

# Dialogue on Dialect Standardization



# Dialogue on Dialect Standardization

Edited by

Carrie Dyck, Tania Granadillo, Keren Rice  
and Jorge Emilio Rosés Labrada

**CAMBRIDGE**  
**SCHOLARS**  

---

**P U B L I S H I N G**

Dialogue on Dialect Standardization,  
Edited by Carrie Dyck, Tania Granadillo, Keren Rice  
and Jorge Emilio Rosés Labrada

This book first published 2014

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Copyright © 2014 by Carrie Dyck, Tania Granadillo, Keren Rice, Jorge Emilio Rosés Labrada and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-6661-X, ISBN (13): 978-1-4438-6661-3

# TABLE OF CONTENTS

List of Illustrations .....	vii
List of Tables .....	ix
Chapter One.....	1
Introduction	
Carrie Dyck, Tania Granadillo, Keren Rice and Jorge Emilio Rosés Labrada	
Chapter Two .....	9
Multidialectal Orthographies: An Approach to Systematically Spelling Differing Dialects	
Peter Unseth	
Chapter Three .....	23
Addressing Writing System Issues in Dialectal Lexicography: The Case of Cypriot Greek	
Spyros Armosti, Kyriaki Christodoulou, Marianna Katsoyannou and Charalambos Themistocleous	
Chapter Four .....	39
The Standardization of a Latin-based Orthography for Podlachian	
Jan Maksymiuk	
Chapter Five .....	55
The Importance of Identity and Affiliation in Dialect Standardization	
Mark E. Karan, Kerry M. Corbett	
Chapter Six .....	63
Orthography Development for the Greek Cypriot Dialect: Language Attitudes and Orthographic Choice	
Aspasia Papadima, Ioli Ayiomamitou, Stelios Kyriacou and Georgios Parmaxis	

Chapter Seven.....	81
Orthography as a Marker of Group Identity in Dialects John M. Clifton	
Chapter Eight.....	99
The Development of a Writing System for Multiple Dialects of Caribbean English Creole Ken Decker	
Chapter Nine.....	117
Negotiating the Roles of Orality and Literacy in Iroquoian Languages Carrie Dyck, Mary Joy Elijah and Amos Key, Jr.	
Chapter Ten .....	127
Standardization in Language Revitalization André Bourcier	
Chapter Eleven .....	145
Individuality versus Unity in Mixean: Challenges in Orthography Design Carmen Jany	
References .....	161
Contributors.....	179
Index.....	187

## LIST OF ILLUSTRATIONS

4-1. Podlachian Province .....	43
6-1. Preferences regarding the Greek word <χέρι> ['çéri] realised as ['jérin] in the dialect (Section 2-Question 1).....	74
7-1. Location of Vanimó Language.....	83
7-2. Location of Kok Borok and Kau Brung Languages .....	86
7-3. Location of Bine Language .....	88
7-4. Location of Tanchangya and Chakma Languages.....	92
10-1 Northern Tutchone dialect variation ©YNLC, reprinted by permission.....	142





## LIST OF TABLES

1-1. Language ISO codes.....	4
2-1. Majang 2nd person imperative prefixes from two dialects .....	18
3-1. Common words between Cypriot Greek and Standard Modern Greek that are missing from all three dialectal dictionaries mentioned .....	33
4-1. Examples of local name misrepresentations in the Orla commune.....	44
4-2. The consonant inventory of Podlachian .....	46
4-3. Comparison of Podlachian notation with Belarusian, Ukrainian and Polish.....	49
4-4. One sentence in different Podlachian orthographies .....	52
5-1. Language of Preference (Corbett 2012, 99):.....	59
6-1. The phonological system of consonants in Greek Cypriot.....	66
6-2. The phonological system of consonants in Standard Modern Greek.....	67
6-3. The frequency with which GCs encounter the written form of their dialect according to the occasion (Section 1-Question 1) .....	71
6-4. The frequency with which GCs use each writing method (Section 1-Question 2). .....	72
6-5. GC participants' orthography preferences for the six words (Section 2-Question 2) .....	75
6-6. The popularity of two writing methods in the case of the words [tʃi'la] 'belly' and [strondzi'los] 'round' .....	76
6-7. The distribution of the participants' answers for three suggested writing methods (Section 2- Question 3). .....	76
6-8. A two dimensional cross-tabulation of the extent of preference for each orthographic convention of the word ['ferin] in Section two, in relation to the extent of preference the typographic conventions GC speakers use in writing GC.....	77
7-1. Chakma Alphabet.....	93
7-2. Tanchangya Alphabet.....	94
8-1. Scalability of attributes as presented by Fishman (2010).....	101
8-2. Some examples of phonological variation .....	109
8-3. Examples of spelling preferences.....	110
8-4. Variation in some Creole pronouns and tense markers .....	112
11-1. Orthographies compared .....	153

11-2. Comparison of graphemes for consonants..... 154  
11-3. Comparison of graphemes for palatalized consonants ..... 155  
11-4. Comparison of graphemes for vowels..... 155

# CHAPTER ONE

## INTRODUCTION

CARRIE DYCK, MEMORIAL UNIVERSITY OF NEWFOUNDLAND

TANIA GRANADILLO, WESTERN UNIVERSITY

KEREN RICE, UNIVERSITY OF TORONTO

JORGE EMILIO ROSÉS LABRADA, WESTERN UNIVERSITY

This collection of papers grows out of a workshop presented at the conference *Methods 14* at Western University in August 2011. The aim of the workshop was to provide a space for the development of dialogue between dialectologists, language community activists, and others working on the development of orthographies regarding the issues that arise during the creation of writing systems in places where there is dialect variation and there is no writing system, or there is a writing system for a national language but not for the particular language.

Recent years have seen increased attention to issues related to the development of writing systems; see, for instance, Lüpke (2011), Sebba (2007), Seifart (2006), and Venezky (2004), among many others. We are, however, not aware of work that focuses on the pressure for standardization when there is dialect variation. In our own experiences in such situations, standardization to a single system of spelling for a language often comes prematurely, and can, in fact, be a barrier to the development of literacy rather than a support.

This collection contains chapters based on many of the talks presented during the workshop, as well as an invited paper. The participants in the workshop left feeling what we might call a sigh of relief, realizing that they were not alone in seeing that complex issues underlie the assumption that standardization is to be valued.

The chapters in this volume address two major themes. First, the imperative for standardization is influenced by non-neutral as opposed to “scientific” factors, including identity, age, ease of use of the language, and familiarity. Second, the assumption of the value of standardization in

many cases leads to overt or covert negotiations or conflicts in the process of language planning and orthography development.

The book consists of ten chapters and is concerned with languages in various parts of the world, including Cyprus, Poland, Canada, the Caribbean, and Mexico, among others. Languages include those for which there have long been writing systems for “standard” dialects (e.g., Cypriot Greek and Podlachian, which is sometimes said to be a Belarusian-Ukrainian variety) and those for which writing has been only recently introduced (e.g., Cayuga and Oneida, Mixean). Table 1-1 below lists the languages and ISO codes for languages discussed in this volume. Many of the authors argue that standardization is problematic in the particular situation; others seek standardized systems, while at the same time recognizing barriers to this. What shines through all of the papers is the importance of considering factors such as individual varieties as well as languages, the role of identity and affiliation, choice of writing systems, and the history of writing in an area in thinking about writing systems.

Language	Code	Location
Anyi	[any, mbt]	Côte d’Ivoire
Arabic	[ara]	Middle East
Awad Bing	[bcu]	Papua New Guinea
Belarusian	[bel]	Belarus, Russia, Ukraine, Poland
Bine	[bon]	Papua New Guinea
Bislama	[bis]	Vanuatu
Bouyei	[pcc]	China
Caribbean English-lexified Creole	[aig], [bah], [bzj], [bzk], [fpe], [gcl], [gpe], [gul], [gyn], [icr], [jam], [svc], [tch], [tgh], [trf], [vic]	Caribbean
Catalan	[cat]	Spain
Cayuga	[cay]	Canada
Chakma	[ccp]	Bangladesh, India
Chinese	[cmn]	China
Czech	[ces] / [cze]	Czech Republic
English	[eng]	USA, Great Britain, Canada, Australia and others

Language	Code	Location
French	[fra]	France
Fur	[fvr]	Sudan
Gbe	[gbe], [gbs]	Togo, Ghana, Benin, Nigeria
German	[deu]	Germany
Greek (Cypriot)	[ell]	Cyprus
Greek (Modern Standard)	[ell]	Greece, Cyprus
Haitian	[hat]	Haiti
Ibo	[ibo]	Nigeria
Indonesian	[ind]	Indonesia
Inuktitut (Inuinnaqtun; Ulukhaqtuuq)	[iku]	Canada
Jamaican	[jam]	Jamaica
Kabuverdianu	[kea]	Cape Verde islands
Kaingang	[kcp]	Brazil
Kalenjin	[kln]	Kenya
Kau Brung	[usi]	Bangladesh
Kichwa	[que], [quw], [qxl], [qvj], [qvi], [qug], [qxr], [qud]	Ecuador
Kok Borok	[trp]	Bangladesh, India
Majang	[mpe]	Ethiopia
Mixe	[mco], [mir], [mto], [mxp], [mxq], [mzl], [neq], [pxm]	Mexico (Oaxaca)
Mohawk	[moh]	Canada, USA
Nenets	[yrk]	Russia
Occitan	[oci]	France
Oneida	[one]	Canada
Oromo	[gax], [orm]	Ethiopia, Kenya
Papiamentu	[pap]	Curacao
Picard	[pcd]	France
Podlachian (East Slavic vernacular)	N/A	north-eastern part of Poland
Polish	[pol]	Poland
Portuguese	[por]	Portugal, Brazil
Russian	[rus]	Russia

Language	Code	Location
Sango	[sag]	Central African Republic
Seselwa	[crs]	Seychelles
Spanish	[spa]	Spain and Latin America
Tanchangya	[tnv]	Bangladesh
Tetun Dili	[tdt]	East Timor
Tok Pisin	[tpi]	Papua New Guinea
Tonga	[toi]	Mozambique and South Africa
Tumbuka, Senga	[tum]	Malawi, Mozambique, and Zambia
Tutchone (Northern)	[ttm]	Canada
Ukrainian	[ukr]	Ukraine
Vanimo	[vam]	Papua New Guinea
Yao (ciYao)	[yao]	Malawi, Mozambique, and Zambia

**Table 1-1. Language ISO codes**  
**(<http://www-01.sil.org/ISO639-3/codes.asp>)**

Chapter 2, “Multidialectal Orthographies: an approach to systematically spelling differing dialects”, by Peter Unseth reviews four approaches to spelling dialects that differ in phonological characteristics and advocates a “systematic multidialectal approach” (p. 9) for creating a unified orthography. Unseth terms this approach “another tool in the tool kit” (p. 21) for developing orthographies. The principles of spelling for one dialect can differ from another dialect, as long as the resulting system is consistent and systematic. Some compromises (over- and under-differentiation) may result. Unseth advocates using either “shallow” historical forms, or morphophonemic representations. One result of this is over-differentiation: for example, the orthography might represent contrasts that are retained in one dialect, but merged in another or represent sounds that are preserved in one dialect (or register) but not in another. Another potential result of the choice of “shallow” historical forms is under-

differentiation: for example, using an historic phoneme to represent its reflexes in all dialects; e.g., in Niger-Congo Gbe using the spelling <ph> (from proto- \*ɸ) to spell /ɸ, p, χ<sup>w</sup>/.

Chapter 3, “Addressing writing system issues in dialectal lexicography: the case of Cypriot Greek”, by Spyros Armostis, Kyriaki Christodoulou, Marianna Katsoyannou, and Charalambos Themistocleous, together with chapter 6, deals with the interesting case of a dialect (Cypriot Greek, CG) which exists in the shadow of a standard dialect (Standard Modern Greek, SMG). Both chapters deal with the problem that the SMG spelling system does not adequately represent CG (in particular, the CG palatoalveolar phonemes). A related problem is the legacy and existence of alternative, informal orthographies for CG. For the authors of chapter 3, orthography standardization is driven by the need to have a standard way of representing CG words in a CG dictionary and related materials (the Syntychieis project). The principles underlying the orthography developed for the Syntychieis project are transferability to and from SMG, and a relatively close adherence to phonological form. The authors also discuss the problem of using characters (especially diacritics) that are not in the Unicode set; while they develop some work-arounds, they ultimately recognize the need to propose new glyphs for Cypriot Greek in the Unicode set (i.e., glyphs that would include diacritics, removing the need for using combining diacritics).

Chapter 4, “The Standardization of a Latin-Based Orthography for Podlachian”, by Jan Maksymiuk addresses issues that have arisen in the attempt to design a writing system for Podlachian, an East Slavic vernacular related to Belarusian and Ukrainian that is spoken in the north-eastern part of Poland. Podlachian does not have a standardized writing system, in part because there is not official recognition of the Podlachian as an entity in Poland. There are at least six alternative orthographies to represent Podlachian; some are Latin-based, and some, Cyrillic-based. The author of chapter 4 and Aleksander Maksymiuk took it upon themselves to develop and promote a Latin- (as opposed to Cyrillic-) based orthography for Podlachian. Maksymiuk describes similarities and differences between Podlachian and related East Slavic (Belarusian and Ukrainian) varieties, and proposes an orthography which accommodates three extra diphthongs and captures related processes of palatalization present only in Podlachian. From a practical perspective, the orthography represents Podlachian phonemically, and uses graphemes that are relatively accessible (for computer users or typists). The chapter also describes a fascinating, grass-

roots effort to promote an ethnic identity and unified orthography for Podlachian.

Chapter 5, “The importance of identity and affiliation in dialect standardization”, by Mark E. Karan and Kerry M. Corbett argues that a particular group’s choice of language, dialect, and (by extension) orthography is not based on purely linguistic principles, but rather on community behaviour or the degree to which the group identifies with a larger group or standard. The authors underscore a common thread in the book, namely that social practices, not linguistic principles alone, help determine the success or failure of standardization efforts (c.f., Sebba 2007).

Chapter 6, “Orthography Development for the Greek Cypriot Dialect: Language Attitudes and Orthographic Choice”, by Aspasia Papadima, Ioli Ayiomamitou, Stelios Kyriacou and Georgios Parmaxis, like chapter 3, addresses Cypriot Greek. In contrast to chapter 3, chapter 6 describes some of the sociocultural factors that influence the development a standardized orthography. For example, the authors point out that graphemes have a cultural meaning for their users, and are not solely an abstract means of representing sounds. The authors designed several sociolinguistic surveys to determine attitudes and preferences towards various ways of representing the palatoalveolar consonants particular to CG (and missing from MSG). They found that orthographic preferences were influenced by the similarity of CG spellings to the orthography of SMG, although this was more true of older speakers than of younger speakers; the transparency (the degree to which the grapheme(s) accurately represent CG sounds); and ease of use. They also uncovered a mismatch between CG user’s “beliefs or attitudes and their actual choices of orthographic conventions” (p. 79).

Chapter 7, “Orthography as a Marker of Group Identity in Dialects”, by John M. Clifton draws on several fascinating, first-hand experiences in Papua New Guinea and Bangladesh. Clifton demonstrates the influence of identity and affiliation on the development of writing systems. The grassroots decision-making processes described in this chapter resulted in quite different solutions to the problem of over- and under-representation of phonemic contrasts in related dialects. Decisions were also greatly influenced by community members’ desire to either show or disavow similarities between mutually-intelligible dialects.

Chapter 8, “The Development of a Writing System for Multiple Dialects of Caribbean English Creole”, by Ken Decker provides some historical context for the development of writing systems in general



(discussing the development of written French). The author points out that writing systems for major European languages evolved organically over centuries. In contrast, with some notable exceptions (see CG in chapters 3 and 6), development of orthographies over the past century has been engineered, sometimes through negotiation; examples are provided. The author then turns to a discussion of English Caribbean Creole varieties. With some exceptions (e.g., Jamaica), English Creole varieties are perceived as low-status and are not used in official domains such as education, factors which inhibit the development of Creole writing systems. Decker advocates a pan-English Caribbean Creole writing system, but points out the many barriers to creating one (or more than one). The chapter includes an interesting discussion about speakers' desires to preserve distinct allophonic pronunciations and even ad-hoc spellings, in order to emphasize the distinctness of their Creole.

Chapter 9, "Negotiating the Roles of Orality and Literacy in Iroquoian Languages", by Carrie Dyck, Mary Joy Elijah, and Amos Key, Jr. discusses the history and legacy of literacy among the Cayuga- and Oneida- (Iroquoian-)speaking peoples in Ontario, Canada. The discussion is framed within an Indigenist Research Paradigm, or holistic paradigm. The authors discuss both community-developed (including community-modified) writing systems and orthographies developed by linguists. The manner in which the writing systems were developed is described as a hindrance to language (re-)vitalization. The authors discuss the tensions caused by developing a written dimension for historically oral cultures. For example, they discuss the problem of sacred oral texts, which, once written down, can theoretically be accessed by anyone, including people for whom the texts are not intended. They advocate developing a more coherent, holistic, and community-oriented approach to creating a written language that supports language vitalization efforts.

Chapter 10, "Standardization in Language Revitalization", by André Bourcier, in contrast to other chapters in this volume, questions whether language standardization is necessary and argues that it may be detrimental to language revitalization efforts. Bourcier argues that at least in a language endangerment context, orthographies should support the acquisition of, and preserve, dialect variation, instead of promoting a standard. He argues that standardization is an instance of language planning or engineering and discusses the political and economic underpinnings of language planning. Bourcier provides an example from the Inuinnaqtun dialect of Inuktitut, showing that the official orthography, which more accurately represents Inuinnaqtun pronunciations, fails to

acknowledge the community's attachment to their traditional orthography (developed by the Anglicans). The result is a somewhat unsatisfactory hybrid system that "created major problems in language acquisition planning" (p. 135). The main problem is that, for Canadian Aboriginal languages in general, the first dialect to be recorded is treated as the "standard", leading to a situation where language teachers who speak a different dialect perceive their own speech to be substandard; this in turn creates confusion for students, particularly in an endangerment context. Bourcier advocates a culturally-appropriate kind of literacy, which is based on knowing the author of the variety and utterance/speech being studied, instead of teaching an anonymous standard. Language speakers are encouraged to learn "...the dialect of their kin Elder, or to "adopt" an Elder as a speaker model" (p. 141). Bourcier presents this model as an intermediate stage in language planning, one which avoids confusion while promoting language revitalization.

Chapter 11, "Individuality versus unity in Mixean: Challenges in Orthography Design", by Carmen Jany discusses a language group (Mixean) which is robust, rich, diverse, and not endangered, a very different situation from those addressed in chapters 9 and 10. The main challenge for orthographic standardization in this case is the hundreds of Mixean dialects, many of which are mutually unintelligible and whose genetic affiliation is not well understood. Like other authors in this volume, Jany claims that orthography design is not based solely on linguistic principles but also on "pedagogical, sociopolitical, and practical" principles (p. 145), including the degree of similarity to Spanish orthography. Depending on how Mixean communities balanced these principles, quite different solutions to orthography design have been implemented, leading to a situation where it is difficult to develop a unified writing system that facilitates standardized reading and writing practices. Jany's chapter is a detailed case study, illustrating these points by describing the phonemic inventory of Mixean varieties, their genetic affiliation, and the orthographies that have been developed for these communities.

We hope that this volume provides more food for thought for those considering orthography development and dialects. If anything is clear from these case studies, it is that there is no simple solution to the problems that arise, and that first and foremost, the practices, ideas and opinions of the users should be sought out to develop a truly practical and useful orthography that will be used by the speakers.

## CHAPTER TWO

# MULTIDIALECTAL ORTHOGRAPHIES: AN APPROACH TO SYSTEMATICALLY SPELLING DIFFERING DIALECTS

PETER UNSETH  
GRADUATE INSTITUTE OF APPLIED LINGUISTICS  
AND SIL, INTERNATIONAL

### 1. Introduction

This chapter reviews different approaches to spelling dialects and related languages. It advocates and illustrates an approach that is informed by both phonology and sociolinguistics: a systematic multidialectal approach that advocates spelling in ways that allow all dialects to read and pronounce literature by systematically applying the rules of their own dialect's phonology. This is written with small but vital language communities in mind, not for revitalization efforts such as described by Bourcier (this volume). This is written for contexts where communities want to write their languages in a uniform way and actively produce literature in their languages, in contrast to the approach described by Dyck, Elijah, and Key (this volume). Also, I am writing this for situations where different dialects actually want to write in a uniform way, unlike the situation in Bangladesh described by Clifton (this volume).

A variety of approaches has been tried to spell languages that have significant dialect variation. These can be broadly classified into four categories.<sup>1</sup> The “Unilectal” approach represents a single, standard dialect and all readers spell it the same way. The “Union” approach is the result of writing different language features according to different dialects, so that all dialects are included in some way, though not systematically. The

---

<sup>1</sup> My work builds on and expands Simons (1994).

“Incomplete” approach lists symbols, without working out the details by which they will be used in spelling dialects. The “Multidialectal” approach works to systematically spell in a way that is psycholinguistically accessible to multiple dialects.

This chapter advocates building on shared linguistic features in orthography development (Simons 1994 and Schroeder 2010), but opposes the idea of creating unified orthographies that are so driven by sociolinguistic motivations that they ignore significant linguistic differences, an approach which sometimes requires readers to learn another language rather than learning to read their own.

Examples will be drawn from a variety of languages, including the author’s field work in Ethiopia. I admit that the examples, the sources cited, and some of the topics addressed will reflect the fact that much of the writing on this topic has dealt with African languages. I trust that the principles I present are clear and equally applicable to readers and language communities around the world.

By way of definition, an orthography is much more than just a set of symbols. It encompasses the rules to map sounds to the symbols, spelling rules for loan words, rules for morphophonemic changes, often spelling some irregular forms, etc.

When language names are cited, they are usually followed by the ISO 639-3 three letter identifying code in square brackets (as best as I can identify it), e.g., Tonga [toi]. I have inserted these into quoted lists of languages, also.

### **1.1 What this paper does NOT address**

This paper specifically does not address two types of situations. First, it does not address the question of methods to use for related, but non-intelligible varieties using the same basic orthography, as envisioned by Prah:

We have found that practically all the South Central Bantu languages can use the same orthography. These include languages in Malawi, Mozambique, and Zambia. They include languages which are as varied as ciTumbuka/ciSenga [tum], ciYao [yao], eLomwe [lon and ngl]... These languages are not all mutually intelligible to any significant degree. They are separate languages, but structurally so similar that they can utilise the same basic orthography. (Prah 2009, 19)

As evidence that such unified orthographies can be done, Prah cites *A Unified Standard Orthography for South-Central African Languages* (Banda et al. 2001), but this book is an incomplete description of an orthography.

If people cannot understand what each other say or write, what is the benefit from adopting a common orthography? If I understand the position correctly, it would be similar to advocating that all Romance languages (e.g., French, Italian, Romanian) use the same orthography. The users of such an orthography would not understand what was written in the other languages, they could only share a mystical orthographic unity.

However, I admit that my opinion on this is at odds with Kwesi Kwaa Prah (1998, 2002a, 2002b, 2009), Phillip Oketcho (2010), and a number of other writers from Africa, many published by the Centre for Advanced Studies of African Society, CASAS. They have made several calls for having a standard set of symbols used across related languages, different writers calling for varied degrees of scope, across various degrees of language relationship. It is not always clear which of these writers are merely calling for similar orthographies among related languages and which writers optimistically assume that a harmonized orthography will enable readers of different languages to read the same materials with comprehension. Those who assume greater widespread intelligibility across speech varieties may mentally minimize the difference between these two categories.

It is worth noting that among Bantu languages, there seems to be more shared phonological similarity than among some other language families, leading to the possibility of greater orthographic similarity than among some other language families (Schroeder 2010). It is not coincidental that many of the CASAS authors have written about Bantu languages.

The second topic that I am not addressing is the harmonising of existing orthographies for single languages that overlap national borders. Nguna and Siteo (2003, 44) lay out plans for harmonizing the spelling of Tonga [toi] across the border of Mozambique and South Africa. Similarly, Borana Oromo [gax] is spelled quite differently in Ethiopia and Kenya. Cross-border coordination on orthography is a worthwhile goal, but is not the topic of this paper.

However, I will briefly point out a complexity some have overlooked, a point that some will see as major but others will see as minor: if the official language on different sides of a border is different, there will likely be pressure (from inside and outside the language community) to spell local languages in ways that are harmonized with the official language.

For example, Chumbow called for harmonising the spelling of languages that straddle the Nigeria-Cameroon border (Prah and King 1998, 58). On the Cameroon side French-based spellings are generally used, but on the Nigerian side English-based spellings are used. Should the vernacular languages be spelled according to French spelling conventions or English spelling? If a unified orthography is used in such areas, it would inevitably reduce the transfer of orthographic and reading skills between the vernacular and (at least one) of the official languages. Similar challenges arise on the border of Mozambique (Portuguese official) and Zambia (English official), and also the border of Equatorial Guinea (Spanish official) and Gabon (French official), and also for the Sámi languages spanning the borders of Norway and Finland. In some places, cross-border spellings are even more complicated by the fact that different scripts are used on each side of the border, such as Belarus and Poland (Maksymiuk, this volume), India and Pakistan, Serbia and Croatia.

## 2. Two general positions

When considering the creation of an orthography to span dialects, there are two general positions. Some orthographers are more strongly influenced by ethnic identity. Others are more strongly influenced by linguistics. When the ethnic identity is consciously shared, there is more pressure to create an orthography that unifies dialects, despite linguistic differences, such as work among the Oromo [orm] groups of Ethiopia. When the ethnic identity is divided, there is pressure to create different orthographies, despite linguistic similarities, such as among the Serbs, Croats, Montenegrins, and Bosnians.

Unfortunately, some *start* with the overwhelming conviction of ethnic unity, believing that a multidialectal orthography can be designed for a certain group of speech varieties before the requisite linguistic research is conducted. However, when these “comprehensive descriptions” are eventually done, might the studies show that the language varieties differ so much that they cannot use a single comprehensive orthography after all?

When a comprehensive description of each of the various varieties has been done, we can approach the issue of standardising the varieties. This can be done through the development of a comprehensive orthography for these varieties. (Eno-Abasi 2002, 25)

Beginning from a similar position, in 1837 Ján Kollár called for “a uniform, philosophic orthography [...] which all Slavs can use” (Maxwell 2003, 135). Then he outlined a three-level plan for education as Slavs learned not only to read their own “dialects”, but to read other Slavic dialects. But his plan was too ambitious, ordinary people could not learn to read these multiple levels of Slavic, but only scholars: “The “nation” remains in this way a living-room plant; only philologists can be real Slavs” (Theodore Locher quoted by Maxwell 2003, 136).

Based in Cape Town, and led by Kwesi Kwaa Prah, CASAS has worked hard to promote the idea of unified and harmonized spellings across dialect and national borders (Prah 1998, 2002a, 2002b, 2009). Much of their work is on harmonising different orthographies for identical languages across the borders of countries, in situations fraught with politics more than phonology. But some of their work aims to create orthographies that will span not just dialects, but even languages. They have produced over a dozen proposed harmonized broad orthographies. These are great goals, but they have not yet demonstrated that they can implement these proposed orthographies.

Another example of overly optimistic hopes for a broad harmonized orthography comes from Botswana:

With various dialectal variants that are so distinct phonetically, the greatest challenge among the speakers has been how to harmonize and create a common orthography that will promote a linguistic convergence for all dialectal tendencies falling within Shekgalagarhi. Without a proper framework to undertake this important codification exercise, speakers are often at loss as to how best to resolve basic writing problems, and without a defined role for linguists, speakers do not understand that outsiders could help them better plan for the development of their language. (Monaka 2005, 2)

It appears that the role envisioned by this writer for linguists is to engineer orthographic unity, not to evaluate whether it is advisable or even possible.

Below, I discuss four approaches to unified orthographies, comparing their weaknesses and calling for systematic multi-dialectal approaches.

## **2.1. Unilectal approach**

In this approach, one dialect is chosen as the standard for use by all dialects. The written form of the language follows the norms of that

standard dialect, and speakers of other dialects must learn the pronunciation of this standard dialect. This has been acceptable to people when there is common agreement that a particular dialect is preferred, for a variety of criteria, such as “It is the old, pure way of speaking”, “It is the most geographically central dialect”, “It is the dialect spoken in the area where important facilities are located, such as government offices, schools, markets, church headquarters, etc.”. In some cases, this has required language learning so that students can read what is said to be the students’ “own language”. The unilectal approach has been implemented in a number of places, including major languages, e.g., French, Chinese, Indonesian, where the dialect that is standardized is often that of the capital city.

The unilectal approach provides the “simplest” solution to lexical variation across dialects; it simply requires speakers of other varieties to learn the vocabulary and relevant phonology of the selected dialect.

## 2.2 Union Approach

In the union approach, the orthography does not reflect any single dialect. “standardization may be [...] *by attempting to create a composite of all the main dialects*” (W. H. Whitely quoted by Ansre 1971, 681; emphasis added by Ansre).

The problem with this sort of union orthography is that it represents nobody’s speech. Everybody has to adjust to another dialect, at least in some parts of the orthography. The results have usually been orthographic orphans, unclaimed by any of the dialects. Also, a union approach makes it difficult for the writers to remember how to write words consistently. In some places, it has been taken to extremes and union languages have been created, e.g., Union Kalenjin in Kenya. The motivation was based on sociolinguistic desires for unity rather than on linguistic similarity. It is an admirable idea, but it generally does not work (Angogo 1982, Karan and Corbett, this volume).

The promoters of the written language have created a synthesis from existing dialects; they by so doing have created a new form of language [...]. This method of standardization has not gained much success. Language speakers did not feel comfortable in the promoted standard language which seemed to them to be a foreign language with no affective and cultural attachment or background as a native language often does. It was as if they were called to learn a new language, under the pretext of



neutrality and unity of the concerned linguistic group. (Sadembouo 1989, 13)

One famous example of a union approach is Union Ibo of Nigeria from the 1800's. Missionaries in Nigeria worked to publish in a form of Igbo [iibo] that could be understood by all speakers of Igbo, an engineered form of the language known as Union Ibo. As people discussed and negotiated, there were strong advocates for certain dialects to be used or preserved. "The delegates from Bonny [...] claimed that their Igbo was the real Izuama, the "Parent Ibo language", and if there were to be one translation for all Igbo-speaking people it had to be theirs" (van den Bersselaar 1997, 283).

The chief proponent of the project "stressed that the endeavour was not to produce a translation in the dialect of one district, but one which would be understood in any district" (van den Bersselaar 1997, 283). They "tried to mingle the Bonny and Onitsha [Scripture] translations in such a way as to make it comprehensible to people from either place [...] Union Ibo was not a written form of Owerri Igbo but a newly created dialect" (van den Bersselaar 1997, 283- 284).

The Union vs. [version] is not the language of any one district, and as the majority of the readers of Ibo are in this [Onitsha] district they do not see why their own language should be displaced and superseded by that which is a compound of nearly all the Ibo dialects. (Letter from S. R. Smith 1910, quoted by van den Bersselaar 1997, 284.)

Despite the goal of a unified Igbo, others eventually produced an Onitsha Igbo version (van den Bersselaar 1997, 285) for the Onitsha area.

### 2.3 Incomplete approach

In some cases, enthusiastic orthography promoters have published lists of symbols and announced an orthography, but have not actually worked out how these symbols will be used by the various language varieties involved. For example, in Guatemala, the *Academia de las Lenguas Mayas de Guatemala*, produced a set of symbols to be used in writing all Mayan languages. It was recognized by the government in "1987 as the official version of the alphabet for Mayan languages. It was made up of 51 graphemes" (French 2004, 253-254). An earlier 1976 attempt at a Pan-Mayan orthography had contained "61 graphemes" (French 2003, 493 fn. 17). But announcing a unified orthography and implementing it are not

the same. “The Academy has done almost no work on standardization beyond the establishment of the alphabet” (Grenoble and Whaley 1998, 112).

## 2.4 Systematic multi-dialectal approach

There have been two emphases when developing orthographies for multi-dialectal situations. Most attempts are either ethnically sensitive or linguistically sensitive. On the one end of the continuum we find linguistically sensitive work (e.g., Fine 2003 and Simons 1994), where the sound differences between speech varieties have been systematically compared and then mapped to standardized sets of symbols. These authors recognized different phonemes in the present speech varieties and specifically set out to represent them in a way that as many speakers as possible would be able to read them, each using their own symbol-to-sound mappings. At the other end of the continuum, we find people doing ethnically sensitive work, attempts to include large number of speakers, ethnic unity overriding linguistics.

Orthographies at both ends of the continuum, and all points in between, are generally done by people of good will. However, the successful implementation of an orthography can be greatly enhanced by doing good linguistics, as argued below.

## 3. Multi-dialectal approach

As mentioned in §2.4, it is also possible to develop an orthography that represents the sounds of different dialects in a systematic way, basically by spelling at different levels for different dialects. That is, the spelling system may be more morphophonemic for one dialect, but more surface level phonemic for another. There may also be some points of over-differentiation and/or under-differentiation in various dialects that result. But the result is (ideally) an orthography that can represent all dialects’ speech patterns in a consistent, systematic way. This is a systematic multidialectal orthography.<sup>2</sup>

Developing such a multidialectal orthography is useful when two or more speech forms are different primarily on the phonological level, but are still adequately similar in lexicon and syntax so that one set of written

---

<sup>2</sup> Some orthographies that are labeled as “union” have actually incorporated some such systematic multidialectal features.

materials is seen to be adequate. A multidialectal orthography cannot eliminate the problems of different affixes, roots, or other elements of the dialect's grammar.

The simplest examples of this are to spell historical forms (historical at a somewhat shallow level, not ancient levels).

Very different dialects may have the same or a very similar system of underlying representations. If this is true, then the same system of representations for underlying forms will be found over long stretches of space and time. (Kreidler 2001, 325)

This seems to be similar to the “neo-language” approach of Capo (1989).<sup>3</sup> Some have misunderstood this, fearing that spelling proto-forms could lead to languages losing identity (Miti 2003, 60), but this objection seems to misunderstand the difference between spelling a common, phonologically accessible underlying form as opposed to learning of proto-forms. This highlights the tension between ethnic distinctions and unity.

Koffi (2006, 8) gives an instructive example of spelling a more historical form to cover present day dialect differences from Anyi varieties [any, mbt, etc.] in Côte d'Ivoire. There is a rule that reduces or deletes vowels in initial syllables before approximants. Therefore, in some dialects, “woman” is pronounced in phonetically different forms, such as [blǎ] and [bərə], the last form also undergoing a change of [l] to [r]. Koffi claims that if the word is spelled <bala>, then all dialects will be able to read it and pronounce it according to their own phonology.

The following data, based on Majang [mpe] (Ethiopia) show how two dialects vary in their pronunciation of the plural prefix in the imperative mood:

---

<sup>3</sup> A specific difference is that Capo proposes that his neolanguage become a spoken form of the language (1989, 56).

gloss	root	northern dialect	southern dialect
“run”	reer-	Irreere	inreere
“swim”	ley-	Illeye	inleye
“return”	mur-	Immure	inmure
“try”	paar-	Impaare	inpaare
“come”	kus-	Iṅkuse	iṅkuse
“believe”	nad-	Innade	innade
“drink”	ut-	inʔute	inʔute
“eat”	dám-	indáme	indáme
“deceive”	nun-	iṅpune	inṅpune

**Table 2-1. Majang 2nd person imperative prefixes from two dialects**

The solution to this orthographic problem is to spell at the morphophonemic level, keeping the visual shape of the prefix consistent. This allows everybody to read one set of materials, but pronounce it in their own way. For example, if the plural imperative of “run” is spelled <irreere>, northern dialect speakers could read it, but southern dialect speakers would have difficulty. But if it is spelled <inreere>, then speakers from both can read and understand it.

Sometimes, a multidialectal orthography uses forms that are close to the shape of morphemes historically, though they may be pronounced differently in the various dialects today. This concept is misunderstood by some.

If we were to harmonise the speech forms of these varieties, we would apply the strategies used in comparative reconstruction, namely majority rules strategy and phonetic plausibility. (Miti 2003, 60)

In some cases, a historical phoneme may be pronounced quite differently in related varieties. Emenanjo (1990, 5) gives a good example from Niger-Congo Gbe (varieties spoken in Togo, Ghana, Benin, Nigeria), where /ɸ/ in some dialects corresponds to /p/ in others, and /χ<sup>w</sup>/ in others. By spelling <aphá> in the “neo-Gbe orthography”, it will be pronounced as [àɸá], [àpá], [àχ<sup>w</sup> á] according to the dialect.

Among the Picard dialects of France, the spelling <qh> is pronounced differently according to the local dialect: [k<sup>w</sup>], [k<sup>j</sup>], [k] (Auger 2011). The one spelling allows each dialect to pronounce it their own way, but still read the same written materials.

### 3.1 Lost contrasts in some dialects

Sometimes a phonemic contrast is lost in some dialects, but preserved in others. Or there may have been a phonemic split, resulting in new contrasts in some dialects. In such cases, if the orthography continues to mark the higher number of contrasts, readers of both kinds of dialects can learn to read and pronounce it in their own way. Readers from the dialects that have lost the contrast will pronounce two spellings the same way, such as with English “which” and “witch”, but they will have to learn to spell them differently.

A basic principle (with many exceptions) for orthography preparation is to represent only phonemes, not allophones. This is worth repeating in a discussion of spelling dialects. In his discussion of multidialectal orthography, Capo (1989, 36) gave some good examples of this for Gbe:

In some dialects [...] /s/ and /z/ are palatalized to [ʃ] and [ʒ] before front close vowels, whereas in other dialects they are not; in addition in Ajá dialects only, they are also optionally palatalized before back close vowels. It is recommended that in the dialects concerned, [s] and [ʃ] on the one hand, and [z] and [ʒ] on the other, be represented with same graphemes, *viz* “s” and “z” respectively... For example, Ajá and Awlan [ʃi] would be written <si> “respect” and [ʒi] would be <zi> “push”.

Some dialects of Kaingang [kgp] of Brazil have merged /a/ and /e/. The same literature is being used in all dialect areas, with the /a/ and /e/ distinction preserved in the orthography. In those dialects that have merged these two vowels, they simply “pronounce them alike” (Wiesemann 1989, 7).

In Bouyei [pcc] of China, some varieties have preserved the contrast between /m/ and /ŋ/, but in the Shuicheng variety the contrast has been lost, leaving only /ŋ/ (Fine 2003, 60). For this situation, Fine recommends spelling the two distinct phonemes, <m> and <ng>, and letting varieties like Shuicheng pronounce them both as [ŋ] while other phonologically conservative varieties will pronounce them as /m/ and /ŋ/. Similarly, some varieties have distinct vowel phonemes /a/ and /ɛ/, but some varieties have lost the distinction. A proposal to spell them based on pronunciation that had lost the phonemic distinction used only the symbol <a>, but the current orthography uses <a> for /ɛ/ and <aa> for /a/, allowing other dialects to have the clue to pronounce words according to their phonology.

### 3.2 Unpronounced segments

In some speech varieties, some segments are not pronounced in casual speech. For example, in Puerto Rican Spanish of Ponce, syllable-final obstruents are deleted, e.g., [komo eta] for ¿cómo estás? “how are you?”. Also, word-final nasals are pronounced as [ŋ]. As a result, Spanish *bien* “well” comes out as [bieŋ]. However, when reading and writing, the consonants are all still spelled in the standard way, so that readers from any dialect can understand Puerto Ricans’ writing, and Puerto Ricans can read other dialects.

Similarly, in Andalusian Spanish, final /s/ is deleted, but there is a phonological change to preceding vowels to indicate its loss.

Thus, instead of the contrast *comes/come* “you (familiar/polite) eat” or *libros/libro* “books/book”, one has the contrast *come/come* or *libro/libro* (where /e̞ o̞/ indicate more open vowels). (Hochberg 1986, 610)

The Andalusians pronounce words in their own way, but spell them according to a universally accepted Spanish orthography so that there is no confusion in reading.

### 3.3 Different morphophonemic rules

Between some speech varieties, there are different morphophonemic rules. By spelling the underlying form, all can read it with full understanding and each can pronounce it their own way. In some Pennsylvania dialects of English, word final /l/ is pronounced as [w]. I was once baffled when my Pennsylvania friend told me to turn at [k<sup>h</sup>anaw] Street. When she finally showed me the written form <Canal Street>, then I understood perfectly.

My former pastor’s accent led him to pronounce word final of [θ] as [f]. He writes by the rules of standard English spelling, but pronounces my last name *Unseth* with a final [f]. In both of these cases, standard English spelling allows us all to spell with the same orthography and still pronounce such words in our own ways.

In Awad Bing [bcu] of Papua New Guinea, “forehead” is *damo-*. Between the Eastern and Western dialects, in the third person possessive the root final vowel alternates before [j]: [damo-j] (Eastern) and [dama-j] (Western) “his/her forehead” (Simons 1994:23). If the orthography follows the Western form of <damay> this could cause some confusion to