographically distinct regions of the continent, and that internal pressures (e.g., regionally dominant languages and large-scale population movements) rather than external factors (e.g., the influence of the languages of former colonizers) drive and unify the pattern of African language endangerment. As such, the state of language endangerment in sub-Saharan Africa is distinct from that in much of the rest of the world.

Unlike Australia, northern Asia, and the Americas, where local languages are threatened and replaced by the nationally dominant languages of colonizers, the most immediate threats to minority African languages are posed by other local or sub-national languages (Brenzinger et al. 1991), barring infrequent and exceptional cases like the threat posed by English in certain regions of Nigeria (Connell 2015) and by national languages like Swahili in Tanzania and Setswana in Botswana (Brenzinger 2007b). Scotton (1982), for instance, concludes that less than 10% of rural Africa has competence in an imported European language, and Traill (1995) notes that the only documented instance of an African speech community abandoning its heritage language for the language of its former colonizers comes from the Khoekoe shift to Dutch around 1700. Brenzinger (2007b) identifies mass migration and cyclic immigration as a second unique internal pressure driving the pattern of language endangerment in Africa, predominantly in eastern Africa. The bottom line is that external threats like colonization have not threatened African minority languages in the way they have in most parts of the world (Grenoble & Whaley 1998). Rather, internal pressures almost exclusively characterize the state of language endangerment on the continent. Thus, because its pattern of endangerment is unique, Africa represents a fertile landscape with great potential to provide fresh perspectives on and valuable new insights into the causes, consequences, and characteristics of human language endangerment.

## 1.2. Researching Africa's endangered languages

Despite their great potential to fill gaps in our understanding of the inner workings of language endangerment, Africa's endangered languages pose several unique challenges to documentation and revitalization efforts.<sup>2</sup> For instance, insufficient infrastructure, scarcity of resources, incomplete and/or inaccurate information,<sup>3</sup> and a general absence of public awareness (both locally and internationally) represent serious hurdles for the documentation and maintenance

of threatened languages and cultures in most regions (Blench 2007; Connell 2007). In addition, because the general threat level of language endangerment in Africa has been perennially characterized as low, research on and funding opportunities for endangered African languages are often perceived as less than urgent (Connell 2007). However, as Brenzinger (2007a:195) puts it, “The fact that African languages replace other African languages does not mean that language endangerment on the African continent is less severe than in other parts of the world.” According to Batibo (2005), there are approximately 308 “highly endangered” African languages (12.32% of all extant languages on the continent) and at least 201 extinct African languages, to say nothing of the countless other less threatened but nonetheless vulnerable ones. This highlights the fact that the state of language endangerment in sub-Saharan Africa is more serious than it is typically perceived to be.

As a consequence of the challenges to documentation and revitalization outlined above, as well as the perceived non-urgent threat level priority status of Africa’s endangered languages, a disproportionately low amount of research and funding is devoted to the study of these languages when compared to any other linguistically threatened region in the world (Blench 2007). More regrettably, even less has been done to create a community of Africanists and concerned linguists who might work on these issues.<sup>4</sup> This volume seeks to stimulate and enhance the visibility of endangered African language research, in the hope of reversing this trend and bringing the unique insights and perspectives afforded by African language endangerment to bear on the burgeoning fields of language documentation and endangered language research. By highlighting the contribution that Africa’s endangered languages have on our understanding of natural language, we advertise their value and increase their visibility in the scientific community.

In today’s climate of economic instability and ongoing regional conflict in many parts of Africa, stimulating and enhancing the visibility of endangered African language research is indeed a formidable task. We believe that one fruitful way to achieve this goal is to promote mutually beneficial synergistic partnerships between documentary and theoretical linguists researching endangered African languages.

### **1.3. The symbiotic nature of theory and documentation**

Despite the fact that practitioners of language documentation and linguistic theory are often perceived as opposing or getting in the way of one another, the symbiotic nature of the two disciplines has been widely recognized (e.g., Everett 2001; Gil 2001; Hyman 2004, 2009; Mithun 2001; Rice 2001; Sells 2010, among others). Linguistic theory informed by marginalized or under-represented languages crucially draws on data unearthed by language documentation and could not proceed without it, while theory in turn guides the documentation process by predetermining the issues investigated, the questions asked, and the data sought (Hyman 2009).

In some cases, awareness of and sensitivity to theoretical concerns can even reveal missing gaps in the documentary record, for instance with research on logophoric pronouns following the seminal work of Clements (1975). The two disciplines, therefore, form a kind of “cycle” which drives linguistics forward. This cyclic interplay suggests that, at the very least, linguistic theory and language documentation are interdependent. Some researchers, though, take an even stronger position, claiming that the line between theory and documentation is a blurred one. Hyman (2004), for instance, argues that description and documentation are essentially indistinguishable from theory. When each is done right, they not only have the same concerns, they have the same results: each mode of inquiry is a vehicle of discovery. Matthewson's semantic fieldwork methodology (Matthewson 2004; Bochnak & Matthewson 2015) and Bruening's (2008a,b) quantifier scope materials illustrate the point nicely, demonstrating that theoretically oriented research can not only yield novel descriptive discoveries but also effectively drive the development of data-collecting techniques for both linguistic theory and language documentation. If theory and documentation are indeed interconnected and complementary, then partnerships between documentarians and theorists or projects that marry theory with documentation are destined to be synergistic affairs. And synergistic affairs are likely to be more visible and impactful than non-synergistic ones.

The keyword in the subtitle of this volume is therefore not “documentary” or “theoretical,” but rather the word “and.” Our aim in this book is to bring together both documentary and theoretical approaches to endangered African language research in order to highlight the respects in which the two methodologies are co-informing, mutually supportive, and equally essential to documentation and preservation efforts. We believe that doing so will not only encourage increased partnerships between these two types of linguists and consequently bolster the net output of research on endangered African languages, but it will also greatly improve the visibility, depth, breadth, and overall quality of that research.

#### **1.4. A brief tour through this book**

Many of the themes introduced in this chapter are taken up in greater detail in chapter 2, where Sands discusses the challenges of documenting Africa's least-known languages and concludes that the level of language endangerment in Africa has been grossly underestimated.

The next four chapters deal with the documentation and theoretical analysis of Nata, an endangered Bantu language of northern Tanzania, by a team of researchers at the University of British Columbia. Existing descriptions and analyses of Nata in the literature are scarce, making the contributions in this book some of the first published materials on the language. Chapter 3 provides an overview of both the language and the Nata research project, outlining the broader issues connected to the interplay

between language documentation and linguistic theory that unify the three subsequent articles. In these three chapters, both theory and documentation converge on a robust partition between nouns and verbs in the language. In chapter 4, Gambarage and Pulleyblank treat this partition by way of an investigation into tongue root vowel harmony that depends crucially on the iterative cycle connecting language documentation, language analysis, and theory development. Anghelescu and colleagues examine nominal and verbal tone in Nata in chapter 5, while Déchaine and colleagues document and analyze deverbal nominalization in chapter 6.

The two chapters that follow are concerned with community-based approaches to African language documentation and revitalization. In chapter 7, Childs discusses two pedagogical frameworks for language revitalization and, on the basis of a case study of Mani, an endangered Atlantic language of Sierra Leone, concludes that the so-called community-centric “busy intersections” model is best suited for success in the African context. In chapter 8, Nash explores a conflict that can arise between the needs of the community and the goals of the researcher in language documentation projects that have both documentary and theoretical aims. Drawing on his experience working with the Ekegusii community of southwestern Kenya, Nash advocates for collaborative community-based documentary research, arguing that it is a pursuit in which community and academic goals are both complementary and mutually beneficial.

Chapters 9 through 13 focus on the interplay between the documentation and theoretical analysis of syntax and morphology in endangered African languages. In chapter 9, Kandybowicz and Torrence investigate intervention effects on in-situ interrogative constructions in Krachi, an endangered Kwa language of eastern Ghana. The significance of the project is that it represents an instance in which the influence of linguistic theory on descriptive fieldwork leads to the discovery (and subsequent remedy) of missing gaps in the documentary record of a language. Jenks and Rose explore a similar theme in chapter 10. Focusing on the documentation, analysis, and theoretical implications of raising and control in the endangered Kordofanian language Moro of the Republic of Sudan, they argue that fieldwork guided by linguistic theory yields insights that would be difficult to establish solely on the basis of the documentary practice of text collection and analysis. They conclude that the documentation of endangered languages is most effective when it has a solid foundation in linguistic theory. Collins echoes this sentiment in chapter 11, on the syntax of the “linker” in five critically endangered non-central Khoisan languages of southern Africa. Collins reveals a number of new and fascinating properties of linkers, particles that introduce or “link” a wide range of expressions in the verb phrase. Because it is highly unlikely that a purely documentary-based approach would have produced similar results, Collins makes a strong case for the ability of formal/theoretical linguists to produce high-quality descriptive work. Bassene and Safir’s contribution (chapter 12) makes a strong case for this as well. In their analysis of verb stem structure in Eegimaa, an endangered

Atlantic language of southern Senegal, Bassene and Safir demonstrate that a set of theoretical challenges posed by Eegimaa morphology led to a series of analytical links which allowed the researchers to go beyond mere descriptions of facts and uncover deeper underlying organizational principles. The value in such work is that this deeper level of understanding can lead future Eegimaa researchers to discover (and fill) hitherto unknown gaps in the documentation of the language's grammar. This is an excellent example of the cyclic interplay of documentation and linguistic theory previously discussed, and it is the focus of McPherson's contribution on Seenku verbal morphology in chapter 13. In her article, McPherson presents the first published description of verbal morphology in Seenku, a threatened and previously undescribed Dogon language of Burkina Faso. McPherson explains how the cyclic and symbiotic interplay of linguistic theory and documentation led to a deeper account of the puzzling nature of verb forms in the language, which surprisingly appear to have two stem forms.

The final six chapters address issues concerning the phonology and phonetics of endangered African languages. In chapter 14, Marlo discusses the symbiotic relationship between linguistic description and micro-comparative typological research. His discussion proceeds by way of two case studies on tone and reduplication in the object-marking systems of Yao and Buguumbé Kuria, two Bantu languages spoken in the Tanzania-Malawi-Mozambique region. Marlo shows that in each case, knowing about analogous patterns in other languages informs the description and analysis of the individual language. Furthermore, each case expands knowledge of the typology of object-marking patterns in Bantu languages, leading to improvements in the quality of descriptions of other languages. In this way, he argues, theoretical approaches (broadly construed) can improve grammatical description. Zsiga and Boyer's contribution in chapter 15 treads similar ground by approaching the problem of the "unnatural" alternation of post-nasal devoicing in Sebirwa, an endangered Bantu language of Botswana, from the vantage point of a similar "unnatural" alternation in the related (yet phonologically distinct) language Setswana. Once again, knowing about analogous patterns in other languages informs the description and analysis of the individual language. In chapter 16, Stirtz discusses the phonology of plosives in Caning, an endangered Nilo-Saharan language of the Republic of Sudan with a four-way plosive series. He examines three analyses of the plosive system in great detail, but concludes that additional documentation is needed to furnish the missing decisive data. This conclusion resonates strongly with the theme of this volume, for it underscores the symbiotic interplay between language documentation and linguistic theory and analysis. In this case, theoretical concerns (which were themselves a product of documentary efforts) will play a catalyzing role in the future broadening of the Caning documentary record. Chapters 17 and 18 investigate properties of the Somali Chizigula (Mushunguli) sound system. Chapter 17 deals with hiatus resolution and its exceptions in the language. Hout provides another striking example of how a project with humble and purely descriptive origins can feed theoretical/analytical inquiry, which in turn

gives way to deeper and more refined characterizations of the data. Hout's article thus embodies the cyclic interplay and symbiosis between language documentation and linguistic theory that lies at the heart of this volume. In chapter 18, Temkin Martinez and Rosenbaum examine the acoustic and aerodynamic properties of Chizigula stops in an effort to complement the description and documentation of stops in the language. Although traditional language descriptions and revitalization efforts have benefited from instrumental approaches utilized in other fields of linguistics, the techniques of instrumental phonetics are infrequently applied to the documentation and analysis of Africa's endangered languages. Temkin Martinez and Rosenbaum's work thus joins a select and highly welcome body of research that helps sharpen the description and documentation of Africa's least studied languages. Chapter 19 closes the book with a critical look at the relationship between orthography and language documentation, as informed by the endangered Bantu languages Nata and Ikoma, as well as Swahili. Gambarage argues that orthographies are "masks" that disguise and often misrepresent the true phonetic qualities of vowels. He discusses current vowel documentation methodologies and theoretical approaches in the context of Bantu, arguing that revisiting the orthographic analyses of the languages that preceded both modern linguistic theory and speech analysis is essential to the documentation and description of endangered Bantu languages. Because unmasking is ultimately a theoretical/analytical endeavor, we are once again face to face with an instance in which the interplay between linguistic theory and language documentation leads to synergistic results in the study of Africa's endangered languages.

## Notes

\* The present collection of articles grew out of the workshop *Africa's Endangered Languages: Documentary and Theoretical Approaches*, which took place at the University of Kansas on April 17–19, 2014, in conjunction with the 45th Annual Conference on African Linguistics. We gratefully acknowledge the financial support of the National Science Foundation (NSF-DEL grant 1360823) for making the workshop possible. Thanks also to the University of Kansas Department of Linguistics for providing logistical support. We would also like to thank our wonderful editors, Hallie Stebbins and Hannah Doyle for helpful feedback, support, and guidance along the way. Finally, we thank the following individuals who served as reviewers for the articles submitted to this volume: Mark Baker, Herman Batibo, Lee Bickmore, Robert Botne, Matthias Brenzinger, Leston Buell, Michael Cahill, Roderic Casali, Anderson Chebanne, Bruce Connell, Laura Downing, James Essegbey, Colleen Fitzgerald, Carol Genetti, Jeff Good, Christopher Green, Heidi Harley, K. David Harrison, John Haviland, Brent Henderson, Larry Hyman, Peter Jenks, Allard Jongman, Raimund Kastenholz, Michael Kenstowicz, Ruth Kramer, Nancy Kula, Fiona McLaughlin, Amanda Miller, Scott Myers, David Odden, Mary Paster, Gérard Philippon, Keren Rice, Sharon Rose, Bonny Sands, Russell Schuh, Anne Storch, Mauro Tosco, Susi Wurmbrand, and Jochen Zeller.

1. See Sands (this volume, chap. 2) for other notable projects.

2. See Sands (this volume, chap. 2) for an in-depth discussion of these unique challenges.

3. At the time of writing, Wikipedia's list of endangered languages in Africa ([https://en.wikipedia.org/wiki/List\\_of\\_endangered\\_languages\\_in\\_Africa](https://en.wikipedia.org/wiki/List_of_endangered_languages_in_Africa)) is severely underpopulated, containing a mere 210 sub-Saharan languages, whose threat levels range from "vulnerable" to "critically endangered." By comparison, the figure reported in *Ethnologue* (Lewis et al. 2013) is 346 ("at risk" languages) and Sands (this volume) estimates that a more accurate figure is closer to 600. Many of the languages featured in this volume do not appear on Wikipedia's list, highlighting the dearth of accurate information publicly available about Africa's endangered languages.

4. Essegbey et al. (2015) represents a recent exception and, we believe, a step in the right direction.

## References

- ANDERSON, GREGORY, & K. DAVID HARRISON. 2006. Language hotspots: Linking language extinction, biodiversity and the human knowledge base. *Living Tongues Institute for Endangered Languages Occasional Papers Series 1*. Online: [http://www.livingtongues.org/docs/Hotspots\\_whitepaper%20copy.pdf](http://www.livingtongues.org/docs/Hotspots_whitepaper%20copy.pdf). Salem, OR: Living Tongues Institute for Endangered Languages.
- BATIBO, HERMAN. 2005. *Language decline and death in Africa: Causes, consequences and challenges*. Clevedon, UK: Multilingual Matters.
- BLENCH, ROGER. 2007. Endangered languages in West Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 140–162. Berlin: Mouton de Gruyter.
- BOCHNAK, M. RYAN, & LISA MATTHEWSON. 2015. *Methodologies in semantic fieldwork*. Oxford: Oxford University Press.
- BREZNINGER, MATTHIAS. 1992. *Language death: Factual and theoretical explorations with special reference to East Africa*. Berlin: Mouton de Gruyter.
- BREZNINGER, MATTHIAS. 1998. *Endangered languages in Africa*. Cologne: Rüdiger Köppe.
- BREZNINGER, MATTHIAS. 2007a. *Language diversity endangered*. Berlin: Mouton de Gruyter.
- BREZNINGER, MATTHIAS. 2007b. Language endangerment in southern and eastern Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 179–204. Berlin: Mouton de Gruyter.
- BREZNINGER, MATTHIAS, BERND HEINE, & GABRIELE SOMMER. 1991. Language death in Africa. *Diogenes* 153.19–44.
- BRUENING, BENJAMIN. 2008a. Quantification in Passamaquoddy. *Quantification: A cross-linguistic perspective*, ed. by Lisa Matthewson, 67–104. Bingley, UK: Emerald.
- BRUENING, BENJAMIN. 2008b. The Scope Fieldwork Project. Online: <http://udel.edu/~bruening/scopeproject/scopeproject.html>
- CHILDS, G. TUCKER, JEFF GOOD, & ALICE MITCHELL. 2014. Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation* 8.168–191.
- CLEMENTS, GEORGE. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10.141–177.
- CONNELL, BRUCE. 2007. Endangered languages in Central Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 163–178. Berlin: Mouton de Gruyter.

- CONNELL, BRUCE. 2015. The role of colonial languages in language endangerment in Africa. *Language documentation and endangerment in Africa*, ed. by James Essegbey, Brent Henderson, & Fiona McLaughlin, 107–130. Amsterdam: John Benjamins.
- ESSEGBEY, JAMES, BRENT HENDERSON, & FIONA MCLAUGHLIN. 2015. *Language documentation and endangerment in Africa*. Amsterdam: John Benjamins.
- EVERETT, DANIEL L. 2001. Monolingual field research. *Linguistic fieldwork*, ed. by Paul Newman & Martha Ratliff, 166–188. Cambridge, UK: Cambridge University Press.
- GIL, DAVID. 2001. Escaping Eurocentrism: Fieldwork as a process of unlearning. *Linguistic fieldwork*, ed. by Paul Newman & Martha Ratliff, 102–132. Cambridge, UK: Cambridge University Press.
- GRENOBLE, LENORE A., & LINDSAY J. WHALEY. 1998. Towards a typology of language endangerment. *Endangered languages: Current issues and future prospects*, ed. by Lenore A. Grenoble & Lindsay J. Whaley, 22–54. Cambridge, UK: Cambridge University Press.
- HAARMANN, HARALD. 2001. *Die Kleinsprachen der Welt-Existenzbedrohung und Überlebenschancen: Eine umfassende Dokumentation*. Frankfurt: Peter Lang.
- HYMAN, LARRY M. 2004. Why describe African languages? *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, ed. by Akinbiyi Akinlabi & Oluseye Adesola, 21–41. Cologne: Rüdiger Köppe.
- HYMAN, LARRY M. 2009. Good things come in small languages: Grammatical loss and innovation in Nzadi. *Proceedings of Conference on Language Documentation and Linguistic Theory 2*, ed. by Peter K. Austin et al., 3–11. London: School of Oriental and African Studies.
- LEWIS, M. PAUL, GARY F. SIMONS, & CHARLES D. FENNIG. 2013. *Ethnologue: Languages of the world*. 17th ed. Dallas, TX: SIL International.
- LEWIS, M. PAUL, GARY F. SIMONS, & CHARLES D. FENNIG. 2015. *Ethnologue: Languages of the world*. 18th ed. Dallas, TX: SIL International.
- MATTHEWSON, LISA. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70.369–415.
- MITHUN, MARIANNE. 2001. Who shapes the record: The speaker and the linguist. *Linguistic fieldwork*, ed. by Paul Newman & Martha Ratliff, 34–54. Cambridge, UK: Cambridge University Press.
- MUFWENE, SALIKOKO. 2002. Colonization, globalization and the plight of ‘weak’ languages. *Journal of Linguistics* 38.375–395.
- RICE, KEREN. 2001. Learning as one goes. *Linguistic fieldwork*, ed. by Paul Newman & Martha Ratliff, 230–249. Cambridge, UK: Cambridge University Press.
- SCOTTON, CAROL MYERS. 1982. Learning lingua francas and socioeconomic integration: Evidence from Africa. *Language spread: Studies in diffusion and social change*, ed. by Martha L. Cooper, 63–94. Bloomington: Indiana University Press.
- SELLS, PETER. 2010. Language documentation and linguistic theory. *Language documentation and description*, ed. by Peter K. Austin, vol. 7, 209–237. London: School of Oriental and African Studies.
- SIMONS, GARY F., & M. PAUL LEWIS. 2013. The world’s languages in crisis: A 20-year update. *Responses to language endangerment: In honor of Mickey Noonan: New directions in language documentation and language revitalization*, ed. by Elena Mihas et al., 3–20. Amsterdam: John Benjamins.

- SOMMER, GABRIELE. 1992. A survey on language death in Africa. *Language death: Factual and theoretical explorations, with special reference to East Africa*, ed. by Matthias Brenzinger, 301–417. Berlin: de Gruyter.
- TRAILL, ANTHONY. 1995. The Khoesan languages of South Africa. *Language and social history: Studies in South African sociolinguistics*, ed. by Rajend Mesthrie, 1–18. Cape Town: David Philip.
- WURM, STEPHEN A. 1996. *Atlas of the world's languages in danger of disappearing*. Paris: UNESCO.

# The Challenge of Documenting Africa's Least-Known Languages

Bonny Sands

## 2.1. Introduction

As Africanist linguists, a major challenge ahead of us is the documentation of endangered African languages.\* Prioritizing the languages most in need of documentation is far from simple, since language vitality can be difficult to measure, and criteria other than language vitality should be taken into account. *Ethnologue's* language vitality labels are one common measure of language endangerment, but these labels may not always reflect a thorough evaluation of a language's threat level. Languages spoken by a marginalized community or by a small population should also be considered potentially under threat. Isolate languages (including sign languages) and languages from families that are sparsely documented should be prioritized for documentation as well.

In documenting an African endangered language, researchers face many challenges. I will survey these challenges, and focus on a few that I think have not been sufficiently emphasized in the recent literature on linguistic fieldwork. I will draw examples from my own fieldwork on Hadza in Tanzania and †Hoan in Botswana, as well as from the experiences of others.<sup>1</sup>

The reasons for language shift in Africa are varied and have been discussed in a number of publications:

- |                         |   |
|-------------------------|---|
| Continent-wide surveys: | Batibo 2005, Brenzinger 1992, 1998a, 1998b, Brenzinger et al. 1991, Brenzinger & Batibo 2010, Dimmendaal & Voeltz 2007, Mous 2003, Sommer 1992, Tamanji 2008, Tourneux et al. 2000. |
| Central Africa:         | Anchimbe 2013, Connell 1998, 2007, Idiata 2009.   |

|                            |   |
|----------------------------|---|
| West Africa:               | Blench 2007a, Childs 2006, Haruna 2012.                                     |
| Northern Africa:           | Brenzinger 2007a, Mezhoud & El Allame 2010, Nyombe 1997, Savà & Tosco 2006. |
| Southern & Eastern Africa: | Brenzinger 2007b, Dimmendaal 1989, Savà & Tosco 2006, Zelealem Leyew 2004.  |

Case studies of language shift in different parts of Africa include Kédérébéogo (1996), Robson (2011), Sommer (1995), and Zelealem Leyew (2003).

## 2.2. Defining the challenge

### 2.2.1. THE MOST THREATENED LANGUAGES

Languages that are in urgent need of documentation include those that may not be spoken by future generations. One way to measure this is by using a vitality scale such as the Extended Graded Intergenerational Disruption Scale (EGIDS) (Simons & Lewis 2013)

The web edition of *Ethnologue* 17 (Lewis et al. 2013) uses EGIDS to label 346 African languages as “vital,” “in trouble,” or “dying,” as shown in Table 2.1. A language is counted as “in trouble” if it falls under 6b or 7 in the EGIDS—that is, it is known by the child-bearing generation but transmitted to only some or none of youngest generation.

Most African languages fall into one of the three EGIDS categories: 6a “vigorous,” 6b “threatened,” and 7 “shifting,” shown in Table 2.2.<sup>2</sup> “Vigorous” languages are those still being learned by children; “threatened” languages are passed on to only some children; and “shifting” languages are no longer passed on to children but are known by the child-bearing generation. A language is not considered “moribund” or “nearly extinct” until the only remaining speakers are of the grandparental generation.

Many languages labeled “vigorous” or “developing” actually show signs of language shift, particularly as populations migrate to urban centers (e.g.

TABLE 2.1. Numbers of African languages at different risk levels (Lewis et al. 2013).

| Region          | # of living languages | Vital       | In trouble | Dying      |
|-----------------|-----------------------|-------------|------------|------------|
| Middle Africa   | 677                   | 555         | 77         | 45         |
| Western Africa  | 889                   | 794         | 49         | 46         |
| Northern Africa | 97                    | 53          | 23         | 21         |
| Eastern Africa  | 433                   | 362         | 57         | 19         |
| Southern Africa | 50                    | 41          | 3          | 6          |
| total:          |                       | <b>1805</b> | <b>209</b> | <b>137</b> |

TABLE 2.2. Levels 6–7 of the Expanded GIDS (Simons &amp; Lewis 2013), adapted from Fishman (2001) (boldface added).

| GIDS Level | Label      | Description   | UNESCO                |
|------------|------------|---|-----------------------|
| 6a         | vigorous   | The language is used orally by all generations and is <b>being learned by children</b> as their first language.                               | safe                  |
| 6b         | threatened | The language is used orally by all generations but <b>only some</b> of the child-bearing generation are transmitting it to their children.    | vulnerable            |
| 7          | shifting   | The child-bearing generation knows the language well enough to use it among themselves but <b>none are transmitting it</b> to their children. | definitely endangered |

Mugaddam 2012). For instance, even a large (7 million speakers) developing language such as Gĩkũyũ is seeing language shift to Kiswahili (Orcutt-Gachiri 2013). If only some of the child-bearing generation are transmitting it to their children, then it would seem that Gĩkũyũ could be labeled “threatened.” One study of 800 Nigerians found that only 40% spoke to their children in their indigenous language (Ohiri-Aniche 2008), but another study saw only 18% retention of mother tongue among urban dwellers (Ndimele 2005). Major Nigerian languages such as Yoruba and Igbo have begun to show signs of shift and attrition (Adéníyí & Bèllò 2009; Ugorji 2005; Fabunmi & Salawu 2005; Fabunmi 2005). *Ethnologue* 18 (Lewis et al. 2015) now defines a “vigorous” language as one “used for face-to-face communication by all generations and the situation is sustainable.” It is unclear how the term “sustainable” is applied; if a language consistently loses speakers from one generation to the next, then its vitality level may not be sustainable over time.<sup>3</sup> The label “threatened” is not such a rare commodity that only the most endangered of languages can be referred to in this way. Indeed, even relatively “vigorous” languages that are losing speakers may benefit from language maintenance and revitalization efforts.

Of course, the level of threat faced by Gĩkũyũ is not as high as it would be for a language with a much smaller population, or for one losing a greater proportion of speakers due to shift. If only 10% of the children of a language group with a million speakers continue to use the language, that language would have 100,000 speakers in the youngest generation, and would therefore be less threatened than a language with 1,000 speakers, whether undergoing some degree of language shift or not. Though the absolute number of speakers and proportion of speakers within a community are considered by UNESCO to be factors that affect language vitality (UNESCO 2003), these criteria are not explicitly used in major publications such as UNESCO’s *Atlas of the world’s languages in danger* (Moseley 2010) or in *Ethnologue* 17. *Ethnologue* 18 (Lewis et al. 2015) does consider speaker population and ethnic population in evaluating language endangerment.

The EGIDS label “shifting” is problematic because it is used to label languages whose communities have already undergone the process of language shift. “Shifted” might be a better label. One language labeled “shifting” is †Hoan, a Kx’a language spoken in Botswana. There are middle-aged speakers of the language and no transmission to children, but the extremely small number of speakers of any age would lead me to consider it not just as “in trouble,” but as a severely endangered language. There are fewer than 50 speakers, and most are between 60 and 70 years old (Gerlach & Berthold 2011).

During my own fieldwork, it was difficult to assemble a group of †Hoan speakers in order to do a phonetic study because there is not an actual community, but rather a few individuals in scattered locations. There were significant differences in pronunciation between the older and younger generations of speakers, with the pronunciations of the older generation more closely reflected in previous work on the language (Gruber 1975). Had I not recorded a number of older and younger speakers at the same time, I might not have even recognized some words as being the same because the surface realizations were so different. For instance, while some younger speakers typically pronounced the word /ŋlɔ́ ʼ:ŋlɔ́lɔ́/ ‘ostrich feathers’ with barely audible clicks but with strong pharyngealization, some older speakers had strong clicks but barely audible pharyngealization. Some †Hoan speakers seem to have phonological inventories that are converging with the G|ui inventory (Gerlach 2015). Had researchers waited for †Hoan to be considered “dying” (i.e., spoken proficiently only by members of the grandparental generation) before prioritizing it for fieldwork, then much of the phonetic variability of the language would have vanished before it had been recorded.

†Hoan was formerly thought to be a dialect of a different language, !Xóǝ, labeled “vigorous” in *Ethnologue* 17. It was only following the targeted fieldwork on the language, notably by Jeff Gruber (e.g. Gruber 1973, 1975; Collins & Gruber 2014; cf. Traill 1973), that there were enough data on †Hoan that it was recognized as not only a distinct language from !Xóǝ, but a member of a different language family altogether (Westphal 1974; Heine & Honken 2010). In fact, †Hoan, along with Sasí, belongs to a cluster now known as †’Amkoe (Collins & Gruber 2014; Güldemann 2014), which together with Ju (!Xuun) constitute the Kx’a family. !Xóǝ, or !Xoon, is considered to belong to the Taa language family.

Some languages labeled “extinct” may actually still be in use. I have done fieldwork in South Africa with fluent speakers of N|uu (cf. Sands et al. 2007), a language that had been considered extinct for several decades (cf. Traill 1999). Birgid (*Ethnologue* code [brk]), a Nubian language of Sudan, was considered to be on the verge of extinction in 1974 (Bell 2006), but in 2003, 29 years later, two elders were found who could recall about 100 Birgid words (Edenmyr et al. 2011). Some words of a language may live on in the memory of people long after the language has ceased to be acquired as a mother tongue. For instance, Kwadi, a Khoe-Kwadi

language of Angola long considered dead (Winter 1981), may instead be considered “dormant.” As of 2014, it was still partially remembered by two people who had used it as a secret language when they were children (Anne-Maria Fehn, personal communication, March 7, 2014).

While *Ethnologue* identifies many threatened African languages, it does not accurately identify all of them. The EGIDS labels cannot be blindly used to determine the vitality of a language or to set language documentation priorities. An “extinct” language may be extant, or a “vigorous”/“developing” language may already have lost a significant number of speakers and shown signs of language obsolescence.

### 2.2.2. MARGINALIZED POPULATIONS

Some of the languages most under threat are those spoken by marginalized communities and/or by populations whose lifestyles are threatened. For instance, the G|ui (*Ethnologue* code [gwi]) and G||ana [gnk] languages of Botswana are listed as “vigorous” in *Ethnologue* even though their populations are estimated to be only 2,500 and 2,000, respectively, and the actual numbers may be quite a bit lower; Brenzinger (2013:19) cites figures of 1,470 |Gui and 1,030 G||ana, while Batibo (2001:315) gives population figures of 500 and 800. Some 1,000 G|ui and G||ana lived by hunting and gathering in the Central Kalahari Game Reserve (Tanaka & Sugawara 2010), but evictions from the park have threatened the traditional linguistic ecology in which the languages have been maintained. We are fortunate that a great deal of work has been done on G|ui by Hiroshi Nakagawa and others, and a dictionary of G|ui is in progress (Nakagawa et al. 2013), but other languages spoken by marginalized populations such as the #Haba, Shua and Cire-Cire have not been so well documented. We certainly cannot blindly follow the labels provided by *Ethnologue* 17 if a language spoken by a thousand hunter-gatherers is not considered “threatened.”

Naro [nhr] is one of the most vital of the Khoe languages, being spoken by some 14,000 speakers according to *Ethnologue* (Lewis et al. 2013), which lists it as a “language of wider communication,” reflecting the fact that many #Au||eisi use Naro in addition to their own Ju language. Ranking 3 on the GIDS scale, this language might not seem to be in as desperate need of documentation as more threatened languages, yet Naro, G|ui and other Remote Area Dwellers in Botswana lack equal access to education (Hays 2004; Sekere 2011) and political representation, and Naro is not what I would consider an entirely “safe” language.<sup>4</sup> Furthermore, although Naro is listed as having several dialects, none of these has been the subject of a linguistic documentation project. Although a few Ts’aokhwe words have been noted in the Naro dictionary (Visser 2001), the Ts’aokhwe reportedly “prefer to see themselves as separate people from the Naro” (Barnard 1985:2). Kango [kzy], a Bantu language of the Democratic Republic of Congo spoken by Pygmies, is listed as “vigorous” in *Ethnologue* 17—with perhaps only 2,000 speakers; very little is known about either dialect of this language.

Groups that tend to be marginalized and whose languages may therefore be under threat include hunter-gatherers, special castes (blacksmiths, leatherworkers, potters, etc.), refugee populations, and the Deaf. Marginalized groups may include African populations that have been targeted for genocide in the 20th and 21st centuries (Hitchcock & Koperski 2008:586), including some pastoralist groups. Some of these groups may speak marginalized dialects of otherwise “vigorous” languages.

An example of a language spoken by blacksmiths is Sombɔ [kgt] (Connell 2010), cited as having only three elderly speakers. The language of the Jeri leatherworkers of Ivory Coast is only spoken in a few villages (Kastenholz 1998). Kastenholz (1998:261) does not consider the Jɔɔ [lig] language of blacksmiths in Ghana to be endangered (population estimate, 19,000; Lewis et al. 2013), even though groups in Ivory Coast have already undergone language shift.

Because marginalized populations typically have unequal access to education, we generally see a lack of native-speaker linguists from these populations. Languages spoken by very small populations are often not targeted for Bible translation efforts, linguistic surveys of them may be lacking, and distinct dialects and languages may be unattested in the linguistic literature. Language development is typically lacking for marginalized groups as well, with the exception of some sign languages. State-funded or NGO-funded sign language development often comes at the expense of local sign languages, however, some of which are severely threatened (Brenzinger & Batibo 2010).

### 2.2.3. POPULATION SIZE

The size of a population that speaks a language may be one indicator of the language's vitality. Batibo (2005:69) suggests considering minority languages spoken by fewer than 5,000 people to be endangered. Because downward population pressures such as disease, climate change, and armed conflict may affect the number of a language's users suddenly (cf. Lüpke & Storch 2013), it would be prudent to consider any language with a population under 5,000 to be threatened. Given the often abrupt nature of language shift, it might be even more prudent to consider languages with populations below 100,000 to be threatened. Certainly, outside of Africa, languages with larger populations are classified as “threatened.” For instance, Navajo is cited in *Ethnologue* as having a population of 171,000 (2010 census) and is considered “threatened.” Studies of language shift are not available for many African language groups, yet we know that the trends of population movement, urbanization, education policies, and so on are in the direction of more, rather than less, language shift.

There is evidence that many of the languages labeled “vital” actually are threatened by language shift. For one thing, many supposedly vital languages have small populations. For another, languages labeled “vital” have actually appeared in the literature as examples of language endangerment, obsolescence, or marginalization. Examples include Qimant [ahg] (Zealelem Leyew 1998), Iko [iki] (Urui 2004),

Oko [oks] (Adegbija 2001) and Vute [vut] (Mutaka 2008). These languages are estimated to have populations of 1,650, 5,000, 10,000, and 21,000, respectively (Lewis et al. 2013, based on studies from 1994, 1988, 1989, and 1997).

In the following sections, I will look at the level of language threat for languages classed as “vital” by *Ethnologue* 17 in northern Africa, Tanzania, Cameroon, and Nigeria. Since population figures are not given for sign languages, they will not be discussed in the country surveys.

### 2.2.3.1. Northern Africa

The northern African region consists of Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, and Western Sahara. There are 53 northern African languages classed as “vital” or not threatened in *Ethnologue* 17. Of these languages labeled “vigorous,” “developing,” or “educational,” 30 are estimated to have fewer than 100,000 speakers, as shown in Table 2.3. *Ethnologue* does not give population figures for Algerian Sign Language [asp], Egypt Sign Language [esl], Libyan Sign Language [lbs], Tunisian Sign Language [tse], or Moroccan Sign Language [xms], but these almost certainly also have fewer than 100,000 users. Likely 35 of 53, or two-thirds of the supposedly “vital” languages of North Africa, have populations below 100,000. Tocho, Ganza and Tulishi, each with fewer than 5,000 speakers, should certainly be considered threatened.

TABLE 2.3. North African languages labeled as “educational” or “developing” by *Ethnologue* 17, with estimated populations less than 100,000 (Lewis et al. 2013).

| Vigorous   | Developing  | Educational  |
|--|---|--|
| Daju, Dar Fur [daj] 80,000 (1983) & Daju, Dar Sila [dau] 63,100 (2000) | Katcha-Kadugli-Miri [xtc] 75,000 (2004)                 | Tira [tic] 40,000 (1982)   |
| Gaam [tbi] 67,200 (2000)   | Ama [nyi] 70,000 (1982)                                 | Koalib [kib] 44,300 (1984)   |
| Midob [mei] 50,000 (1993)  | Uduk [udu] 22,000                                       | Laro [lro] 40,000 (1998)   |
| Senhaja Berber [sjs] 40,000 (2011)                                     | Krongo [kgo] 21,700 (1984)                              | Moro [mor] 30,000 (1982)   |
| Dagik [dec] & Ngile [jle] 38,000 (1982)                                | Gbaya [krs] 16,000 (1987)                               | Lumun [lmd] 30,000 (2012)<br>Home area population 19,000 (2011 census) |
| Tegali [ras] 35,700 (1984)   | Otoro [otr] 10,000 (2001)                               | Burun [bdi] 18,000 (1977)  |
| Siwi [siz] 30,000 (2006)   | Acheron [acz] 9,830 (2006)<br>Ethnic population: 20,000 |  |
| Jumjum [jum] 25,000 (1987)   |   |  |
| Komo [xom] 18,530  |   |  |
| Ghomara [gho] 10,000 (2008)  |   |  |
| Wali [wl] 9,000 (2007)   |   |  |
| Kanga [kcp] 8,000 (1989)   |   |  |
| Tocho [taz] 3,800 (1989)   |   |  |
| Ganza [gza] 3,000  |   |  |
| Tulishi [tey] 2,500 (2007)   |   |  |

Population estimates in *Ethnologue* vary widely in their reliability. The more reliable figures are those based on recent sociolinguistic surveys and/or census figures that indicate linguistic affiliation, but these are a minority. The population estimates for many languages are based on studies which are several decades old, and data on current patterns of language shift may be lacking for many languages. For instance, a recent study of Siwi (Naumann 2012) gives the population as approximately 15,000 speakers, or half as many as in the *Ethnologue* estimate. I think a substantial fraction of the 30 northern African languages listed as “vital” might more accurately be considered “shifting” or “threatened.”

### 2.2.3.2. Tanzania

There are 26 Tanzanian languages labeled “vigorous” or “developing” with fewer than 100,000 speakers (in all countries combined), shown in Table 2.4. These relatively small languages are in addition to those 39 Tanzanian languages listed as “in trouble” according to *Ethnologue* 17. This includes 26 “threatened” and seven “shifting” languages (not counting languages considered “vigorous” in another country). Tanzania is also listed as having five “moribund” languages (Gweno [gwe], Ngasa [nsg], Okiek [oki], Segeju [seg], Zaramo [zaj]), and one “nearly extinct” language (Aasáx [aas]).

Table 2.4 compares the population estimates from *Ethnologue* 17 and *Ethnologue* 13 (Grimes 1996). *Ethnologue* 17 did not have updated population estimates for nine languages, and two languages, Cutchi Swahili and Tanzanian Nyika, were not listed in *Ethnologue* 13. Changes in populations could therefore only be estimated for 15 of 26 languages (58%). Keeping in mind that the quality of population estimates cannot be controlled for and may vary greatly, we have evidence for population increases for only nine of the languages in Table 2.4, even though population growth in Tanzania might have led us to expect population increases for all “vigorous” or “developing” languages that are not losing speakers due to language shift. Six of the languages showed significant population decreases, even including two languages labeled “developing,” Datooga<sup>5</sup> [tcc] and Vwanji [wbi]. So, for the 15 languages with estimated population changes, 40% had significant population declines while 60% had population increases.

It is possible that some of these population changes are spurious, with better data replacing poorer data. This is certainly the case for Hadza, a “threatened” language isolate spoken in Tanzania. It was estimated as having 200 speakers in *Ethnologue* 13 and 800 in *Ethnologue* 17, but the earlier number was simply inaccurate (Blurton Jones et al. 1992).

In another study of Tanzanian languages, Legère (1992:103) cites Bende, Ikizu, and Pimbwe as groups which have had striking declines in population figures, strongly suggestive of language shift, yet Table 2.4 shows Bende and Ikizu as two of the languages with apparent increases in population. Since language shift can happen in a single generation, and is widespread in Tanzania owing to the influence of Kiswahili, the use of decades-old population data to consider a language

TABLE 2.4. 26 Tanzanian languages labeled “vigorous” or “developing” with population estimates 100,000 or fewer in *Ethnologue* 17 (Lewis et al. 2013).

| Language name & Ethnologue code | Status     | Population estimate & year of estimate (from Ethnologue 17 (Lewis et al. 2013)) | Population estimate & year of estimate (from Ethnologue 13 (Grimes 1996)) | Change in speaker population estimates | % Change     | Year of most recent grammar or grammar sketch |
|---------------------------------|------------|---|---|--|--------------|---|
| <b>Bende</b> [bdp]              | Vigorous   | 27,000 (1999)   | 20,000 (1987)   | +7,000                                 | +35%         | 2006  |
| Bungu [wun]                     | Vigorous   | 36,000 (1987)   | same survey   | —                                      | —            | —   |
| <b>Burunge</b> [bds]            | Vigorous   | 13,000 (2002)   | 31,000 (1987)   | -18,000                                | -58%         | 1994  |
| Cutchi-Swahili [ccl]            | Vigorous   | 45,000 (2002)   | n/a   | n/a                                    | n/a          | —   |
| Gorowa [gow]                    | Vigorous   | 50,000 (1999)   | 30,000 (1987)   | +20,000                                | +67%         | —   |
| <b>Ikizu</b> [ikz]              | Vigorous   | 55,000 (2005)   | 28,000 (1987)   | +27,000                                | +96%         | —   |
| Kerewe [ked]                    | Vigorous   | 100,000 (1987)  | same survey   | —                                      | —            | 1909 sketch <sup>6</sup>                      |
| Konongo [kcz]                   | Vigorous   | 51,000 (1987)   | same survey   | —                                      | —            | 1905 sketch                                   |
| Machinga [mvw]                  | Vigorous   | 36,000 (1987)   | same survey   | —                                      | —            | 1876 sketch                                   |
| Malila [mgq]                    | Vigorous   | 65,000 (2003)   | 52,000 (1987)   | +13,000                                | +25%         | —   |
| Manda [mgs]                     | Vigorous   | 22,000 (2002)   | 18,000 (1987)   | +4,000                                 | +22%         | —   |
| Matumbi [mgw]                   | Vigorous   | 72,000 (1978)   | same survey   | —                                      | —            | 1912 sketch                                   |
| Mpoto [mpa]                     | Vigorous   | 80,000 (1977)   | same survey   | —                                      | —            | —   |
| Ndendeule [dne]                 | Vigorous   | 100,000 (2000)  | 79,000 (1987)   | +21,000                                | +27%         | 1999 sketch                                   |
| Ngoreme [ngq]                   | Vigorous   | 55,000 (2005)   | 32,000 (1987)   | +23,000                                | +72%         | —   |
| Nyika [nkt]                     | Vigorous   | 25,000 (2007)   | n/a   | n/a                                    | n/a          | —   |
| <b>Pangwa</b> [pbr]             | Vigorous   | 95,000 (2002)   | 177,000 (1987)  | -82,000                                | -46%         | 1983  |
| <b>Pimbwe</b> [piw]             | Vigorous   | 29,000 (1987)   | same survey   | —                                      | —            | —   |
| Rwa [rwk]                       | Vigorous   | 90,000 (1987)   | same survey   | —                                      | —            | 2009 sketch                                   |
| <b>Sandawe</b> [sad]            | Vigorous   | 40,000 (2000)   | 70,000 (1987)   | -30,000                                | -43%         | 2012  |
| Temi [soz]                      | Vigorous   | 30,000 (2002)   | 20,000 (1995)   | +10,000                                | +50%         | 1991/92                                       |
| <b>Tongwe</b> [tny]             | Vigorous   | 13,000 (2001)   | 22,000 (1987)   | -9000                                  | -41%         | —   |
| Vinza [vin]                     | Vigorous   | 10,000 (1987)   | same survey   | —                                      | —            | —   |
| Zanaki [zak]                    | Vigorous   | 100,000 (2005)  | 62,000 (1987)   | +38,000                                | +61%         | —   |
| <b>Datooga</b> [tcc]            | Developing | 87,800 (2000)   | 150,000 to 200,000 (1993)   | -62,200 to -112,200                    | -41% to -56% | 1983 sketch                                   |
| <b>Vwanji</b> [wbi]             | Developing | 28,000 (2003)   | 60,000 (1987)   | -32,000                                | -53%         | —   |

“safe” is risky. It would be more prudent to consider a language to be at risk unless evidence proves otherwise.

It is important to point out that the majority of Tanzania's small “vigorous” languages are almost completely undocumented. As can be seen in the rightmost column of Table 2.4, fewer than one-third of these languages have been the subject of a grammar or grammar sketch in recent years (based on Maho & Sands 2002 and Nordhoff et al. 2014). Although we cannot know how endangered these languages are, we can state with certainty that their current level of documentation is pitiful.

Some of the larger, better-documented Tanzanian languages are themselves showing signs of language shift and attrition. For instance, younger speakers of Sukuma [suk], a language with more than 5 million speakers, know only a fraction of the vocabulary of flora and fauna known to older speakers (Batibo 2013). Special elicitation methods had to be used with speakers of Ngoni [ngo] in order to minimize the large number of Swahili code-switches (31.4%) produced by consultants (Rosendal & Mapunda 2014), and the language is regarded as “shifting” rather than “developing” in another study (Matiki 2009). Although particular lexical fields may undergo contraction owing to changes in lifestyles without necessarily entailing language shift, the loss of traditional vocabulary is a lost opportunity for language documentation.

### 2.2.3.3. Cameroon

Cameroon has many more small “vigorous” languages than does Tanzania. In order to have a table of manageable size for Cameroon, I will focus on very small languages. Table 2.5 shows the 36 languages of Cameroon labeled “vigorous,” “developing,” or “wider communication” with 5,000 or fewer speakers in *Ethnologue* 17 (Lewis et al. 2013). Three of these languages had population estimates only for Cameroon but might have more than speakers in other countries: Jukun Takum [jbu], Akum [aku], and Dzodinka [add]. In addition to these very small “vigorous” Cameroonian languages, there are also more than 60 Cameroonian languages “in trouble” or “dying.”

In Table 2.5, I do not include languages with fewer than 5,000 speakers in Cameroon if the language has a large number of speakers in other countries. For instance, Koma [kmy], with 3,000 speakers in Cameroon but with 32,000 speakers in Nigeria, is not listed in Table 2.5. The majority of languages in Table 2.5 either had no population estimates in *Ethnologue* 13 (Grimes 1996), or had estimates based on the same studies cited in *Ethnologue* 17. Whether a language is labeled “developing” or “vigorous” seems to correlate with the presence or absence of translations of parts of the Bible. Even with limited language development, it is unclear that a language with only a few thousand speakers is best considered safe. Language shift may happen rapidly if the ecology of language use changes, or because of factors such as drought, disease, and conflict (see Chia 2006).

TABLE 2.5. Languages of Cameroon labeled “vigorous,” “developing,” or “wider communication” with 5,000 or fewer speakers in *Ethnologue* 17 (Lewis et al. 2013).

| Language name & Ethnologue code | Status              | Population estimate & year of estimate (from <i>Ethnologue</i> 17 (Lewis et al. 2013)) | Grammar (Nordhoff et al. 2014) | Bible portions |
|---------------------------------|---------------------|--|--------------------------------|----------------|
| Jukun Takum [jbu]               | Wider communication | 2,440 in Cameroon (2000)   | 1980 ✓                         | ✓              |
| Cuvok [cuv]                     | Developing          | 5,000 (1983), increasing   | 2003 phonology                 | ✓              |
| Duupa [dae]                     | Developing          | 5,000 (1991)   | —                              | ✓              |
| Jimi [jim]                      | Developing          | 3,500 (1982)   | —                              | ✓              |
| Kolbila [klc]                   | Developing          | 2,500 (1997)   | —                              | ✓              |
| Yambeta [yat]                   | Developing          | 3,700 (1982)   | —                              | ✓              |
| Yasa [yko]                      | Developing          | 2,400 (2000)   | 2004 sketch ✓                  | —              |
| Akum [aku]                      | Vigorous            | 1,400 in Cameroon (2002)   | —                              | —              |
| Ambele [ael]                    | Vigorous            | 2,600 (2000)   | 2001 phonology                 | —              |
| Bamenyam [bce]                  | Vigorous            | 4,000 (1994)   | 2003 sketch                    | —              |
| Bassossi [bsi]                  | Vigorous            | 5,000 (2004)   | —                              | —              |
| Bati [btc]                      | Vigorous            | 800 (1975)   | —                              | —              |
| Befang [bby]                    | Vigorous            | 2,980 (2000)   | 2004 sketch                    | —              |
| Bonkeng [bvg]                   | Vigorous            | 3,000 (2000)   | —                              | —              |
| Cung [cug]                      | Vigorous            | 2,000 (2001)   | —                              | —              |
| Dek [dek]                       | Vigorous            | 2,980 (2000)   | —                              | —              |
| Dugwor [dme]                    | Vigorous            | 5,000 (2001)   | —                              | —              |
| Dzodinka [add]                  | Vigorous            | 2,600 in Cameroon (2000), increasing   | —                              | —              |
| Gimme [kmp]                     | Vigorous            | 3,000 (1982)   | —                              | —              |
| Gimnime [gmn]                   | Vigorous            | 3,000 (1982)   | —                              | —              |
| Ipulo [ass]                     | Vigorous            | 2,500 (1990)   | —                              | —              |
| Kemedzung [dmo]                 | Vigorous            | 4,500 (2001)   | 2010 sketch                    | —              |
| Kwa' [bko]                      | Vigorous            | 1,000 (2000)   | —                              | —              |
| Kwaja [kdz]                     | Vigorous            | 2,980 (2000)   | —                              | —              |
| Laimbue [lmx]                   | Vigorous            | 5,000 (1994)   | —                              | —              |
| Longto [wok]                    | Vigorous            | 2,400 (1982)   | —                              | —              |
| Malimba [mzd]                   | Vigorous            | 2,230 (2001)   | 1974 phonology                 | —              |
| Mbonga [xmb]                    | Vigorous            | 1,490 (2000)   | —                              | —              |
| Naami [bzv]                     | Vigorous            | 2,500 (2001)   | —                              | —              |
| Ndaktup [ncp]                   | Vigorous            | 2,980 (2000)   | —                              | —              |
| Njen [njj]                      | Vigorous            | 1,800 (2002)   | —                              | —              |
| Nubaca [baf]                    | Vigorous            | 4,500 (2007)   | 1990 phonology                 | —              |
| Nyokon [nvo]                    | Vigorous            | 3,900 (1956)   | 2011 phonology                 | —              |
| Tibea [ngy]                     | Vigorous            | 1,400 (1992)   | —                              | —              |
| Tsuvan [tsh]                    | Vigorous            | 2,300 (2000)   | —                              | —              |
| Yangben [yav]                   | Vigorous            | 2,300 (1994)   | —                              | ✓              |

Languages with small numbers of speakers tend to have low rates of language documentation. Surely the lack of previous descriptive work should be considered when prioritizing languages to be documented, and not just the total number of speakers or degree of vitality. The majority of very small languages in Cameroon in Table 2.5 are almost entirely undocumented. Most are known only through wordlists and lack a grammatical description. Larger Cameroonian languages tend to be better described than do smaller ones, but since most Cameroonian languages are fairly small, the documentation problem is a very large one.

#### 2.2.3.4. Nigeria

Nigeria is similar to Cameroon in that it has a large number of very small languages. Table 2.6 shows the 103 “vigorous” or “developing” Nigerian languages with 5,000 or fewer speakers according to *Ethnologue* 17 (Lewis et al. 2013). This number is in addition to more than 60 that are “in trouble” or “dying.”

TABLE 2.6. 103 Languages of Nigeria labeled “vigorous” or “developing” with 5,000 or fewer speakers in *Ethnologue* 17 (Lewis et al. 2013).

| Total = 103          | Duhwa [kbz]        | Kir-Balar [kkrr]        | Obulom [obuj]       |
|----------------------|--------------------|-------------------------|---------------------|
| Abon [abo]           | Dungu [dbv]        | Koenoem [kcs]           | Okodia [okd]        |
| Abureni [mgj]        | Ebughu [ebg]       | Kofa [kso]              | Oruma [orr]         |
| Acipa, Eastern [acp] | Eki [eki]          | Kpasham [pbn]           | Panawa [pwb]        |
| Aduge [adu]          | Fam [fam]          | Kugama [kow]            | Pe [pai]            |
| Àhàn [ahn]           | Firan [fir]        | Kugbo [kes]             | Piya-Kwonci [piy]?  |
| Ajiya [ide]          | Fungwa [ula]       | Kumba [ksm]             | Sha [scw]           |
| Ake [aik]            | Gengle [geg]       | Kutto [kpa]             | Shama-Sambuga [sqa] |
| Alege [alf]          | Gudu [gdu]         | Kyak [bka]              | Shuwa-Zamani [ksa]  |
| Ambo [amb]           | Gwa [gwb]          | Laka [lak]              | Sur [tdl]           |
| Áncá [acb]           | Hasha [ybj]        | Lamja-Dengsa-Tola [ldh] | Tala [tak]          |
| Asu [aum]            | Horom [hoe]        | Laru [lan]              | Tambas [tdk]        |
| Ayere [aye]          | Ibuoro [ibr]       | Leelau [ldk]            | Teme [tdo]          |
| Baan [bvj]           | Idere [ide]        | Lopa [lop]              | Tha [thy]           |
| Bali [bcn]           | Iko [iki]          | Mághdì [gmd]            | Tita [tdq]          |
| Basa-Gurmana [buj]   | Ito [itw]          | Mashi [jms]             | Toro [tdv]          |
| Biseni [ije]         | Itu Mbon Uzo [itm] | Mbongno [bgu]           | Tumi [kku]          |
| Bo-Rukul [mae]       | Janji [jni]        | Mingang Doso [mko]      | Ubang [uba]         |
| Bukwen [buz]         | Jiru [jrr]         | Moo [gwg]               | Uzekwe [eze]        |
| Burak [bys]          | Ju [juu]           | Mundat [mmf]            | Vono [kch]          |
| Buru [bqw]           | Kaivi [kce]        | Ndunda [nuh]            | Waka [wav]          |
| Cakfem-Mushere [cky] | Kam [kdx]          | Nggwahyi [ngx]          | Wannu [jub]         |
| Cineni [cie]         | Kami [kmi]         | Nkari [nkz]             | Wāpha [juw]         |
| Ciwogai [tgd]        | Kariya [kil]       | Nkoro [nkx]             | Wom [wom]           |
| Cori [cry]           | Kholok [kct]       | Nkukoli [nbo]           | Yotti [yot]         |
| Dong [doh]           | Kinuku [kkd]       | Nnam [nbp]              | Zumbun [jmb]        |

The most up-to-date source on Nigerian languages is Blench (2012). Blench notes, for instance, that Abon is spoken in just one town and Abureni in only 3 villages. He notes that many Plateau languages were still flourishing in his 1993–1999 survey (2012:vii). To say that a language is thriving is different, however, from claiming that it is not threatened.

There are many other languages in Nigeria with small populations that might be best considered “threatened.” While Ukaan [kcf] is described as “vigorous” with more than 5,000 speakers, a recent documentation project describes it as a cluster of four dialects spoken in only five villages (Salffner 2009). There are perhaps only 600 to 700 speakers of the Ikaan dialect, which has already shown signs of obsolescence among young adults and children (Salffner 2009:29). Another language, Ichen [cen] has seen almost two-thirds of the population shift away from their mother tongue (Fakuade 2000–2001:18). It may well be that the majority of Nigerian languages are endangered (Agbedo & Omeje 2008:541).

The majority of small Nigerian languages are very poorly documented. At the ACAL 45 meeting, only a few attendees were familiar with five or more of the languages listed in Table 2.6. It is hard to imagine a group of Australian or North American language specialists being equally unfamiliar with 100 languages from those regions.

#### 2.2.4. LANGUAGES BELONGING TO UNDER-DOCUMENTED FAMILIES

A language that is an isolate or a member of a sparsely documented family can potentially contribute a great deal to linguistic knowledge. Africa is often perceived as having only four language families, though the linguistic genetic diversity of African languages is actually much higher (cf. Sands 2009). The groups considered most important to describe from the perspective of a typologist are given in Table. 2.7.

Work on Shabo, or Chabu, spoken by some 400 people, shows that it is most likely an isolate (Schnoebelen 2009). Little is known about Mpra, or Mpre apart from a few lexical items (Blench 2010). Kujarge is one of the highest-priority languages to be documented. It has only an estimated 1,000 speakers and the only descriptive material available for the language is a 200-item wordlist (Hammarström 2010:184). Its initial classification as a Chadic language is uncertain (Blažek 2013). Further documentation of the language would shed light on the internal structure of Afroasiatic (Blench 2013), and on whether Kujarge is actually Afroasiatic, or had contact with Afroasiatic languages.

Grammars of some of these languages or language clusters have appeared very recently, showing that linguists have risen to the challenge of language documentation. Examples are Palayer & Yaya (2011) on Daju, Stirtz (2011) on Gaahmg (Jebel), Kibebe (2015) on Chabu (Shabo), Hantgan (2013) on Bangime,

TABLE 2.7. Sparsely documented African language families (Hammarström 2010)

| Language name                      | location/status             | Grammar sketch           |
|------------------------------------|-----------------------------|--------------------------|
| Kujarge [vkj]                      | (Chad/Sudan border area)    | —                        |
| Mpra (Mpre)                        | (Ghana), nearly extinct     | —                        |
| Birri [bvq]                        | CAR                         | √                        |
| Daju [byg, djc, daj, dau, njl]     | Chad                        | √                        |
| Eastern Jebel [soh, xel, zmo, tbi] | Sudan                       | √                        |
| Kresh group [krs, aja]             | Sudan, South Sudan          | √                        |
| Shabo [sbf]                        | Ethiopia                    | √                        |
| Tama [mgb, sjg, tma]               | Chad                        | √                        |
| Temeinan [teq, keg]                | Sudan                       | √                        |
| Warnang [wrn]                      | Sudan                       | possibly forth. Quint    |
| Tegem/Lafofa [laf]                 | Sudan                       | possibly forth. Thelwall |
| Mao [myf, gza, hoz, sze]           | Sudan, Ethiopia             | √                        |
| Bangime [dba]                      | Mali                        | √                        |
| Jalaa [cet]                        | (Nigeria) probably extinct  | —                        |
| Dume                               | (Ethiopia) probably extinct | —                        |

and Ahland (2012) on Northern Mao. In most cases, however, only one variety has been described. Further documentation of the languages listed in Table 2.7 remains a high priority.

Other languages that should be investigated because of the unique window they may provide on issues of classification and linguistic history include, for example, Siamou [sif], Mansoanka [msw], Sere [swf], Mbatl [mdn] and Kam [kdx] (Hammarström, personal communication). Since “at least two-thirds of the Chadic languages remain virtually undescribed” (Zaborski 2013:378), research on those languages should be prioritized as well.

Some of Africa's linguistic diversity may reside in what we may term “hidden languages.” As mentioned before, †Hoan was thought to be a dialect of !Xô. Similarly, Bangime had been classified as Dogon rather than as an isolate (see Blench 2007b), and Dogon itself was not always recognized to be a group of languages rather than a dialect cluster. Guébie, a language with only about 1,000 speakers, is misclassified in *Ethnologue* 17 as two different “vigorous” languages: as a dialect of Bété-Gagnoa [btg] and as another name for Dida-Lakota [dic] (Sande 2017). The documentation of Africa's unique and least-known languages may come about only with the investigation of poorly known dialects.

### 2.2.5. SIGN LANGUAGES

In addition to true isolates and isolates within language phyla, we must not forget that there are many African sign languages, particularly sign languages of local

TABLE 2.8. Some undocumented West African Sign language isolates (cf. Nyst 2010, Orie 2013).

---

|   |
|---|
| Bura Sign Language (Nigeria)                                      |
| Mbour Sign Language (Senegal)                                     |
| Burkina Faso Sign Language/Langue des Signes Mossi (Burkina Faso) |
| Nanabin Sign Language (Ghana)                                     |
| Yoruba Sign Language (Nigeria)                                    |

---

origin. Adamarobe Sign Language (Nyst 2007) and Algerian Jewish Sign Language (Lanesman 2013) are examples of sign language isolates that have been described. The languages listed in Table 2.8 are surely only a few of the locally developed sign languages for which there is little published documentation.

### 2.2.6. VARIATION IN TYPES OF STUDIES

Recent funding trends appear to favor certain types of single-language documentation projects over cross-linguistic and other studies. But in order to know more about Africa's least-known languages, it is important to remember that studies of contact patterns (e.g. Good 2013) and patterns of language use (e.g. Lüpke & Storch 2013; Childs et al. 2014) are an important part of linguistic description. Cross-linguistic studies may be the best way to gain a thorough understanding of rare sounds (labial flaps, clicks, etc.), rare syntactic patterns (linkers, nominative case, etc.), anaphors (see the Afranaph anaphor project <http://www.africananaphora.rutgers.edu/>) and micro-typological patterns (e.g. Marlo 2013). Currently, some of these types of studies are considered "investigative" rather than "descriptive" and may therefore not be thought suitable for funding under a program designed to support endangered language documentation. Studies focused on topics such as intonation and prosody, ethnobotany (e.g. Legère 2009), or language socialization are valuable, but similarly might be considered too narrow in scope for some funding agencies.

## 2.3. Particular challenges

### 2.3.1. FUNDING LEVELS

Most of us who wish to conduct fieldwork on languages in Africa require funding. The U.S. National Science Foundation (NSF) has funded several projects, including some of my own, and they have also sponsored the meeting "Africa's Endangered Languages: Documentary & Theoretical Approaches," at which this article was first presented. Despite the large number of endangered African languages, not enough research is being funded, in large part because not enough proposals are being submitted. I believe that the dearth of earlier linguistic work on so many endangered African languages means that there is little awareness of the interesting linguistic

structures that these languages may have. It is hard for linguists to have their curiosity piqued about a language if they have never even heard of it. Besides this, there is a perception that African languages are not as threatened as those of other continents (e.g. Simons & Lewis 2013).

### 2.3.2. DIFFICULTY IN ACCESSING LANGUAGE COMMUNITIES

In many instances, there are logistical issues which make the documentation of some of Africa's least-known languages difficult. Some languages are spoken in areas that are hard to travel to, or are spoken across national borders, which can make it time-consuming to obtain necessary research permits and/or travel documents. Disease and conflict can impede access to research sites; between September 27, 2013 and March 27, 2014, the U.S. State Department issued warnings and alerts for 17 different African countries. Poor roads and unreliable electricity can affect the choices researchers make about where to work, as can lack of access to health care, fresh produce, or a quiet working space.

I faced a number of logistical difficulties during own fieldwork on Hadza [hts], an isolate spoken in Tanzania. I worked primarily with speakers to the southeast of Lake Eyasi in Mang'ola, and a bit in the Tlhi'ika dialect area south of the lake. In 1991, when I conducted much of my research, Mang'ola had no electricity, running water, telephone service, shops, or health clinic. To obtain research permission, I had to appear at government offices at national, regional, and district levels, in Dar es Salaam, Arusha, Mto wa Mbu, and Mbulu, taking one and one-half to two weeks of time that might otherwise have been spent doing research. To have investigated the Dunduhina dialect of Hadza, spoken to the north and west of Lake Eyasi, would have required obtaining research permission to work in Shinyanga region. Travel to the Shinyanga regional capital would have been difficult given the lack of reliable fueling points and good roads at the time. The areas west of the lake are those that have experienced the greatest loss of speakers due to shift to Sukuma. It would be interesting to see how different the dialects are, given different levels of contact with Sukuma and Isanzu, but to my knowledge no linguist has conducted research in those areas because of the logistical difficulties.

### 2.3.3. Difficulties in working on obsolescent languages

#### 2.3.3.1. PRACTICAL ISSUES

In addition to the logistical issues that affect African language documentation, there are practical difficulties in working on poorly documented, obsolescent languages. For one, finding and identifying proficient language users may be challenging. People may not live in a single community but be scattered across a wide area, making it hard to work with more than one or two people for an extended period. Another problem may be that semi-speakers or rememberers may be upset if they

are left out of a language documentation project. The most competent users of a very endangered language may be elderly, yet elderly consultants often have health problems that affect their ability to participate in language documentation. They may have limited vision or hearing, or be missing many teeth, which can affect their pronunciation of certain consonants. Problems with sight may make it hard for them to identify flora and fauna through pictures in field guides, and problems with stamina and balance may make it difficult to identify plants and natural features by taking a bush walk. On a more positive note, however, it is a joy to work with people who become rejuvenated as they get together with others and speak the language of their youth.

Marginalized communities may be disproportionately affected by alcohol abuse, poor health, and domestic violence. These issues certainly can affect the ability of consultants to participate fully in a language documentation project.

With larger languages, it may be possible to find a consultant with a great deal of metalinguistic awareness, or one with university education. With very small languages, one may not be able to find any fluent speaker who can annotate texts, provide grammaticality judgments, or even have the patience to help a linguist record a full paradigm.

Very small language communities may lack ties to scholars within their own countries or abroad. Obtaining funding to work on a language where you have no ties can be difficult, but connections between interested linguists and endangered language communities may be facilitated through the Internet, as with Roger Blench's YouTube channel (<http://www.youtube.com/user/rogerblench?feature=mhee>).

#### 2.3.3.2. Linguistic issues

Obsolescent languages can be especially difficult to document because of the level of idiosyncratic variation found across speakers (see Dorian 2010), particularly in pronunciation (Connell 2002) and in the lexicon (Sands et al. 2007). This variation may be difficult to see if one is only able to work with a very small number of speakers in a single location. Aspects of tone systems, morphology, and syntax may be challenging to characterize because of inconsistencies across speakers (Childs 2012).

The lexicon of obsolescent languages in particular, and of shifting languages generally, may be challenging to document. Even with the best translation assistance, there are mismatches in lexical semantics between contact language(s) and the target language. People from marginalized communities may not speak a standard version of a contact language and may not be aware of differences between the way they use a contact language and the way it is understood by the linguist and/or the translator. Often a linguist will work with younger speakers who are bilingual and have been to school, but these may be the very people who have a reduced lexicon in comparison to other language users. Even a language that is still being learned by children, such as Gizey<sup>8</sup> (Ajello 2007), may have seen loss of competence in the domain of color terms, for instance. Khwe has

seen changes in folk taxonomies, the numeral system, and the spatial orientation system (Brenzinger 2003). Kinship terminology may differ from one speaker to another depending on the different languages spoken by each person's relatives (Boden 2014).

Since obsolescent languages are used in limited domains, the linguist may be restricted in the types of discourse he or she may be able to document. Special registers, ideophones, and certain genres may be lost. Grammatical changes will likely be in evidence for shifting languages (see Schaefer & Egbokhare 1997), and certainly for obsolescent languages. Certain linguistic structures may disappear or become less frequent as patterns adopted from dominant languages become more prevalent.

## 2.4. Discussion

Ideally, much fieldwork can be done on languages before they are at the stage where they are already considered “dying,” when fieldwork becomes most urgent. The number of languages that need to be documented is in the hundreds, but I hope that we are not so short of linguists willing to go and do fieldwork that we are constantly chasing after the last few speakers, who may or may not be fluent. The challenges of working on an obsolescent language are such that it might be best if graduate students new to fieldwork first worked on languages that are still in daily use. And, as Lenore Grenoble (2013:54) states, “If we actually care about revitalization and maintenance, we would be well advised to study languages with smaller numbers of speakers that are not (yet) shifting.” Tucker Childs has similarly called for “the reprioritization of the study of endangered languages, focusing on languages that have a present” (Childs 2014:9).

Given the highly competitive funding environment that exists today, the chances of funding a project may be reduced if a language is not considered to be under great risk of language shift. Population figures and data documenting language shift for African languages are likely to be inaccurate or lacking, particularly in comparison with figures for North American and European minority languages.<sup>9</sup> The least-known African languages may continue to be less known because of the fact that we do not have much accurate information about them. The problem is not a shortage of linguists, but rather in a lack of funds connecting linguists with languages that are in need of documentation.

For more of Africa's least-known languages to be documented, it may be necessary for scholars to have access to mentoring and/or workshops on proposal writing. More funding for African language research will become available only if there is proposal pressure put on funding agencies—if dozens of extremely high-ranked, well-written proposals are turned down for funding, then this effectively demonstrates the need for additional funding.

Anecdotal evidence suggests that the *Ethnologue* vitality labels can play a determining role in whether a particular language documentation project is chosen for funding. *Ethnologue*'s primary purpose is to serve the needs of missionaries and Bible translators, and it is not ideally suited for playing a gatekeeper role in the funding of academic linguistic research. Data on languages in *Ethnologue* is often out of date, and it provides little information on the numbers of speakers of endangered dialects of otherwise vigorous languages. EGIDS, as it is currently used, does not take into account important factors such as trending patterns in language use, proportion of speakers undergoing language shift, population movement, language attitudes, and the absolute number of speakers (see Brenzinger, Yamamoto, et al. 2003)

## 2.5. Conclusion

The number of African languages in urgent need of documentation appears to be closer to 600 than to the 346 considered at risk in *Ethnologue* 17 (Lewis et al. 2013). Since there is a growing consensus that we should document languages before they are critically endangered, the number of languages in need of documentation is certainly much higher than 600. Current levels of documentation for most African languages with fewer than 100,000 speakers are quite low, and many of these languages are at great risk of language shift, especially in countries such as Tanzania. Nigerian languages are among the world's least-known languages, but all parts of the continent contain undescribed threatened languages.

There is a shockingly large number of African languages that lack even basic lexical and grammatical descriptions. In order for linguists to rise to the challenge of documenting Africa's least-known languages, we cannot simply target only the most critically endangered ones. The vitality level of a language as reported by *Ethnologue* or UNESCO should not be considered the most important criterion used in setting priorities for language documentation.

## Notes

\* Thanks to my consultants Muchuwaiko Tsaidla, Basenane Mosupatsela, Mosekathoshe Molutwane, Suag||ai N||aleciexo, Talela Mosupatsela, and Justa Mosekatoshe of Duthwane and Tshwaane, Botswana and to Gudo Bala in Tanzania. Thanks also to the organizers of this conference for funding my travel, and to the National Science Foundation for supporting research on endangered African languages.

1. This material is based upon work supported by the National Science Foundation under Grant #BCS-0726200: "Collaborative Research: Phonetic and Phonological

Structures of Post-velar Constrictions in Clicks and Laterals” and SBR-9514214 “Research on Hadza Demography and Language.” Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

2. *Ethnologue* 18 (Lewis et al. 2015) has slightly revised some of the EGIDS definitions.

3. Even if the heritage language has been transmitted to children, once they attend school, the local lingua franca (such as Kiswahili in Tanzania, or Setswana in Botswana) may be so dominant across communicative domains that complete mastery and retention of the mother tongue may be compromised.

4. Naro does not have anything like the number of speakers or level of governmental, political, financial, and educational support of a language such as Irish, which is considered by UNESCO to be “definitely endangered” (Moseley 2010).

5. Several dialects of Datooga are endangered (Griscom 2014).

6. David Odden is working on a Kikerewe grammar (<http://www.ling.ohio-state.edu/~odden/>).

7. Piya-Kwonci was labeled “developing.” All other languages in the table were labeled “vigorous.”

8. Gizey would not be considered “dying” or “in trouble” as it is used by approximately 12,000 people, including children (Ajello 2007). In fact, *Ethnologue* does not list Gizey as a separate language, but has “Gizay (Guissey)” listed as dialects of Masana [mcn], with a population of 103,000 (Lewis et al. 2013).

9. Furthermore, as one reviewer points out, sociolinguistic surveys that do not use proficiency tests can overestimate speaker populations. Respondents asked about “their” language may indicate the language of the ethnic group they belong to, even if this is not the language they speak.

## References

- ADEGBIJA, EFUROSIBINA. 2001. Saving threatened languages in Africa: A case study of Oko. *Can threatened languages be saved? Reversing language shift, revisited, a 21st century perspective*, ed. by Joshua A. Fishman, 284–308. Clevedon, UK: Multilingual Matters.
- ADÉNYÍ, HARRISON, & RACHAEL BÉLLÒ. 2008. Language use in the Badagry regulated market. *Critical issues in the study of linguistics, languages and literatures in Nigeria: A Festschrift for Conrad Max Benedict Brann*, ed. by Ozo-mekuri Ndimele, Imelda Udoh, & Ogonna Anyanwu, 219–232. (Nigerian Linguists Festschrift Series 7). Port Harcourt, Nigeria: Linguistics Association of Nigeria, & Munich: Lincom Europa.
- AGBEDO, CHRIS U., & GRACE O. OMEJE. 2008. Endangered languages in Nigeria: Linguistic strategies for safeguarding and reviving them. *Critical issues in the study of linguistics, languages and literatures in Nigeria: A Festschrift for Conrad Max Benedict Brann*, ed. by Ozo-mekuri Ndimele, Imelda Udoh, & Ogonna Anyanwu, 541–552. Port Harcourt, Nigeria: Linguistics Association of Nigeria, & Munich: Lincom Europa.
- AHLAND, MICHAEL BRYAN. 2012. A grammar of Northern Mao (Màw's Aas'è). Eugene: dissertation, University of Oregon.
- AJELLO, ROBERTO. 2007. The importance of having a description of the endangered languages: The case of Gizey (Cameroon). *Undescribed and endangered*

- languages: *The preservation of linguistic diversity*, ed. by Amedeo de Dominicis, 8–20. Newcastle: Cambridge Scholars.
- ANCHIMBE, ERIC A. 2013. Crossing the survival line: Factors that protect indigenous languages from extinction. *Language policy and identity construction: The dynamics of Cameroon's multilingualism*, ed. by Eric A. Anchimbe, 77–96. Amsterdam: John Benjamins.
- BARNARD, ALAN. 1985. *A Nharo wordlist*. (Occasional Publications 2.) Durban, South Africa: Department of African Studies, University of Natal.
- BATIBO, HERMAN M. 2001. The endangered languages of Africa: A case study from Botswana. *On biocultural diversity: Linking language, knowledge, and the environment*, ed. by Luisa Maffi, 311–324. Washington, DC: Smithsonian Institution Press.
- BATIBO, HERMAN M. 2005. *Language decline and death in Africa: Causes, consequences and challenges*. Clevedon, UK: Multilingual Matters.
- BATIBO, HERMAN M. 2013. Preserving and transmitting indigenous knowledge in diminishing bio-cultural environment: Case studies from Botswana and Tanzania. *African Study Monographs* 34(3).161–173.
- BELL, HERMAN. 2006. Birgid Nubian: Recording a nearly extinct language. *Insights into Nilo-Saharan language, history and culture: Proceedings of the 9th Nilo-Saharan Linguistic Colloquium, Institute of African and Asian Studies, University of Khartoum, 16–19 February 2004*, ed. by Al-Amin Abu-Manga, Leoma Gilley, & Anne Storch, 75–83. (Nilo-Saharan Linguistic Analyses and Documentation 23). Cologne: Rüdiger Köppe.
- BLÁŽEK, VÁCLAV. 2013. Kujarge wordlist with Chadic (Afroasiatic) cognates. *Topics in Chadic linguistics VII: Papers from the 6th Biennial International Colloquium on the Chadic Languages, Villejuif, September 22–23, 2011*, ed. by Henry Tourneaux, 43–52. Cologne: Rüdiger Köppe.
- BLENCH, ROGER. 2007a. Endangered languages in West Africa. *Language diversity endangered*, ed. by Matsthiias Brenzinger, pp. 140–162. Berlin: Mouton de Gruyter.
- BLENCH, ROGER. 2007b. Bangi Me, a language of unknown affiliation in Northern Mali and its affinities. *Mother Tongue* 12.147–178.
- BLENCH, ROGER. 2010. Recovering data on Mpra [=Mpre], a possible language isolate in North-Central Ghana. Ms. circulated for the workshop “Language Isolates in Africa,” Lyon, December 3–4, 2010. Online: <http://www.ddl.ish-lyon.cnrs.fr/projets/clhass/PageWeb/ressources/Isolats/Mpra%20%20Blench%202007.pdf>.
- BLENCH, ROGER. 2012. *An atlas of Nigerian languages*. 3rd ed. Online: [www.languagesampler.com/alt/docs/Atlas\\_of\\_Nigerian\\_Languages.pdf](http://www.languagesampler.com/alt/docs/Atlas_of_Nigerian_Languages.pdf)
- BLENCH, ROGER. 2013. Links between Cushitic, Omotic, Chadic and the position of Kujarge. *Proceedings of the 5th International Conference on Cushitic and Omotic Languages, Paris, 16-18 April 2008*, ed. by Marie-Claude Simeone-Senelle & Martine Vanhove, 67–80. (Cushitic and Omotic Studies 3). Cologne: Rüdiger Köppe.
- BLURTON JONES, NICHOLAS G., LARS C. SMITH, JAMES F. O'CONNELL, KRISTEN HAWKES, & C. L. KAMUZORA. 1992. Demography of the Hadza: An increasing and high density population of savanna foragers. *American Journal of Physical Anthropology* 89.159–181.
- BODEN, GERTRUD. 2014. N||ng kinship terminology—salvage documentation with the last speakers. *Southern African Khoisan kinship systems*, ed. by Alan Barnard & Gertrud Boden, 63–84. (Quellen zur Khoisan-Forschung/Research in Khoisan Studies 30). Cologne: Rüdiger Köppe.

- BREZZINGER, MATTHIAS, ed. 1992. *Language death: Factual and theoretical explorations with special reference to East Africa*. Berlin & New York: Mouton de Gruyter.
- BREZZINGER, MATTHIAS, ed. 1998a. *Endangered languages in Africa*. Cologne: Rüdiger Köppe.
- BREZZINGER, MATTHIAS. 1998b. *Sprachwechsel afrikanischer Minoritäten aus soziolinguistischer Sicht*. (Afrikanistische Monographien 9.) Cologne: Institut für Afrikanistik, Universität zu Köln.
- BREZZINGER, MATTHIAS. 2003. Language and conceptual diversity under threat: language endangerment on the African continent. *Linguistic cultural identity and international communication*, ed. by Johann Vielberth & Guido Drexler, 59–77. Saarbrücken: AQ Verlag.
- BREZZINGER, MATTHIAS. 2007a. Language endangerment in northern Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 123–139. Berlin: Mouton de Gruyter.
- BREZZINGER, MATTHIAS. 2007b. Language endangerment in southern and eastern Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 179–204. Berlin: Mouton de Gruyter.
- BREZZINGER, MATTHIAS. 2013. The twelve modern Khoisan languages. *Khoisan languages and linguistics (Proceedings of the 3rd International Symposium July 6–10, 2008, Riezlern/Kleinwalsertal)*, ed. by Alena Witzlack-Makarevich & Martina Ernszt, 1–31. (Quellen zur Khoisan-Forschung/Research in Khoisan Studies 29.) Cologne: Rüdiger Köppe.
- BREZZINGER, MATTHIAS, & HERMAN BATIBO. 2010. Sub-Saharan Africa. *Atlas of the world's languages in danger*, 3rd ed., ed. by Christopher Moseley, 20–25. Paris: UNESCO.
- BREZZINGER, MATTHIAS, BERND HEINE, & GABRIELE SOMMER. 1991. Language death in Africa. *Endangered languages*, ed. by R. H. Robins & E. M. Uhlenbeck, 19–44. Oxford: Berg.
- BREZZINGER, MATTHIAS, AKIRA YAMAMOTO, NORIKO AIKAWA, DIMITRI KOUNDIOUBA, ANAHIT MINASYAN, ARIENNE DWYER, COLETTE GRINEVALD, MICHAEL KRAUSS, OSAHITO MIYAOKA, OSAMU SAKIYAMA, RIEKS SMEETS, & OFELIA ZEPEDA. 2003. *Language vitality and endangerment*. Paris: UNESCO Expert Meeting on Safeguarding Endangered Languages. Online: <http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf>.
- CHIA, EMMANUEL N. 2006. Rescuing endangered Cameroonian languages for national development. *African linguistics and the development of African communities/Linguistique africaine et le développement des communautés africaines (Papers Presented at the 23rd Congress of the West African Linguistic Society, at the University of Buea, Cameroon, 7–9 August 2002)*, ed. by Emmanuel Chia, 115–128. Dakar: Council for the Development of Social Science Research in Africa.
- CHILDS, G. TUCKER. 2006. Language endangerment in West Africa: Its victims and causes. *The joy of language: Proceedings of a symposium honoring the colleagues of David Dwyer on the occasion of his retirement*. Online: <https://www.msu.edu/~dwyer/16-Childs.doc>.
- CHILDS, G. TUCKER. 2012. How to pretend you speak a dying language when you don't really know how to speak it: Methodological worries in documenting dying languages. *Proceedings of the 6th World Congress of African Linguistics, Cologne, 17–21 August 2009*, ed. by Matthias Brenzinger & Anne-Maria Fehn, 45–56. Cologne: Rüdiger Köppe.
- CHILDS, G. TUCKER. 2014. Problems for the salvage linguist: Picking up the pieces after asymmetrical contact. *In and out of Africa: Languages in question in honour of Robert*

- Nicolai*; vol. 2, *Language contact and language change in Africa*, ed. by Carole de Féral, Maarten Kossmann & Mauro Tosco, 73–93. Leuven: Peeters.
- CHILDS, TUCKER, JEFF GOOD, & ALICE MITCHELL. 2014. Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation* 8.168–191.
- COLLINS, CHRIS, & JEFF GRUBER. 2014. *A grammar of †Hǝǝ*. (Quellen zur Khoisan-Forschung/Research in Khoisan Studies 32). Cologne: Rüdiger Köppe.
- CONNELL, BRUCE. 1998. Moribund languages of the Nigeria-Cameroon borderland. *Endangered languages in Africa*, ed. by Matthias Brenzinger, 207–225. Cologne: Rüdiger Köppe.
- CONNELL, BRUCE. 2002. Small languages and small language communities, 39: Phonetic/phonological variation and language contraction. *International Journal of the Sociology of Language* 157(1).167–185.
- CONNELL, BRUCE. 2007. Endangered languages in Central Africa. *Language diversity endangered*, ed. by Matthias Brenzinger, 163–178. Berlin: Mouton de Gruyter.
- CONNELL, BRUCE. 2010. Language ecology and language endangerment: An instance from the Nigeria-Cameroon borderland. *Journal of West African Languages* 37(1).3–32.
- DIMMENDAAL, GERRIT J. 1989. On language death in eastern Africa. *Investigating obsolescence: Studies in language contraction and death*, ed. by Nancy Dorian, 13–31. Cambridge, UK: Cambridge University Press.
- DIMMENDAAL, GERRIT J., & F. K. ERHARD VOELTZ. 2007. Africa. *Encyclopedia of the world's endangered languages*, ed. by Christopher Moseley, 579–634. London: Routledge.
- DORIAN, NANCY C. 2010. *Investigating variation: The effects of social organization and social setting*. Oxford: Oxford University Press.
- EDENMYR, NIKLAS, HÉLÈNE FATIMA IDRIS, & KARSTEN LEGÈRE. 2011. Endangered languages: Examples from Tanzania and Sudan. *North-south contributions to African languages*, ed. by Christina Thornell & Karsten Legère, 81–94. (Topics in Interdisciplinary African Studies 19.) Cologne: Rüdiger Köppe.
- FABUNMI, FELIX ABIDEMI. 2005. Dialectological perspectives on language attrition among the Yorùbá. *Trends in the study of languages and linguistics in Nigeria: A Festschrift for Philip Akujuboobi Nwachukwu*, ed. by Ozo-mekuri Ndimele, 261–269. Aba, Nigeria: National Institute for Nigerian Languages.
- FABUNMI, FELIX A., & AKEEM SEGUN SALAWU. 2005. Is Yorùbá an endangered language? *Nordic Journal of African Studies* 14(3).391–408.
- FAKUADE, GBENGA. 2002. Language endangerment in the north-eastern Nigeria: Instances and strategies for averting it. *Annals of Borno* 17/18:13–23.
- FISHMAN, JOSHUA A., ed. 2001. *Can threatened languages be saved? Reversing language shift, revisited: A 21st century perspective*. Clevedon, UK: Multilingual Matters.
- GERLACH, LINDA. 2015. Phonetic and phonological description of the N!aqriaxe variety of †Amkoe and the impact of language contact. PhD thesis, Humboldt Universität zu Berlin.
- GERLACH, LINDA, & FALKO BERTHOLD. 2011. The sociolinguistic situation of †Hoan, a moribund 'Khoisan' language of Botswana. Online: <http://www.afrikanistik-aegyptologie-online.de/archiv/2011/3164>.
- GOOD, JEFF. 2013. A (micro-)accretion zone in a remnant zone? Lower Fungom in areal-historical perspective. *Language typology and historical contingency: In honor of*

- Johanna Nichols*, ed. by Balthasar Bickel, Lenore A. Grenoble, David A. Peterson & Alan Timberlake, 265–282. Amsterdam: John Benjamins.
- GRENOBLE, LENORE A. 2013. Unanswered questions in language documentation and revitalization: New directions for research and action. *Responses to language endangerment: In honor of Mickey Noonan*, ed. by Elena Mihás et al., 43–58. Amsterdam: John Benjamins.
- GRIMES, BARBARA E., ed. 1996. *Ethnologue: Languages of the world*. 13th ed. Dallas, TX: Summer Institute of Linguistics (SIL).
- GRISCOM, RICHARD. 2014. Isimjeeg: A distinct and endangered variety of Northern Tanzania. Paper presented at the 45th Annual Conference on African Linguistics, April 17, 2014, University of Kansas.
- GRUBER, JEFFREY S. 1973. #Hòã kinship terms. *Linguistic Inquiry* 4.427–449.
- GRUBER, JEFFREY S. 1975. Bushman languages of the Kalahari: #Hoan—vocabulary—recorded utterances. Technical project report. Washington, DC: National Endowment for the Humanities.
- GÜLDEMANN, TOM. 2014. ‘Khoisan’ linguistic classification today. *Beyond ‘Khoisan’: Historical relations in the Kalahari Basin*, by Tom Güldemann & Anne-Maria Fehn, 1–41. Amsterdam: John Benjamins.
- HAMMARSTRÖM, HARALD. 2010. The status of the least documented language families in the world. *Language Documentation and Conservation* 4.177–212.
- HANTGAN, ABBIE. 2013. Aspects of Bangime phonology, morphology, and morphosyntax. Bloomington: Dissertation, Indiana University.
- HARUNA, ANDREW. 2012. Language shift in Northern Nigeria: The precarious situation of the minority languages. *Celebrating H. E. Wolff's contribution to Chadic linguistics: Proceedings of the 6th World Congress of African Linguistics, Cologne, 17–21 August 2009*, ed. by Matthias Brenzinger & Anne-Maria Fehn, 577–588. Cologne: Rüdiger Köppe.
- HAYS, JENNIFER. 2004. Indigenous rights in education: The San of southern Africa in local and global contexts. *Indigenous peoples' rights in southern Africa*, ed. by Robert K. Hitchcock & Diana Vinding, 228–248. (IWGIA Document 110). Copenhagen: International Work Group for Indigenous Affairs.
- HEINE, BERND, & HENRY HONKEN. 2010. The Kx'a family: A new Khoisan genealogy. *Journal of Asian and African Studies (Ajia Afuriku gengo bunka kenkyu)* 79.5–36.
- HITCHCOCK, ROBERT K., & THOMAS KOPERSKI. 2008. Genocides of indigenous peoples. *The historiography of genocide*, ed. by Dan Stone, 577–617. London: Palgrave.
- IDIATA, DANIEL FRANCK. 2009. *Langues en danger et langues en voie d'extinction au Gabon: Quand la génération des enfants se détourne des langues vernaculaires ou quand les parents détournent leurs enfants de la langue de la communauté*. Paris: L'Harmattan.
- KASTENHOLZ, RAIMUND. 1998. Language shift and language death among Mande blacksmiths and leatherworkers in the diaspora. *Endangered Languages in Africa*, ed. by Matthias Brenzinger, 253–266. Köln: Rüdiger Köppe.
- KÉDRÉBÉOGO, GÉRARD. 1996. Language maintenance and shift in communities of oral tradition: The case of the Sillanko in Burkina Faso (West Africa). Urbana-Champaign: Dissertation, Department of Linguistics, University of Illinois.
- KIBEBE TSEHAY TAYE. 2015. Documentation and grammatical description of Chabu. PhD thesis, Addis Ababa University.

- LANESMAN, SARA. 2013. Algerian Jewish Sign Language: Its emergence and survival. Preston, UK: MA thesis, University of Central Lancashire.
- LEGÈRE, KARSTEN. 1992. Language shift in Tanzania. *Language death: Factual and theoretical explorations with special reference to East Africa*, ed. by Matthias Brenzinger, 99–115. (Contributions to the Sociology of Language 64). Berlin & New York: Mouton de Gruyter.
- LEGÈRE, KARSTEN. 2009. Plant names in the Tanzanian Bantu language Vidunda: Structure and (some) etymology. *Selected proceedings of the 38th Annual Conference on African Linguistics: Linguistic Theory and African Language Documentation*, ed. by Masangu Matondo, Fiona McLaughlin & Eric Potsdam, 217–228. Somerville, MA: Cascadilla Proceedings Project.
- LEWIS, M. PAUL, GARY F. SIMONS, & CHARLES D. FENNIG, eds. 2013. *Ethnologue: Languages of the world*. 17th ed. Dallas, TX: SIL International. Online: <http://www.ethnologue.com>.
- LEWIS, M. PAUL, GARY F. SIMONS, & CHARLES D. FENNIG, eds. 2015. *Ethnologue: Languages of the world*. 18th ed. Dallas, TX: SIL International. Online: <http://www.ethnologue.com>.
- LÜPKE, FRIEDERIKE, & ANNE STORCH. 2013. Language dynamics. *Repertoires and choices in African languages*, ed. by Friederike Lüpke & Anne Storch, 267–344. (Language Contact and Bilingualism 5.) Berlin: Mouton de Gruyter.
- MAHO, JOUNI FILIP, & BONNY SANDS. 2002. *The languages of Tanzania: A bibliography*. (Orientalia et africana gothoburgensia, 17). Göteborg: Göteborg University.
- MARLO, MICHAEL R. 2013. Verb tone in Bantu languages: Micro typological patterns and research methods. *Africana Linguistica* 19.137–234.
- MATIKI, ALFRED JANA. 2009. Re-examining language shift cases in Malawi in the context of Fishman's GIDS. *Journal of Multilingual and Multicultural Development* 30.535–546.
- MOUS, MAARTEN. 2003. Loss of linguistic diversity in Africa. *Language death and language maintenance: Theoretical, practical and descriptive approaches*, ed. by Mark Janse & Sijmen Tol, 157–170. Amsterdam: John Benjamins.
- MEZHOUD, SALEM, & YAMINA EL KIRAT EL ALLAME. 2010. North Africa and the Middle East. *Atlas of the world's languages in danger*, 3rd ed., ed. by Christopher Moseley, 26–31. Paris: UNESCO.
- MOSELEY, CHRISTOPHER, ed. 2010. *Atlas of the world's languages in danger*. 3rd ed. Paris: UNESCO. Online: <http://www.unesco.org/culture/en/endangeredlanguages/atlas>.
- MUGADDAM, ABDEL RAHIM HAMID. 2012. Language use attitudes in Sudanese urban centers. *De l'exploration du multilinguisme dans les villes africaines = Exploring multilingualism in African urban cities*, ed. by Gabriel Mba & Etienne Sadembou, 113–127. Paris: L'Harmattan.
- MUTAKA, NGESSIMO M. 2008. Ecosystem of the Vute-Banyo area and language endangerment. *Explorations into language use in Africa*, ed. by Augustin Simo Bobda, 95–109. Frankfurt am Main: Peter Lang.
- NAKAGAWA, HIROSHI, KAZUYOSHI SUGAWARA, & JIRO TANAKA. 2013. [Gui-English dictionary (a preliminary edition)]. Ms., June 1, 2013.
- NAUMANN, CHRISTFRIED. 2012. *Acoustically based phonemics of Siwi (Berber)*. (Berber Studies 36). Cologne: Rudiger Köppe.

- NDIMELE, ROSELINE IHUOMA. 2005. Sociolinguistic motivation for language shift in multilingual cities in Nigeria. *Trends in the study of languages and linguistics in Nigeria: A Festschrift for Philip Akujooobi Nwachukwu*, ed. by Ozo-mekuri Ndimiele, 153–159. Aba, Nigeria: National Institute for Nigerian Languages.
- NORDHOFF, SEBASTIAN, HARALD HAMMARSTRÖM, ROBERT FORKEL, & MARTIN HASPELMATH. 2014. *Glottolog 2.2*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: <http://glottolog.org>.
- NYOMBE, B. G. V. 1997. Survival or extinction: The fate of the indigenous languages of the Southern Sudan. *International Journal of the Sociology of Language* 125.99–130.
- NYST, VICTORIA. 2007. *A descriptive analysis of Adamorobe Sign Language (Ghana)*. (LOT Dissertation Series 151.) Leiden: CNWS.
- NYST, VICTORIA. 2010. Sign language varieties in West Africa. *Sign languages*, ed. by Diane Brentari, 405–432. Cambridge, UK: Cambridge University Press.
- OHIRI-ANICHE, CHINYERE. 2008. A survey of awareness of language endangerment in Nigeria. *Critical issues in the study of linguistics, languages and literatures in Nigeria: A Festschrift for Conrad Max Benedict Brann*, ed. by Ozo-mekuri Ndimiele, Imelda Udoh, & Ogbonna Anyanwu, 15–38. (Nigerian Linguists Festschrift Series 7.) Port Harcourt: Linguistics Association of Nigeria.
- ORCUTT-GACHIRI, HEIDI. 2013. How can a language with 7 million speakers be endangered? *The persistence of language: Constructing and confronting the past and present in the voices of Jane H. Hill*, ed. by Shannon T. Bischoff et al., 229–255. Amsterdam: John Benjamins.
- ORIE, ỌLANIKÉ ỌLA. 2013. From conventional gestures to sign language: The case of Yoruba sign language. *Proceedings of the 42nd Annual Conference on African Linguistics: Linguistic Interfaces in African Languages*, ed. by Ọlaniké Ọla Orie & Karen W. Sanders, 244–251. Somerville, MA: Cascadilla Proceedings Project.
- PALAYER, PIERRE, & CHAÏBO YAYA. 2011. *Grammaire du dadjo d'Eref (Tchad)*. (SELAF 453, Afrique et Langage 13). Louvain: Peeters.
- ROBSON, LAURA. 2011. The documentation of the language ecology of Njanga, a moribund language of Cameroon. Canterbury: Dissertation, University of Kent.
- ROSENDAL, TOVE, & GASTOR MAPUNDA. 2014. Is the Tanzanian Ngoni language threatened? A survey of lexical borrowing from Swahili. *Journal of Multilingual and Multicultural Development* 35.271–288.
- SALFFNER, SOPHIE. 2009. Tone in the phonology, lexicon and grammar of Ikaan. London: Dissertation, Department of Linguistics, School of Oriental and African Studies, University of London.
- SANDE, HANNAH. 2017. Classification of Guébie within Kru. *African linguistics on the prairie: Selected papers from the 45th Annual Conference on African Linguistics*, ed. by Jason Kandybowicz, Travis Major, & Harold Torrence (in press). Berlin: Language Science Press.
- SANDS, BONNY. 2009. Africa's linguistic diversity. *Language and Linguistics Compass* 3.559–580.
- SANDS, BONNY, AMANDA L. MILLER, & JOHANNA BRUGMAN. 2007. The lexicon in language attrition: The case of N|uu. *Selected proceedings of the 37th Annual Conference on African Linguistics*, ed. by Doris L. Payne & Jaime Peña, 55–65. Somerville, MA: Cascadilla Proceedings Project.

- SAVÀ, GRAZIANO, & MAURO TOSCO. 2006. La mort des langues en domaine chamito-sémitique. *Faits de Langues* 26/27.279–290.
- SCHAEFER, RONALD P., & FRANCIS O. EGBOKHARE. 1997. Assessing language endangerment in Africa. *Mid-America Linguistics Conference: Proceedings*, ed. by Clifton Pye, 396–406. Lawrence: University of Kansas.
- SCHNOEBELN, TYLER. 2009. (Un)classifying Shabo: Phylogenetic methods and results. *Proceedings of the Conference on Language Documentation and Linguistic Theory 2*, ed. by Peter K. Austin et al., 275–284. London: School of Oriental and African Studies.
- SEKERE, BIHELA. 2011. Secondary education for San students in Botswana: A new Xade case study. *Diaspora, Indigenous and Minority Education* 5(2).76–87.
- SIMONS, GARY F., & M. PAUL LEWIS. 2013. The world's languages in crisis: A 20 year update. *Responses to language endangerment: In honor of Mickey Noonan*, ed. by Elena Mihas et al., 3–19. Amsterdam: John Benjamins.
- SOMMER, GABRIELE. 1992. A survey on language death in Africa. *Language death: Factual and theoretical explorations with special reference to East Africa*, ed. by Matthias Brenzinger, 301–417. Berlin & New York: Mouton de Gruyter.
- SOMMER, GABRIELE. 1995. *Ethnographie des Sprachwechsels: Sozialer Wandel und Sprachverhalten bei den Yeyi (Botswana)*. Cologne: Rüdiger Köppe.
- STIRTZ, TIMOTHY M. 2011. *A grammar of Gaahmg: A Nilo-Saharan language of Sudan*. (LOT Dissertation Series 296). Utrecht: LOT.
- TAMANJI, PIUS N. 2008. Globalization and African languages: Regression in linguistic diversity. *Explorations into language use in Africa*, ed. by Augustin Simo Bobda, 71–94. (Duisburger Arbeiten zur Sprach- und Kulturwissenschaft 70). Frankfurt am Main: Peter Lang.
- TANAKA, JIRO, & KAZUYOSHI SUGAWARA, eds. 2010. *An encyclopedia of |Gui and ||Gana culture and society*. Kyoto: Laboratory of Cultural Anthropology, Graduate School of Human and Environmental Studies, Kyoto University.
- TOURNEUX, HENRY, EDEMA ATIBAKWA BABOYA, MARCEL DIKI-KIDIRI, AND MARIE-CLAUDE SIMEONE-SENELLE. 2000. Langues en danger en Afrique noire. *Les langues en danger*, 45–55. (Mémoires de la Société de Linguistique de Paris, new series, vol. 8.) Louvain: Peeters.
- TRAILL, ANTHONY. 1973. “N4 or S7”: Another Bushman language. *African Studies* 32(1). 25–32.
- TRAILL, ANTHONY. 1999. *Extinct: South African Khoisan languages*. Compact disc & booklet. Johannesburg: Department of Linguistics, University of the Witwatersrand.
- UGORJI, C. U. C. 2005. The Igbo language: Endangerment and empowerment. *Trends in the study of languages and linguistics in Nigeria: A Festschrift for Philip Ajujuobi Nwachukwu*, ed. by Ozo-mekuri Ndimele, 161–170. Aba, Nigeria: National Institute for Nigerian Languages.
- UNESCO AD HOC EXPERT GROUP ON ENDANGERED LANGUAGES. 2003. *Language vitality and endangerment*. Document submitted to the International Expert Meeting on UNESCO Programme safeguarding of Endangered Languages, Paris, 10–12 March 2003. Paris: UNESCO.
- URUA, ENO-ABASI E. 2004. Language marginalization: The Lower Cross experience. *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, ed. by Akinbiyi Akinlabi & Oluseye Adesola, 559–571. (World Congress of African Linguistics/Congrès Mondial de Linguistique Africaine 4). Cologne: Rüdiger Köppe.

- VISSER, HESSEL. 2001. *Naro Dictionary: Naro-English, English-Naro*. 4th ed. Ghanzi, Botswana: Naro Language Project & SIL International.
- WESTPHAL, E. O. J. 1974. Correspondence: Notes on A. Traill: "N4 or S7?" *African Studies* 33.243–247.
- WINTER, J. C. 1981. Die Khoisan-Familie. *Die Sprachen Afrikas*, ed. by Bernd Heine, T. Schadeberg, & E. Wolff, 329–374. Hamburg: Helmut Buske.
- ZABORSKI, ANDRZEJ. 2013. What type of typology of Afroasiatic? *Folia Orientalia* 50. 377–386.
- ZELEALEM LEYEW. 2003. *The Kemantney language: A sociolinguistic and grammatical study of language replacement*. (Kuschitische Sprachstudien 20). Cologne: Rüdiger Köppe.
- ZELEALEM LEYEW. 1998. Small languages and small language communities, 27: An Ethiopian language on the verge of extinction: K'emant, a preliminary sociolinguistic survey. *International Journal of the Sociology of Language* 134.69–84.
- ZELEALEM LEYEW. 2004. The fate of endangered languages in Ethiopia. *Endangered languages and linguistic rights: On the margins of nations; Proceedings of the Eighth FEL Conference, Barcelona (Catalonia), Spain, 1–3 October, 2004*, ed. by Joan A. Argenter & R. McKenna Brown, 35–44. Bath: Foundation for Endangered Languages.

# The Nata Documentation Project

## AN OVERVIEW

Joash J. Gambarage, Andrei Anghelescu,  
Strang Burton, Joel Dunham, Erin Guntly,  
Hermann Keupdjio, Zoe Wai-Man Lam,  
Adriana Osa-Gomez, Douglas Pulleyblank, Dayanqi Si,  
Yoshiko Yoshino, and Rose-Marie Déchaine

### 3.1. Goals of the project

The Nata documentation project started as a year-long Field Methods course at the Department of Linguistics, University of British Columbia (UBC) in 2012–2013 (September 2012 to April 2013). A collaborative research group—which includes faculty members, graduate and undergraduate students, and off-campus volunteers—called the Nata Working Group (NaWoG) was established in June 2013. NaWoG has two goals: first, to produce and maintain a record of the Nata language via descriptive grammatical sketches<sup>1</sup>; second, to analyze language phenomena in Nata that challenge or shed light on current linguistic theories. Integration of data collection with theoretical analysis is the cornerstone of this project.

### 3.2. The Nata language: an introduction

Nata, called Ekinata by speakers, is an under-described Eastern Bantu language mainly spoken in Nata Mbiso, Nata Motukeri, and Makondose (Serengeti District, Mara Region, northwestern Tanzania), with some speakers in Mugeta and Kyandege (Bunda District), as indicated by the star in Figure 3.1. In Guthrie's

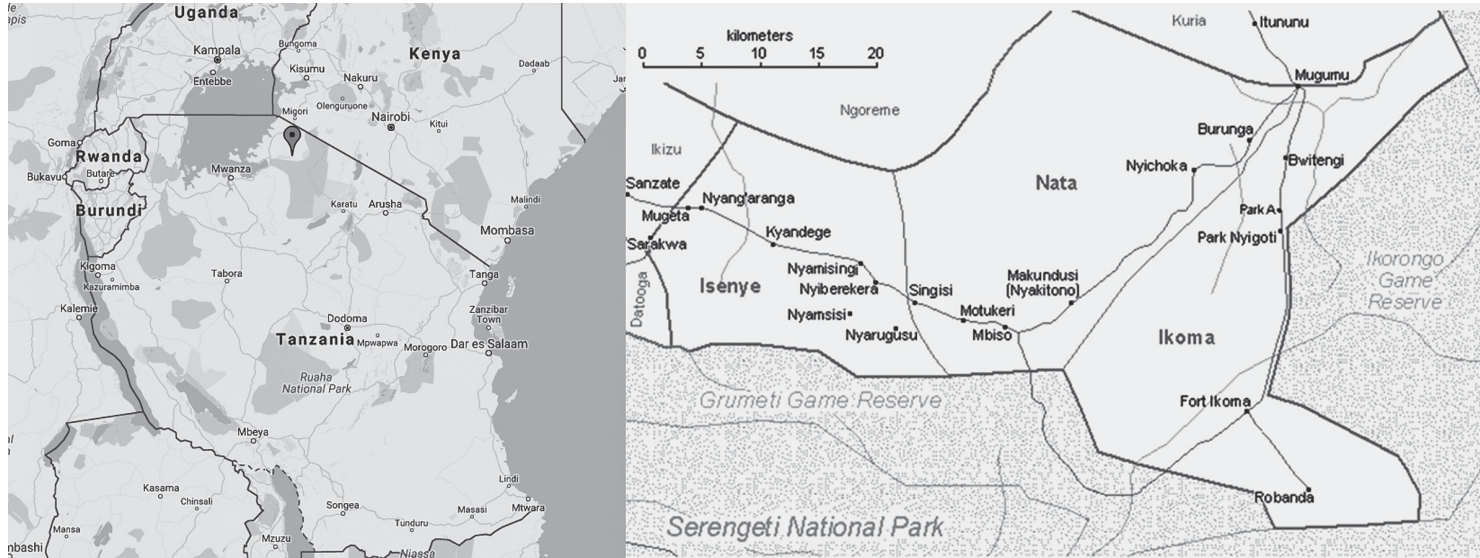


FIGURE 3.1. *Geographical location of the Nata speech community.*