Introduction

This document is intended to guide a linguist through the process of developing an orthography for a Bantu language. It contains steps and principles you should follow. It is distributed together with other accompanying documents. The document, xx_orthography_statement.doc, is a blank template for an orthography statement for a language. An orthography statement is a technical document written for linguists and education officials that records decisions made about the orthography. The document, Xx_transition_primer.doc, is a blank template for a transition primer whose target audience is educated speakers of the language who have learned to read in the national language. These documents are based on the experience of numerous SIL teams as well as orthography statements developed for Bantu languages throughout the southern half of Africa.

You should not attempt to develop an orthography unless you are a linguist or literacy specialist. If you are not a linguist or literacy specialist, you should get the help of a consultant from your national university or SIL. The consultant should have experience in analyzing the phonology and developing an orthography for a Bantu language. A good orthography is based on linguistic insights into the phonological system of a language. A good orthography is also based on sound pedagogical principles. Orthographies sometimes fail because they are not linguistically sound and do not correspond to the phonological system in the reader’s mind. They also fail because they do not take into consideration sociological pressures from government and education authorities, or the appearance of the orthography in contrast with the orthographies of surrounding languages. Unless you take these factors into consideration, your orthography will very likely fail.

Steps in developing an orthography.

There are four phases in the development of the orthography.

Stage 1: Preliminary research phase.

[from ECG Language Programme Planning Guidelines] Early in the first year of this phase, interaction with a consultant is needed to discuss options for recording linguistic data and developing a phonological analysis of the language. Any previous written materials on or in the language need to be looked over carefully and language speakers consulted about the problems they find with the orthography used. During the first year in the project, a team should do a phonological analysis of at least 1000 words and discuss this with a phonology consultant. During this stage, a team can record and type their linguistic data and analysis, write up preliminary drafts of language learning lessons, and devise test items intended to gather feedback on linguistic and sociolinguistic questions while linguistic decisions are still pending. Together with the consultant, the team should make some preliminary decisions about the basic alphabet, and make initial decisions on writing tone and word boundaries. It

1 The 'xx’ in the name should be replaced by the name of your language (e.g. Lugungu_orthography_statement.doc).
is desirable that the alphabet chosen harmonize in main points with other alphabets used in
the area, if appropriate.

During this phase, copies should be made only for developing test items for
orthographic, phonological, sociolinguistic and grammatical hypotheses. Distribution should
only be to people who are important to this testing, not for public relations purposes.
(Normally no more than 20 copies should be made, but reasonable exceptions may be
allowed by the Supervisor.)

1. **Collect 10,000-12,000 words in a dictionary database.**

The Dictionary Development Program\(^2\) can help you to do this. A large dictionary will give
you examples of all sounds and patterns in the language. If you find homographs that are
pronounced differently, this will indicate you have a case of underdifferentiation. The
dictionary will also provide examples for your literacy materials. As the orthography
develops, periodically update the citation form of the words in the dictionary to keep them
consistent with the orthography. Doing this will ensure that the orthography correctly handles
all the words of the language. Once the orthography is reasonably stable, it would be good to
publish the dictionary. Users frequently consult a dictionary to determine the correct spelling
of a word. Consequently one effect of a dictionary is to standardized spelling.

2. **Produce a phonology statement covering the segmental phonology.**

It is essential that an orthography be based on a thorough understanding of the phonological
system of the language. In the early stages in the development of an orthography you should
concentrate on understanding the segmental phonology\(^3\) of the language. In Bantu languages
the segmental issues are relatively simple and straightforward. The suprasegmental issues,
such as phonological word breaks, intonation, and tone, are far more difficult to resolve.

The Bantu Initiative Phonology Template can help you investigate the segmental issues. It
will enable you to produce a phonology statement that covers segmental issues. This will
include identifying all the phonemes of the language.

Once you have described the segmental phonology, you should go on to describe the
suprasegmental features. Even during the early stages of a language development project, you
can begin collecting data on the suprasegmental issues. If you never investigate the
suprasegmental issues, you will not have an adequate basis for making orthography decisions
in those areas. However it will take some time to completely investigate and describe the
suprasegmentals. You should expect to do the bulk of this research during the trial
orthography phase.

3. **Produce a grammar statement covering morphology, morphophonemics, clitics and phrase structure.**

There are a number of grammatical features of Bantu languages that have an impact on the
orthography. One of these is the verb phrase. Many Bantu languages have verbal auxiliaries
that need to be written as separate words. Understanding the structure of the verb word and

---

\(^2\) You can obtain a copy by sending an email to ron_moe@sil.org.

\(^3\) By ‘segmental’ we mean issues related to consonants and vowels, as opposed to phonological features like
tone that operate on the syllable, word, or phrase level.
verb phrase is essential in identifying where to make these word breaks. Tone is sometimes affected by syntax. The same word will be pronounced with different tones depending on if it is said in isolation, in the middle of a phrase, or at the end of a phrase. So understanding the syntax of the language is essential. It is also necessary to understand the behavior of clitics. They will sometimes need to be written together with the word they attach to. Or they may need to be separated with a space, hyphen, or apostrophe.

**4. Review the principles of orthography development.**

Get a copy of *The Alphabet Makers* and read it. This book is an overview of the history of the alphabet and gives examples of alphabets from around the world. It is an excellent resource to use in an orthography workshop to give the participants a broader perspective. Also read the document orthography_principles.doc.

**5. Research the orthographies of the national language and neighboring languages.**

Identify all the languages that might influence orthography choices for your language. If possible find an orthography statement for each language. Make a list of all the phonemes of each language and what characters are used to represent them. Put all these in a single chart. Identify all rules used in writing, including word breaks, capitalization, punctuation, and type styles used for emphasis. It would be good if this could be done for each country and the results distributed. The book *Alphabets of Africa* is a good place to start.

**6. Collect and analyze any examples of literature and any previous literacy materials.**

Obtain copies of anything published in or about the language. Look especially for orthography statements and primers. If anything like this exists, analyze it for orthography choices and issues that need to be dealt with. Look for inconsistencies in writing that might reveal problems. For instance beginning writers often are not consistent in writing vowel length.

If you find that there are competing orthographies being used by various groups of people, you are very likely in for a hard time. You need to carefully research the history of each orthography, how extensively each is used, how much has been published in each, the personalities and status of the proponents, if there have been conflicts between the camps, and any other factors that might bear on the situation. You can easily get sucked into a raging battle before you are aware of it. The best thing to do in this case may be to get all the parties together and try to get them all to agree to participate in an orthography development program. Outline the steps involved and get them to commit to helping with the research, decision making, and testing.

**7. Talk to people about how they write.**

Find people who write in the language and talk to them about how they write. Ask them what problems they have. Ask if anything puzzles them. Sometimes they will be aware of something, but not understand why it seems strange. For instance a man was marking surface tone and was sometimes aware that there was something funny about a tone. He didn’t know what it was, because it didn’t fit what he had been taught. It turned out it was downstep, something he had heard of but never understood. Here are some sample questions to ask:
• What problems do you have writing your language?
• Are there any words that you don’t know how to spell?
• Are there any differences between your language and the national (or regional) language?
• Are you using one letter to write two different sounds? (Some languages use the letter b for both a plosive b and a fricative b.)
• Do you write one sound in two different ways? (This may reveal cases where two phonemes in the national language are allophones in the vernacular, e.g. l and r.)
• Do you ever have trouble reading and understanding what you have written? (Ask for examples.)
• Are there any pairs of words that sound different, but you write the same? (You are looking for minimal pairs.)
• Are there any pairs of words that sound the same, but you write differently?

8. Make a list of orthography issues in the language.

These issues and problems will be presented to the local language committee for discussion and decision. Collect examples of each problem. Determine what solutions might be used for each. Produce documents that illustrate each problem, explain the problem, and portray how each potential solution would look.

9. Form a language committee.

The decisions on how to write a language should be made by the speakers of the language. For this reason a committee should be formed of educated speakers who are opinion leaders, education officials, and government officials whose responsibility includes literacy and education. The committee should include those who have authority and influence over the teaching of literacy, those who have experience in education, linguists, writers, and any others who are involved in literature production. In many languages the church is the primary producer of literature.

10. Hold an orthography workshop with the local language committee.

The workshop should last two or three weeks depending on how much you need to cover, whether there is already an orthography, and how much exposure the committee members have had to linguistics and principles of orthography development. During this workshop you should cover the following topics:

• The history of the alphabet.
• Basic linguistic principles and insights.
• Principles of orthography development.
• The phonology of your language.
• Specific orthography problems.

Each problem should be discussed. First seek to raise the committee’s awareness of the phonological feature or contrast. You should present each potential solution without trying to persuade the committee. Present the pros and cons of each solution. When discussing the relative merits of the options, keep three factors in mind:

• Does this solution fit the linguistic facts, as they are understood?
• Does this solution make good practical sense? Is it easy to teach, easy to write (both printing and cursive), easy to read, easy to print (by typewriter if they are still being used in the area, and by computer if special characters are required)?

• Does this solution fit sociolinguistic and political factors? Do people like the look of the solution? Do people want the orthography to look like the national language or do they want it to look different? Does this solution agree with government policy?

The third factor often overrules the other two.

If you are not a native speaker of the language, you can help in the discussion, but let the people make the decisions. Only native speakers are qualified to vote or otherwise participate in the decision making procedure. Outsiders may be given a voice, but no vote. Since in many Bantu cultures it is difficult to oppose the most senior member, it might be good for decisions to be made by secret ballot. The participants in the Lugungu workshop proposed that decisions be made in this way and it worked well.

Note any decisions that seem to run counter to what you know of the phonology. Sometimes this is due to a mistake on the part of the committee, but sometimes it reveals errors in the phonological analysis.

The goal of the workshop should be to produce a draft of an orthography statement. This first workshop should at least cover the segmental issues.

11. Produce a draft of the orthography statement.

Use the document xx_orthography_statement.doc as a pattern. An orthography statement is an official record of decisions made concerning the orthography. It is written for linguists and education officials. It is not intended to teach people to read. Therefore it can and should use technical language and IPA symbols. It can be shared with other linguists developing orthographies for other languages. Therefore it should be written in a language of wider communication. It is the basis for primers and other literacy materials. It should be revised whenever a decision is changed or when new issues arise that need to be resolved and incorporated into the orthography. It may never be published, since its distribution is likely to be very limited.

You should not attempt to combine the orthography statement and transition primer into one book. The purposes of the two are too different and the needs of the target audiences are too different. You will end up with an overly simplistic orthography statement or an overly technical primer. The orthography statement is a place to record decisions and the reasons for those decisions. The transition primer is a pedagogical tool to teach people to read. The subject matter may be the same, but the presentation ought to be very different.

Stage 2. Trial orthography phase.\(^4\)

[from ECG Language Programme Planning Guidelines] Based on the preliminary phonological analysis and an initial orthography statement, a consultant may recommend to the AAD that "trial orthography status" be granted. Based on this recommendation, the AAD may grant this status to the project. With the trial orthography status, trial materials such as calendars, alphabet charts, stories, songs, trial literacy materials and trial translation materials may be developed. Trial literacy programs may begin. These trial materials will be used in gathering feedback for writing a full phonological analysis and for further refining the linguistic analysis to the point that all linguistic and sociolinguistic requirements are met for

\(^4\) This phase is sometimes called the “tentative orthography” phase.
approved orthography status. With a trial orthography status, materials must be checked by a linguistic consultant for correct application of the orthography in the materials as well as by a consultant appropriate for the materials produced. The Academic Affairs Director may choose to delegate one consultant to perform both the linguistic and the content check.

12. **Obtain ‘trial orthography’ status.**

*Trial orthography* status is a stage in orthography development established by SIL. It requires review and approval of the orthography statement by a linguistics consultant and a literacy consultant. The consultants will often recommend further research. Trial orthography status is required by SIL in order to begin publishing vernacular materials. Generally only a limited number of copies is printed during this stage. The reason is that the orthography will very likely undergo changes once people begin to use it. A lot of money can be wasted in printing materials that are soon obsolete. In fact obsolete materials should be destroyed. During the trial orthography phase, materials should be carefully checked to ensure that they conform to the orthography statement.

13. **Publicize the trial orthography.**

If the language is large and covers a large geographic area, it is important to publicize the trial orthography throughout the language area. This is to ensure agreement and support from all stakeholders. You may want to hold seminars with leaders in several areas to present the orthography and get input. This is especially important if the workshop that produced the trial orthography was small or was not fully representative of the whole speech community. The orthography statement should also be presented to the ministry of education, the national university, and any other official bodies.

14. **Conduct a dialect survey of the language.**

The survey should look for differences that would impact the orthography. These may include a different phoneme inventory and differences in the tone system. You will need to work with the committee to agree on a reference dialect. The reference dialect will be used as the standard for writing purposes. The orthography and literacy materials would be based on the reference dialect, but dialect differences would need to be discussed in the orthography statement and dealt with in literacy materials.

15. **Describe the morphophonemics.**

Identify all the morphophonemic rules of the language. You will need to do an analysis of morphophonemic processes, covering such things as labialization, palatalization, prenasalization, and other processes that occur at morpheme boundaries. If you are working in a structuralist theory, you will need to decide to what level the language will be written phonemically or morphophonemically. Identify the underlying forms of all morphemes that may cause problems in the orthography. If you are working within lexical phonology, work out what is the output of the lexical level. This involves determining which rules are lexical and which are post-lexical, so that the output of the lexical level can be determined for any given word. We have found that the output of the lexical level is a good starting point for writing a language.
16. **Describe word and phrase boundary phenomena.**

Begin by recording a text. Transcribe the text carefully in what you think would be a good orthography. Then listen carefully to the text and look for examples of vowel elision and other phenomena at word breaks. Mark all pauses. Try to mark and describe intonation contours.

Determine where the phonological word differs from the grammatical word. This would include identifying all clitics. Identify all word break and word juncture issues that will impact the orthography.

17. **Describe the tone system.**

Begin by marking tone in your dictionary. Collect examples of lexical minimal pairs. Identify how many tone classes you have for nouns and verbs. Put each tone class into frames to determine what kinds of tone changes are caused by syntactic boundaries and the tones of neighboring words. Fill out verb paradigms, paying special attention to any minimal pairs where tone signals a change in meaning. Make a note whenever tone causes a reader confusion in pronunciation or comprehension. Compile and organize these notes until you have a reasonably good idea where tone is going to be an issue.

18. **Produce a draft of the transition primer.**

Use the document xx_transition_primer.doc as a pattern. A transition primer is intended for educated speakers of the language who have learned to read in the national language. Therefore it should be written in the national language. This seems better than writing it in the vernacular. It should not contain technical language or IPA symbols, since these would not be understood. It should be completely understandable by an educated speaker. It should cover similarities and differences between the national language and the vernacular. It should contain copious examples and exercises. It is based on the orthography statement. Before publication, check the draft with a linguistics consultant, literacy consultant, and several speakers of the language.

19. **Publish and test the transition primer.**

The transition primer should be published as a stand-alone book. In other words a person who can read the national language should be able to read the book, do the exercises, and be able to write their language correctly by the end of the book. However if there are many differences between the two languages, it may be helpful or even necessary to use the primer in a classroom or one-on-one setting.

In any case you will want to teach large numbers of people to read using the primer. You may want to hold a number of one day courses throughout the language area. Observe how the primer works and listen to people’s responses. You will get feedback both on the primer and the orthography. A statement in the primer should invite the users to submit problems and suggestions to the language committee.

20. **Hold a writers workshop or story writing contest.**

The purpose of these exercises is to get people to begin writing and to collect stories that can be published. Encourage people to write stories, letters, and other materials in the language.
21. **Analyze the way people write.**

Look for systematic errors and inconsistencies that would indicate problems. Pay close attention to patterns that appear to contradict your phonemic analysis. Look for sociolinguistic reasons for these inconsistencies, but also reexamine your analysis of the phonology. Get people’s reactions to the look of the orthography and what problems they have in writing. For instance they may have trouble forming new letters while writing in cursive or they may forget to put in diacritics.

22. **Test people’s ability to read the orthography.**

Once you have collected several stories, choose several first person accounts that most people would not be familiar with. Type and edit these so that they conform to the proposed orthography. Then have someone read the stories aloud and record them on tape. Listen carefully to the recording and mark each time the reader pauses, hesitates, makes a mistake, gets the tone wrong, or makes a false start (begins a word, stops, then starts the word over again). Mark each kind of error differently. Repeat this exercise with several people and collate the mistakes. (Mark each person’s mistakes with a different colored pen on a single copy of the story.) If there are any problems with the orthography, the readers should all make mistakes in the same places. Test both new readers and experienced readers. New readers will make more mistakes, but they should also make different kinds of mistakes because they don’t read ahead.

If you do enough of these tests, patterns should emerge. For instance if the tone pattern of a word is affected by the following word, new readers will often get the tone wrong and have to repeat the word. If there are distinctions of verb tense and aspect that are signalled by tone, the reader may read the tone wrong, or start with the wrong tone pattern, recognize his error half way through the word, and start the word over. If tone is affected by syntactic junctures (and therefore signal those junctures), the reader will have trouble disambiguating the syntax and putting the correct tone on the word. If word junctures are not written correctly, the reader may not correctly apply vowel elision rules, or may repeat a phrase to correct an error. It is especially important to note when the reader makes a mistake and does not realize it, when he misunderstands what he is reading, or when he gets lost and can’t figure out what something means. This will be apparent when he pauses for a considerable time and his eyes move back and forth searching for a clue to disambiguate the meaning. These kinds of errors will reveal significant deficiencies in the orthography. Failure to mark tone in any way will most likely be a significant deficiency.

If you have two orthographic options, you can prepare two versions of a text, each using one of the options. Have people read each version and see which is read most smoothly.

23. **Produce a calendar and alphabet chart.**

A calendar can be printed with a Scripture quotation for each month. The Lunyole project produced a calendar with the alphabet chart as the decoration. Some projects print the alphabet chart on card stock. Other projects have printed it on cloth. The purpose of these publications is to get the language up on walls where people will see it.

24. **Publish easy reading material.**

A book of traditional stories is a big favorite and makes for easy reading if the stories are familiar to people.
25. **Hold a second orthography workshop.**

The purpose of this second workshop is to review and revise the orthography. Review the results of the tests. Evaluate all feedback. Make necessary corrections. Make decisions concerning morphophonemics, vowel elision, word breaks, clitics, tone, punctuation, dialect differences, loan words, and any other issue that has arisen. Revise the orthography statement accordingly. The revised orthography statement should again be presented to the ministry of education, the national university, and any other official bodies.

**Stage 3: Approved orthography phase.**

[from ECG Language Programme Planning Guidelines] In the approved orthography phase, a linguistic consultant check of the materials produced is no longer necessary though a check for content and mechanics of materials is still required. The AAD may grant "approved orthography" status to a project when the team has:

1. written up a detailed dialect survey of the language
2. satisfactorily answered all questions concerning the "Trial Orthography" previously raised by the consultant
3. has presented the suggested orthography to the church and language community leaders for their reaction, modification and approval
4. adequately tested the "Trial Orthography" in a trial literacy programme and in writers' workshops and has made necessary revisions
5. evaluated the reactions of the speakers of the language to the suggested orthography
6. written up a full phonological analysis which includes tone and morphophonology
7. identified and written-up grammatical subjects which will affect the orthography
8. written an orthography statement including elision and word breaks
9. there is little likelihood of serious demands for future change in the orthography

**Stage 4. Established orthography phase.**

[from ECG Language Programme Planning Guidelines] After further use of the orthography and before the consultant checks more than 70% of the NT, any revisions to the orthography will be made. Upon recommendation of the Academic Affairs Committee, the AAD may grant "Established Orthography" status to the project.

26. **Obtain ‘established orthography’ status.**

Approved orthography status requires testing and feedback on the orthography from speakers of the language. An SIL linguistics consultant will have expected you to have done the following:

- Write up a detailed dialect survey of the language. This is to ensure that there are no serious dialect differences that will cause trouble.
- Satisfactorily answer all questions concerning the trial orthography previously raised by the consultant
- Present the orthography to church, education, and community leaders for their reaction, modification, and approval.

---

5 This phase is sometimes called the “working orthography” phase.
6 This phase is sometimes called the “approved orthography” phase.
Adequately test the trial orthography in a trial literacy program and in writers’ workshops and make necessary revisions.
Evaluate the reactions of the speakers of the language to the trial orthography.
Write up a full phonological analysis, which includes tone and morphophonology.
Identify and write-up grammatical subjects which will affect the orthography.
Include sections on elision and word breaks in the orthography statement.
Determine that there is little likelihood of serious demands for future change in the orthography.

Established orthography status is granted after a period of time during which the orthography is used without any serious issues being raised or changes being made. In an SIL project this status is required by the time 70% of the New Testament is translated. Once established orthography status has been achieved, you should continue monitoring reading and writing practices for indications of problems. One thing that sometimes happens is that mature writers begin taking shortcuts. For instance, they may begin leaving off diacritics. This may indicate that mature readers don’t need them or that writers simply find them a nuisance.

Another thing to watch for is the fluency achieved by mature readers. If people want to practice before reading in public, this may indicate that mature readers still make mistakes due to deficiencies in the orthography. A mature reader using an adequate orthography should be able to read new material fluently and with few mistakes. Those mistakes should not be caused by the orthography, but by things like inherent ambiguity or unfamiliar words.7

Failure to mark tone appears to result in a deficient orthography. Even mature readers in a language like Luganda cannot read fluently, and this in spite of the fact that Luganda has been written for a hundred years, has had a Bible for a hundred years, and has daily newspapers and other literature. In cases like this you have two choices: either devise a way to mark tone, or accept that even mature readers will never be able to read with full fluency.

References

---

7 It must be admitted that no orthography is perfect. The voice is more versatile and can carry nuances through variations in pitch, amplitude, pacing, and pauses. To this we could add gestures and facial expressions, which cannot be captured in an orthography.