



2013

The ABD of orthography testing: Integrating formal or informal testing into the orthography development process

Elke Karan
SIL-UND

Follow this and additional works at: <https://commons.und.edu/sil-work-papers>

 Part of the [Linguistics Commons](#)

Recommended Citation

Karan, Elke (2013) "The ABD of orthography testing: Integrating formal or informal testing into the orthography development process," *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*: Vol. 53 , Article 2.

DOI: 10.31356/silwp.vol53.02

Available at: <https://commons.und.edu/sil-work-papers/vol53/iss1/2>

This Article is brought to you for free and open access by UND Scholarly Commons. It has been accepted for inclusion in Work Papers of the Summer Institute of Linguistics, University of North Dakota Session by an authorized editor of UND Scholarly Commons. For more information, please contact zeineb.yousif@library.und.edu.

The ABD of orthography testing:

Integrating formal or informal testing into the orthography development process

Elke Karan

Literacy and Education consultant, SIL International

Developing an orthography for a previously unwritten language or reforming an existing writing system is a complex process. It calls for consideration of all the factors that play into the process and sensitivity to stakeholder expectations. The desired outcomes are acceptance of the orthography by intended users, ease of learning, and general efficiency. Integrating evaluation processes and testing events into the orthography development process and making appropriate adjustments based on the findings will maximize chances of attaining those outcomes.

This article provides a rationale for orthography testing and recommendations for what to test and who to test. It includes a literature review, pointing readers toward helpful resources.

1. Introduction

Field experience clearly shows that orthography development is not only about rigorous linguistic analysis and consequent mapping of contrastive sounds to symbols. A variety of factors affect the orthography development and implementation process.¹ Sociolinguistic identity, for instance, is a weighty factor in script and symbol choices and may determine whether suggested orthographic conventions are applied, ignored, or resisted (see Unseth 2005; Sebba 2007 and 2009). An optimal orthography would not only represent a language well; it would be acceptable to the local population and also be simple and easy to learn. Orthography development thus needs to be a multi-disciplinary process, giving attention to and balancing linguistic, sociolinguistic, and educational factors, among others.

Since orthography is closely linked to allegiance, solidarity and identity, and individual egos may at times come into play, emotions can be very much involved. David Crystal (2000:147) comments: “Linguists...often remark on how emotionally stressful this sociopolitical context can be. Traditionally, there is nothing in a linguist’s training which prepares for it.”

The good news is that publications on principles of orthography development are increasing and becoming more readily available.²

However, few helps are currently available on the specific topic of *orthography testing*. This article explains how orthography testing fits into the larger orthography development plan and provides a brief literature review. Differentiating between formal and informal testing, it offers some guidelines for orthography testing, based on existing publications, as well as on unpublished training materials currently used in SIL courses and workshops.

My objectives are to encourage the integration of orthography testing in the orthography development process and to help equip field workers to do this effectively.³ The expectation is that

¹ These factors are listed and/or expounded upon in various publications. Cahill and Karan (2008) provide a summary.

² See Berry 1958; Smalley 1959; Smalley 1964 (ed.); Sjoberg 1966; Powlison 1968; Venezky 1970; Gudschinsky 1973; Fishman 1977 (ed.); Berry 1977; Wiesemann et al. 1983; Van Dyken and Kutsch Lojenga 1993; Rogers 1995; Baker 1997; Coulmas 1999; Snider 2001; Robinson and Gadelii 2003; Unesco 2004; Karan 2006; Seifart 2006; Cahill and Karan 2008; Schroeder 2010.

research results will reveal if the orthography being proposed is indeed acceptable to community members, is easily learned for reading and writing, and works efficiently for all the speech varieties it is to serve.

2. Orthography testing as a component of the participatory language development process

Increasingly, language development agents recognize community participation in the orthography development process as “good practice.” The intended users of the orthography need to be favorably disposed toward it in order to subscribe to it. Popular favor may be elusive if the language community itself is not involved in the decision-making process. If outsiders, or some governmental agency, create (or reform) an orthography in a vacuum and then present it to the community for adoption, it might be perceived as an imposition. Attempts to raise awareness after the fact will never result in the same excitement and loyalty as community involvement throughout the whole process.

Christina Eira (1998:174) recognizes that a participatory approach could turn into a “painstaking, if tedious, process of inclusive community decision-making,” yet stresses the unacceptability of excluding the local stakeholders from the process.

It would not be practical to involve a thousand people in orthography discussions and the decision making process, regardless of the size of the language group. Usually, only key representatives get involved in discussions and in drawing up an orthography proposal. By evaluating the proposed orthography with the help of community members who were not previously involved in the original design process, the number of people who participate in the decision-making process can be greatly increased.

It is to be expected that those who were involved in the process prior to testing already have a favorable bias toward the system. Hopefully they are open-minded and flexible and will heed feedback and test results from individuals and groups who may not have the same bias or mindset. Orthography testing will bring to light favorable or unfavorable reactions, as well as strengths or shortcomings of the system itself. Testing will reveal if it is time for implementation or time to return to the drawing board to rethink some of the decisions.

If orthography testing is neglected, much time and other precious resources can be wasted. If the community resents not having been involved in discussions, is not pleased with the orthography, or finds it difficult to learn, efforts to promote it are likely to fail. (Orthography standardization efforts which failed in Micronesia in the 1970s illustrate this. See Yunick 2000 and Rehg 2004.)

3. Why test

Apart from inviting popular participation in the orthography development process, and hopefully building anticipation and enthusiasm along the way, orthography testing will help discover

- if the system enhances reading and writing fluency and comprehension,
- if the system can be easily learned,
- if it lines up with community preferences and if there is readiness to accept and implement it,
- if the system can be expected to work effectively across dialects.

Orthography testing is thus expected to answer some very pertinent questions.

The question *Will it work?* relates to readability and general efficiency, i.e., issues such as fluency and ambiguities that may interfere with comprehension. In this, field workers might be too easily satisfied. The closest associates they personally trained obviously have no problems reading and writing in the system, so they conclude that “it works.” They might be tempted to neglect testing in the wider

³ Testing procedures and the design of testing tools will be treated in more detail in an upcoming article.

community. But an orthography needs to be tested on a larger number of people to assure it can be read and written, and not only by a few elite individuals. If it does not work well with the average person, it may indicate that the linguistic analysis may be incorrect in some point. Or else, it may show that a symbol choice or the decision to underdifferentiate a phonemic contrast is problematic.

The question *Can a non-literate individual of average intelligence learn how to read and write using this system from a mediocre or average teacher within a reasonable amount of time?* relates to learnability. It takes time, good pedagogical material, and trained teachers to discover the answer to this question. Finding teachers who are willing to commit to using the local language in the classroom and are capable of discerning if learner errors might be due to shortcomings in the orthography may not be easy.

Do people not in the “inner circle” and “not on the pay-roll” of the project like the system? Would they be motivated to learn it or would they prefer something slightly or totally different? These questions relate to community preferences. The answers may be easily discovered in some cultures but not in others. In a highly stratified social system people who do not hold a position of influence may not feel that their opinion counts for much. High levels of trust may have to be established and local mediators trained to get honest feedback.

The question *Can this orthography be learned and used effectively by people who speak different varieties of the language?* relates to the extendibility of the proposed orthography to the various dialects. Testing needs to be done on location of the dialects in question. It should not exclusively be done with people who have moved about and are therefore accustomed to the various dialects. Reading and comprehension scores of such individuals can be expected to be considerably higher than for those who are mono-dialectal, and therefore such scores are not valid for determining extendibility.

Orthography testing does more than answer pressing questions like those listed above. It also serves to test the actual linguistic analysis. If orthographies work well, the analysis may be confirmed. If they do not work well, linguists should not assume that the symbolization choice is at fault. They may need to revisit parts of the phonological, morphophonemic or grammatical analysis.

Orthography testing may facilitate the decision making process. It need not be carried out after the fact, i.e., once decisions have been made. If there is disagreement about whether one system is better or more likeable than another, or if discussions and negotiations are at a standstill, data resulting from quality testing may help choose between proposed options.

Testing may also indicate that the options being considered are equally poor and that a different solution needs to be found.

Linguists recognize that writing system development is highly political. It is therefore important to document the process carefully from the beginning, including factors such as details of all meetings, decisions made, the rationale, and action points. Different parties may have very strong, opposing opinions on certain symbolization issues. Some may already be implementing their favorite option. If at all possible, the different stakeholders should participate in the dialogue, planning, and implementation process. Hopefully all will see the wisdom of and the need for doing orthography testing.

Orthography testing should not be undertaken by one party to prove the other party wrong. If testing is undertaken by one stakeholder without involving other stakeholders in the planning, it could be perceived as combative. It may result in a conflict situation rather than a cooperative working together for the best solution. Transparency and clear communication are essential, even if certain stakeholders do not wish to be directly involved in orthography testing.

4. Review of the Literature

4.1 *Perceived need for testing*

Earliest publications discussing orthography development did not often consider orthography testing. If they did, they simply stated that it needed to be done (Smalley 1959; Powlison 1968; Berry 1977).

In the 1960s, Sarah Gudschinsky encouraged orthography testing. She was concerned about some field workers advocating for underdifferentiation on the basis of low functional load of phonemic contrasts. She argued that redundancy in the written context was not adequate criteria for such decisions (1968). Taking exception to some of Powlison's 1968 article, she raised the question: "How many phonemes with low functional load can be omitted from an orthography before their combined functional load becomes critical?" (1970:23) As an educator, she feared that readers would pay "a very high price for orthographic simplicity and economy" (ibid., 24) and stated, "We urgently need a large number of perceptive and sensitive tests to find out to what extent the orthographies currently in use are in fact adequate" (ibid.). She critiqued unfounded claims that certain simplified orthographies "worked" since such claims could not be backed up by test results or published evidence. She offered her article in "the hope that succeeding articles on orthography may be supported by reports of *careful testing and/or detailed case histories*" (ibid., 25, emphasis mine).

Subsequently, Gudschinsky included orthography development and orthography testing modules in training courses she developed. In addition, her *Manual of Literacy for Preliterate Peoples* (1973), used as a textbook for training SIL International field workers for many years, included a chapter on orthography testing. The chapter is quite short, but it touches on what types of linguistic features might need testing for readability and suggests some possible testing procedures. Instructors of various literacy training courses, also perceiving the need, included this topic and over the years expanded and improved the training material.

Relatively few published articles to date describe informal orthography testing procedures. The dearth of write-ups about informal orthography testing does not necessarily mean that it was neglected, however. Such testing was simply par for the course: Checks were put into place; observations of individuals stumbling while reading or struggling while writing were noted; the orthography was tweaked to improve it to make it more efficient. Busy fieldworkers most likely did not consider such low-key informal testing worth writing about. For example, SIL member Curtis Cook tested orthography options with 28 Zuni [zun] in New Mexico to determine mother tongue speaker symbolization preferences. The result was a popular orthography and "an unexpected amount of native interest" (Walker 1969:164). Cook himself never wrote about Zuni orthography testing. Its significance is noted by Willard Walker, however, who reports on it based on personal communication with Cook. Richard Venezky (1970) and Jack Berry (1977) subsequently cite Walker in reference to the same testing event.

Case study reports, though not intended as "how-to" instructions, thanks to their procedural descriptions and concluding remarks, can be instructive for others wanting to follow suit—or wanting to avoid the same mistakes. The following section of this paper reviews articles and case studies that consider orthography testing. Publications focusing primarily on tone orthography are listed separately.

4.2 *General works*

Venezky (1970) focuses his discussion on psycholinguistic factors of writing system development. He gives attention to the needs of persons acquiring reading skills for the first time, rather than those who are looking to transfer their reading skills to an additional language. He refers to learning to read in a second language as "a relatively trivial task of acquiring new relationships for symbols" (ibid., 46) and points out that accomplished readers "can read English fairly rapidly even when all of the vowel are deleted or when the bottom halves of the letters are covered over" (ibid., 50). He expresses that testing spelling preferences with literate individuals will only reveal their bias, not the relative readability of the proposed system. He states, "To use subjects who are literate in one language for judgments about a

writing system for a different language is highly suspect — unless most of the people who will be reading the new system are at the same level of literacy” (ibid.). The testing procedure he deems reliable would be an experiment in which two groups are taught and develop their skills in the alternate systems. Their skills would be evaluated and compared over a period of one or two years. Yet he warns that if there is a chance that only advanced readers would be able to handle one of the systems, “such an experiment is not justified” (ibid.).

Easton and Wroge developed a training manual (2002) for SIL Alphabet Design Workshops (ADW), held in Papua New Guinea since the mid-1990s. One component of these successful workshops trains Papua New Guineans to test the preliminary orthographies in various speech communities after leaving the workshop location. The manual, detailed for the sake of those who help run such workshops, has been revised over the years as needed. Easton (2003) describes ADWs and their objectives. Testing, as recommended, is not an elaborate procedure but mostly consists of applying the orthography, i.e., writing stories, and keeping track of difficulties encountered. Community testing takes the form of seeking feedback from people who have read such stories and from individuals who examined the alphabet guide that was drawn up during the workshop.

Robinson and Gadelii (2003) include a paragraph on orthography testing, mostly focusing on the need to involve a variety of people in the sample of respondents. They mention that both reading and writing need to be tested.

Moe (2005) focuses on orthography development for Bantu languages. He includes more guidelines for orthography testing than most, and suggests these activities:

- interviewing people about how they write (He includes a list of questions that could be asked of interviewees),
- keeping a list of issues that still need to be discussed,
- collecting “evidence” that these are indeed problematic issues,
- using certain investigative questions that might be addressed at a workshop,
- analyzing actual writing,
- observing people’s reactions to the orthography,
- using reading tests with recording and subsequent analysis of errors.

Unlike others, Moe suggests including experienced as well as new readers in the testing since they might make different types of mistakes. He takes it for granted that after testing, revision of the tentative orthography will take place.

Karan (2006) includes a brief section on orthography testing covering four basic topics: 1) What needs to be tested, 2) How to test, 3) The test population, and 4) Ongoing feedback from users.

Cahill and Karan (2008) mention the need to test an orthography, but few directives are offered.

One of the most comprehensive tools for orthography development, the *Bantu Orthography Manual*, was compiled by Schroeder (2010). It proposes a strategy for participatory orthography development via a workshop approach, providing practical worksheets which take Bantu language characteristics into consideration. Chapter 3 is dedicated to orthography testing. Because of its brevity at six pages, however, untrained field workers who do not have the opportunity of participating in such workshops will want to consult additional sources.

One helpful resource for testing procedures is the Literacy Bookshelf of SIL’s *LinguaLinks Library*. Since it was first released in 1996 it has included “how-to” modules on orthography testing, as well as full-text articles from various SIL journals, some of which focus on orthography development and/or testing.

4.3 Case studies

Presentations on orthography development for specific languages are becoming more common at professional conferences. Such case studies may or may not necessarily focus on orthography testing, but they do illustrate its value. A discussion of select case studies follows.

Lauck (1987) reports on reading tests conducted in Patep [ptp] in Papua New Guinea, to test potential orthography modifications. But the proposed change of not writing a predictable schwa [ə] in unstressed syllables introduced a complication: not writing the schwa resulted in ambiguous consonant sequences, namely <Cy> and <Cw>, which could be interpreted as palatalized and labialized consonants respectively instead of the intended Consonant+ ə +y/w. Alternative symbolizations for palatalization and labialization were tested along with not marking the schwa. The test had two inherent weaknesses: 1) all test subjects were already used to the earlier version of the orthography; 2) only 20 individuals were tested. Lauck insightfully comments, “the difficulty of reading was not so much determined by the orthography used as it was by how predictable the particular word was in a given context. If there was a high degree of predictability, it could be read regardless of how it was written” (ibid., 156). The postscript to the article by editor John Clifton gives tips for better test design and interpretation of results.

Gravelle (1990) reports on orthography testing for a tentative orthography of Meyah (a.k.a. Meah) [mej] of Indonesia. The research addressed three different investigative questions. The first question posed was if it would be better to write semi-vowels with <y, w>, or follow Indonesian writing conventions <i, u>. The number of errors and hesitations in reading tests indicated that <y, w> would serve Meyah readers better. However, in terms of preference of one notation option over the other, there was no consistency. The second question was if the practice of not marking pitch or stress in texts would cause difficulties for readers. It did not. The third question was if the various dialects could be served by a single orthography, and, if this were possible, which of the dialects should be used as a reference dialect. A variety of testing instruments were used to discover the answer to these questions: To determine symbolization preferences researchers used multiple choice tests and interviews. To determine readability and extendibility, they conducted word-level and text level reading tests. The difference in dialects did not present an obstacle for readability. Acceptance of material written in various dialects was facilitated by dialect boundaries not being geographically distinct, no one dialect being more prestigious than others, and a high level of mutual intelligibility due to mixed settlements and frequent travel. A choice of reference dialect was made based on interview responses.

Hasselbring (1996) describes the procedure used to test which word-breaking option (conjunctive vs. disjunctive writing of certain morphemes) might prove more efficient for writing Kalanga [kck] of Botswana. Although more than 350 individuals participated in a reading test, results were inconclusive.

Kosonen (2003) describes the procedure and results from informally testing the Chong [cog] orthography, which uses an adapted Thai script. The report includes the actual texts used in testing. A Powerpoint presentation on the larger Chong Revitalization Project (Premsrirat 2004) also describes the orthography testing process.

Bos, Bos and Page (2008) describe the participatory orthography development and testing approach used among the Kuy [kdt] in Cambodia, using the Khmer script. Testing results fed directly into symbolization decisions. The authors mention additional benefits of the participatory approach: higher awareness in the community; broader participation in the orthography development process; and that “decisions were not unilaterally imposed by the researchers” (ibid., 12).

Boswell (2008) reports on a 1980s attempt by linguists to improve upon an existing Cheke Holo [mm] orthography in the Solomon Islands. The reform was motivated by linguistic factors, desired simplicity of production, and easy transfer to English. A change was envisioned for eleven out of the 31 consonant symbolizations in use at the time. A dictionary with the revised orthography was already under preparation. However, promotion of the changes and workshops to get the proposed revision accepted were not successful. The population was satisfied with the system that had worked for them for about 100

years, and the changes were rejected. Orthography reform decisions had been made in a vacuum and sales pitch efforts followed. Orthography testing had, for the most part, been neglected. Indications that the changes did not take effect were, however, observed and heeded. The dictionary was therefore revised to reflect mother tongue speaker preferences before going to press.

Markowski (2009) reports on the participatory orthography revision process employed for the Sô language [sss] in Thailand. This case study illustrates how a partnership approach, with the community, a language committee, and linguists who take a facilitator role rather than decision-maker role, can result in community ownership of the process and the orthography itself. The paper refers to two stages of orthography testing. The first stage involves preliminary testing by orthography committee members, who not only serve as decision makers but also as initial test subjects. While using the preliminary orthography in reading and writing exercises, they took note of difficulties and met together to discuss problematic symbolizations. The second stage involves testing in the wider community. This second type of testing had not yet occurred at the time of publication. Markowski only mentions that this needed to be done and describes the testing procedures she envisioned.

4.4 *Works with a focus on tone*

Essien (1977) describes a tone orthography experiment carried out in Efik [efi] of Nigeria. It illustrates that the practices of 1) zero marking of tone, and 2) marking tone on every syllable both pose problems for readers. The data indicates that *partial* tone marking on nouns and verbs was by far the most beneficial both for perception time and vocalization time. The groundbreaking results from this experiment caught the attention of Bird (1999b) and Roberts (2008b), who cite this study in their papers. The weaknesses of the experiment were: 1) the small sample, that is, fifteen highly educated Efik speakers, and 2) that testing was only done on the sentence level. Sentences provided few contextual clues that could have helped to disambiguate readings. Test subjects commented that they had to make a choice of the possible readings, but they could not identify the criteria they used to venture their guesses when tone was not marked at all.

Duitsman (1986) reports on testing two tone marking systems for Western Krahn [krw] of Liberia. The efficiency of the use of punctuation marks before syllables was compared with that of more conventional diacritics over vowels. The results showed no advantage for either system with those already literate in the language of schooling, in this case, English. The report illustrates the difficulty of coming up with an adequate number of suitable readers to determine if findings are statistically significant. Roberts (2008b) reports that the punctuation-based tone notation was not adopted for Eastern Krahn.

Mfonyam's doctoral thesis (1989) reports on tone experiments in Bafut [bfd] and Limbum [lmp] of Cameroon. Deemed exemplary by Wiesemann (1995), Bird (1999b), and Roberts (2008b), this work would serve as a good model for formal testing. Roberts comments that Mfonyam "tackled a thorny issue with clarity" and that "no other researcher to date has done so much preparatory background research before undertaking experimentation" (*ibid.*, 211). Mfonyam compared the effectiveness of four different orthographies against one another—none of them zero tone marking since he already knew it was problematic. He invested heavily in training test subjects in advance and tested reading on the sentence as well as the text level. The Bafut experiment included some creative writing exercises.

Bird (1999b) reviews tone orthography experiments done by other researchers, namely Essien (1977), Mfonyam (1989), and Bernard et al. (1995; 1997)⁴. He then reports on an original experiment he conducted in Cameroon in Yemba (a.k.a. as Dschang) [ybb], which illustrates how tone marking can get in the way of efficient reading and writing. He points out that the question is not whether or not to mark tone, but how much and at what level. Clearly, testing is key, and test results should guide decision making. The one weakness of this experiment was Bird's small sample of eleven individuals. Yet,

⁴ Bernard et al. 1997 consisted of unpublished test results in Bird's reference. The report on this later experiment was published in 2002.

Roberts (2008b:215) highlights the importance of this study: “[W]hat this experiment lacks in sample size, it more than makes up for in its statistically rigorous approach.”

Bernard et al. (2002) report on a second tone experiment conducted in Kom [bkm] in Cameroon. The earlier one, conducted in 1995 and reviewed by Bird (1999b), had used only one test subject. This second one, conducted with thirteen subjects, was designed to investigate if the original findings could be validated and generalized to the larger population. The second experiment confirmed the findings of the first, i.e., that accent tone notation did not improve accuracy in reading. However, longer sentences, which provided more context, were found to contribute to higher reading accuracy rates. Roberts praises Bernard’s team because they “combined their forces [in anthropology, statistics and linguistics] and the quality of the result shows” (2008b: 239).

Roberts focused his own research on comparative efficiencies of tone marking systems as well.⁵ His PhD. thesis (2008a) reports on a quantitative experiment he conducted in Kabiye [kbp] in Togo.

In an article published the same year (2008b), Roberts reviews ten different tone orthography experiments conducted in West Africa between 1977 and 2007, including the aforementioned ones. He compares them according to different parameters, such as investigative focus, test design, sample profile and size, the type of experience test subjects already had or training they were given, test materials used, and other variables. He not only reports on the experiments but also comments on their strengths and weaknesses, providing guidelines and principles for those designing future experiments. His article is generous in helpful tables, charts, and diagrams. One of Roberts’ motivations, in his own words, is “to identify whether any consensus is emerging about the profile of an optimal tone orthography” (ibid., abstract). This article is essential for anyone considering doing formal orthography testing. Its practical remarks go beyond tone orthography issues.

Roberts and Walter (2012) report on a more recent experiment in Kabiye, which, although tonal, is traditionally written with zero tone notation. The experiment compares the learnability and efficiency of two experimental tone marking strategies, one phonologically based, noted with diacritics, and one grammatically based, disambiguating text through the use of punctuation marks, medial capitals and superscript letters. The experiment was unique, not only in the use of unusual symbolization:

First, it was not intended to determine which of two orthographies should be promoted. In fact, the authors state that the grammar-based orthography used in the experiment would come across as “outlandish,” —one that “no language committee would ever authorize” (ibid., 242). Instead, the researchers tested the principle of marking grammar instead of tone. It seems that such an approach works well since the “grammatical representation” test group outperformed the “tone representation” test group both in accuracy and speed in writing activities. Secondly, unlike most other tone experiments, a major emphasis was placed on writing activities such as dictation and original composition. The authors state that their paper “helps build a case for teaching grammatical tone as grammar, without any reference to the tones themselves at all” (ibid., 247). They point out that such an approach may allow a wider coverage of various dialects.

Hoover (2012) reports on tone orthography testing in Maninka Mori, a dialect of Eastern Maninkakan [emk] in Guinea, Africa. It is unique in that it compares the efficiency of two different scripts for the language: the traditional Nko script and the commonly used Roman script. Hoover briefly reviews Bird’s work (1999b) before commenting on a variety of tone notation recommendations made in the past for languages in the Manden family. He then proceeds to describe the experiment he conducted. In the experiment, ten individuals read the Nko script text, and ten others read the same text in the more commonly used Roman script orthography. While the Nko script marks tone fully, the Roman script only marks tone on one grammatical element. Recordings of the readers provided the data for miscue analysis

⁵ Roberts has written many useful articles in French and in English and presents regularly at conferences. His website is a valuable resource. Readers can access the site at <http://www.nyinyu.com> .

and documented the time it took for each reading. Although those reading in Nko with full tone notation took longer to read the texts, the total number of miscues were considerably lower. Nko script readers also had more self-corrections, which suggested that they had a better comprehension level of the text. Hoover hypothesizes that the differences in performance by Roman script text readers were rooted in their formal education experience in French. He suspects that the habit of not correcting errors, which in this instance resulted in some readings that made no sense, might be due to “a tolerance for understanding merely the gist of a text.” He states that “[t]his poor reading habit may transfer to their reading of Maninka texts” (2012: 56). In retrospect, Hoover wishes that he had tested a larger number of individuals, so his research would be statistically valid. He also regrets not having included comprehension questions, which would have given further insight into the results. In his opinion, the high number of miscues⁶ suggests that even a perfect orthography will not solve the problem. He proposes that what is needed is better reading instruction, more reading material, and more reading practice. As a literacy specialist I agree that this might solve part of the problem, but I suspect that the two orthographies still need some serious work. Hoover does propose some changes for the Roman script tone orthography, likely based on the miscue analysis of the Roman script recordings. It is noteworthy that one reader of the Roman script text only registered eight miscues⁷ and was one of the fastest readers. No explanation is offered for this superior performance.

Wright (2012) reports on testing a proposed tone orthography for Chrambo (a.k.a. Bambalang) [bmo] in Cameroon. Testing did not compare efficiency of the new system with the older zero-tone marking system. Instead, it focused on evaluating the learnability of the new. The new orthography was designed to reduce ambiguity in text. Yet it aimed to keep diacritics to a minimum by only noting lexical tone on nouns and verbs and uniquely on the first vowel to represent the word’s tone melody. Grammatical tone, to differentiate aspect and mood, is also marked and needed to be taught. Four Chambro speaking groups were tested:

1. three individuals in a Chrambo literacy class who were not yet readers and had only covered a fraction of the alphabet,
2. seven individuals who were readers of English and were familiar with the Chrambo zero-tone marking orthography,
3. a literacy class of three individuals who had nearly completed the second-level primer, and
4. six individuals who were fluent readers of English and had a higher education level than those in test group two.

Wright believes the writing system is simple to learn and recommends that the tone notation, as proposed, be adopted. She acknowledges that since only nineteen subjects participated fully in the testing, additional testing with a wider range of Chrambo speakers will need to be done. Efficiency of the system seems to largely depend on users’ capacity to differentiate between the various parts of speech.

5. Formal vs. informal testing

There are two basic types of testing: *formal testing* and *informal testing*. The choice of which to use depends partly on the technical resources available to support the testing effort and partly on how urgent it is that one be able to make a strong and compelling argument on the matter under consideration. (See section 5.2 for more detail on this point.) Both types qualify as “research” and are therefore designed around a basic investigative question.

⁶ 1172 total miscues for 10 readers on the 1620 word Nko script text; 1402 total miscues for the same text in Roman script (Hoover 2012:44), averaging 117 and 140 miscues per person respectively.

⁷ Three of these eight miscues were caused by typographical errors.

5.1 Characteristics of formal and informal testing

The term *formal testing* suggests a scientific approach, i.e., the testing will be rigorous and tightly controlled. The results will be quantitative and need to be analyzed for statistical significance. All of these factors are true of formal testing. The expectation, however, that *informal testing* need not be rigorous, that it will not be labor intensive, that there is no need to control for variables, and that informal testing does not involve number crunching, but rather reports only anecdotal and qualitative data, is erroneous.

For informal and formal testing alike, testing procedures need to be carefully planned, the investigative question(s) clearly stated, appropriate test instruments designed, and assistants trained to administer the test(s). Permissions need to be sought from the appropriate national and local authorities. Most orthography testing activities constitute “human subject research.” Respect for individuals’ human rights and international research ethics standards dictate that evidence of prior informed consent be sought from test subjects themselves. If researchers do not comply with this requirement, no reputable agency will publish the findings.⁸

For both types of testing, an adequate number of individuals should participate so the test results can be shown to be statistically significant and not the result of chance. Ideally, participants should consist of a random sample yet be representative of a wide variety of backgrounds to represent both genders and diversity in characteristics such as age, level of education, location of residence and origin, economic activity, and religious affiliation, if pertinent. Such background information, plus any additional information that might provide insight during the analysis, should be collected in a preliminary interview. For instance, which languages a person is already literate in might affect the range of symbolizations they find acceptable; age or religion might make a difference in degree of attachment to an older writing tradition.

For both types of testing the test subjects have to be put at ease. This usually involves a clear explanation of the procedure and what to expect, as well as a practice exercise before actual testing begins. It needs to be made clear that they themselves are not being evaluated – the orthography is. By participating, test subjects are helping decision makers figure out what the best way to write the language might be. They need to be assured that their names will not be used in any reporting.

For formal and informal testing alike, subjects need to be screened to ensure they are qualified to participate; otherwise, results may be skewed. If someone stumbles throughout a simple reading test, they should probably not be included in the sample. It would be difficult to determine in the analysis if their results are due to poor reading skills or due to the orthography issue being tested.

We note that the two approaches have much in common. So what are the *differences* between formal and informal testing? Steve Walter⁹, experienced in formal research design and analysis, clarified this for me. The difference between formal and informal testing is only partially a matter of rigor. Attention to rigor will not convert an informal test into a formal test. It is possible for rigor to be lacking in formal testing. When this is the case, it raises questions concerning the validity of the findings.

Formal testing typically exhibits the following characteristics:

1. The hypotheses to be tested are stated explicitly.
2. The criteria for sampling are carefully and explicitly stated and implemented.
3. Formal testing typically includes experimental and comparison groups in an effort to control for the influence of possible confounding variables.

⁸ Research involving human subjects needs to adhere to a strict code of ethics. University websites usually provide guidelines and standard forms. Here is one such example, that of the University of Edinburgh: <http://www.lel.ed.ac.uk/facilities/howto/ethics/1.shtml>

⁹ Dr. Steve Walter is an associate professor at the Graduate Institute of Applied Linguistics in Dallas and Chair of the Department of Applied Anthropology. He is a Senior Literacy and Education Consultant with SIL International and served as the SIL International Literacy Coordinator from 1988–1999.

4. Testing materials and procedures are carefully and explicitly designed to provide the data needed to address the initial hypothesis.
5. A conscious effort is made to identify and control (i.e., gather data on) relevant independent variables, such as age, gender, level of education, reading skill level, etc.
6. Statistical methods are used to analyze the results.

The application of formal testing normally entails some specialized training or assistance from someone with the requisite technical knowledge, especially in the areas of sampling, testing designs, and statistical analysis. As a result, informal testing is much more common. Testing is always considered “informal” if any of the following conditions apply: those orchestrating the testing are not in a position to 1) control variables well enough, 2) claim a high level of test accuracy, or 3) have access to the number or variety of test subjects that formal testing requires.

5.2 *Formal or informal testing: making a choice*

Informal testing will be adequate and appropriate in most situations. However, if one must satisfy a critical audience or help resolve an issue over which there is much disagreement or controversy, it may be necessary to do formal testing and seek assistance with test design and the analysis of important data.

Formal testing may well be needed when

1. One is writing for or needs to satisfy an academic audience;
2. The orthography issue under consideration is especially sensitive or controversial and one needs to provide compelling “objective” evidence for a given solution or course of action;
3. The issue to be resolved is especially subtle and needs more powerful methods to tease out an answer;
4. There are significant issues of policy or strategy having long term consequences which need to be resolved especially at a regional or national level;
5. When there are many poles of opinion on a matter and these tend to cut across typical demographics.

When none of these conditions hold, informal testing is usually sufficient to provide one with an answer to the question(s) at hand. So, it is not necessarily the complexity of the investigative question that determines which approach should be taken, but rather who the audience is for the report on the findings and what the stakes are.

Investigative questions like “What causes newly literate individuals to stumble when they read never-before-seen texts?” or “Will people prefer the doubling of the vowel or the use of <: > for denoting vowel length?” can typically be addressed by informal testing.

Readability questions, such as “Will symbolizing phoneme /ŋ/ as <ng> while symbolizing nasal vowels with digraphs with the vowel followed by <n> result in ambiguities and interfere with reading fluency?” might call for either informal or formal testing. It is not the investigative question that determines this, but rather this question: “Who needs to be convinced and how difficult will it be to convince them?”

In contrast, the investigative question “How much tone needs to be marked in language X and what symbolization might work best?” may require formal testing. It is a much more complex issue and the findings may have implications for additional languages and will be of interest to the wider academic community. In fact, most of the articles published to date reporting on formal orthography testing in lesser-studied languages deal with tone writing issues.

Researchers in the fields of Educational Psychology, Psycholinguistics, and Applied Linguistics use formal testing approaches as they aim to find answers to very complex questions. When their investigative questions relate to orthography issues, the focus is frequently on the comparative efficiency

of majority language writing systems and/or the link between the reading process and orthography.¹⁰ Reports on this type of research can serve as models for formal orthography testing since they illustrate how, with formal testing, painstaking care needs to be taken in designing testing instruments, controlling for various variables, and analyzing results.

6. What to test

Readability of an orthography needs to be checked, of course. The time it takes to read texts and the number and type of errors readers make serve as the primary evaluative metrics. Ease of writing with the orthography also needs to be tested. If underdifferentiation of segments or partial tone notation is being considered, tests checking on resulting ambiguity may be required. In such a case, reading tests need to include comprehension questions. Errors in reading, writing, and comprehension will reveal problematic issues.

Not *all* symbolizations need to be tested since not all would pose a challenge. Representation of most consonants or common syllables found in syllable-based writing systems would be quite straight forward, based on the accepted, de facto norms established by other languages in the context. For instance, a language which has both /r, l/ will not likely need to test <r, l>. A language which has free variation between [r, l] may, on the other hand, need to test for symbolization <r> vs. <l>. For items that do need testing, both acceptability and effectiveness will need to be investigated. A list of issues which commonly need closer scrutiny follows.

6.1 *Script preference*

In the context of multilingualism and multiplicity of scripts, the first question that needs to be settled is script choice. This is usually an identity issue. It may be prudent to avoid a binary choice, i.e., *not* “one vs. the other(s).” It may be better to discuss preferences in terms of first choice and second choice and the reasoning, as well as which script is preferred in which context or domain and for which audience. This may set the stage for negotiation if there are divergent opinions.

6.2 *Symbolization of certain contrastive sounds*

Which symbolizations may prove problematic depends on the specific language in question and which script is being adapted to suit its needs. The following issues commonly need testing:

- vowel symbolization beyond the normative set provided by the chosen script,
- any sounds or syllables not included in the signary of the language of wider communication,
- unusual digraphs, trigraphs, or ligatures,
- nonconventional uses of letters, diacritics, punctuation symbols or numbers,
- vowel length (<VV vs. V: vs. \underline{V} vs. V.> etc.) and nasalization (<Vn vs. \tilde{V} vs. $\underset{\sim}{V}$ vs. V’> etc.),
- labialization (<w vs. u>) and palatalization (<y vs. i>),
- complex sounds, such as lengthened and pre-nasalized consonants and affricates,
- departure from written norms in languages with register, for instance, combining first and second series consonants with vowels that are not usually associated with the series (see Bos et al. 2008).

¹⁰ These topics are beyond the scope of this paper. Here are a few select publications in these fields: Tzeng and Singer 1981; Taylor and Taylor 1983, 1995; Henderson 1984; Frost and Katz 1992; Taylor and Olson 1995; Cook and Bassetti 2005; Joshi and Aaron 2006. In addition, the journals *Written Language and Literacy* and *Writing Systems Research* regularly publish papers on research related to dominant languages and their writing systems.

6.3 *Proposed underrepresentation*

Underdifferentiating vowels is a strong temptation. It is often done for the following features: tense vs. lax; short vs. long; plus vs. minus advanced tongue root (ATR). Those who strongly dislike the use of diacritics or digraphs to augment the basic Roman alphabet repertoire of <a, e, i, o, u> may not recognize the impact that underdifferentiation frequently has on learnability and readability of texts. It may be wise to postpone decision making until some testing has been done. There may also be additional options the committee or community hadn't considered: i.e. not only the digraph option <fit, feet>, or use of a phonetic symbol <fit, fit>, or punctuation <fit, fi:t>, but also <fitt, fit>. This option may be more acceptable to the community since the doubling of consonants is quite common in international official languages. It also presents less of a challenge in sending text messages, which has become an important consideration in more recent years.

Minority language speakers educated in the official national language script may not be aware of adaptation options. They only know what is done in the language of schooling. Therefore, if the possibility of underrepresentation is being considered, testing is necessary, but so is awareness raising concerning additional potential options, which may already be used in less familiar languages.

6.4 *Diphthongs and semi-vowels*

Vowel sequences may be suspect in the linguistic analysis and should be tested.

6.5 *Cross-syllable ambiguities*

Complex syllable structures can present decoding challenges for readers of alphabetic systems. For instance: Depending on which syllable structures occur in the language, a series such as CVCCVC could represent CVC.CVC or CV.CCVC with CC representing a consonant blend or sequence. (ex. *ma.stiff* vs. *mas.tiff* in English.) Experienced readers do not usually find this problematic.

The use of certain digraphs may, however, increase visual ambiguity and thus complicate the reading process. If the components of a digraph also occur as a sequence of individual letters, the eye may treat them as a digraph at the wrong time. The CVCCVC sequence could thus additionally be read as CV{CC}.VC or CV.{CC}VC. If the first attempt at orthographic segmentation does not result in a meaningful word, a regression will be required to seek a meaningful option. Baker (1997:135) offers the examples *hothead* and *hogshead* from English. Karan (2006:135) mentions *lighthouse*, *uphill*, and *mishap* and some examples from German. Again, experience usually helps readers overcome such challenges. However, a high frequency of ambiguous consecutive occurrences of digraph component letters might indicate that the digraph symbolization is a poor choice. Testing may be required.

6.6 *Stress*

Some languages rely more heavily on stress to distinguish meaning than others. Whether or not stress needs to be noted, to what extent, and which notation system would be best should be investigated. There is a risk in testing such a feature: it is possible to build testing instruments with contrived texts which would give the impression that noting stress would be extremely important. Ambiguity in testing instruments should not be higher than in naturally occurring texts. (Note that I say "naturally occurring texts," not just "natural texts." The latter suggests grammatical naturalness; the first requires that it be embedded in a naturally occurring context.)

6.7 *Tone marking options*

Unless there is a national alphabet or script which dictates how and to what extent tone should be written, this will likely be a major issue for any tonal language. Tone testing could seem like an overwhelming task. But, it may not be as complex as imagined. Those who have gone before and done extensive tone orthography testing have made discoveries that can serve as a point of departure. All who

are involved in the development and testing of a tone orthography would do well to study what has already been written. (See works mentioned in 4.4, as well as Bird 1999a; 2001.)

6.8 *Word breaks and white space*

Word breaks need to be tested. They should match mother tongue speaker intuition. Long words may present a challenge for beginning literacy instruction, but segmentation into decodable chunks can be taught. It just takes practice. The use of hyphens, if proposed, should be tested. Some helpful word break guidelines have been written up for Roman scripts. (See Van Dyken and Kutsch Lojenga 1993; Kutsch Lojenga 2014.)

Some non-Roman scripts, such as Thai, Burmese, and Vai, a syllabary of West Africa, do not use spaces between words. Regarding Vai, Singler (1996:596) comments that “the lack of division between words makes reading the script difficult.” Although Burmese does not use white space between words, it does use space to denote phrases and sentences (Wheatley 1996). In contrast, the Khmer script in Cambodia does not indicate word breaks nor the end of sentences. It only uses a special marker to indicate a change of topic in continuous text (Schiller 1996). If a writing system which does not use spaces between words is adapted for additional languages and proves difficult to learn, it may be judicious to compare readability of texts with and without spaces to determine if extra spaces add benefit or not. If readability is enhanced by extra spaces, ease of learning will surely be enhanced as well. Acceptability of such a deviation from the norm will need to be investigated.

6.9 *Elision and unusual consonant sequences*

Whether the full form of words should be written rather than the contracted form needs to be tested. When certain syllables do not follow the normative syllable structure, the suspect sequence should be tested. For instance, do members of the language community want to write [ps] as <ps> or do they perceive a vowel in the underlying form that they feel needs to be written?

6.10 *How to write loan words*

Should loan words be written as monolinguals pronounce them, or should the spelling reflect the spelling of the source language, or something in between? Should letters not needed to write the local language be included in its alphabet? These are questions that need to be settled. Preferences and efficiency of various options need to be tested.

7. **Who to test**

In general, it is desirable to test a wide sample of people so the claim that an orthography works well can be generalized to a wider clientele, not just a small subset of the population. For formal testing, it is especially important to test an adequate number of people so that findings can be shown to be statistically significant.

Although a random sample is heralded as ideal in human research, this is often not practical when conducting orthography tests. To participate in a reading and writing test, test subjects need to be somewhat literate. If groups are to be compared with each other, it is more important that the groups be even in skill level and other independent variables than that they be randomly assigned. This is why personal information is gathered during interviews and test subjects are given a reading evaluation before actual orthography testing activities begin.

The number of people to test depends on the specific investigative question and test design. Roberts (2008b:211) speaks of a major division in methodological designs, i.e., a “within-subject” design, and a “between-groups design.” In the *within-subject* design, the performance of a subject using one orthography is compared with the same person’s performance using a different orthography, or orthographies. In the *between-groups* design, the performance of one group trained in one system is compared with that of a group trained in another system. This latter type of test needs stricter controls so

that differences in performance are not the result of undue previous familiarity with the system or some imbalance between the groups. Pretesting will help prevent potential imbalances.

A pre-literate person likely has a script preference, and a semi-literate person probably has symbolization preferences based on aesthetics or familiarity. Such individuals need therefore not be excluded from testing activities which do not require advanced reading skills.

As already stated in section 6, it is important to test people from a variety of backgrounds. There should be monolingual as well as multilingual individuals, varying in gender, age, educational and professional backgrounds, religious affiliations, from various locations and dialects, etc. Testing six male literacy teachers in the one orthography option with which they are familiar would not qualify as a valid once-for-all test.

Yet, informal testing with a small uniform test group need not be excluded. Testing such a group, or even an individual, can be revealing and identify shortcomings of the orthography. Only: testing should not end there. After adjustments, further testing should be conducted with a wider population sample.

One subgroup to test is highly educated speakers of the minority language, who read a language of wider communication fluently but have never read or written anything in their mother tongue. Such a test focuses on ease of skill transfer and system interference issues.

Another group could be those who have already learned to read the mother tongue with the experimental system or in an earlier orthographic tradition. The test might focus on challenges readers encounter and identify patterns in their difficulties.

An additional group of individuals would be those who either have never learned to read or never mastered reading in any language. This kind of test requires a major investment of time and resources. It involves developing basic instructional materials and training teachers who could lead such a test class for a year or even longer, depending on the amount of time learners can devote per week. It takes an exceptional teacher to take note of reader errors and recognize that they might be due to orthography issues rather than lack of learner competence or something inherent in the instructional method or story content. Teacher education should include a module on orthography testing with training in recognizing what might trigger learner errors. This type of investment is worthwhile, since the writing system needs to be designed so that it can be learned by an average person, not only a highly skilled elite.

8. Conclusion

This article has made a case for the importance of orthography testing as an integral part of the orthography development process. Hopefully, readers have discovered helpful resources as well as practical tips in these pages. The most significant points may well be 1) to make the process as painless as possible for test subjects, 2) to carefully document the orthography development process, and 3) to use orthography test results in a way that builds consensus rather than aggravating a potential conflict.

Certain types of orthography testing, such as discovering script and symbolization preferences, would be done early in the development process. Other types, such as discovering problematic issues and testing solutions might be done during the decision making process. After the fact, testing can reassure stakeholders that the right decisions were made. Orthography testing can be instrumental in stirring up excitement in a speech community about reading and writing and promoting the local language.

Although this article addresses questions like *Why test?*, *What to test?*, and *Who to test?*, it only barely touches on the topic of *How to test*. This will need to be treated in more depth in a separate article.

References

- Baker, Philip. 1997. Developing Ways of Writing Vernaculars: Problems and Solutions in a Historical Perspective. In: Andrée Tabouret-Keller, Robert B. Le Page, Penelope Gardner-Chloros, and Gabrielle Varro (eds.), *Vernacular Literacy: A Re-Evaluation*. Oxford: Clarendon Press, 93–141.
- Bernard, H. Russell, George Ngong Mbeh and W. Penn Handwerker. 1995. The tone problem. In: A. Traill, R. Vossen, and M. Biesele (eds.), *The Complete Linguist — Papers in Memory of Patrick J. Dickens*. Cologne: Rüdiger Köppe, 27–44, cited in Bird 1999.
- Bernard, H. Russell, George Ngong Mbeh and W. Penn Handwerker. 1997. Does tone need to be marked? Unpublished manuscript, University of Florida, cited in Bird 1999.
- Bernard, H. Russell, George Ngong Mbeh and W. Penn Handwerker. 2002. Does marking tone make tone languages easier to read? *Human Organization* 61(4), 339–349.
<http://nersp.osg.ufl.edu/~ufruss/publications.htm> (accessed Sept. 12, 2013)
- Berry, Jack. 1958. The Making of Alphabets. Paper presented at the Proceedings of the 8th International Congress of Linguistics. In: Joshua A. Fishman (ed.), *Readings in The Sociology of Language*. 1968. The Hague: Mouton, 737–747.
- Berry, Jack. 1977. ‘The Making of Alphabets’ revisited. In: Joshua. A. Fishman (ed.), *Advances in the Creation and Revision of Writing Systems*. The Hague: Mouton, 3–16.
- Bird, Steven. 1999a. Strategies for representing tone in African language systems. *Written Language and Literacy* 2(1), 1–44.
- Bird, Steven. 1999b. When marking tone reduces fluency: An orthography experiment in Cameroon. *Language and Speech* 42(1), 83–115. <http://cogprints.org/2173/> (accessed Sept. 12, 2013)
- Bird, Steven. 2001. Orthography and identity in Cameroon. *Written Language and Literacy* 4(2), 131–162.
- Bos, Kees Jan, Mirjam Bos, and Christina Page. 2008. Community based orthography development: Experiences from the Kuy in Cambodia. Paper presented at the 2nd Int’l Conference on Language Development, Language Revitalization and Multilingual Education in Ethnolinguistic Communities, 1-3 July 2008, Bangkok Thailand. http://www.seameo.org/_ld2008/document.html (accessed Sept. 12, 2013)
- Boswell, Freddy. 2008. Cheke Holo Orthography: Ethnic language community decision-making and the role of trained linguists. Paper presented at the 2nd Int’l Conference on Language Development, Language Revitalization and Multilingual Education in Ethnolinguistic Communities, 1-3 July 2008, Bangkok Thailand. http://www.seameo.org/_ld2008/document.html (accessed Sept. 12, 2013)
- Cahill, Michael, and Elke Karan. 2008. Factors in Designing Effective Orthographies for Unwritten Languages. SIL Electronic Working Papers. 2008-001. SIL International.
<http://www-01.sil.org/silewp/2008/silewp2008-001.pdf> (accessed Sept. 12, 2013).
- Cook, Vivian, and Benedetta Bassetti (eds.). 2005. *Second Language Writing Systems*. Second Language Acquisition 11. Clevedon, England: Multilingual Matters Ltd.
- Coulmas, Florian. 1999. Development of Orthographies. In: Daniel Wagner, Richard L. Venezky and Brian V. Street (eds.), *Literacy: An International Handbook*. Boulder, Colorado: Westview Press, 137–142.
- Crystal, David. 2000. *Language Death*. Cambridge: Cambridge University Press.
- Duitsman, John. 1986. Testing two systems for marking tone in Western Krahn. *Notes on Literacy* 49, 2–10.

- Easton, Catherine. 2003. Alphabet Design Workshops in PNG: A Community-Based Approach to Orthography Development. Paper presented at the 1st Int'l Conference on Language Development, Revitalization and Multilingual Education in Minority Communities in Asia. 6–8 November 2003. Institute of Language and Culture for Rural Development, Mahidol University, Bangkok. http://www.sil.org/asia/ldc/parallel_presentations.html (accessed Sept. 12, 2013).
- Easton, Catherine, and Diane Wroge. 2002. *Manual for Alphabet Design through Community Interaction*. Ukarumpa, Papua New Guinea: SIL.
- Eira, Christina. 1998. Authority and discourse: Towards a model of orthography selection. *Written Language and Literacy* 1(2), 171–224.
- Essien, Udo E. 1977. To end ambiguity in a tone language. In Paul François Amon Kotey and Haig Der-Houssikian (eds.), *Language and linguistic problems in Africa: proceedings of the 7th conference on African linguistics*, Gainesville, 1976. Columbia SC: Hornbeam Press, 155–167.
- Fishman, Joshua A. (ed.). 1977. *Advances in the Creation and Revision of Writing Systems*. The Hague: Mouton.
- Frost, R., and L. Katz (eds.). 1992. *Orthography, Phonology, Morphology, and Meaning*. Advances in Psychology 94. Amsterdam: North-Holland.
- Gravelle, Gilles. 1990. Orthography testing in Meah: using Meah or Indonesian semi-vowel rules. *Irian: Bulletin of Irian Jaya [Development]* 18, 125–142.
- Gudschinsky, Sarah C. 1968. A test for orthographic ambiguity. *Notes on Literacy* 3: 1–2
- Gudschinsky, Sarah C. 1970. More on formulating efficient orthographies. *The Bible Translator* 21(1), 21–25.
- Gudschinsky, Sarah C. 1973. *A manual of literacy for preliterate peoples*. Ukarumpa, PNG: SIL.
- Hasselbring, Sue. 1996. Orthography testing in Botswana. *Notes on Literacy* 22(2), 34–37.
- Henderson, Leslie (ed.). 1984. *Orthographies and Reading: Perspectives from Cognitive Psychology, Neuropsychology and Linguistics*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Hoover, Joseph D. 2012. Testing orthographies in the Nko and Roman scripts. *Mandenkan* 48, 39–58. <http://llacan.vjf.cnrs.fr/PDF/Mandenkan48/48Hoover.pdf> (accessed Nov. 1, 2013).
- Joshi, R. Malatesha, and P.G. Aaron (eds.). 2006. *Handbook of Orthography and Literacy*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Karan, Elke. 2006. Writing system development and reform: A process. M.A. thesis. University of North Dakota. http://arts-sciences.und.edu/summer-institute-of-linguistics/theses/_files/docs/2006-karan-elke.pdf (accessed Sept. 12, 2013).
- Kosonen, Kimmo. 2003. Testing Chong orthography in Thailand. Paper presented at the 1st Int'l Conference on Language Development, Revitalization and Multilingual Education in Minority Communities in Asia. 6–8 November 2003. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University. http://www.sil.org/asia/ldc/parallel_presentations.html (accessed Sept. 12, 2013).
- Kutsch Lojenga, Constance. 2014. Basic principles for establishing word boundaries. In Michael Cahill and Keren Rice (eds.), *Developing Orthographies for Unwritten Languages*. Publications in Language Use and Education. SIL International.
- Lauck, Linda M. 1987. Orthography testing in Patep. In: John M. Clifton (ed.), *Studies in Melanesian orthographies*. Data Papers on Papua New Guinea Languages 33. Ukarumpa: SIL, 151–159. http://www-01.sil.org/pacific/png/pubs/928474523780/Orthography_Testing_Patep.pdf (accessed Sept. 12, 2013).
- Markowski, Linda. 2009. So Orthography Committee and Revision Process. Working Paper. Chiang Mai: Payap University. http://ic.payap.ac.th/graduate/linguistics/papers/So_Orthography_Development.pdf (accessed Sept. 12, 2013).

- Mfonyam, Joseph. 1989. Tone in Orthography: The Case of Bafut and Related Languages. PhD thesis, University of Yaoundé. <http://www.silcam.org/folder030401/page.php> (accessed Sept. 12, 2013).
- Moe, Ronald. 2005. Developing an orthography for a Bantu Language. unpublished manuscript.
- Powlison, Paul S. 1968. Bases for formulating an efficient orthography. *The Bible Translator* 19 (2), 74–91.
- Premrsirat, Suliwai. 2004. Chong Language Revitalization Project (PowerPoint Presentation). Institute of Language and Culture for Rural Development. Bangkok: Mahidol University. http://www.unescobbkk.org/fileadmin/user_upload/aims/Pattaya_SCB_Jun05/Chong_Language_-_Part_1.pdf (accessed Sept. 12, 2013).
- Rehg, Kenneth L. 2004. Linguistics, literacy, and the law of unintended consequences. *Oceanic Linguistics* 43(2), 498–518.
- Roberts, David. 2008a. L'orthographe du ton en Kabiye au banc d'essai (Kabiye tone orthography on the testing bench). Doctoral thesis. Paris : INALCO.
- Roberts, David. 2008b. Thirty years of tone orthography testing in West African languages (1977–2007). *Journal of West African Languages* 35 (1-2), 199–242.
- Roberts, David, and Stephen L. Walter. 2012. Writing grammar rather than tone — an orthography experiment in Togo. *Written Language and Literacy* 15(2), 226–253.
- Robinson, Clinton and Karl Gadelii. 2003. Writing Unwritten Languages: A Guide to the Process. UNESCO.
- Rogers, Henry. 1995. Optimal Orthographies. In: Insup Taylor and David R. Olson (eds.), *Scripts and Literacy: Reading and Learning to Read Alphabets, Syllabaries and Characters*. Dordrecht, Netherlands: Kluwer Academic Publishers, 31–43.
- Schiller, Eric. 1996. Khmer Writing. In: Peter T. Daniels and William Bright (eds.), *The World's Writing Systems*. New York: Oxford University Press, 467–473.
- Schroeder 2010. Bantu orthography manual. E-book. SIL International. <http://www.sil.org/resources/archives/9241> (accessed Sept. 12, 2013)
- Sebba, Mark. 2007. *Spelling and Society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.
- Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1(1), 35–49.
- Seifart, Frank. 2006. Orthography Development. In: Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.), *Essentials of language documentation*. Trends in Linguistics 178. Berlin: Walter de Gruyter, 275–299. http://anthro.ucsd.edu/~jhaviland/AudVid/AudVidReadings/GippertEtAl/LangDoc_11_Seifart.pdf (accessed Sept. 12, 2013)
- SIL International. 2003. Testing an Orthography. *LinguaLinks Library* version 5 plus. Dallas: SIL International. (CD-ROM) at: Develop an orthography/Developing an orthography/Establishing a working orthography.
- Singler, John Victor. 1996. Scripts of West Africa. In: Peter T. Daniels and William Bright (eds.), *The World's Writing Systems*. New York: Oxford University Press, 593–598.
- Sjoberg, Andree F. 1966. Sociocultural and linguistic factors in the development of writing systems for preliterate peoples. In: *Sociolinguistics*, W. Bright (ed.). The Hague: Mouton. 260–276.
- Smalley, William A. 1959. How shall I write this language? *The Bible Translator* 10(2), 49–69. Reprinted in Smalley et al., 31–52.
- Smalley, William (ed.). 1964. *Orthography studies: Articles on new writing systems*. Helps for Translators 6. London: United Bible Societies.

- Snider, Keith L. 2001. Linguistic Factors in Orthography Design. In: Nguessimo M. Mutaka and Sammy B. Chumbow (eds.), *Research Mate in African Linguistics: Focus on Cameroon*. Köln: Rüdiger Köppe Verlag, 323–332.
- Taylor, Insup, and M. Martin Taylor. 1983. *The Psychology of Reading*. New York: Academic Press.
- Taylor, Insup, and M. Martin Taylor. 1995. *Writing and Literacy in Chinese, Korean and Japanese*. Amsterdam: John Benjamins.
- Taylor, Insup, and David R. Olson (eds.). 1995. *Scripts and Literacy: Reading and Learning to Read Alphabets, Syllabaries and Characters*. Neuropsychology and Cognition 7. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Tzeng, Ovid J.L., and Harry Singer (eds.). 1981. *Perception of Print: Reading Research in Experimental Psychology*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- UNESCO. 2004. *Manual for developing Literacy and Adult Education Programs in Minority Language Communities*. Bangkok: UNESCO. <http://unesdoc.unesco.org/images/0013/001351/135164e.pdf> (accessed Sept. 12, 2013).
- Unseth, Peter. 2005. Sociolinguistic parallels between choosing scripts and languages. *Written Language and Literacy* 8(1), 19–42.
- Van Dyken, Julia R., and Constance Kutsch Lojenga. 1993. Word Boundaries: Key Factors in Orthography Development. In: Rhonda L. Hartell (ed.), *Alphabets of Africa*. Dakar: UNESCO-Dakar Regional Office and Summer Institute of Linguistics, 3–20.
- Venezky, Richard L. 1970. Principles for the Design of Practical Writing Systems. *Anthropological Linguistics* 12(7), 256–270. Reprinted in 1977: Joshua. A. Fishman (ed.), *Advances in the Creation and Revision of Writing Systems*. The Hague: Mouton, 37–54.
- Venezky, Richard L. 2004. In search of the perfect orthography. *Written Language and Literacy* 7(2), 139–163
- Walker, Willard. 1969. Notes on Native Writing Systems and the design of native literacy programs. *Anthropological Linguistics*. 11(5), 148–166
- Wheatley, Julian K. 1996. Burmese Writing. In: Peter T. Daniels and William Bright (eds.), *The World's Writing Systems*. New York: Oxford University Press, 450–456.
- Wiesemann, Ursula, Etienne Sadembouo, and Maurice Tadadjeu. 1983. *Guide pour le développement des systèmes d'écriture des langues africaines*. Collection PROPELCA No. 2. Yaoundé: SIL and University of Yaoundé.
- Wiesemann, Ursula, Etienne Sadembouo, and Maurice Tadadjeu. 1995. Tone orthography and pedagogy. *Notes on Literacy* 21(3), 25–31.
- Wright, Jennifer. 2012. Testing and revising a tone orthography proposal for Chrambo (Bambalang). Yaoundé, Cameroon: Ministry of Scientific Research and Innovation and SIL. <http://www.silcam.org/languages/languagepage.php?languageid=49> (accessed Nov. 1, 2013)
- Writing Systems Research*. Oxford Journals. Oxford: Oxford University Press.
- Written Language and Literacy*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Yunick, Stanley. 2000. Linguistics, TESL, and language planning in Micronesia. *Studies in the Linguistic Sciences* 30(1):183–200.

Elke Karan

elke_karan@sil.org