

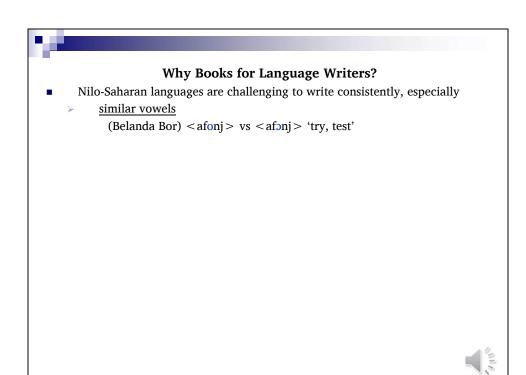
Documenting Language Structures for Language Writers

Timothy M. Stirtz

SIL International



My talk is about making books that help language users write more consistently.



So why do language communities need help writing their language? Well, we all familiar with many complexities in Nilo-Saharan languages, and these complexities make writing a challenge, even when a well-chosen writing system has been decided. I want to mention 5 common challenges for writers of Nilo-Saharan languages. The first is similar vowels. I've seen Belanda Bor writers go back and forth trying to decide if certain words are plus or minus ATR. Sometimes the sound actually differs from one speaker to the next, but usually it's just a matter of the writer learning which letter goes with which sound.

```
Why Books for Language Writers?

Nilo-Saharan languages are challenging to write consistently, especially

similar vowels
(Belanda Bor) < afonj > vs < afonj > 'try, test'

tone representation
(Jur 'Bëli') < dɔ-gbe > 'on the son' vs < dɔ gbe > 'head of the son'
[dɔ̄ gbɛ̄]

[dɔ̄ gbɛ̄]
```

Many Nilo-Saharan languages need at least some marking for tone in the writing system in order to distinguish meaning. Beli has a few prepositions derived from body parts, like the word 'on, above' that comes from the word 'head', and some of these are only distinguished by tone. So the Beli decided to hyphenate all prepositions to make this distinction clear in writing.

```
Why Books for Language Writers?

Nilo-Saharan languages are challenging to write consistently, especially

similar vowels

(Belanda Bor) < afonj > vs < afonj > 'try, test'

tone representation

(Jur 'Bëli') < do-gbe > 'on the son' vs < do gbe > 'head of the son'

[dō gbè]

word-breaks

(Bongo) < bilehi ji > 'our worshiping' vs < bi lehi ji > 'he worships us'

/bi-lehi ji / NMLZ-verb 1PL

[bìlèhì jī]

[bílèhì jī]
```

Word breaks are another challenge for writers. In Bongo, there is a nominalizer prefix that is identical to the third singular pronoun 'he' except for tone. For this and other reasons, the Bongo choose to attach the nominalizer and separate the pronoun, but it is challenging for writers to remember which to separate.

```
Why Books for Language Writers?
Nilo-Saharan languages are challenging to write consistently, especially
   similar vowels
     (Belanda Bor) <afonj> vs <afonj> 'try, test'
   tone representation
      (Jur 'Bëli') <do-gbe > 'on the son' vs <do gbe > 'head of the son'
                 [dɔ̄ gbɛ̀]
                                           [ádg cb]
   word-breaks
      (Bongo) < bilehi ji > 'our worshiping' vs < bi lehi ji > 'he worships us'
              /bi-lehi ¡i/ NMLZ-verb 1PL
                                              /bi lehi ji/ 3SG verb 1PL
              [bìlèhì jī]
                                               [bílèhì jī]
   sound changes at morpheme boundaries
      (Gəəmg) < wirí > 'he kills' vs < wirrí > 'in order for him to kill'
                /wir-i/ verb-3S
                                    /wīr-da-î/ verb-SBJV-3SG
                [wīrî]
                                    [wírî]
```

Sound changes at morpheme boundaries can also be a challenge for writers. In Gaahmg, the subjunctive suffix –da sometimes assimilates to a root-final /r/. Then with vowel elision, the verb 'in order for him to kill' is the same as 'he kills' except for tone. The Gaahmg decided to write all subjunctives with two medial consonants to make this distinction clear, but this rule is still a challenge to remember.

```
Why Books for Language Writers?
Nilo-Saharan languages are challenging to write consistently, especially
    similar vowels
     (Belanda Bor) <afonj> vs <afonj> 'try, test'
    tone representation
      (Jur 'Bëli') <do-gbe> 'on the son' vs <do gbe> 'head of the son'
                                           [ádg cb]
                 [dɔ̄ gbɛ̀]
    word-breaks
      (Bongo) < bilehi ji > 'our worshiping' vs < bi lehi ji > 'he worships us'
              /bi-lehi ji/ NMLZ-verb 1PL
                                              /bi lehi ji/ 3SG verb 1PL
              [bìlèhì jī]
                                              [bílèhì jī]
    sound changes at morpheme boundaries
      (Gəəmg) < wirí > 'he kills' vs < wirrí > 'in order for him to kill'
                                    /wīr-da-î/ verb-SBJV-3SG
                /wir-i/ verb-3S
                [wīrî]
                                    [wírî]
    similar consonants
  (Belanda Bor) 'only' <gbɛl> vs <ngbɛl> 'only'
```

Similar consonants can also be a challenge for writers, such as at the beginning of the word 'only' in Belanda Bor. As with vowels, this is often a matter of the writers learning which letter to write for which sound.

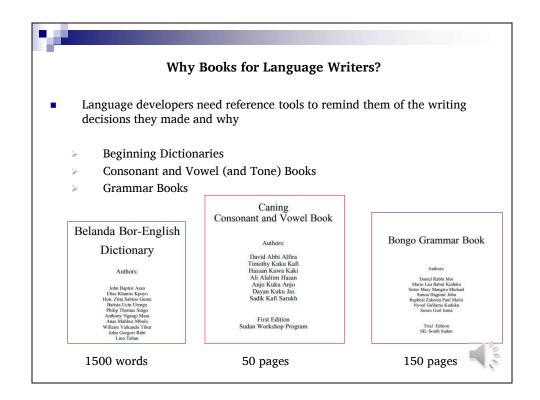
Why Books for Language Writers?

- To be true owners of the language, speakers must
 - decide their writing system (Stirtz 2015)
 - consistently use their writing system (Stirtz 2018)

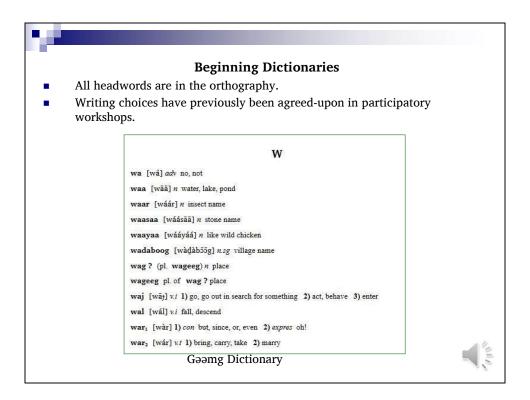




We want language speakers to be the owners of their language. But in order to be the true owners, they need to both decide their writing system and consistently learn to use it. I've facilitated participatory workshops for 10 language projects, and written about this in a 2015 paper. In these workshops, the community representatives become well-informed of their language patterns and then decide writing choices accordingly. But I've learned that in order to consistently write their language, they also need to have ongoing guidance after the workshops. So, as reading books and other materials are drafted, I check each word and remind them of the writing rules they decided to use. Through this checking process, their writing skills significantly improve in a relatively short time. I've written about in a 2018 paper.



But in order to remind community representatives of how they decided to write and why, there needs to be some documentation. So everything that we learn and decide in the participatory workshops is presented in three books—a beginning dictionary of about 1500 words, a consonant and vowel book of about 50 pages, and a grammar book of about 150 pages. On the last slide of this PowerPoint, I list a link for the books I will be showing examples from the rest of this talk.



The beginning dictionaries use the orthography for all headwords. These are spelled the way speakers decided in the participatory workshops. Since the dictionary is uploaded to Webonary, the words can easily be searched, and writers can check spelling as they produce new reading books.



Beginning Dictionaries

• The introduction explains how to read entries and what each word means

How to Read an Entry in this Dictionary

When you look in the dictionary, you will see many lines of words. Some look similar to the example below:

bukutu n pot for sauce ka bukutu

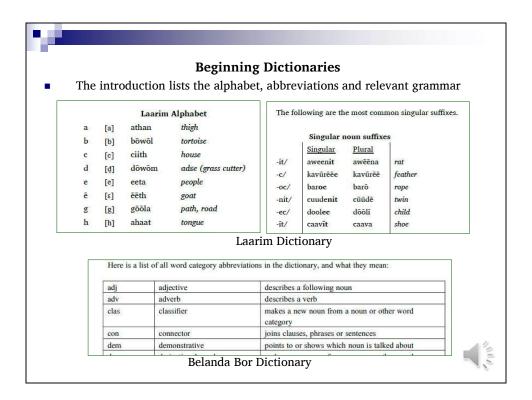
Here's what the words mean:

- The word bukutu in bold is called the 'entry'. All the words after it give information about this
 word. All entries are listed according to the order of the Belanda Bor alphabet letters. See the
 Belanda Bor Aphlabet Order below for the list of letters in order. Entries that are nouns are in
 the singular form (only one of the noun). Entries that are verbs are in the incomplete form of the
 verb (following na 'be, is').
- Next, the letter <u>n</u> in italics shows the word category (type of word) of the entry is a noun. See
 the Word Category Abbreviations below for a list of all possible word category abbreviations.
- Next, the words pot for sauce is the definition of the entry and shows its meaning in English.
- The next words ka bukutu in bold is the plural of the entry. It is how to say more than one of the
 noun. Only entries that are nouns have a plural form. The words ka, nying, or nyi before a noun
 show it is plural. Many nouns do not have a plural form.

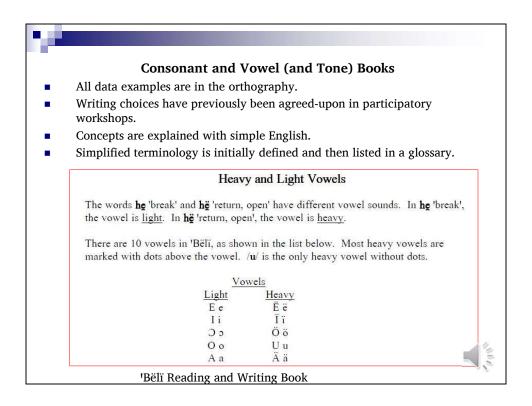
Belanda Bor Dictionary



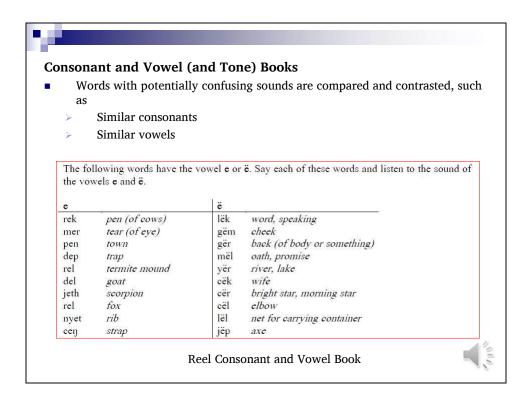
The dictionaries also have an introduction that explains how to read the entries and what each word means.



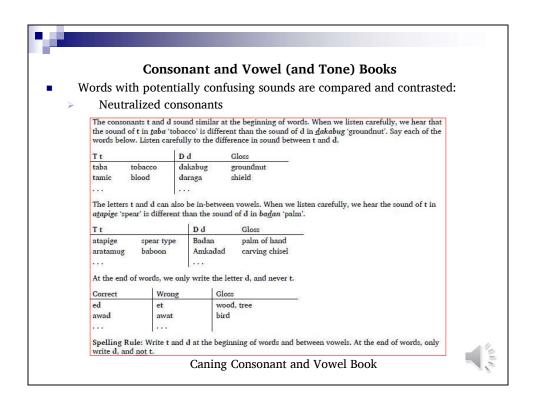
The introduction also includes the alphabet, with letters paired with the IPA. There is also a list of abbreviations for the word categories and their definitions. And some relevant grammar is also included, such as the singular and plural noun suffixes that are listed in some entries.



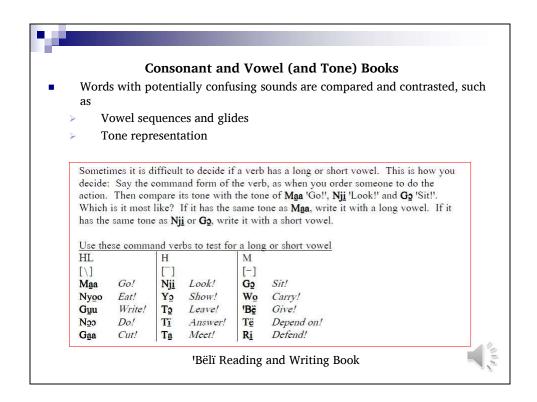
Now, let me talk about consonant vowel books, which also may includes tone, depending on the language. These books present the root level phonology relevant to writers. Again, all data examples are in the orthography, and speakers decided the spelling of all words in previous workshops. Concepts are explained with simple English and terminology is simplified. For example, the words "heavy and light" is used instead of plus and minus ATR. These words underlined here in this lesson are also listed in the glossary with a definition and example.



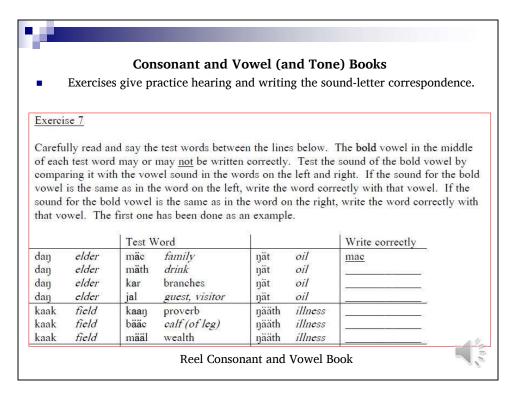
The main reason for the consonant and vowel book is to teach potentially confusing sounds, such as similar consonants and vowels, including ATR and consonant and vowel length. We try to list at least 10 words for each similar sound in a table like this. Then, as writers read down the list, the sound becomes more clearly cemented in their minds. They can also read across to more clearly hear the difference in the two sounds. When they find new words with one of the sounds, they can compare it with these lists and choose the correct letter for that sound.



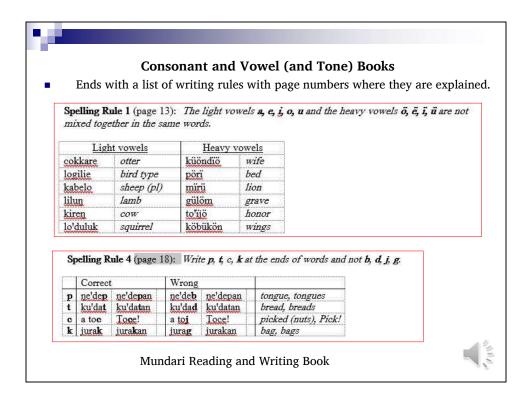
Other lessons deal with the neutralization of consonant contrast. Here the lesson shows the contrast at the beginning and middle of words, but not at the end. So the writer should only use one of the letters in this final position.



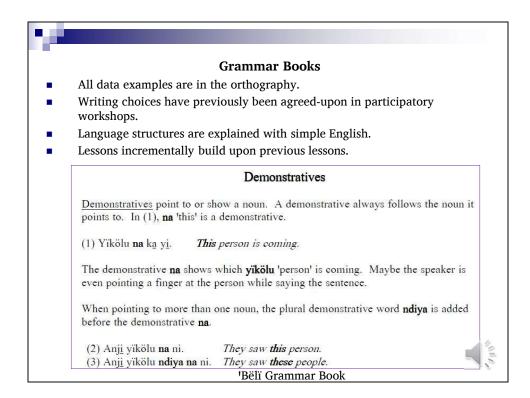
There are also lessons for distinguishing vowel sequences from glides and how to represent tone. In Beli, there are is no vowel length distinction, so long vowels are used in the orthography to show falling tone instead of level tone. (Many of the tone minimal pairs are between verbs and other word categories, so all verbs have an underlined vowel to reduce this ambiguity.) Here, the verbs are grouped according to their tone pattern. So, when writers don't know the tone of a new verb, they just compare it with these groups to decide.



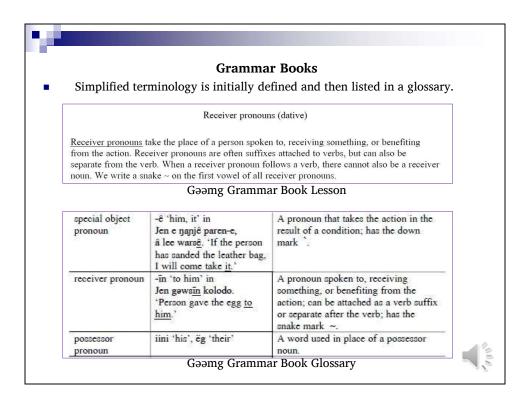
At the end of each lesson, there is an exercise for practice hearing and writing the sound-letter correspondence.



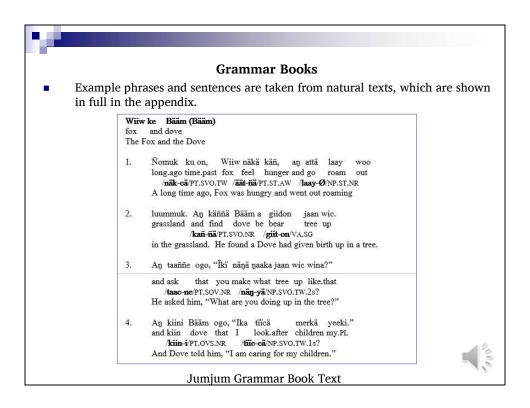
The book ends with a list of writing rules presented earlier in the book, with page numbers where they are further explained. Here there is a rule for how plus and minus ATR vowels are not mixed in the same word, and another rule for only writing certain consonants at the end of words because of neutralization.



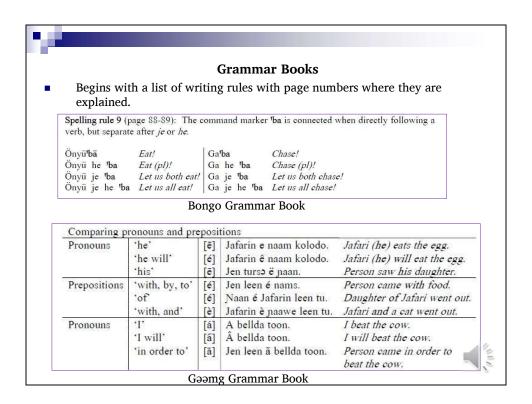
Now, let me talk about grammar books, which are the most extensive of the three books. Again, all data examples are in the orthography with previously chosen spelling. The morphemes, word categories and other structures are explained in simple and brief English, followed by examples of the morpheme in context. Lessons incrementally build upon previous lessons and only a few technical terms are introduced in each lesson.



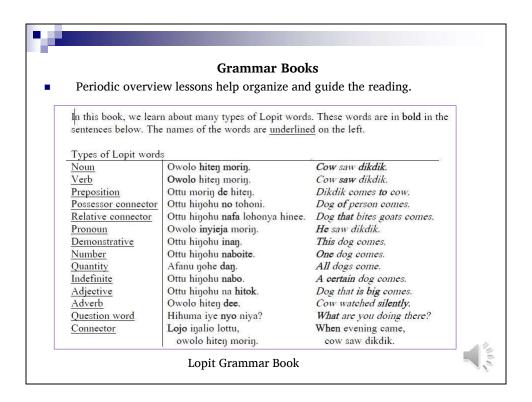
Terms are sometimes simplified. For example, the words "receiver pronoun" is used instead of "dative". They are initially defined and underlined when introduced in the lesson, then in the glossary are again listed with a definition and example.



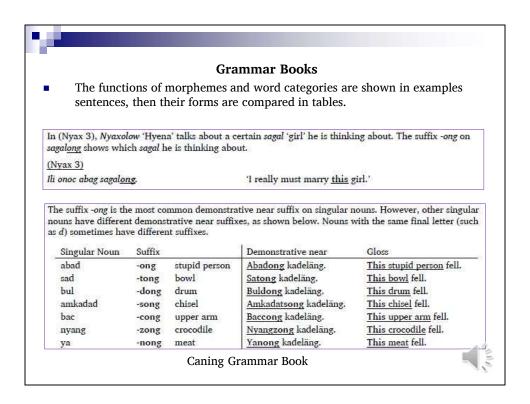
The examples of each lesson are often taken from natural texts. Although there is just a free gloss of the example in the lesson, the full interlinear of the entire text is presented in the appendix.



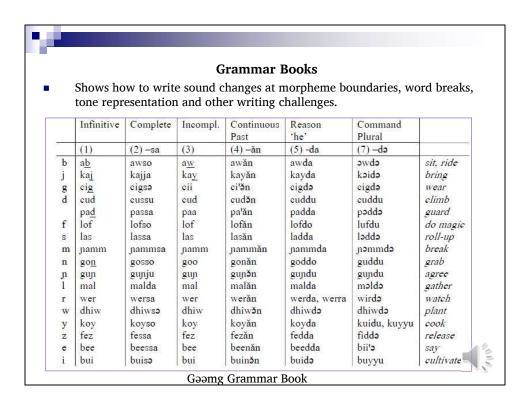
The grammar book also has a list of writing rules with page numbers where each rule is further explained. Here we see a writing rule that shows when the Bongo imperative morpheme is bound and when it is free, based on ATR spreading. The Gaahmg table is not a writing rule, but shows the importance for several writing rules following this table that specify grammatical tone marking. The table lists several sentences where the pronoun or preposition could be confused if the marking wasn't written.



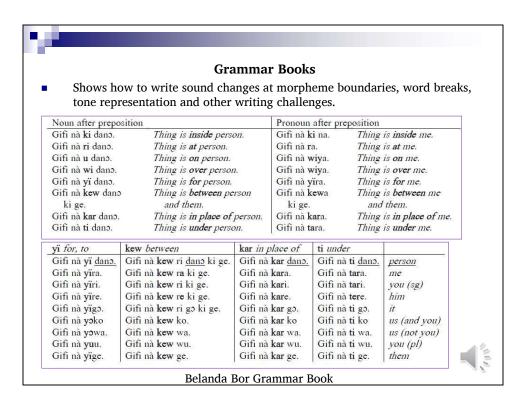
To make the grammar books a bit more readable, there are occasional overview lessons. Here we see a list of Lopit word categories with examples sentences. Next in this grammar book, there is a lesson for each of these categories and in this order.



Many lessons begin by explaining the function of a morpheme in its natural context, such as the demonstrative suffix —ong in Caning. Then immediately afterwards in the same lesson, the alternations of the morpheme are presented. If you look closely at the second box, you can see that three of the singular nouns end with the same consonant d, but each take a different unpredictable suffix. So, these alternations are also important for writers to remember.



The main purpose of grammar books is to show how to write sound changes at morpheme boundaries, word breaks, tone representation and other challenges for writing. Here the Gaahmg verb table shows how to write the irregular assimilation of root-final segments in various verb forms. Writers often struggle to remember how to write a certain verb when drafting a reading book. But referring back to this table can help them see the patterns and why they chose to write it in that way.



Here the Belanda Bor grammar book shows where they decided to join and separate pronouns following prepositions. Although all nouns are separate, the first three pronouns are often joined to the preposition because of assimilation, elision or other processes. Although speakers often make wise writing choices when seeing the full paradigm, the choices are difficult to remember when they are drafting reading books without tables such as these.



Grammar Books

Exercises give practice identifying morphemes and various constructions.

Exercise 12

In the sentences below, underline all prepositions. Do not underline any words that are not prepositions. The first sentence is done as an example.

Mämb 3-4

Loʻngu moko üdü nga <u>'ba</u> nyihi <u>dongara</u>
ye ngara höli monoo ngori
ka Mämbillingänjä ye <u>na</u> Mämbillingämä.

A big fight began <u>at</u> moon <u>between</u> two
great birds called Mämbillingänjä
with Mämbillingämä.

(Mämb 7-8)

Ye na ng nyere hi ji ye raa

Those with authority in their hands
nabi ga'dee ye gu'bu di'ba nyihi.

(Mämb 9)

Ye hölï ngori na roo, ye raa nabi mbi'dee ro dohii na 'bagajee.

Then the two birds came to our land.

(Mämb 14)

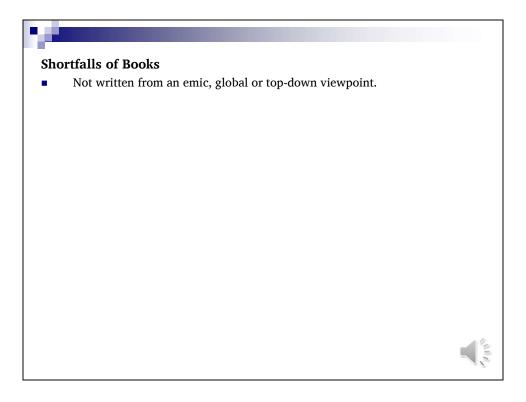
Mbili Mämbilingänjä na dikori do hu gewegewe mile mbili mehe yama.

Mambilinganja's ears hang around her head like oversized spears.

Bongo Grammar Book



The grammar books also have exercises after each lesson to give practice identifying morphemes and various constructions.



In closing, I'd like to mention a few drawbacks and benefits of these books for language writers. First, the books are not ideal in that they are not written from an emic, global or top-down viewpoint. That is, the books don't represent the thinking about the language patterns of the speakers themselves. Rather, the presentation is from my viewpoint as an outsider. As much as I'd like for the speakers themselves to draft the books, I haven't found a way to guide them in that process.



Shortfalls of Books

- Not written from an emic, global or top-down viewpoint.
- Not generally accessible for language communities
- Used by non-speakers more often than speakers.



Second, the books are too academic for many in the language community to engage. The books are not really intended for the general communities, but for writers who want to produce reading books for their communities. Often, even the writers are not able to effectively use the books on their own if they are not accustomed to using reference tools. They usually need assistance from outsiders to know which page to look at for relevant data as they draft books.



Shortfalls of Books

- Not written from an emic, global or top-down viewpoint.
- Not generally accessible for language communities
- Used by non-speakers more often than speakers.

Benefits of Books

- Comprehensive enough to cover 80% of language constructions involving the writing system.
- Contain quality data that can often be used in academic publications.



But having said this, the books do have benefits. They cover 80% or more of the language constructions involved in writing decisions. Because the data has been discussed and agreed-upon by languages speakers in a group setting, it is often more accurate than elicited data from a single speaker. So, much of the data can be used as a basis for academic writing.



Shortfalls of Books

- Not written from an emic, global or top-down viewpoint.
- Not generally accessible for language communities
- Used by non-speakers more often than speakers.

Benefits of Books

- Comprehensive enough to cover 80% of language constructions involving the writing system.
- Contain quality data that can often be used in academic publications.
- Adequately demonstrate the agreed-upon writing system to language developers.



Finally, the books do fulfill their main purpose as reference tools, which is that writers can use then for improved writing consistency. Academic publications typically include illustrative examples in IPA and the morphemes are often shown in only a few contexts. But these books show tables where all possible sound changes are represented in the orthography for each morpheme. They show where morphemes are bound and when they are free. They show how tone is represented in writing. They compare word lists with similar sounds so that writers can decide the sounds of other words. In short, they can enable writers to produce reading materials with improved writing consistency. And the improved consistency gives the writing system a chance to become standardized, even challenging Nilo-Saharan languages with all their

complexities. Thank you.



Grammar Books

- Laarim Grammar Book www.sil.org/resources/archives/58705
- *Didinga Grammar Book* <u>www.sil.org/resources/archives/58715</u>
- *Tennet Grammar Book* www.sil.org/resources/archives/58732
- 'Beli Grammar Book www.webonary.org/beli/language/grammar/
- Mundari Grammar Book www.webonary.org/mundari/overview/copyright/grammar/
- Caning Grammar Book www.webonary.org/caning/language/grammar/
- Lopit Grammar Book www.webonary.org/lopit/language/grammar/
- Belanda Bor www.webonary.org/belandabor/language/grammar/
- Bongo Grammar Book www.webonary.org/bongo/language/grammar/
- Gaahmg Grammar Book https://www.webonary.org/gaahmg/language/grammar/
- Jumjum Grammar Book https://www.webonary.org/jumjum/language/grammar/
- Reel Grammar Book https://www.webonary.org/reel/language/grammar/



References

- Stirtz, Timothy. 2015. Rapid Grammar Collection as an Approach to Language Development. SIL Electronic Working Papers. http://www.sil.org/resources/publications/entry/62483
- Stirtz, Timothy. 2018. Linguistic Checks: The Missing Link in Modern Language Development. SIL Electronic Working Papers.
 - https://www.sil.org/resources/publications/entry/76614