

9: Participatory Research in Linguistics for language and orthography development: A practical guide

Constance Kutsch Lojenga

Abstract: Over the years, I have had many opportunities to run one-language and multi-language orthography development workshops in quite a few countries in Africa. This experience has led me to develop an alternative to the traditional way of analysing the phonology of a language, which is faster and often more precise than the traditional approach. In this methodology, native speakers participate in the research process in close collaboration with an outside linguist, with the result that they are the first beneficiaries of the outcome of the research in terms of a resultant orthography. When working by means of this approach, people are collaborating, sharing the work, using their experience in their own language against the background of the sound system of another language, in which they first learned to read.

The methodology turns various aspects of the traditional approach to discovering the sound system of a language ‘upside down’. Instead of using narrow IPA transcriptions, I advocate broad transcriptions, based on people’s awareness of the sound system of the language of education in their area. This ‘upside-down’ approach aims to find the sound contrasts first, through a Same-Different test which also raises the speakers’ awareness of the sound contrasts in their own language, and only after that doing further detailed phonetic research where necessary. Thirdly, I ‘reverse’ the burden of proof: contrast is not primarily established by minimal pairs, but rather by the exclusion of allophony, resulting from a thorough phonotactic study of sounds and combinations of sounds in the same structural position – root and syllable – when comparing sets of words.

Keywords: Africa, Bantu, native-speaker intuition, collaboration, tone, vowel contrasts, workshops, same-different tests, awareness-raising process, orthographic transcription, exclusion of allophony

1 Participatory research and language development¹

Language, the mother tongue, forms a very important and, in fact, an inalienable part of every human being. It is one of the fundamental rights of people of all ethnic backgrounds to have a well-developed system for writing their language, in order to strengthen their identity, including a basic description and documentation for the preservation of their language. Language development² is concerned with creating a writing system for a language, or, as it used to be called: ‘reducing a language to writing’ (Pike 1947). In our modern day and age, we would like to see speakers of the language involved by actively participating in the linguistic

¹ The author has used the method discussed in this paper with many different African languages, among others Central-Sudanic, Nilotic and other Nilo-Saharan languages, Bantu, Kwa, Gur, Ubangian, as well as Austronesian languages from Madagascar. Others have also used it in the Chadic subgroup of Afro-Asiatic and in some languages spoken in Asia. The methodology is flexible enough to handle languages of any type: languages with different root and syllable structures, isolating as well as morphologically complex languages.

² Orthography development is one aspect of the broader concept of language development. It consists of developing an alphabet, the study of morphophonological processes, as well as establishing word boundaries in order to make a comprehensive writing system to ensure that people can learn to write their language consistently. In this chapter, I am using both terms interchangeably.

research necessary to create a suitable orthography which represents all the relevant sound contrasts found in their language. In this chapter, I present a global overview of a methodology for researching the phonology, and following that, the morphology and basic syntax of a language in order to create a complete writing system with which the speakers can develop written materials in their language, whether it concerns their cultural literature, folktales, their history, the Bible, writing letters or WhatsApp messages to communicate with their family members who live elsewhere. The researcher can also use the language data to pursue further linguistic goals.³ This chapter contains a description of and instructions for the methodology I developed over the years in one-language and multi-language workshops in a number of African countries.

Language development requires linguistic research. Traditionally, this was mostly done by a non-native-speaker linguist together with a ‘language helper’, more recently also called a ‘language consultant.’ In many cases, the linguist would do the analysis and propose an alphabet to the native speakers, whereby his/her analysis would guide him/her in further issues like morphophonology and word boundaries for the writing system. Often the outside linguist would then explain and teach the writing system to the speakers of the language. There are many instances of languages having been analysed following this traditional approach which are unfortunately ‘lacking’ some vowels in the analysis and in the orthography (Casali 2008: 509); in others, some of the choices of graphemes are not very ‘intuitive’ in that they are based on frequency or linguistic markedness rather than on speaker awareness, raised through collaborative research and informed by how the sounds of their language are mapped onto the language used in education.

I use the term ‘participatory’ in the context of research for language and orthography development for a linguist working in close collaboration with mother-tongue speakers, researching their language together. This close collaboration between linguist and native speaker is necessary given the highly scientific character of the discipline. Many elements of language and orthography development are outside the experience and scope of native speakers, so they need a linguist to guide them through the process. Similarly, the linguist needs to bring the speaker awareness of the sound system of a language to the surface in order to get the best results in their collaboration, namely an alphabet and writing system that fits the language. This model is adaptable in scope whether one is starting analysis from scratch or focusing on one or more elements, such as vowel harmony, or tone if vowels and consonants have already been discovered. In the methodology as presented in this chapter, I aim to see native speakers participate in the whole research process. Unlike often was the case in the past, there should be not just one, but several speakers of the language participating in the process. Their interaction enhances the quality and outcome of the research.

My personal experience is in African languages of all different phyla, as well as Austronesian languages of Madagascar. In this chapter, I describe the main aspects of the methodology I developed for such participatory language-development workshops. I am exemplifying the various steps of the discovery procedure from the perspective of morphologically rich Bantu languages with predominantly disyllabic root structures, open syllables only, a vowel inventory of five or more vowels, a noun-class system, as well as verbal derivational and inflectional affixes. Other languages may have more complex structures, inventories, or morphophonological processes, such as other root structures and/or closed syllables. The researcher has to take additional and different factors into consideration, though

³ I have earlier documented my experience of participatory research in a complex Central-Sudanic language (Kutsch Lojenga 1996), and after finishing developing the writing system with a group of speakers of the language, used the data as the source of my dissertation (Kutsch Lojenga 1994).

in principle, one follows the same procedure. The reader will have to adapt and modify the basic principles of the methodology for use in such cases.

In §2, I talk about background research for the linguist/facilitator as researcher.⁴ Following that, in §3, I treat the preliminaries of language-development workshops: the choice of participants and other pre-workshop preparations. §4 is entirely devoted to the contents of the workshop itself. In §5, I present some pedagogical considerations related to participatory group research. Finally, in the Conclusion, I focus on some benefits of doing the analysis for language development by using this participatory approach.

2 Typological background research

When getting involved in language development, it is important to explore – to the extent possible – anything written on the phonology or the grammar of the language in question and/or related languages. In this way, one develops a typological background knowledge, which will extend more and more when one gets in touch with other languages. Such a background study will be most helpful in making hypotheses for what may or may not be found in the language, and it may give ideas as to what domains to research with the speakers in particular detail.

It is equally important to familiarise oneself with other orthographic systems which are in use in the area or among typologically similar languages by consulting reading primers or other publications written in the language itself to compare orthographies or other writing conventions used in the past. This will help in ascertaining if there are cases of underrepresentation, overrepresentation or inconsistencies in the orthography as used in the past or present.

For the sake of beginning language development, this concerns particularly:

- Vowel systems in the target language or in related languages, including different possible types of vowel harmony, or features of length and nasalisation with their orthographic symbols, if they exist.
- Consonant systems, specifically unusual consonant sounds or contrasts, and how they are symbolised in the orthography of related or surrounding languages.
- Basic tone systems and particular tone rules, such as downstep, high spreading and high shifting, Meeussen's Rule, etc. operating in languages of the area.
- Nominal and verbal morphology of related languages (noun-class systems and verbal derivational categories and inflectional systems in Bantu languages).
- Word boundaries and (in)consistency of representing morphophonological processes in the orthography.

3 Preliminaries

Participatory-research workshops can be held for one language only or for a number of languages together.

⁴ In this chapter, I use the terms 'linguist' and 'researcher' interchangeably. In the context of language-development workshops for orthography development, I also use 'workshop leader' for someone heading up a one-language or multi-language workshop, and 'facilitator' or 'linguist' for someone who facilitates the work in one language during a multi-language workshop. In these contexts, all four terms are used for someone who is linguistically trained and who works with a group of speakers of a particular language in order to analyse various aspects of the language for the purpose of creating a writing system or for a descriptive-linguistic publication on that language.

A one-language workshop can be held in a central place in the area where the language is spoken. This has the advantage that word will get out among the language community that their language is going to develop an ‘alphabet,’ i.e., a written form. Working with more than one or two people may be a challenge and a learning process at first for some linguist-facilitators, but in the context of participatory research, it is best to work with a (small) group of people for the sake of the interaction, so people may come up with additional ideas based on what their fellow participants are contributing. The size of the group depends very much on the capacity and capability of the linguist-facilitator. When working with one language group only, the linguist can ask the language community to choose a smaller or larger number of people to participate according to what s/he considers realistic in his/her situation. I consider it not impossible to work with 10 participants, or even more, if the linguist-facilitator has enough experience.

In a multi-language workshop, speakers of several language communities come to a central town to take part in the workshop together, with around four native-speaker participants per language. A linguist-facilitator needs to be available for each language to assist and guide the small group of participants on a full-time basis, under the overall guidance and leadership of the workshop leader.

The workshop leader should have his/her hands free to organise and lead the workshop, give direction and introduce the topics to be studied and researched each day. S/he should also have time to join each of the language groups regularly and verify the results of their work, preferably on a daily basis. This leadership should be based on extensive typological background knowledge and a capacity to direct a large group of people from a variety of language backgrounds.

Ideally, one would like to see a group of related languages in a multi-language workshop, rather than languages belonging to different language families.⁵ The workshop leader may want to go through the various steps of the discovery procedure for all languages at roughly the same pace. This will give him/her the possibility of having sessions in which s/he can introduce topics for the whole group together.

The question of how many languages should participate in a multi-language workshop mainly depends on the experience and capability of the workshop leader. Under ideal circumstances, one would wish to have no more than 3 – 5 languages represented, each with about four participants, and a linguist/facilitator. Multi-language workshops with more languages are inevitably very intense for the workshop leader. It is more important to get the basics for the alphabets correct in the individual languages than to have a larger number of languages represented in these workshops.

The next question is what kind of participants one would like to see involved in the development of an orthography for their language. I do not normally choose the participants myself. The choice is given to the local stakeholders and language-community members. However, I generally give some basic principles or criteria to the stakeholders for their choice of participants:

- Proficient first-language speakers from different age groups, who speak their mother tongue well, and who love speaking and learning how to write their language. This is mostly the case when people have grown up in their area and where both parents speak the same language so that their language was spoken in their parental home. People from linguistically mixed marriages or those who have lived in different environments during

⁵ The reason for this is that it is harder to work with a group of languages which are structurally very different: what one language expresses through morphological means, another may express syntactically – even through different word-order configurations - or languages with different vowel systems (in as far as one can make preliminary hypotheses for the vowel systems).

their childhood may appear to have some gaps in their mastery of the language or simply not speak one variety completely fluently.

- A good level of education, preferably at least a high-school diploma, or even beyond that, a university education, though not necessarily in linguistics. In the majority of cases, this is a reasonable prerequisite. However, this is not always possible in places where the homeland of certain language communities is remote and has difficult access. In such cases, one works with the best option in the situation, knowing that the research will likely take more time.

4 The actual workshop

This section begins with the introduction on the first day of the workshop (§4.1) followed by collecting vocabulary as a basis for the phonology research (§4.2). Then, basic linguistic concepts are presented and introduced to the participants (§4.3). Following that, I treat each topic of the research, first the most crucial one, the analysis of the vowel system (§4.4); then the consonant system (§4.5), and the discovery of the basics of the tone system (§4.6). In the last subsection, (§4.7), I present a brief overview of the sequence and contents of further research.

4.1 Initial questions

After a short opening ceremony, I address the entire group. My introductory talk includes the questions: ‘In which language did you learn to read in school?’, followed by: ‘And how did they teach you to read the vowels?’ This gives me as workshop leader insight into the background of the participants with respect to their ‘*surface intuition*’ of the number and parameters of the contrastive vowels they are aware of at this point. This will then serve as a starting point for the discovery of the complete vowel system of their mother tongue, trying to bring their ‘*underlying intuition*’ or ‘*deeper*’ or ‘*subconscious awareness*’ to the surface.

Another couple of questions I frequently ask the first day are: ‘Does your language have a grammar?’ and ‘Does your language have a dictionary?’ The participants are often taken aback by that question. Some will say ‘yes’, others will say ‘no’. I often get the answer ‘no’, because in most languages in which we start orthography development, these do not exist.

In addition, people often think that they speak their language all ‘mixed up’, i.e., without any regularity or system. They think – pejoratively – that their language is ‘just a dialect,’ compared to others which have a grammar book and a printed dictionary, which are ‘real’ languages, like the international languages, or African languages of wider communication. I try to steer them away from that thought by talking about the fact that their language has indeed a grammar and a dictionary, but it is all still hidden in their heads. We need to get it out of their heads and on paper. After a while, when they see the patterns in the structure of their language, they become aware that their ‘dialect’ is more than just irregular speech, and even that their language is ‘rich’ in that it may have elements which are not found in the LWC, topics which they will discover as the work continues.

4.2 Collecting vocabulary

The first activity for the participants of the workshop is to collect vocabulary, writing the words according to the writing system they were taught in school. This can be done by using a wordlist or a semantically-based list of topics or a mixture of both. Picture alphabets from

other languages are also a good source, because those may produce words that are usable for a picture alphabet poster in their language, once the system has been properly analysed.

I will tell the participants at the beginning of their language-development project: 'This is the beginning of your dictionary!' They get excited with that prospect, because they realize that their 'dialect' is on its way to become a 'real language.' In addition, they will write all items themselves on small cards (I generally use A4-size paper cut into 8 equal pieces). They will work together as a group, with one person as their 'scribe'.

This becomes a social activity in which the participants interact, creating a bond among themselves. The facilitator is there as an encourager, sometimes to let them know how to handle certain issues. For example, if they don't know the French or English equivalent, they should write 'tree, sp.' (species) instead of wasting a lot of time searching for the exact term. At other times, s/he actively participates in the assignment of the day. The focus here is on writing and not on having endless discussions on the precise meanings of the words. Sometimes this aspect is not easy because speakers of a language are naturally very much oriented towards meaning and like discussing rather than writing.

Concerning nouns, one should ask the participants to write the singular form with the plural beneath it, which will help to identify the noun roots. When a language has no separate forms for singular and plural, one has no choice but to write only one form.

Similarly for verbs, one should ideally also have two forms on each card, e.g., an infinitive and an imperative if these forms exist in the language. Even if these forms may be segmentally the same, one should still write them, since one may later find that they are tonally different. If, however, there is no separate infinitive form, one may try to find another form to accompany an imperative.

When collecting vocabulary this way, no attempt should be made to write tone on the words, because they cannot link it to something they already know. Many people do not realise that they speak a tonal language or that certain pitches they utter are 'higher' or 'lower' than others. Basic tone analysis will be done after their vowel and consonant systems have been established.

I also advise participants to write in 'print,' instead of cursive, as they may have learnt in school. They should not use capitals, as capitals are never used to list words in dictionaries.

One should also discuss with the participants in what languages the words should be 'translated' ultimately, for their bi-, tri-, or quadrilingual dictionary. I ask them to write these 'translations' in the form of glosses on the cards.

They should also write 'N' or 'V' in the top right-hand corner of each card for 'noun' or 'verb', and keep these two categories separate. In addition, everything which is not a noun or a verb should be kept separate. Nouns in the target language should be glossed with a noun and verbs with verbs. This is not always easy. Sometimes the 'translation' of a verb in English or French is not a bare verb in the target language, but a combination of a verb and a complement. Participants should be made aware of these things and learn to separate them. For example, 'to fish' would be translated 'to kill fish' in the Central-Sudanic Ngiti language from Eastern D.R. Congo. It is not necessary to have the Ngiti combination of verb-plus-complement written on one card at this point in the research, since they are just preparing for the analysis of the sounds.

I often use resealable plastic bags: one for nouns, one for verbs. The bag of nouns should preferably not contain compound nouns, loanwords, or derived nouns. All of these can be treated at a later stage. Even if they were not asked to write anything else but nouns and verbs, there will inevitably be a number of cards with words from other grammatical categories: numerals, adverbs, prepositions etc., some of which can only be labelled with a grammatical category when studied in context later. These can be kept in a separate bag, labelled 'other'.

This work of writing down their vocabulary with a view towards constructing a dictionary in their own language fascinates and motivates the speakers in such a way that sometimes, they will work enthusiastically for several days. This is just as well, because one wants to aim at a good number of lexical items, if at all possible, something in the range of 1000 words initially.

All of this has a purpose in our approach. Our analysis will focus first on the lexical roots of nouns and verbs, because it is expected that nearly all phonological features will be found in the analysis established with noun and verb roots.⁶

4.3 Syllable, root, and word structure

Since collecting around 1000 words may take a couple of days, the workshop leader and the facilitators could interrupt the work from time to time to get people used to hearing some ‘linguistic’ terms explained as simply as possible.

First, one can practise with the basic structural concepts *vowels* and *consonants*, until, hopefully, all understand what these two categories of sounds are and can give examples from their language.

Next, we discuss the concept of ‘*syllable*.’ One can ask participants to count syllables, take a word and break the flow of speech after every syllable, or clap them, so as to determine how many syllables there are in a word, first in the LWC and then in the target language. One can also ask the participants to produce words in their language with two syllables, three syllables, four syllables etc. All this should be presented as a kind of ‘game.’ People learn to isolate, recognise, and count syllables in their language in this way. If a language has both *open* and *closed* syllables, the participants should learn to recognise these and find examples from the words they are collecting.⁷

A slightly more difficult (but vital) concept is that of the *root*.⁸ Since the speakers of the language were asked to write singular and plural forms of nouns, one can ask them which part is different between singular and plural, and what is the same. They can learn that the part that is the same is the root. The same type of exercise can be done while comparing the verbs written on the cards: e.g., comparing infinitives with imperatives. It is not always possible to identify the root precisely, especially when the root begins with a vowel, and morphophonological processes obscure the precise boundary between prefix and root. But it is important to fix the general concept in the participants’ minds so they will grow in their understanding. As will be discussed in the following sections, when analysing the vowels, consonants, and tones, they will need to be able to sort nouns and verbs according to their root vowels or according to their root-initial consonants.

All of these exercises may be fun to do and will raise people’s awareness of basic concepts of the structure of their language!

⁶At most, one may find some different and often unusual sounds in ideophones which haven’t surfaced in the analysis of the nouns and verbs, for example labial flaps, which in some languages predominantly occur in adverbs or ideophones.

⁷ The facilitator should know this from his/her background research or notice it while the participants are collecting vocabulary.

⁸ This does not mean that we ask the participants to pronounce the root without its obligatory affixes. They should be able to identify the root from within the word.

4.4 Vowel research

4.4.1 Introduction: In many African countries, children learn to read and write in a language other than the mother tongue: an international language like French, English or Portuguese,⁹ or an African language of wider communication (LWC). Their own mother tongue often has a different sound system from the language in which they learnt to read, which is especially the case for the vowels. Native speakers are generally not consciously aware of all the contrastive vowels of their own language. However, if they have learnt to read and write in the LWC of their country or area, they do know the representation of the basic vowels. They know how to pronounce them based on what they were taught in school. In many African countries, these are the five basic vowels < a – e – i – o – u >.¹⁰ This is the outer limit of their experience. If their mother tongue has more vowels, they will map these in a systematic way onto the five (or seven) basic vowels they are aware of, probably thinking that these are the only ones universally found in all languages. This is even more so the case because people are generally very focused on the written language. People need to be told that the oral form, i.e., the pronunciation of the vowel sounds in their language is more basic than the written form. I call this knowledge of the LWC vowels their '*surface psycho-linguistic awareness*'. The system of contrastive vowels in their mother tongue is hidden at a much *deeper, subconscious* level. They are not (yet) aware of these 'extra' vowels in their own language. The linguist should take this *surface awareness* as the point of departure in order to dig deeper to get to the *underlying, subconscious* level, and bring that to the surface.

After analysis, the mother tongue may be found to have more distinctions in the vowel system than the writing system of the language of education. If an orthography for the language under study were to employ only the orthographic vowel symbols of the LWC, certain contrastive vowels will be underdifferentiated. Readers of underrepresented orthographies have to guess regularly, which leads to interruptions, rereading of words or whole sentences, whereby part of the meaning may be lost. Subsequently, people may get discouraged because they are not capable of reading their own language as easily and fluently as they expected. By way of example, the Ngitu people mentioned above had tried to write their language using the five vowels of Swahili, the LWC of the area, but they were not able to correctly read or write their own language with that system. After they became aware of this, they looked actively for outside help. After researching their language with them, there appeared to be nine contrastive vowels, which they have since represented in their alphabet.

Establishing the number of contrastive vowel sounds needing to be symbolised in the alphabet is the most important part of the initial research for orthography development. Speakers should develop an awareness of all contrastive vowel sounds of their language, especially if it is expected that their language has a different vowel inventory from the local LWC. Until they have become consciously aware of all the vowel contrasts, they will ignore extra vowels in reading and keep guessing, based on the vowel system of the LWC. Learning to read and write using more vowels takes time, and for some, more than for others. It is an awareness-raising process which takes practice pronouncing and reading the extra vowels correctly.

⁹ International languages like French, English, and many others are written with only five vowel symbols, even though they often have more than five contrastive vowels.

¹⁰ In other African countries, the LWC in which children learn to read in school may contain seven vowels. The linguist should try to find out how native speakers map the vowels of their own language on to the seven vowel symbols of the LWC in which they have learned to read.

Our pedagogical methodology for discovering the sound system of a language using a participatory approach is also appropriate and effective as a basis for further research aiming at descriptive as well as theoretical concerns.

Traditionally, a linguist would study articulatory phonetics, and on that basis s/he would do a narrow phonetic transcription of the (words of the) language. However, it is good to be aware of the fact that articulatory phonetics is subjective (see also Cucchiarini 1993, Chapter 1), and that the traditional approach using narrow IPA transcriptions as perceived by an (outside) linguist tends to lead to random errors. By letting mother-tongue speakers do the initial writing (on the basis of their '*surface awareness*'), the research is founded on their experience of their own language, and subsequently, any errors will be systematic, because mother-tongue speakers map the sounds of their own language in a systematic way on the alphabet letters they were taught in school. Vowel research should therefore not be done by listening to each vowel separately and trying to assign an appropriate phonetic symbol based on what the linguist has learnt in articulatory phonetics.

The first step in vowel research should take into consideration the root and syllable structures of the language. Vowels should be analysed within identical structures. In a language with predominantly disyllabic roots, investigating the combinations of vowels is also very important. Attaching IPA symbols to the contrastive sounds can be done afterwards for descriptive-linguistic purposes.

As stated above, the point of departure for vowel research is the vowel system in which they have learnt to read in their basic school education. This typically means a broad 'phonetic' transcription with underrepresentation.¹¹

A correct vowel analysis will greatly benefit the discovery of vowel-harmony patterns within the root or in the entire word, including its affixes. In addition, it is also important to discover whether the language has a systematic length contrast in the vowel system, which is especially important for tone research later.

Secondly, the vowel research in the group should *not* primarily focus on proving contrast by trying to find *minimal pairs*. Undoubtedly several minimal pairs will come to the surface spontaneously, as speakers think of words which form a minimal pair with another word they have just read on a card. However, in most languages, it is very difficult to find minimal pairs for all vocalic contrasts. Rather, the research should focus on trying to *exclude allophony*, which would then immediately *confirm the opposite, namely that sounds are in contrast* (see also §4.4.3).

Thirdly, identifying the contrastive vowel sounds should not be based on acoustic measurements either. Vowels which are auditorily close may also have close F1 and F2 formant structures, and it is hard to determine *contrastivity* of vowels in that manner. Similarly, a set of plottings of one vowel phoneme often partially overlaps with the set of plottings of another vowel phoneme (see e.g., Casali 2008). These plottings, therefore, cannot clearly and unambiguously be separated into phonemes in any language before the set of contrastive vowels has been established. Acoustic measurements can be done afterwards to confirm contrastivity and to study further the amount of acoustic overlap of certain vowels through plotting the measurements in the vowel space.

Rather, establishing vowel contrasts should be based on a SAME-DIFFERENT listening test of identically-written (most likely underdifferentiated) vowels in identical root structures, such as on cards read by speakers together with their facilitator, as described below.

¹¹ This approach is the opposite of what is generally being taught about how to discover the vowel phonemes of a language, starting with a narrow phonetic transcription, with all the fine details that the researcher can detect with his/her ears.

4.4.2 Card sorting: Participatory research of the vowel system revolves around discovering the number of contrastive vowel qualities, contrastive vowel length, nasalisation, or any other vowel feature that may be contrastive in the target language. It is of utmost importance that these contrasts be determined from the roots of nouns and verbs.

First, the participants sort the *nouns* on root structures, with the *assistance* of their facilitator. Disyllabic roots and monosyllabic roots are kept separate. Open and closed syllables should also be kept separate. Later on, the *verbs* also need to be sorted on their root structures, beginning with the canonical root structure.¹²

Within this primary sorting process, the participants and their facilitator also need to sort the nouns with disyllabic roots into piles with two identical vowels in the roots, so that there is a separate pile for each of the vowels so far written. Similarly, they need to sort the nouns with monosyllabic roots, and, again the verbs of the most common root structure. In all these sorted piles, the aim is to find the number of contrastive vowels within that pile.

4.4.3 Listening and establishing the phonemes: After the cards are sorted, it is time to listen to the words using the SAME-DIFFERENT test. One speaker of the language reads the words on the cards of one pile one by one, and everyone else carefully listens to the $V_1 = V_2$ combination of identical vowels to hear if the pairs of vowels sound the same for every word. I ask the participants not to read the meaning, nor to give explanations, but to exclusively read and listen to the sounds, without any interruption.¹³

If, while reading, it becomes immediately clear that there are two distinct vowels in the <i - i> pile and the facilitator knows from his/her background knowledge that this may be a case of 'extra' vowels (as e.g., in a vowel-harmony system), s/he should help the team to make two groups of cards according to how they sound. The listening process needs to be repeated and repeated, until each pile of cards is consistent in the set of identical vowels it represents. The two piles of words the speakers have both written <i - i> as root vowels may well split out into [i - i] and [ɪ - ɪ], resulting in two different phonemes: /i/ and /ɪ/.¹⁴ Some speakers will quickly 'feel' that these two sounds are different, even though they have been writing them the same so far. Ideally, both the facilitator and the speakers should agree – not by using one's ear for determining which vowel it is, but rather to ascertain whether the vowels sound the same on every card in the pile. The same process is used to distinguish any vowel differences which may or may not occur in what the participants have so far written as the other known vowels, <u>, <e>, <o> or <a>. For example, the original pile of words written with <u - u> may contain /u - u/ and /ʊ - ʊ/, which the speakers had interpreted as the vowel <u> from what they learnt in school. Similarly, the vowels they wrote as <e> and <o> may or may not split up into two vowel phonemes each: /e/ and /ɛ/, and /o/ and /ɔ/. So far, the speakers may have been unaware of a phonemic contrast between them. This listening test for SAME – DIFFERENT is the beginning of a process of awareness raising of these contrasts in their mother tongue.

In the listening test, both the speakers and the facilitator should give their opinion on whether the vowels on one card are the SAME or DIFFERENT from those on the other cards. This should divide one original pile of cards with words written with two identical vowels into

¹² In Bantu, this will often be px-CVC-a, though the final vowel (FV) may undergo changes based on the root vowel.

¹³ For languages which have a singular/plural distinction in the nominal system (e.g., most languages in the Bantu subgroup), it would be good to read both forms. Other languages, like Central Sudanic, do not have a singular/plural distinction in [-human] nouns and so there will only be one form to read.

¹⁴ Many Niger-Congo and most Nilo-Saharan languages have an ATR vowel-harmony system with 7, 9, or 10 contrastive vowels, divided in two sets, [-ATR] and [+ATR] which cannot cooccur within roots; affixes normally harmonise morphophonologically to the root so that the whole word contains either [-ATR] or [+ATR] vowels. Therefore, vowel combinations like */i - ɪ/, or */u - ʊ/ within roots are impossible.

two (or at most three) different piles. They should repeat checking these piles until everybody agrees and each new pile only contains words with audibly the same vowel. Once the speakers (and hopefully the facilitator/linguist) have become aware of these contrasts, they can start perfecting their ears to hear and determine the ‘extra’ vowel qualities on isolated words by ear.

The process of listening to $V_1 = V_2$ combinations will ultimately reveal how many contrastive vowel qualities the language contains, and thus how many more vowels there are than those which were used for the initial transcription by the native speakers.

Because one is dealing with identical vowels in the root, it is impossible that there could be allophonic variation caused by a different vowel. By excluding this possibility, the vowel of each pile is *automatically in contrast* with the vowel of every other pile, i.e., the conclusion will be that they are all phonemes, each in their own right.

When the SAME-DIFFERENT exercise for the vowels has been done with each pile of cards, the final result is the number of vowel phonemes, proven by the *exclusion of complementary distribution*. As mentioned, the SAME-DIFFERENT exercise should be done separately for roots of different syllable structures, and separately for noun and verb roots.¹⁵

If there are more than the five basic vowels, the facilitator needs to present a way of writing the extra vowels. The facilitator does not need to write in IPA symbols for the speakers, unless the symbols are likely to be used in the orthography.¹⁶ Rather, it would be ideal for the symbols to conform to the orthographic usage in languages with a similar vowel system, something which has proven to function well for learning to read and to write in other languages, something that is closest to the speakers’ intuition and does not contradict with the system they were taught in school. The team can then work on ‘correcting’ the cards just listened to for the ‘new’ vowels.

Once the participants have revised the piles of cards as needed with the new vowel symbols, the facilitator can note down several words with each vowel as the only vowel in the root (and if possible, affixes), to be used for an alphabet poster and a vowel-teaching booklet.

The facilitators can do small reading and writing exercises from time to time as they work together with the group, in order to gradually raise awareness, and assist speakers in correctly writing the ‘new vowels’ orthographically.

While using this method to determine ‘extra’ vowel phonemes in one of the languages we worked with, we came across a word written <mili>. As soon as one participant pronounced the word /mili/ ‘root(s)’, another participant came up spontaneously with the word /mɪli/, which we wrote immediately in the new orthography as <mɪli> ‘tree(s)’. The two words containing this minimal contrast are used in their vowel-teaching booklet exemplified in Figure 5 in §4.7 below. Even though the tones are different, it was very clear to the participants that these words had different vowels. Such spontaneous minimal pairs are the best examples for awareness raising, and also for teaching the contrast between the two sounds in reading and writing. In fact, I would advise the facilitator and the participants to keep a record of any minimal pairs they come across.

In conclusion, the idea is not to teach the participants articulatory phonetics,¹⁷ including the official IPA symbols, nor analytical procedures for phonology either. The linguist-facilitator can later ‘transform’ the orthographic symbols into the official IPA symbols if s/he wants to use the material for a linguistic publication.

¹⁵ Occasionally one comes across a language in which the nominal system has a different number of contrastive vowels than the verb system.

¹⁶ Of course, the facilitator can write the words phonemically using the IPA symbols for his/her own linguistic purposes – possibly even later, but the native speakers do not have to go through that exercise: they should immediately write in what will most likely be their orthography.

¹⁷ I have found that people do not normally respond to sounds that are different from their own language.

4.4.4 Step two in vowel research: In languages with canonical disyllabic roots, a second step in vowel research is the study of the combinations of vowels in roots. Once again, it is not advisable to just listen to each vowel sound separately in order to decide what these vowels are to the ear of the native speaker or the facilitator. Rather, each vowel combination should be considered as a unit in order to establish them as SAME or DIFFERENT from other vowel combinations.

As was the case with the $V_1 = V_2$ combinations, so also these $V_1 \neq V_2$ combinations are typically written with five vowels on their cards. The result of the first round of checking will have revealed the real number of contrastive vowels in the language, often between five and nine, or sometimes ten, but rarely more or fewer.

For this second step of vowel research, participants and facilitator now need to sort the remaining cards of the same root structure according to their combinations. The cards containing words with two identical vowels have already been dealt with, so in the original five-vowel writing of the participants, this could yield up to 20 combinations in which V_1 and V_2 are different.

Most African languages with nine or ten vowels, as well as many languages with seven vowels, have some kind of vowel harmony, in which certain vowels may cooccur and others not. The facilitator should investigate what the possible combinations in such a language may be, and work on that hypothetical system with the speakers of the language. Within ATR-harmony systems as found in Africa, the two vowels *within a root* must ‘belong’ together in one harmony set, [+ATR] or [-ATR]. They form a ‘harmonic pair’, i.e., the whole root is [-ATR] or [+ATR].¹⁸

If the language shows a phonemic contrast between seven vowels in the $V_1 = V_2$ SAME-DIFFERENT test, each $V_1 \neq V_2$ combination also needs to be tested through a SAME-DIFFERENT test: what was originally written as px-CoCu may result in two different possibilities, e.g., px-CoCu [+ATR] and px-CɔCu [-ATR], and what was at first written as px-CuCe may turn out to yield two possibilities, namely px-CuCe [+ATR] and px-CuCɛ.

Following that, the team – facilitator and native speakers – can now fill in two cooccurrence charts, one for nouns with [+ATR] combinations in the root, and one for nouns with [-ATR] combinations in the root, accompanied by examples for each existing combination. This will show the reality of a vowel-harmony system. Any ‘holes’ in the system will show where the systematic and accidental gaps are. If there are unexpected combinations, especially between what seems to be [-ATR] and [+ATR] vowels in one word, one needs to take a closer look at such ‘exceptions’ to find out if they might involve a compound or loan word.

The native speakers can gradually learn how the vowel-harmony system of their language functions, and which vowels are [+ATR] and which are [-ATR], and this will help them when they write and proofread texts in their language. In the beginning of the workshop, we use the non-technical terms ‘heavy’ for [+ATR] vowels and ‘light’ for [-ATR] vowels, and in a more advanced stage we may introduce the linguistic terms. As a shortcut, I often write a small ‘plus’ sign in the bottom corner of the card for [+ATR] and a small ‘minus’ sign for the [-ATR] vowels.

Languages with open and closed syllables need extra checks. In order to discover the full vowel system, words with different root-syllable structures should be checked separately. This is particularly true because some syllable-final consonants may influence the pronunciation of

¹⁸ Normally the prefixes and suffixes follow the same harmony pattern as the root, but those should be studied separately, because the number of contrastive vowels in prefixes or suffixes may be different from that in roots and because of the vowel-harmony morphophonology.

the preceding vowel, thereby creating several allophonic realisations of vowels compared to those in open syllables.

Analysing the vowel system of a language, discovering its vowel inventory and a vowel-harmony system based on noun and verb roots can often be done within a couple of days. The facilitator gets a complete overview of the vowel system, and the speakers will begin to write and read with their newly-discovered vowel system, the system they were ‘born’ with, but which was hidden deep down in their subconscious, because it was overridden by the system they learnt to read in school. The more they practise from now on, the more the deep-seated ‘underlying awareness’ of the system of their mother tongue will come to the surface.

4.5 Consonants

The next topic is analysing the consonants, i.e., finding proof for the phonemic status of each consonant in the most efficient way. We focus on the *root-initial syllable*, the consonant together with the vowel. As long as we stick to this same structural position, we can prove contrast between consonants when each of the consonants is followed by the same vowel, such as /a/. These ‘minimal syllables’ would be enough to prove contrast. However, when one wants to be more complete, one can take each consonant followed by each vowel in that position. The result is a phonotactic table which will reveal if there are cases of *complementary distribution* between consonants, based on their cooccurrence with the following vowel, and if not, this will prove the contrast between consonants, again, by *exclusion of complementary distribution*.

So, after the vowel system has been established and all cards (nouns and verbs) have been looked at and corrected – including vowel length, vowel harmony, vowel nasalisation, or any other contrastive vowel feature – the cards are resorted on the first root consonant. There are always some vowel-initial roots lacking an initial C₁ consonant; the consonant following that vowel does therefore not count as C₁, but rather as C₂. Such cards are laid aside at this point. Once this primary sorting is done, a secondary sort is done on the vowel following C₁.

The next step, once again, is listening for SAME or DIFFERENT. One language speaker reads the words out loud that are written with <ba...> root-initially, and every speaker as well as the facilitator listens if the sound written sounds always the same before <a>. Does the language have one b-sound, or two b-sounds: e.g., an implosive /ɓ/ and an egressive /b/? Often the participants will have already written the egressive one as and the implosive one as <bh>. If so, attention needs to be paid to all b-initial nouns and verbs to establish whether they are all the same or whether there are two contrastive b-sounds. It is best to stay as close as possible to people’s original ‘intuition’ for their orthography,¹⁹ though, as with the vowels, if they have so far merged two consonants in one symbol, they now need to choose two symbols, one for each contrastive consonant. This principle applies to checking all consonants. Consonants seem to be more ‘concrete’ in people’s minds, so usually, there are not many problems.

There may be a few consonants, however, that may not combine with the most general vowel /a/. In quite a few languages, the fricatives /f/ and /v/ may only ever occur followed by a high vowel /i/ or /u/. One would want to contrast these with some other root-initial consonant(s) differing by only one distinctive feature, and which could therefore conceivably be in complementary distribution. However, in the example of Figure 1 from Bangubangu [bnx] (D.27), they appear to be contrastive: /f/ with /p/ and /v/ with /b/, because they are both followed by the same high vowels /i/ or /u/.

¹⁹ The ultimate choice of orthographic symbols for every phoneme should be as close as possible to the system with which the speakers have ‘grown up’ in their school system and/or what seems to work well with native-speaker intuition or in neighbouring languages.

u- <u>ful</u> -a	to dig	tu- <u>vi</u>	faeces
u- <u>pup</u> -a	to winnow	bu- <u>ĩ</u>	evil

Figure 1: Bangubangu words contrasting /f/ - /p/ and /v/ - /b/ (data courtesy of Ginger Boyd).

Another example which can be found in many languages is that [l] and [r], or [l] and implosive [ɖ], may turn out not to be contrastive but allophones, based on combinatory restrictions with preceding or following vowels. Both French and English have /l/, /r/ and /d/ as separate phonemes, and they are written each with their own symbol. As a result, people in newly-to-be-developed alphabets often want to write separate letters for each, even though they may be allophones. One needs to make sure they can write these consonants in the respective environments consistently. Such issues come up during the presentations and the checking sessions, and one needs to make immediate decisions, work with the participants as to what seems best in their eyes, but also challenge them and evaluate how they are handling the situation.

After having established the contrastive consonants, one can present the results in tables like the one in Figure 2, either arranged according to place of articulation or in alphabetical order. Figure 2 shows contrast by means of ‘minimal syllables’ in the same root position in Vanuma, where <bh> represents /b/ and <wh> represents /β/.

b	ka- <u>bà</u> há	to clear fields again
bh	ka- <u>bhá</u> lá	to sprout
wh	ka- <u>wha</u> ɖà	to dialog
gb	ka- <u>gbà</u> ká	to gather up
mb	ka- <u>mba</u> mbalaka	to change colour
ngb	ka- <u>ngba</u> ngbanya	to be of the same dimension
p	ka- <u>pa</u> ka	to congratulate
kp	ka- <u>kpá</u> ndá	to last

Figure 2: C₁ bilabial and labial-velar root consonants followed by the vowel <a> in Vanuma [vau] D.331 (personal data).

In many Bantu languages, prenasalised consonants cannot occur root-initially in verbs. In nouns, initial prenasalisation of obstruents is frequently found, but it represents a noun-class prefix. When a prenasalised consonant occurs elsewhere in an inflected verb, it may also be an object prefix placed before the verb root. However, prenasalised consonants frequently occur in C₂ position in noun and verb roots. If such is the case, one can show contrast by choosing a contrastive set with all prenasalised consonants in C₂ position followed by the same vowel, as exemplified in Figure 3 in Zimba (in orthographic notation).

/mb/	ko- <u>amba</u>	to pay
/nd/	ko- <u>kanda</u>	to put together
/nj/	ko- <u>anja</u>	to spread out a mat
/ngy/	ko- <u>bengya</u>	to accept
/ng/	ko- <u>bánga</u>	to bend

Figure 3: Prenasalised consonants in C₂ position in Zimba [zmb] (D.26), personal data.

Many languages have exclusively open syllables. On the other hand, there are also those which have both open and closed syllables. In such cases, a separate consonant chart should be made for syllable-final consonants, often a subset of the consonant inventory of syllable-initial consonants. Sometimes, in syllable-final position, a consonant may be found which does not exist in syllable-initial position and which may be an allophone of a root-initial consonant. The cards which contain words with such syllable structures in the root need to be sorted on their final consonant, followed by a secondary sorting based on the vowel preceding the final consonant.

It is important to be aware of the fact that syllable-final consonants may influence the phonetic realisation of the preceding vowel. In fact, before each different syllable-final consonant, the vowel may sound somewhat different – and these vocalic variants would all be allophonic realisations of one and the same vowel phoneme. It may be difficult to determine these just by ear, but using the SAME-DIFFERENT test, participants will normally group all such variants as the same vowel.

4.6 Tone

All languages use melodic prosodies like tone, stress, and intonation as an integral part of the word or longer utterances in the speech chain, like the phrase, clause, or sentence. If any prosodic feature is used contrastively in a language, it needs to be analysed. Because this chapter is written from an African perspective, I will focus on tone. But the principles are the same for languages which have other melodic prosodies. Most languages spoken in Africa are tone languages, and tone forms an integral part of each utterance. Words consist of vowels, consonants, and tone or a tonal melody.

In our approach to language development, participants can begin writing their language on the segmental level as soon as the contrastive vowels and consonants have been determined and symbols chosen for them. The case for tone is different. It needs to be analysed and its function established, before one can decide whether, and if so how, it should be written, so as to make sure people do not have to guess while attempting to read.

To have to do tone analysis in a language-development workshop may be a daunting prospect. At this point it is not necessary to do a complete tone analysis of the language in all its detail. For language and orthography development, one can stick to some basic aspects of tone analysis which will give an answer to the question whether one should consider writing tone in the language or not. As with vowels and consonants, one wants to listen to nouns in isolation and verb infinitives in isolation, focussing on the surface tonal melodies belonging to the noun and verb roots.

Even more so than with vowels and consonants, one should not attempt to interpret and establish the number of tones by listening to individual examples. The SAME-DIFFERENT test should be done to group words with identical tonal melodies together before attempting to establish what might be the underlying tonal melody of each of the groups of cards.

To start the basics of tone analysis, the cards need to be re-sorted by the structure of the root. It is important that long vowels (if they have been established within monomorphemic noun and verb roots) be kept separate from short vowels, because they may or may not consist of two morae (tone-bearing units).

When the words have been sorted in this way, one participant reads the first word, and then s/he or another speaker should whistle it. That is the beginning of the first pile. S/he then continues with each word in the same way, and puts all the cards with the same surface melody on the root in the same pile, and different melodies in different piles. This work should be done carefully and systematically, aiming for consistent piles in the end. Once the initial

sorting into piles has been done, one wants to go through this procedure again with each pile, in order to verify that indeed each separate pile shows internal consistency. Any discussion of specific tone heights should be avoided as much as possible in this early stage. One should first establish *how many* different piles with different melodies there are, before attempting to interpret and assign labels like high, low, and mid to certain pitch levels in a melody. Any interpretation of the surface tones or melodies remains a hypothesis until, in the following steps of the research, it gets gradually confirmed as a solid analysis of the underlying inventory of tonal melodies on lexical items.

In many cultures, it is not acceptable for women to whistle. It is therefore advisable not to put pressure on women to do the whistling part of this exercise. However, I have often been in situations where the women were so fascinated that they wanted to try it anyhow, saying that they were doing this in the context of their language ‘laboratory’. They appeared to be quite good at it.

One wants to observe whether each vowel can only carry a level tone, or whether there are also combinations of two tones possible in the form of rising and falling tones on one (short) syllable, or only on long syllables. In most cases, this is relatively easy to hear.

The whistling exercise will reveal the number of tonal melodies belonging to noun roots and to verb roots. For example, in many Bantu languages,²⁰ nouns may carry one of four different tonal melodies on the root, whereas verbs may only be divided into two tonal classes. In other languages, disyllabic noun roots may yield nine different combinations. In such cases, it is easy to establish the number of contrastive tone levels in a language. However, there are also many languages in which this is less straightforward, e.g., when there are six or seven possible tonal melodies which seem to present three different tone levels. Such languages may have high-tone spreading, downstep, or any other tonal process which obscures the number of underlying tones for a while, until further research has been done.

Based on the number of tonal melodies belonging to the root, one should attempt to interpret the different melodies into a possible system of as few contrastive level tones as possible. Quite often at this point, one can already establish the number of basic level tones in the language, whether the language has a two-tone system (with all possible combinations of L and H on disyllabic roots) or three (and the possible combinations of L, M, and H on disyllabic roots), four, or (rarely) five tone levels in its inventory.²¹

During the process of listening to the tonal melodies on words in isolation, it would be good to make a list of all tonal minimal pairs or triplets one comes across. Sometimes the vocabulary cards contain words written the same, but which differ in tone alone. At other times, pronouncing and whistling a word may remind someone in the group of another word which has the same vowels and consonants, but a different tonal melody. One should use such an opportunity to write cards for those words, too.

A list of tonal minimal pairs is very useful for two reasons. First, one wants to know whether the language has many of such lexical minimal pairs, or just a few. Secondly, in case tone needs to be written in the orthography, these minimal pairs are the best examples for practical exercises to learn to read according to whatever symbols are chosen for a tone orthography.

This first step in tone research gives the facilitator and the native speakers some insight into the functional load of tone in the lexicon. Since one is looking for some pointers towards answering the question of whether to mark (some) tone in the orthography, one also wants to know more about the functional load of tone in the grammar of the language.

²⁰ In other languages also, the number of tonal melodies on nouns and verbs may differ.

²¹ The latter two, 4T and 5T languages are very rare, and mostly found in languages with predominantly monosyllabic roots, see also Kutsch Lojenga 2014:62 and Kutsch Lojenga 2018:82

Gradually, when doing more research in the language, examples of the function and functional load of tone in different morphological or syntactic structures will come to the surface. In fact, these are often more important when deciding to establish a tone orthography. This may concern sets of pronouns which differ by tone alone, singular-plural distinctions in nouns, a definite-indefinite distinction in nouns, TAM distinctions in the verbal system, or a distinction between main clauses and relative clauses, which can all be expressed exclusively by tonal contrasts. At the first signal of a grammatical tone distinction, it would be wise to make a note of the example just discovered, which can then be elaborated upon at a later stage.

After the basic analysis of the sound system and an initial hypothesis for the tone system, the next topics will be noun morphology, noun-phrase morphology and syntax. At each point in the morphological or syntactic research, one should pay attention to tonal issues as well. In this way, the tonal analysis is not one big daunting topic as a whole, but is interspersed by researching other issues in the grammar of the language.²²

4.7 Results and next steps

Establishing an alphabet using the approach of participatory research as outlined above could be done in a time period of a couple of weeks. We normally document the results in an alphabet poster (see Figure 4), laying out all the graphemes (including digraphs, trigraphs) representing simple and complex phonemes²³ with a picture of a word containing that grapheme.











































NGÒRÒSÉ-KÍHÁ MÙGÙ KÈBHÚ-TU KÈ					ALPHABET KEBHU				
a  ándá	b  ibí	bh  ìbhí	d  díí	dh  ídhé	dr  drúkó	dy  kódyá	e  gègèliyá	ε  kéle	
f  ufú	g  gílo	gb  gbòrúwò	h  húhú	hw  hwégé-tu	'  kó'í	i  isí	i  mímí	k  kerú	
kp  kpàkpà	l  lìbhíìbhí	m  mènzì	mb  mbugù	mv  umvu	n  niní	nd  ndéndé	ndr  ándrú	ng  àngù	
ng'  ng'alú	ngb  dángbé	ny  unyl	nz  unzú	o  òmbó	ɔ  òrómbo	p  pòrò	r  ara	s  sìmbá	
t  ttò	tr  trútrú	u  usù	tt  súngúrú	v  vàrà	w  wèrì	y  aya	z  zógó		

Figure 4: A first-time alphabet for the Ndo / Kebhu [ndp] language in D.R. Congo.

²² For more practical information about tone analysis in a participatory way as well as more typological background, see Kutsch Lojenga (2018).

²³ These include digraphs for simple phonemes such as ny for [ɲ] or digraphs and trigraphs for complex phonemes such as prenasalised consonants and affricates.

The vowels are the most crucial aspect of the analysis. However, for some speakers, it takes time to become aware of the ‘extra’ vowel sounds which may not appear in the LWC. To help with this, we also produce a vowel-teaching booklet (see Figure 5). This is a small document for practising the vowel graphemes, primarily intended to raise their awareness of the contrastive vowel sounds, and thereby a simple teaching tool.

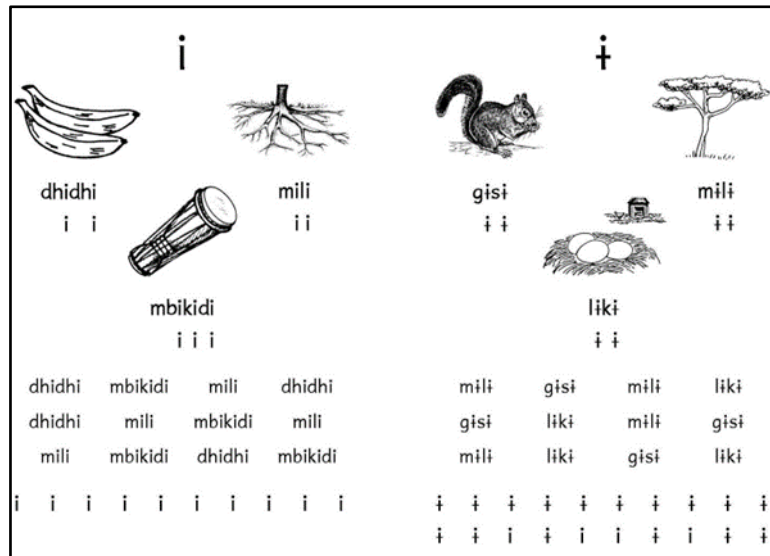


Figure 5: Two pages of a vowel-teaching booklet: one known vowel <i> and one ‘new’ vowel <i>

Once the vowel and consonant awareness has come to the surface, the participants should be encouraged to begin to correct all the words on the cards and try to collect more vocabulary, writing them straightaway in the newly-developed alphabet. This activity, as all others before, ideally takes place under the guidance of the facilitator. It is hoped that the participants can now identify and correctly write the sounds of their language without reverting back to the SAME-DIFFERENT checking procedure as they did earlier on.

The next step concerns participatory research of the morphology. I normally divide this into the following subtopics:

- noun morphology
- noun-phrase morphology
- verbal derivational morphology
- verbal inflectional morphology

The first two can be done in the next couple of weeks. In these topics, the aim is to discover the obligatory affixes to the noun (e.g., noun-class marking), then other affixation to the noun (e.g., plural formation, diminutives and augmentatives), then the nouns along with noun-phrase elements, which often include the closed classes (e.g., numerals, demonstratives, possessives etc.), which may have their own internal morphology. We aim to discover the underlying roots of these noun modifiers with their affixes, along with their various allomorphs. For each and every affix, the facilitator guides the participants in writing systematic examples, full sets of noun-class affixes and later full verb paradigms with verbs containing all variables, investigating each morpheme as well as its possible allomorphs.

At this point in the process, the participants should be encouraged to try to write a simple folktale, even though the verb morphology has not yet been investigated. This means

connected text, in which word boundaries must be observed. Certain word boundaries will be clear, e.g., those around nouns and verbs. Others will gradually develop by practical usage and discussion with the members of the group.

In addition, participants should aim to collect enough vocabulary and work towards a small lexicon, putting the words in a dictionary database now that they have been checked and corrected for spelling, tone, vowel length, ATR or other vowel-harmony phenomena, noun class and any other feature that is important in a particular language.

The topics on verbal derivational and inflectional morphology are often treated in a second workshop of 3 - 4 weeks. Again, everything revolves around the discovery of the basic morphemes and their (mostly phonologically-conditioned) allomorphs. Tone is often an important factor, particularly in verb research, because one needs to discover its functional load in that area.

Of course, one can add more workshops, e.g., for discourse analysis or to train writers or develop a dictionary. All these can be run by delegating much of the language work to the speakers of the language, who will gradually become more fluent in writing their language.

5 Pedagogical considerations

How to run such participatory-research workshops with the aim of putting a language to paper, to develop an alphabet and writing system, is an art in itself: the art of communicating, collaborating, and drawing everybody in to the work in the discussions and on paper. The relationship between linguist and language speaker²⁴ should be mutually beneficial. Each one has their own expertise. The speakers have innate knowledge of their language, albeit that their conscious knowledge remains within the limits of what they have learned and have been exposed to in their lives so far. On the other hand, the linguist contributes with the knowledge acquired in his/her training in linguistics, and s/he needs to establish a method and sequence of topics for the research together with the speakers of the language, since they may not have an overview of the subject matter.

In this style of collaboration, the two parties should work together, communicating and discussing every issue together. Each one should have a chance to contribute according to their knowledge and language competence. The linguist should give guidance in the sequence of activities. S/he should share with the group each time *what* they are going to do and *why*, e.g., 'we are going to collect vocabulary in order to use these words to study the sounds of your language ...', or asking them in what language(s) they want the translation of their words in their dictionary. S/he will steer the thoughts of the participants to a certain degree whenever helpful, but during the research, s/he should not go off on a tangent into other topics that happen to come up spontaneously and that may fit better into the program at a later stage in the discovery process.

At any point in the discovery process, the linguist wants to ask questions, so the speakers of the language can express their observations, comments, or questions, to which s/he then can respond or filter out something which may be helpful to discuss at a later stage. S/he may share his/her observations, as simple as they may be, so as to raise the participants' awareness on the topic they are researching together, using simple language. Many people are not used to observing patterns in structures spontaneously, and an outsider drawing their attention to such

²⁴ Even if the linguist is a speaker of the language under study (which we would not advise), the methodology presented in this chapter should be followed. A native-speaker linguist would still need to guide the participants from his/her language in the discovery process. In fact, s/he may discover some new aspects of his language him-/herself.

patterns will gradually make them more aware in this domain, and they may slowly develop such a skill, too.

The linguist should not necessarily use the official linguistic terminology, although s/he may move into that in a later stage, when the speakers have already understood the concept. Even if s/he thinks that explaining topics like vowel harmony, morphophonological rules, tonal downstep or floating tones may be too difficult for the participants, s/he can draw their attention to one aspect at a time, or repeat bits and pieces of the system in simple, 'popular' terms. The speakers will gradually 'absorb' the details of their sound system and later their grammatical structure. It is hoped that, in every domain, researcher and language speakers ultimately come to the same conclusions about these topics. Language speakers may use the results for their writing system; a linguist could also use the results for a publication.

This approach can be viewed as research-cum-training, whereby the *training* happens through *awareness raising* of sound contrasts with the on-the-spot assistance of a researcher. This is different from the outside linguist first analysing the language and then teaching the mother-tongue speakers the outcome of his/her findings of the language structure. This is also different from the outside linguist teaching the participants courses on various topics, e.g., phonetics, phonology, ahead of time, and then expecting the speakers of the language to have enough insight to do their own analysis predominantly themselves.

The participants should write as much as possible themselves, even though in the early stages they may still make quite a few 'mistakes', since they are only gradually learning the new realities of their mother tongue. The facilitator does not need to correct every 'mistake', because, as the people become more aware and grow in the knowledge and understanding of how their language functions, they will automatically improve. In fact, when starting to collect vocabulary, writing themselves will increase their feeling of responsibility for their language as well as the idea that the language is really theirs!

Each day, or sometimes each morning and each afternoon, the workshop leader or facilitator introduces a new topic, according to the sequence of activities in the planned methodology. If at all possible, this session should not be a lecture about the topic, but a presentation inviting input from the participants themselves. The list of topics (see Appendix) is more or less fixed from the beginning. Only if the results of the research of the previous day make me aware of something that needs to be looked at before the planned new topic will I amend my overall list of topics for research.

The topic of the day needs to have a connection with something already known, a point of departure. The known element comes from the language they have in common, an LWC or an official language. In a country like the Democratic Republic of the Congo, that means one (or more) of the four LWCs: Swahili, Lingala, Kikongo, Tshiluba. It is important for the workshop leader, but also the facilitator to be very familiar with the morphological structures of the relevant LWC for the area. The starting point may also be the international language, be it French, English, or Portuguese.

During each session that introduces a new topic, the workshop leader should try to elicit with the entire group examples of the topic they need to grasp before working on it in their own language, be it *phonological* (e.g., words in an LWC with two identical vowels in the root), *morphological* (e.g., diminutive marking), or *syntactic* (word order possibilities between noun and numeral in the LWC and French). One can even highlight a difference between languages known to the participants. After seeing examples in the LWC, national or official language, they may start providing examples in their mother tongue. As they write them on the board, the workshop leader or facilitator can try to interpret the example on the spot if at all possible, showing similarities or differences with the language used for the introductory presentation.

For example, early on, when one wants to elicit verbs in their infinitive forms, one can ask the group for infinitives in Swahili, and ask someone to write these on the board. Next, one can ask for infinitives in Lingala and write these on the board, too. Then one asks what the difference is (a *ku-* prefix in Swahili and a *ko-* prefix in Lingala). These are very close. Then one asks for infinitives in French, and writes these on the board, too. Discussing the differences between French and the two LWCs shows a substantial difference: the infinitive marking in French is at the end, and there are four infinitive endings in French. In the same way, one can ask for imperatives in Swahili, Lingala, and French. This type of exercise will raise their awareness of grammatical categories as well as of the element they will have to research in their group. One can then ask a participant from each language to write one infinitive and one imperative on the board before starting the research in their own language group.²⁵

Early on, when planning to give an assignment of sorting noun roots with two identical vowels, one can ask a person from each group to write two words on the board containing two identical vowels in the root. It helps the people when they see examples on the board, and a facilitator who confirms that this is the type of example they need to find.

Later on, the presentations with examples will become more complex, such as finding noun-class prefixes in singular-plural pairs, and later agreement prefixes on adjectives, numerals and other constituents. If possible, the topics move from easy to more difficult.

Some facilitators may be afraid to lose control over what data they will get, or the precision they would like to see in their results. They may think that the results would be of higher quality if they themselves keep everything under their control. This is not the case. If the facilitator sticks to the sequence of activities as set up for the work, following roughly the steps as laid out in this chapter and adapting these according to the root structures of the lexical items, only digressing when something new and relevant turns up, they should not lose control at all. The SAME-DIFFERENT test, when done with precision, is a means to produce quality data. When progressing like this, one need not be afraid of losing control; rather, the comments of the participants will enrich the facilitator's understanding of the language structures and may ultimately produce more precise data.

6 Conclusion and benefits

Language constitutes an inherent and inalienable part of people's identity. Their mother tongue is close to their hearts. In company, people will speak an international, official language or LWC only when necessary. As soon as the opportunity arises, when surrounded by those of their own ethnic and linguistic background, they will typically speak their mother tongue with their interlocutors.

Africa has always been largely an oral society, so less attention has been given to writing and reading local languages. In most language communities, however, there have always been a few people who have tried to create a written form of their language, some with an amount of success, while others have dropped the idea because it was beyond their capabilities. Those who have persevered have mostly written their language with a five-vowel system, since that is all they have ever seen. Only a few will have known that there are more vowels in their language.

In many African countries, people are asking for help to write their language. They do not want their heritage to be lost, though often they are nearly too late. They need help. In this chapter, I have proposed a methodology for this task by way of participatory research, a linguist doing the task together with a group of speakers of the language. This takes place as a

²⁵ At this early stage, this will still be done in the five-vowel system with which they started their research.

positive social event, whereby the contents of the research raise native-speaker awareness of vowel and consonant contrasts, as well as morphological and grammatical structures. Such awareness is necessary for writing and reading their language, and above all, it strengthens their identity. My expectation is that if one follows this approach, the analysis of the basic aspects of the language will go faster, and in the end, it will also be more accurate.

Finally, this approach to basic language analysis is not only constructive for creating a practical alphabet and informative for all other aspects of a writing system, it also lays a basis for academic linguistic description, such as a descriptive sketch, a dissertation, or historical and comparative studies.

APPENDIX

The following is a sequence of contents of the linguistic analysis in a participatory-research workshop intended for the development of an orthography of a language.

WORKSHOP 1 (3-4 weeks)

Part 1: Basic phonology

- A. Collect vocabulary; discuss the concept of vowel, consonant, syllable, and root.
- B. Vowel research:
 - Determine the inventory of contrastive vowels by comparing identical vowels in the roots.
 - Confirm the vowel inventory and do a phonotactic study of the vowel combinations in disyllabic roots in order to determine any type of vowel harmony based on cooccurrence restrictions.
- C. Consonant research: Determine the set of contrastive consonants by comparing all root-initial consonants preceding the same vowel.
- D. Tone: Determine the number of underlying tone levels from the contrastive tone melodies in roots.

Part 2: Morphology and Syntax: Determine the morphophonology of bound morphemes and the word order of independent morphemes.²⁶

- A. Morphology of the noun word: Determine the order and underlying form including tone of the following:
 - obligatory morphemes such as noun-class or gender marking
 - non-obligatory morphemes such as plural, diminutive, augmentative and locative marking
 - verb-noun derivational morphemes such as in agentive nouns and abstract nouns
- B. Elements of the noun phrase: Determine the order and underlying form, including tone, of the following:
 - noun + numeral (including the full numeral system used for counting)
 - noun + possessive pronoun (including alienable/inalienable distinctions)
 - noun + demonstrative
 - noun + noun in associative construction
 - noun + adjective
 - (possibly also) noun + (subject and object) relative clause

²⁶ Depending on the language, it may be more important to establish the underlying form of each bound morpheme or the independent word.

WORKSHOP 2 (3-4 weeks)

Verbal morphology and syntax: Determine the morphophonology of bound morphemes and the word order of independent morphemes.

- A. Determine the infinitive and imperative forms.²⁷
- B. Verbal derivation: Determine the infinitive form of each verbal derivational category.
- C. Verbal inflection: Determine the order and underlying form, including tone, of all TAM morphemes in relation to the root.²⁸ Establish the following for each:
 - Paradigms for each root vowel (to observe how root vowels influence affix vowels or vice versa) and tone class (to observe and analyse tonal behaviour)
 - Subject marking
 - Object marking
 - Negative marking

References

- Casali, Roderic F. 2008. ATR harmony in African languages. *Language and Linguistics Compass* 2: 496–549.
- Cucchiariini, Catia. 1993. *Phonetic transcription: a methodological and empirical study* Doctoral dissertation. Nijmegen. Radboud Universiteit.
- Kutsch Lojenga, Constance. 1994. *Ngiti. A Central-Sudanic Language of Zaire*. Köln. Rüdiger Köppe Verlag. Doctoral dissertation. 517 pp.
- Kutsch Lojenga, Constance. 1996. Participatory research in linguistics. *Notes on Linguistics* 73: 13–27.
- Kutsch Lojenga, Constance. 2014. Orthography and Tone: A Tone-System Typology and its Implications for Orthography Development. In Michael Cahill and Keren Rice (eds.), *Developing orthographies for unwritten languages*. Dallas: SIL International, pp. 49-72.
- Kutsch Lojenga, Constance. 2018. Tone and Tonology in African Languages. In Augustine Agwuele and Adams Bodomo (eds.), *The Routledge Handbook of African Linguistics*, pp. 72-92. London: Routledge.
- Pike, Kenneth L. 1947. *Phonemics: A technique for reducing languages to writing*. Ann Arbor, MI: University of Michigan Press.

²⁷ Infinitives and imperatives should be introduced in workshop 1 but discussed more thoroughly in workshop 2.

²⁸ These TAM categories may initially be labeled with non-technical or even arbitrary terms. The morphemes may be determined through elicitation or texts, if there are texts at this point in the research.