Journal of West African Languages

Volume 51.2 (2024)

DAVE ROBERTS & STEVE L. WALTER (EDS.): 2021. *TONE ORTHOGRAPHY AND LITERACY: THE VOICE OF TEN NIGER-CONGO LANGUAGES*. AMSTERDAM: JOHN BENJAMINS PUBLISHING COMPANY. XXII + 375 PP. HARDBOUND: ISBN 9789027208439 E-BOOK: ISBN 9789027260147 EUR 105.00 USD 158.00

Keith L. Snider
SIL International
keith_snider1@sil.org

1. Introduction

To quote from Professor David L. Share's (University of Haifa, Israel) foreword, this book is an "eye-opening tour de force" (xiii). As stated in the introductory chapter, "most reading research is conducted on what Henrich, Heine, and Norenzayan (2010a, 2010b) term WEIRD participants (Western, Educated, Industrialized, Rich, Democratic)." The book under review, however, takes one refreshing leap outside of "normal" by reporting on the impact of "full tone marking" on the reading of non-WEIRD participants from ten Niger-Congo language communities from five countries in sub-Saharan Africa. The authors use the term full tone marking "to describe the African tradition of marking one fewer tones than the number of phonemic level tones in the language, which is the case in all ten focal languages" (21). Many tone orthographies mark one tone less than the number of phonemic levels to minimize the number of tone marks in a text.

Following Chapter 1 (Introduction, authored by David Roberts, Johannes Merz, JeDene Reeder), the book is divided into two parts. Part 1, which consists of chapters 2 through 9, is titled "Tone orthography in ten Niger-Congo languages" and describes the linguistic and orthographic backgrounds of each of the ten languages: Chapter 2 (Tem [kdh];[1] David Roberts, JeDene Reeder, Andy Weathers), Chapter 3 (Nateni [ntm]; David Roberts, JeDene Reeder), Chapter 4 (Mbelime [mql]; David Roberts, Johannes Merz, JeDene Reeder), Chapter 5 (Eastern Dan [dnj]; David Roberts, JeDene Reeder, Valentin Vydrin), Chapter 6 (Yoruba [yor]; Matthew Harley, JeDene Reeder), Chapter 7 (Idaasha [idd]; David Roberts, JeDene Reeder), Chapter 8 (Ife [ife]; David Roberts, JeDene Reeder), and Chapter 9 (Elip [ekm], Mmala [mmu], and Yangben [yqv]; David Roberts, Ginger Boyd, JeDene Reeder). Part 2, titled "Tone orthography and literacy outcomes," begins with Chapter 10 (The tone orthography experiment design; David Roberts, Stephen Walter), the title of which is self-explanatory. The chapters that follow then present the results of the experiment for each language in relation to: Chapter 11 (Tone orthography and oral reading speed; David Roberts, Stephen Walter), Chapter 12 (Tone orthography and general oral reading accuracy; David Roberts, Stephen Walter), Chapter 13 (Tone orthography and tonal oral reading accuracy; David Roberts, Stephen Walter), Chapter 14 (Tone orthography and oral reading comprehension; David Roberts, Stephen Walter), and Chapter 15 (Tone orthography and tone writing accuracy; David Roberts, Stephen Walter). Chapter 16, the final chapter, and aptly titled "Epilogue" (David Roberts, JeDene Reeder, Stephen Walter) concludes with thoughts on tone pedagogy, methodological limitations, and prospects for future work.

This review is organized as follows. It begins with brief descriptions of Chapter 1, the introduction, and Chapter 10, which describes the tone orthography experiment design. With some regret, I omit descriptions of the linguistic and orthographic backgrounds of each of the ten languages, as well as the results of the experiments for each language. While I found those interesting and relevant, they may not be so for everyone, so I encourage those who are interested in one or more of the specific languages studied to read about them in the book. Instead, I summarize the broad results of the experiment as set forth in Chapters 11 through 15, and I conclude the review with my own impressions and evaluative comments.

[1] At the first mention of each language name, its ISO 639-3 code is provided between brackets. In the book, this code appears in a footnote at the beginning of each language chapter in Part 1.

2. Book summary

Chapter 1 (Introduction) proclaims the aim of the book as follows: "Our research carries no assumption that zero and full tone marking are the only available options. It is probable that for many African languages partial marking would be more suitable, or that, in others, it would be more appropriate to mark grammatical tonal contrasts than to mark tone phonographically. We do not investigate those possibilities here. Instead, we limit ourselves to a more general research question which is exploratory in nature: *To what extent does full tone marking contribute to oral reading fluency, comprehension, and writing accuracy, and does that contribution vary from language to language?*" (4)

To answer this question, the authors begin by discussing linguistic, literacy, and social differences between the WEIRD societies, where most previous experimental orthography research has taken place, and the non-WEIRD societies in the African context where the present experiments were situated. They then go on to introduce the ten focal languages with respect to their location, genealogical affiliation, language vitality, previous research, and orthography.

Chapter 10 (The tone orthography experiment design) describes how the setup of the experiment generally followed that of Bird's (1999:95–97) tone orthography experiment design for Dschang [ybb] (also known as Yemba), a Grassfields Bantu language spoken in Cameroon. Although the participants knew the experiments were evaluating the orthographies, as opposed to their own abilities, they did not know the experiments were focused on tone. The ten languages and the number of participants from each language included in the study are as follows (in alphabetical order): Eastern Dan (57), Elip (17), Idaasha (32), Ife (31), Mbelime (32), Mmala (20), Nateni (40), Tem (38), Yagben (14), and Yoruba (27), making for a total of 308 participants. To ensure the languages were as representative as possible of tone languages in Africa, the languages chosen did not all have the same number of phonemic level tones. While Tem, Elip, Mmala, and Yangben each have two phonemic tone levels; Ife, Idaasha, Yoruba, Nateni, and Mbelime each have three; and Eastern Dan has five.

Four Dschang narrative texts from Bird's experiment were translated into each of the ten languages and then used as indicated in the following quotation.

> In each focal language, two "booklets" were prepared, each containing the four texts, half with ZERO TONE marking (Z), and half with FULL TONE marking (F), the latter being the standard orthographies of the ten focal languages. Both booklets contained the same texts, but differed in terms of which texts were marked for tone: one presented the texts in the order Z1, Z2, F3, F4; the other in the order Z3, Z4, F1, F2. The four texts were also prepared separately with interlinearized morpheme-by-morpheme glosses for the researchers' use. A fifth text – labeled scorpion – was prepared in the official language (English in Nigeria; French elsewhere), deviating from Bird's (1999) design which did not include an L2 oral reading task. (190)

The research participants read out loud the texts that were assigned to them in their mother-tongues (L1), which included texts with full tone marking as well as different texts with zero tone marking. In addition, they read out loud texts in the official languages of their countries (L2) and then answered questions about the texts they had read to test their oral reading comprehension. There was also a writing exercise that involved adding tone diacritics to those texts unmarked for tone (viz. zero tone texts). The oral reading tests were evaluated with regard to speed, accuracy, and comprehension, and the writing tests with regard to accuracy. Unlike most orthography tests conducted prior to the present century, which report only averages, the results of the current research experiments benefitted from a rigorous statistical methodology. Moreover, this methodology also took into consideration not only expected independent variables like the participants' age and gender, but also variables highly

relevant to societies with less-than-ideal health care and education opportunities (e.g., quality of eyesight and level of formal education).

Chapter 11 (Tone orthography and oral reading speed) notes that overall, L1 reading speed was only slightly faster with texts that were marked for tone, but that the difference was only statistically significant in Nateni and Ife. This suggests that other factors might affect L1 oral reading speed more than full tone marking. Indeed, once the independent factors were included in the modelling, it became clear that L1 reading speed was influenced more by level of education than by anything else. Comparing L1 oral reading speeds with their L2 counterparts, most participants read faster in their L2 than in their L1, regardless of whether tone was marked or unmarked in their L1. This was attributed to the fact that in the African context, most people learn to read first in their L2, and only later, often only in adulthood, in their L1, giving them considerably more practice reading in their L2. Moreover, the L2 reading speed turned out to be a strong predictor of the L1 reading speed, such that the faster a participant's L2 speed, the faster the L1 speed.

Chapter 12 (Tone orthography and general oral reading accuracy) defines accuracy as "the ability to recognize known words and decode unknown words correctly" (221). Together with speed, accuracy is one of the major predictors of comprehension. In fact, it is an even better predictor since "children who read accurately but slowly usually have good comprehension so long as their oral language comprehension skills are also good (Adlof et al, 2011, pp. 197–198)" (221). Like the results for oral reading speed, the results for oral reading accuracy suggest that full tone marking is only incidental to observed differences in reading accuracy error rates in nine of the ten languages tested, the exception being Nateni, which showed a statistically significant increase in accuracy.

Chapter 13 (Tone orthography and tonal oral reading accuracy) focuses exclusively on the contribution full tone marking makes to reading tones accurately (i.e., to pronouncing the tones correctly when reading), as opposed to general reading accuracy. In this respect the present study differs from most previous studies that test the contribution of full tone marking on reading accuracy. Like the present study, many previous studies also compared the reading accuracy of texts fully marked for tone with texts unmarked for tone, but the previous studies did not distinguish between general reading errors and tone errors. The present study sets forth tonal oral reading errors as a sub-set of the general reading errors described in Chapter 12. Similar, however, to the results for general reading accuracy, the results for tonal reading accuracy suggest that for most of the languages investigated, while tone marks do not seem to make any significant difference to tone reading accuracy, they also do not seem to cause any harm, which may be an indication that the readers for whom this is true are simply ignoring the marks entirely. However, for Nateni, Ife, and marginally for Mbelime, tone marking did improve tone reading accuracy to a degree that was statistically significant.

Chapter 14 (Tone orthography and oral reading comprehension) deals with what is ultimately, of course, the primary aim of reading. The authors claim that "the present chapter… represents a new departure by introducing a comprehension measure as a major component into an African tone orthography experiment for the first time" (252). Unfortunately, the present study only reports on the comprehension results of seven of the ten focal languages due to participants of the three Bantu languages, Elip, Mmala, and Yangben, having inadvertently been given access to the texts while undergoing the test. Nevertheless, mean comprehension on the languages included was greater on the L1 full tone texts (zero tone 69.8%, $SD = 24.08$; full tone 77.48%, SD = 22.27, and this difference was statistically significant ($t = 3.67$, $p = 0.000$). Although statistically significant, this effect is nevertheless very modest (Cohen's $d = 0.33$). As explained in footnote 6, on p. 255, "A Cohen's D value of 0.2 indicates a weak effect size, while 0.5 is moderate and 0.8 is strong." Comparing the L1 comprehension results with those of L2, a paired sample *t*-test showed them to be almost equivalent, although comprehension of the L1 texts with full tone marking was significantly higher than that of the L2 texts, again confirming the modest advantage obtained from full tone marking.

The authors note that of the three oral reading measures (speed, accuracy, and comprehension), comprehension was the only one that measured higher in L1 than in L2, which they attribute to the participants having a better knowledge of L1 than of L2. As they remark, "while knowing the language of a text does not guarantee comprehension, not knowing it does guarantee lack of comprehension" (260), and undoubtedly the participants' better understanding of the L1 languages gave them an advantage over the L2 languages in these tests. The authors also note that L2 comprehension is "strikingly independent of speed, error rate, and education level, which demonstrates that while "a higher level of formal education leads to faster L2 oral reading…, worryingly, this does not contribute to L2 comprehension" (261).

While full tone marking improved the comprehension level for many participants, it also made no difference for some participants and even led to a decline in comprehension for others. This suggests that for those whose comprehension was not improved by having tone marked, they may not have been paying attention to the tone marks.

Chapter 15 (Tone orthography and tone writing accuracy) points out that writing accuracy is dependent on numerous sub-skills such as motor dexterity and cognitive skills such as the ability to reason, remember, and concentrate. Understandably, Chapter 15 focuses only on the ability to correctly place tone diacritics on unmarked texts.

Sadly, only 35% of the participants actually completed the tone writing task, and 14 participants did not even attempt it. This suggests that most participants were either unable to perform the task within the allotted 20 min. time limit or thought they were unable to perform it. Moreover, the results obtained were also not encouraging. Overall, participants were only able to write some 60% of the tone marks correctly, although these results had an unusual distribution. "Although no one scored 100% and only eight participants scored 97.5–100%, about a third of the sample scored between 87.5–97.5% success" (270). The remaining scores were fairly evenly distributed from 2.5–82.5%. All this suggests that full tone marking with respect to the orthographies investigated and the pedagogical methods employed are failing so far as writing tone is concerned.

Factoring into the analysis other independent variables produced some interesting information. For example, the distribution of the scores indicates that the more formal education participants had, the lower their tone writing scores. "This suggests that immersion from a young age in a non-tonal L2 (whether it be French or English) trains learners later in life to ignore tone marks when, as adults, they learn how to read and write the L1" (276–277). Also, the ability to read tone well does not guarantee the ability to write tone well.

Chapter 16 (Epilogue) pulls things together by providing an overview of the essential outcomes of the experiments, summarizing the impact of independent variables, discussing pedagogical issues that arise from the results, and discussing experimental design limitations. Since the essential broader outcomes of the experiments are already summarized above, as well as some of the impacts of the independent variables, we turn our attention to summarizing some of the pedagogical issues discussed, as well as some of the methodological limitations and shortcomings.

The authors make several observations and recommendations for those involved in designing literacy primers. "Two of the five languages that provide experimental evidence in favor of full tone marking – Nateni and Ife – are also those in which the published teaching materials place the strongest emphasis on learning to read tone. They introduce tone early and provide ample opportunities for tone practice in both exercises and stories" (295). Secondly, most of the primers of the focal languages teach tone from the perspective of contrasting the tones associated with individual TBUs, as opposed to contrasting the tone patterns associated with words, the latter being a practice advocated in Snider (2018:98). Of those that demonstrated some effort to teach contrastive tone patterns, only a few patterns were taught and inadequately taught at that. The authors recommend teaching pairs of all major contrastive tone patterns systematically. Thirdly, despite the functional load of grammatical tone being relatively heavy in most of the focal languages, little effort was made in the primers to focus on

grammar. The authors encourage primer developers of languages with significant grammatical tone to focus more on grammar. Finally, those teaching the literacy classes were mostly not well taught how to teach people to read tone marking, and the authors recommend that more time and materials be devoted to tone pedagogy.

Moving on to methodological limitations and shortcomings, the experiment design presented participants with toneless texts first, and it was felt that this might have influenced the results negatively. In the future, the authors recommend that any testing similar to that of the current experiments should counterbalance this potential negative effect by mixing up the order of text presentation. During the testing, the authors realized that ethno-literacy variables had much more to do with the variation in performance between focal languages than did their linguistic and orthographic profiles, so the authors recommend that all future testing improve and expand the list of ethno-literacy variables employed in the current experiments. "Similarly, the statistical analysis did not control for pedagogical variables, so any potential link between tone pedagogy and tone reading fluency has not yet been adequately explored or quantitatively demonstrated" (299). The researchers therefore encourage future researchers to factor pedagogy into their analyses. Regarding comprehension, the researchers feel that better questions and more than three questions should be used to evaluate reading comprehension. They also recommend inclusion of a variable that considers how consistently the orthographies represent the tone systems. For example, while some orthographies represent all tones faithfully, in some of the languages, "the absence of an accent means either L tone, or H tone on some class proclitics" (302).

In their final remarks, the authors raise several issues. They note that the availability of print in African languages is nowhere near sufficient for speakers to become truly proficient readers. And they conclude by saying that while tone analysis has its place, it must be complemented by assembling "persuasive evidence from the classroom, in the hope that decision makers, in turn, will adopt evidence-based strategies" for orthography development (304).

3. Impressions and evaluative comments

This book impressed me with its breadth, depth, and professional quality, and it is exceptionally well-written, a fact that also did not go unnoticed by Professor David Share in his foreword. There is simply no other book like it, and the authors are to be commended for attempting to formally test tone orthographies on this scale in Africa. The research was well thought out and implemented, and it serves as a superior model for future orthography testing. One of its greatest strengths is the employment of serious statistical analysis in a manner that is accessible to both readers who have only a rudimentary acquaintance with statistics as well as to those who are well-trained in statistics. The book includes an appendix, which details the independent variables factored into the analyses, as well as a topical index and a very useful language index. In the language index, numerous sub-entries appear under each of the study's ten focal languages that direct one to the locations where topics relevant to each of the languages are discussed. For example, under Yoruba, the sub-entries include age of participants, ATR vowels, L1 reading accuracy, L1 reading comprehension, L1 reading speed, lexical tone, and tonal processes.

Naturally, a book of this breadth has some shortcomings. As much as testing the efficacy of full tone marking of phonemic tone orthographies is valuable, *how* an orthography represents what it represents is probably not as important as representing the right thing to begin with (e.g., grammatical tone vs. lexical tone). Personally, I would have preferred they tested which level of orthographic depth is better (i.e., underlying, phonemic, lexical, or surface), like the study carried out for Kabiye [kpb] in Roberts et al (2016), but on a broader scale. Or perhaps whether representing phonemic tone is better than representing grammatical constructions uniquely that are differentiated solely by tone? But as the authors acknowledge (301), researching issues like these would have required more extensive phonological analysis of each language than was available, especially considering the desired scope of the project.

I found the authorship of the volume strange. The book was edited by David Roberts and Stephen Walter. Of the sixteen chapters, Roberts was the lead author for all but Chapter 6, the only chapter he did not co-author. JeDene Reeder was a co-author for Chapters 1 through 9 and for Chapter 16, and Walter was a co-author for Chapters 10 through 16. Three of the four other contributors co-authored only one chapter each, and the other one co-authored only two chapters. The result is an edited volume that reads with the flow and unity of a non-edited volume. This, of course, is a good outcome, but it makes one wonder why it is an edited volume, as opposed to a volume simply authored by all of the authors, or perhaps authored by Roberts, Walter, and Reeder, with acknowledged input from the others?

As to be expected in a volume of this size, there are a few errata, but most will escape all but the most ardent of errata searchers. Not finding any reference to the actual number of research participants for each language, I wrote to the corresponding editor for this information and was promptly rewarded not only with the requested information, but also with all the errata the editors had collected for future editions! It was not a long list.

Minor deficiencies aside, this book offers a lot. Reading it confirmed several things that just made sense. For example, the better one knows a language, the better one will be able to read in that language and reading and writing performance increases with more and longer training. It also confirmed some things I had suspected. I had previously observed with testing of certain tone orthographies that the difference between reading with tone-marked texts and reading with zero-marked texts was often not statistically significant, and I had long suspected this lack of difference was because readers weren't looking at the tone marks in either test. In the better controlled experiments under review, the authors came to a similar conclusion, although it applied mainly to those who had previously been taught to read non-tone languages. Having first been taught to read without tone marks, those readers now ignored tone marking when reading tone-marked texts. I had also suspected that marking lexical tone in languages with longer word structures might not improve reading performance due to the increased segmental information available for decoding the meaning, and this also was confirmed.

More important to me than having my suspicions confirmed, however, was providing the answer to a long-standing question: Does tone marking actually improve reading and writing? And the answer is a qualified "yes" for the languages investigated. It was only advantageous, however, for tonal reading accuracy in Nateni and Ife, and for reading comprehension in Nateni, Ife, and Mbelime. Moreover, when the independent variables were factored into the analyses, it showed that when tone marks were added, those with less formal education had better comprehension than did those with more formal education. And those with less formal education also wrote tone marks more accurately than did those with more formal education. From this, the authors concluded that tone marking only benefitted those participants not previously trained to ignore tone marks by first learning to read non-tone languages.

This is important because it nuances Bird's (1999) conclusion that full tone marking reduces fluency in reading in Dschang. In the Dschang experiment, Bird conflated comprehension and accuracy, a design weakness he acknowledged and encouraged others to avoid. This highlights the need to separate comprehension from accuracy when testing orthographies. It also highlights the need to provide training that compensates for the tendency of better educated speakers to ignore tone marking.

I trust that by now, readers of this review will understand that this volume is a most needed and valuable addition to the library of anyone involved in orthography development. It is a must read, not only for its conclusions, but also for the testing methodology employed. Newly minted orthographies must be tested and the results published. The volume under review presents an excellent model for how to carry this out. Finally, this volume underscores the need for better tone orthography pedagogy and for more dialogue between linguistic researchers, literacy specialists, and reading researchers.

References

Adlof, Suzanne M.; Charles A. Perfetti; and Hugh W. Catts. (2011). Developmental changes in reading comprehension: Implications for assessment and instruction. In S. J. Samuels & A. E. Farstrup (eds.) *What research has to say about reading instruction, 4th edition*, 186–214. Newark, DE: International Reading Association.

Bird, Steven. (1999). When marking tone reduces fluency: An orthography experiment in Cameroon. *Language and Speech*, 42, 83–115. https://doi.org/10.1177/00238309990420010401

Henrich, Joseph; Steven J. Heine; and Ara Norenzayan. (2010a). Most people are not WEIRD. *Nature*, 466, 29. https://doi.org/10.1038/466029a

Henrich, Joseph; Steven J. Heine; and Ara Norenzayan. (2010b). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–135. https://doi.org/10.1017/S0140525X0999152X

Roberts, David; Keith Snider, and Stephen L. Walter. (2016). Neither deep nor shallow: Testing the optimal orthographic depth for the representation of tone in Kabiye (Togo). *Language and Speech* 59(1), 113–138. https://doi.org/10.1177/0023830915580387

Snider, Keith L. (2018). *Tone Analysis for Field Linguists*. Dallas: SIL Publications.