Phonological competence in the design of a new orthography

Paul Tench
Centre for Language and Communication Research
Cardiff University

Africa has a population of 797 million, speaking 2,092 languages, across 53 different countries. A sixth of Africa's population live in Nigeria, approximately 130 million, and between them they speak about a quarter of those languages, 516 to be precise. Nigeria has the third largest number of languages within national borders in the world, following Papua New Guinea and Indonesia; India has 427 languages, and China 241 (Gordon & Grimes 2005). Of those 516, 510 are indigenous; Hausa, Igbo and Yoruba are classified as major regional languages ('languages of wider communication'); Edo, Efik, Fulfulde, Ibibio, Idoma, Itsekiri, Izon, Kanuri, Tiv, Urhobo, among others, are classified as 'medium' languages, leaving about 490 as 'minority' languages, spoken by about half of Nigeria's population (see Igboanusi & Peter 2005: 6).

Minority languages are largely unwritten, largely ignored in education and government, but nevertheless provide an essential factor in local social cohesion and the means of maintaining the integrity of a local, distinctive, culture; in other words, although minority languages may not mean a great deal to outsiders, they provide the social bonding among their native speakers. Minority language speakers have their own pride and aspirations for themselves and their people, just as much as the speakers of medium, major and international languages have. Among the minority groups in Nigeria, none have greater pride and sense of ambition than the Tera people.

The Tera

The Tera are one of the ethnic groups that occupy territory between the eastern border of the Hausa and the western border of the Kanuri speaking people in Northeast Nigeria. They number over 100,000 and live mainly in the northern part of Gombe State and the eastern part of Borno, They are mainly agriculturalists, specializing in guinea corn, millet, maize, rice and wheat, and orchards; other major occupations include fishing and weaving. Their traditional dancing is well known in Nigeria. Their mother tongue is used in family and village life and in local markets; they use Hausa as their language for wider communication, but increasingly, English features in higher levels of education and in new business evolving around computing. Hausa is generally used in their education, worship and in the city life of Gombe. Although

there is some population drift into Gombe, the majority of the Tera remain a rural population, whose area is fertile, but whose transport infrastructure is precarious.

The Tera call themselves Nyimatli /ɲimaɬi/; their language is Chadic like Hausa and the string of other languages that lie across the border between Hausa and Kanuri speakers. There was a brief period in the 1930s when literature appeared in the Tera language. The British and Foreign Bible Society published a 'tentative' translation of the Gospel of John in 1930 in an alphabet that included many letters with a subscript dot. A catechism and songbook followed shortly afterwards in stencilled form, but the typing omitted all subscript dots. A change in missionary organization policy downgraded the use of local languages like Tera in preference for developing Hausa as a lingua franca; this facilitated the mobility of personnel, not only in mission, but also in government. But this meant stunted development of Tera literature.

In the 1990s a partial revival of interest in Tera literature was promoted by a local academic, Ayuba Nyagham, who introduced a number of changes to the letters of the Tera alphabet, matching it closer to the Hausa, which the whole educated population could read. His untimely death closed this potential development prematurely, until another local teacher took action. Jauro Maila broadcast news in Tera on the radio in the 2000s and issued a number of papers in an alphabet that resembled Nyagham's, although it was constructed quite independently. The urge to establish new orthographies was fuelled by a new drive to assert the Tera people's distinctive culture and language, their separate identity as a people in the midst of political and religious conflicts, and their determination not to allow the domination of Hausa

language and culture. They were afraid of losing their language and with it their sense of identity, heritage and dignity.

A remarkable young lady stepped on to the stage, a graduate student named Isioma Jideonwo who worked among the Tera on a placement in the national, postgraduate, Youth Corps programme. This enterprising young lady published a book in English, "Let's Develop Nyamatli Language", in 2004, the result of a good deal of research into the history, culture and language of the people. The alphabet she uses bears close resemblance to those of Nyagham and Maila.

The final actor in this tale of development is the Bishop of Gombe who sought to act as a catalyst for the production of Bible translations in the local languages of Gombe State, including Tera. His action eventually resulted in an orthography workshop held in 2004, at which Tera was represented by four men chosen by local communities. The objective of this workshop was the production of a 'working' orthography. The methodology used and the theoretical underpinning are described below, but first the main principles of orthography are discussed.

Orthography

Orthography represents words. This is clear from humanity's earliest forms of writing and from humanity's modern writing systems, whether we consider logographic systems like Chinese characters, syllabaries or alphabets, or indeed mixtures of them. Writing is for meaning, and words and morphology are the basic elements of meaning. Words represent our experience of all the things, actions, qualities and relationships that we perceive in the world around us and within us.

That an alphabetic orthography represents words is clear from the observation that blank ‘slots’ either side of a string of letters are called ‘word spaces’; they mark the beginning and end of words. Symbols in syllabaries are perceived as grouping together to represent words. Also modern icons on domestic articles, charts, mechanical or electronic equipment represent messages that can easily be expressed as single words, eg ‘cloudy’ on a weather map; ‘non-iron’ on a shirt label; ‘print’ on a computer, etc.

Orthography also represents grammar. Sentences and clauses can be marked, eg with capital letters and full stops or commas; relationships between clauses can be marked by other marks of punctuation, including dashes and brackets. Sequences of written words follow the same sequence as spoken. Cultures may also have idiosyncratic features for indicating some grammatical information; for example, in English, possessive <s> is marked with an apostrophe to distinguish it from plural <s> in nouns; and in German, an initial capital serves to identify nouns.

Orthography also represents discourse. Paragraphs, as significant sections of text, are marked by beginning them on a new line, often indented; chapters likewise, by beginning them on new pages. Question marks and quotation marks indicate discourse functions; exclamation marks and typefaces often provide paralinguistic information.

In short, orthography represents language, whether dynamically, i.e. in actual use, or statically, as in dictionaries, telephone directories, etc.

Whereas orthography represents language – its words, grammar and discourses – an alphabet reflects phonology. It reflects phonology at the level of word; phonological resources for representing words include consonantal and vowel systems and their distributional criteria, phonotactic and syllable structures, prosodic features and syllable counts. An alphabet also may contain the means for indicating rhythm patterns in sequences of words, eg the use of hyphens to distinguish compounds from a sequence of separate words, and the use of apostrophes to indicate missing syllables in informal colloquial speech. Intonation can also be marked in an alphabetic writing system, through punctuation marks, underlining or changes in typeface.

The Roman alphabet has 26 letters at its disposal in both upper and lower case and a range of punctuation marks. Each letter is distinguished by distinctive features of its shape, but varies considerably in type (fonts) and handwriting. Some cultures allow additional letters like German <β>; others do not employ the full range, eg Welsh does not use <j, k, q, v, x, z>. Punctuation marks likewise have distinctive features, and likewise vary considerably in print and handwriting. Cultural variation is more widespread in the case of punctuation marks: consider the shape of quotation marks in English, German and French culture, and inverted question and exclamation marks in Spanish.

An alphabet is subject to a scale of complexity (Sgall 1987: 17-18), ie a complexity of *design*. A letter may be

(i) simple (Sgall calls this a ‘protographeme’)

(ii) complex with a regular diacritic (‘subgrapheme’)

(iii) complex with an irregular diacritic

(iv) combined (‘string of protographemes’; ‘bigrapheme’, ‘trigrapheme’)

(v) combined but one element with a regular diacritic

(vi) combined but one element with an irregular diacritic

(vii) as (vi) but corresponding to an additional phoneme

A regular diacritic (see ii above) is a diacritic that marks one phonological feature only and consistently; Sgall gives the example of <´> over Czech vowels to indicate length. An example of an irregular diacritic (iii) might be the subscript dot <̣> in the old Tera alphabet, which means implosive with <b, d, k>, velar with <n>, voice with <x>, and fronting with <u> (at best, it means ‘some kind of phonetic variation from a conventional value’). Combinations of letters like <th> in English (a ‘bigrapheme’), <sch> in German (a ‘trigrapheme’) are common. Combined letters can be accompanied by diacritics; Sgall gives <dě, tě, ně> as an example of (vi) and “Czech ě corresponds to two phonemes after <m, p, b, v, f>” (Sgall 1987: 14), as an example of (vii).

An alphabet is also subject to a scale of univocality (Sgall 1987: 18), ie *function*. This refers to the degree of double consistency between a letter (simple, complex or combined) and a phoneme: that a phoneme is represented by one ‘grapheme’ only; and a ‘grapheme’ represents one phoneme only. The scale of univocality is as follows:

(a) absolute bi-uniqueness

(b) relative bi-uniqueness

(c) regular deviations

(d) regular deviations without corroborating phonemic alterations

(e) irregular deviations

(f) irregular deviations that ensure a single graphemic shape for a morpheme

(g) irregular deviations with no functional justification

Sgall manages to illustrate each level (apart from d) from Czech. In English, absolute and relative bi-uniqueness are hard to demonstrate. A case of regular deviation is the 'bigrapheme' <qu> representing /kw/ except in words of French origin. Irregular deviations are available aplenty – all homophones and homographs (e). *Do* and *say* represent examples of (f): their graphemic shape is maintained despite differences in inflected forms (*do, doing, does, done*; *say, saying, says, said*). Another case of (f) is <-ed> as the regular past tense marker, which keeps its graphemic shape but represents phonological alterations. The idiosyncratic cases (g) are exemplified by the insertion of (silent, quasi-etymological) letters in words like *debt, ghost, salmon, isle*, and by the well known case of <-ough> words.

Despite the many degrees of variation in these two scales, an alphabet has a distinctive advantage over other orthography systems: its repertoire is generally of the same kind of order as the phonology of a language. This means that the letters/graphemes are relatively few in number, compared to the number of units in syllabaries and certainly in logographic systems. In other words, the total output of communication by all who use a language like English is possible with just 26 letters and its range of punctuation marks. Its phonology comprises a vowel system of 20 or so phonemes (Standard Southern England), a consonant system of 24, three degrees of stress, a choice in rhythmical grouping and a few intonation patterns, a total of no

more than 50 or so items – again, enough for the total output of communication by all who use the language.

Naturally, the most efficient alphabet for a language engages the top end of the two scales. One might question whether Sgall is right to consider a letter with a regular diacritic (his ii) as less complex – after all, he is Czech and used to it – than a 'bigrapheme' (his iv) like English < th, sh, ch, ph > - but then, I am English and used to them. One might revise Sgall's scale of complexity, as for instance

i) simple

ii)	simple with regular diacritic	ii)	bigrapheme
iii)	simple with irregular diacritic	iii)	bigrapheme with regular diacritic
		iv)	bigrapheme with irregular diacritic
		v)	as iv) but corresponding to an additional phoneme

The efficiency of the alphabets used for Welsh and Spanish is a good deal greater than that of the alphabets used for English and French; that measure of efficiency is directly related to the two scales. Scheerer (1986) and Coulmas (1989) distinguish between 'shallow', 'intermediate' and 'deep' orthographies. A 'shallow orthography' reflects closely the phonemic distinctions of the language, as the Welsh and Spanish do; their orthographies are a reasonably good guide to pronunciation. An 'intermediate orthography' reflects the phonemic distinctions of the language by and large, but also incorporate some lexical and morphological information, like Dutch and German. A 'deep orthography' contains "a significant amount" of lexical and

morphological information (Coulmas 1989: 169), like English and French. It is not relevant here to justify a 'deep orthography' beyond the observation that they tend to be long established and unrevised and thus do not reflect historical changes in pronunciation (like the Great Vowel Shift in English) and the importation of loan words. (A less 'metaphorical' set of terms might be more transparent: a 'shallow' orthography is highly '*phonemic*'; an 'intermediate' orthography is typically '*morphophonemic*'; and a 'deep' orthography can be said to typically '*lexicophonemic*'.)

Vachek (1964) formulated the two main requirements of an alphabet as *transparency* and *learnability*. Transparency means that the written form should be easily processed as a word or a string of words in their appropriate morphological shape: "the path from the graphemic form of the text to its meaning should be straightforward" (Sgall, 1987: 15). Learnability refers to the simplicity and regularity of the rules for spelling and pronunciation. (It is well known, for instance, that Welsh orthography is more learnable for Welsh speaking children than that of English is for English speaking children.)

However, when a new orthography is created for a language, there are other considerations to be taken into account. (The metaphor of "reducing" a language to writing is no longer appropriate, with the current meaning of *reduce*. Rather, a language is "reproduced" in writing, on the understanding that a current spoken form is being matched with an appropriate written form.)

Creating a new orthography

Smalley (1963) proposed five major criteria for the development of an optimal writing system, which he listed in order of importance (p 30) as follows:

1 maximum motivation for the learner

2 maximum representation of speech

3 maximum ease of learning

4 maximum transfer

5 maximum ease of reproduction

Criteria 2 and 3 match Vachek's 'transparency' and 'learnability' respectively, and represent core applied linguistic concerns. However, it is sociolinguistic concerns that lie at the heart of criteria 1 and 4, and technology that is the basis of criterion 5. What follows is a brief discussion of the five criteria, Barnwell's re-formulation of them and Coulmas's succinct summary.

1 *Maximum motivation for the learner*

What Smalley means by 'learner' is the native-speaker community who will be 'learning' to use the new writing system. It is *their* language; the new orthography is for *their* use. Their acceptance of it is crucial for the success or failure of an entire project. Experienced linguists may well be engaged in the project but their perspectives are likely to be quite different from the perspectives of the native-speaker lay person. A linguist may well wish to promote an as explicit orthography as possible, accurate in detail, elegant in practicals, but the local community might have other priorities. They may wish to align their orthography as much as possible to

another language, or for political reasons may wish, on the other hand, to distance their orthography from a neighbour as much as possible.

Winter (1983) supplies an interesting account of this very dilemma. He acknowledged his priorities as a linguist: unambiguity, maximum explicitness, IPA values, but he also conceded that he was an outsider and was writing for other outsiders. The local Walapai community in Arizona, the native speakers, would not be dependent on the total unambiguity or explicitness of the orthography for literacy; Semitic alphabets, for instance, underprovide on vocalization, Russian on stress placement (p 233). They refused to accept IPA symbols and preferred letters familiar from their language of wider communication, namely English. They wanted <b, d, g> for their lenis voiceless stops, they did not want diacritics to make the distinction between stressed and unstressed <a, i, u>, even though sets of homophones were thereby created. They did not want 'special' letters or letters with diacritics; so <θ> gave way to <th>, <ń> to <ny>. Their younger generation wanted to acknowledge in the new orthography the phonological transition from the speech of their elders. Thus, in many respects, local opinion leant towards conformity to the language of wider communication and towards one generation in particular, their younger, who would make the most extensive use of the new orthography. The factors that weighed heavily were cultural and political rather than linguistic.

In Gombe State in Northeast Nigeria, neighbouring languages within the same Adamawa-Ubangi language family came to opposite decisions on the question of special letters. The two languages, Dadiya and Yebu (Awak) both have a 10 vowel system with two harmony sets. Phonetically and phonologically, the vowel systems

are virtually parallel. Yet the two communities have made different decisions. The Dadiya mark the opener variety of each pair with underlining as a diacritic, thus <i> belongs to one harmony set and <i̲> to the other. The Yebu, however, prefer special (IPA) symbols to mark the opener set, viz <ɪ, ɛ, ə, ɔ, ʊ>. (The differences reflected the different approaches of the linguists originally involved in the analysis of each language.) The Yebu are proud of their distinctive symbols for their 'distinctive' language, whereas the Dadiya are glad to align themselves with their other neighbouring languages. The factors that led to their differing decisions were cultural and, possibly, political, rather than linguistic. Furthermore, linguists may not always be able to predict how these decisions will go.

Maximum motivation for the learner, therefore, refers to the preferences of the local, native speaker community.

2 *Maximum representation of speech*

Smalley's second criterion corresponds to Vachek's principle of transparency. What Smalley advocated was a basically phonemic orthography that displays bi-uniqueness. However, other linguistic considerations might be relevant since orthography represents words rather than phonemes. Thus many languages tolerate letters in word-final position that conventionally reflect voiced values, even though the final phoneme is voiceless. For example, German *Bild* is pronounced /bɪlt/, but when the "final" <d> is followed by inflections, it 'regains' its /d/ value, as in *Bilder*, *Bildes*, *bilden*, etc. Thus the word keeps a single graphemic shape despite its phonemic variations. That may well be perceived as a gain over a purely phonemic 'transcription'.

The question of dialects might also interfere with the application of this criterion. Where there is a social consensus, the question of which accent to base a new orthography on is settled. The phonology of a dispreferred accent may well, therefore, not enjoy maximum representation. (Disagreement among dialect speakers can paralyze a project, and if a linguist has to make a choice, it should not be seen as arbitrary.)

The Roman alphabet was not designed for universality. It was designed for Latin, and was adapted from the Greek alphabet – which, in turn, was adapted from the Semitic alphabet. A new orthography requires adaptation according to the phonological and broader linguistic characteristics of a given language. This has been exactly the case also with the Arabic alphabet, which has been adapted in different ways for Farsi, Urdu, Afghan, Sindhi and Malay.

Adaptations might take the form of

1 new letters (e.g. IPA symbols)
2 letter combinations ('bigraphemes/trigraphemes')
3 new values for otherwise superfluous letters
4 letters with diacritics
5 alternative typefaces and sizes
6 combinations of the above

The adaptation of an alphabet is the direct application of the criterion of maximum representation of speech.

3 *Maximum ease of learning*

To ensure maximum ease of learning, Vachek's 'learnability', an orthography should be as simple and as consistent ('bi-unique') as possible. Sgall's two scales of complexity and univocality belong to this criterion. This generally means that a 'shallow' (phonemic) or an 'intermediate' (morphophonemic) orthography is easier to learn. This also means, generally, that an alphabet is easier to learn than a syllabary, since the latter contains many more symbols ('syllabagraphemes'?) than the former. However, phonologies do often contain plurisegmental features like nasalization, pharyngealization, breathy quality, etc that affect whole syllables, and suprasegmental/prosodic features that also affect whole syllables; but alphabet makers have the adaptation possibilities mentioned above at their disposal.

This criterion applies differently to reading and writing. Different cognitive processes apply, which can be illustrated by the case of homophones in English. Reading *there*, *their* and *they're* presents less of a problem than writing them, because the former represents somebody else's choices and the latter requires the individual to make the choice. Furthermore, reading involves the recognition of semantic and syntactic contexts provided by another mind, whereas writing demands decisions about the appropriate spelling for semantic and syntactic contexts that the writer provides. The writer has to decide what graphemic shape has to be provided for /ðɛ:/, a task which seems beyond the capabilities of many British students!

Since far more reading happens than writing, maybe the criterion of maximum ease of learning should tilt in that direction. Maximum ease of learning could perhaps better be interpreted as maximum ease of processing.

4 *Maximum transfer*

The desire for a language to be 'reproduced' in writing might grow from many different concerns. There might be a concern to assert a people's distinct identity and their pride in a distinctive culture; or to enhance a people's sense of respect, dignity and worth; or perhaps, to capture an endangered source of oral tradition and literature. Often there is a desire to promote primary education in the mother tongue, either for its own sake, or as a bridge to literacy in the written form of a language of wider communication. Bible translation has been a major motivation, as has the availability of translated documents of a political nature.

Because access to a relevant language of wider communication is a major factor in orthography projects, reference to the linguistic characteristics of that language should figure prominently in detailed decisions in the creation of the new orthography. This affects the choice of script, and in the case of alphabets, the choice of values assigned to letters. At times, this will raise problems where the language of wider communication is either English or French with their 'deep' orthographies, but linguists apply IPA values to letters to resolve some of those problems. Languages of wider communication with 'shallow' or 'intermediate' orthographies like Spanish, Portuguese and Bahasa Indonesian, provide much greater accessibility.

Many who speak minority languages become polyglots through education, trading and travel. Although they may be illiterate in their mother tongue, they may well be literate in the language of wider communication. Maximum transfer, therefore, is a most relevant criterion in such situations.

5 *Maximum ease of reproduction*

This was an important criterion in 1963 (Smalley) before the invention of computers. The typewriter did place a constraint on the development of an orthography as they were designed and manufactured primarily for European languages with established orthographies. Unconventional letters and diacritics were cumbersome and were therefore often ignored, as in the case of early Tera literature. The typewriter placed a constraint on the freedom of pen and paper.

Computers have certainly freed alphabet makers from such constraints where computers are available – their availability is thus now the only major constraint. Ease of typing/keyboarding, however, does remain a relevant criterion in terms of economy of effort. Sgall's scale of complexity applies here just as much as in the criterion of maximum representation of speech.

Maximum ease of reproduction for the compositor is relevant in ensuring also the maximum comfort of the consumer.

6 *Balancing the criteria*

Final decisions on a new orthography require a careful balancing of these criteria, some of which may be perceived as conflicting. It might be, for instance, the case

that a community wishes to be as distinct from a dominant culture as possible and yet feel the value of maximum transfer.

The very complexity of a phonological system itself might conflict with certain criteria. A relevant consideration in this respect is the notion of 'functional load'. Functional load refers to the number of systemic contrasts a linguistic unit makes in the language as a whole. This is partly a reflection of the relative frequency of a unit in a system, but also of the relative frequency of the structures in which the unit occurs. For instance, Gimson points out the significance of the relative frequency of contrasts in minimal pairs in English, ie the functional loads of contrasts. "By this measure the contrasts of /θ/ vs /ð/ and /ʃ/ vs /ʒ/ carry a very low functional load, with minimal pairs being almost non-existent" (see Cruttenden 2001: 217), but the frequency of the grammatical items with /ð/ in running text means that its own functional load is relatively high.

Seifart (2006) illustrates functional load by reference to word stress contrasts in English; although pairs of words like *(a) convert* and *(to) convert* are distinguished by the location of stress, the contrast is not marked in the orthography – they are homographs. However, the number of such pairs is relatively low in the total English lexicon; hence, the functional load of contrastive stress is accordingly low. He expresses the implication for orthography in this way: "Thus, while for the phonologist one minimal pair in a list of isolated words may be sufficient to identify a certain feature as contrastive, for the purpose of developing a practical orthography it is crucial additionally to evaluate the functional load of a potentially contrastive feature in connected texts. And if there are no, or only very few, instances where a

given feature (eg stress) in fact disambiguates utterances in a sufficiently large text corpus, then the need to represent the distinction is highly diminished" Seifart (2006: 280). Thus the criterion of maximum representation will need to be balanced against maximum ease of learning.

It will also need to be balanced against maximum ease of reproduction, particularly in the case of diacritics such as might be required with word stress. Both Seifart (2006: 292-4) and Grenoble & Whaley (2006: 149-51) draw attention to the question of the functional load of tone in the lexicon and grammar of tone languages. A major consideration is the additional processing burden on reader and writer that tone diacritics might bring without due reward. Each criterion is important and relevant – the socially and technologically oriented ones as much as the linguistic ones.

Barnwell (2004) summarizes the five criteria as follows, listing the linguistic principles first, as the general basis, then the technical for practical purposes, and finally the social obligations, as the source of final decisions. Barnwell's list follows the procedure in which a new orthography is created:

Accuracy: The writing system should reflect the sound system of the language, so that all the important sound differences are recognized and written in a distinctive way.

Consistency: The same sound should always be written in the same way. The same symbol always represents the same sound. There should be no 'silent' letters (unless they have a clearly defined function.)

Convenience: Any special symbols used should be easy to type and keyboard on a typewriter or computer.

Conformity: As much as possible, follow the writing system of the language of wider communication in the area. This will make it easier for people who can already read

in that language to read this language also. Also consider how other languages of the same language family or spoken in the same region are written.

Acceptability and Agreement:

It is important that the proposals are presented to interested leaders and others in the area for discussion so that agreement can be reached on how to write the language. It will take time and discussion to achieve consensus.

Finally, here is Coulmas's succinct summary of the requirements of a new orthography. It should be:

(1) based on a variety of the language which is acceptable to the majority of the speech community;
(2) easy to learn;
(3) easy to write;
(4) easy to read;
(5) founded on a phonemic analysis of the language while affording access to the morpho-phonemic and lexical levels;
(6) transcending the limitations of the sign inventory of the orthography of the respective major contact language as little as possible; and
(7) in as much agreement with the available printing technology as the internal consistency of the system and the requirement of indicating the basic repertoire of phonemes will permit.

(Coulmas 1989: 238)

Phonological competence

What do speakers of an unwritten language bring to the task of creating an orthography for their language? They have all their language – the systems in phonology, lexis, grammar and discourse – in their minds; this includes their total mental lexicon, including the phonological composition of each word. The mental lexicon is organized in a complex way, in semantic, syntactic and phonological networks (Aitchison 2003: 224-6). This means, among other things, that they know how to pronounce each of the thousands of words in their own lexicon – or, at least, practically each word, as there are often words on the periphery of their experience which they are unsure how to pronounce in a way that would satisfy fellow speakers, eg unusual, technical, archaic words, loan words from another language, etc.

They also know the components of their pronunciation systems, in the sense that they are fully competent, even if they are not aware of them objectively. They recognize what is acceptable and what is not, even if they cannot explain why that is so. They are expert in producing and perceiving their

1 consonant systems
2 vowel systems
3 phonotactic systems
4 syllable systems
5 prosodic systems and
6 morphophonological systems.

For instance, a speaker who uses a standard Southern England accent of English produces and perceives those consonants which belong to the word-initial system, those that belong to the word-final system and those that belong to word-medial

systems. They produce and perceive restrictions in the systems, for instance, /ð/ at the beginning of grammatical words, and /θ/ at the beginning of lexical items. They distinguish a vowel system for stressed syllables, and a different system for unstressed syllables; for instance, the former includes /ʌ/, but the latter excludes it, whereas the former excludes /ə/, but the latter includes it. They produce and perceive the phonotactic permutations that are allowed in their accent; for instance, they know that /str-/ is a permissible combination in word-initial position, but that /-str/ is not permissible in word-final position; and that only short vowels may precede word-final /-mp/ and /-ŋ/, but long vowels (and diphthongs) as well as short vowels may precede word-final /-nd/. They produce and perceive closed and open syllables, strong and weak syllables; they never produce word-final strong open syllables with short vowels, but they do produce and perceive them with long (and diphthongal) vowels, with up to three consonants in initial position in permissible combinations. They produce and perceive three degrees of stress in words, and know how to distinguish, for instance, compound words from identically worded noun groups, like *blackbird* and *black bird,* and *English teachers* as either teachers of English or teachers from England. And they produce and perceive inflections appropriately, such that *–ed* is either a full syllable, or a voiced or a voiceless consonant depending on the quality of the preceding sound.

These are all instances of phonological *competence*, *but* not necessarily of phonological *awareness*; the speakers operate these systems without necessarily being aware objectively of their existence. The distinction between *competence* and *awareness* can perhaps best be demonstrated by reference to abilities in the production of allophonic variations of a phoneme; the speaker of a standard Southern

England accent of English produces a ‘dark’ [ɬ] in word-final position and before /w/, without being aware of its articulatory difference to the ‘clear’ [l] in word-initial position and before /j/. They produce and perceive them both competently, without necessarily being aware that they do so.

They also know the simplification systems that operate in the colloquial styles of speech. Simplification systems occur in close sequences of words like noun, verb and prepositional phrases. They consist of processes like assimilation, dissimilation, elision, epenthesis and liaison; they vary from language to language, being part of the phonology of a given language. They are the means by which fluency is achieved in colloquial speech. Most speakers, however, are quite unaware of these adjustments to the pronunciation of words, but they manifest their competence by expert performance. Sometimes literate speakers ‘misspell’ words as a result, such as the relatively new word in English *input*, which is usually pronounced as /ɪmpʊt/, with /n/ assimilating to the bilabial position of the following /p/ as *imput*.

Speakers also know how to highlight relevant words as elements in clauses; in English this is manifested in phonological prominence being placed on full lexical items, at the ‘expense’ of grammatical items and lexical items with a low semantic contribution. The system is rhythm; it also works by grouping words together semantically through timing, stressing and pausing. Thus a relatively long, stressed *five* with a following pause can be distinguished from a relatively short *five* with no pausing in order to distinguish between two renderings of *five day old chicks*, as either ‘five chicks which are one day old’ or as ‘an unspecified number of chicks which are

five days old'. Again, speakers are highly competent in the production and perception of rhythm systems even if they are unaware of them.

Speakers are also competent in the phonological composition of clauses and clause parts in discourse. This system is one of the functions of intonation. By segmenting discourse into intonation units ('tonality'), speakers indicate their management of the units of information in their message; by highlighting a particular word in an intonation unit ('tonicity'), they indicate new and given information, broad and narrow focus, and any contrast in their message; and by choosing a pitch movement or level ('tone'), they indicate, on the one hand, the relative status of the particular piece of information in a given intonation unit as major, minor, or incomplete, or as implying additional but unspoken information which the addressee is expected to perceive; and, on the other hand, the communicative force of their message by indicating their intention to tell or ask something, to require or request action from the addressee, or to engage in social interaction.

They also know how to structure whole discourses in speech. They know how to group intonation units into phonological paragraphs, sometimes referred to as 'paratones' or 'pitch sequences'. This discourse structuring is another function of intonation and is typically manifested in relatively high registers of pitch at the beginning of a phonological paragraph and relatively low registers at its conclusion. Speakers use this 'textual' system to indicate the introduction of a new topic and its development before a new topic again is introduced in their management of discourse.

Finally, in the production and perception of discourse, speakers know the prosodic composition of different genres of discourse. British speakers, for instance, know what constitute basic elements in the telling of ghost stories in their culture, in contrast to other narratives; weather forecasts and news reading have distinctive prosodic compositions, and so have racing and football commentaries; preaching sounds different from praying, and so does comedy from tragedy, etc. Prosodic composition consists of relative proportions of tones and lengths of intonation units, relative variations of tones, speed, loudness, pitch range, rhythm and other voice qualities (eg 'breathy', 'husky', 'resonant', etc). These are all subconsciously controlled by the speaker in the production and perception of appropriate discourse, and as such are part of the speaker's phonological competence.

In these respects, phonology matches orthography in representing the whole range of linguistic units from words and morphemes to whole discourses. Yet phonological competence embraces other issues too:

1. *regional and social accents*. People vary in the extent to which they can recognize other accents, and reproduce them, but it is part of their phonological competence that they can do so to some extent.
2. *generational differences*. People can often detect incipient changes in the language, like the wider use of [ʔ] and new forms of intonation like the raised rising tone (otherwise known as the 'high rising terminal'). These are changes to the phonological systems that are characteristically different between the generations. Many may frown on such changes, but the fact that they detect them shows that they impinge on their phonological competence.

3. *stylistic variation*. People know how to sound formal and informal, stiff and casual in speech; this indicates their ability to manipulate phonology for stylistic effects. They also recognize variations in style and assess their appropriateness – all part of phonological competence.
4. *children's and learners' acquisition*. Allowance is made for children and learners acquiring a phonology, with varying degrees of toleration of differences between native speaker and 'interlanguage' phonology. Comparisons between fully acquired and developing phonology indicate another aspect of phonological competence.
5. *speech disorders*. Similar allowance is made for disorders in articulation and hearing; comparison and interpretation indicate yet another aspect of phonological competence.
6. *cultural values*. People may make assessments of
    a) euphony, the degree of pleasantness of a sound or a word or a text
    b) femininity, the degree to which the sound of a word suggests a girl's name rather than a boy's
    c) onomatopoeia, the degree to which a sound or word represents natural sound
    d) sound symbolism, the degree to which a sound or word suggests a quality, eg smallness
    e) rhyme, alliteration and other 'poetic' effects
7. *attitudinal expression*. People know how to exploit phonology to express a complete range of attitudes and emotions, eg vowel lengthening, consonant clipping, pitch broadening, etc. Such 'paralinguistic' features belong also to phonological competence. People know how to produce these effects and to interpret them, and this kind of variation is stored in our minds too.

The above 7 issues represent a much broader view of phonological competence than is usual in linguistic description, but they are real nevertheless in a speaker's mind. They are also, quite clearly, subject to individual variation. But for someone who is literate in a language that uses either an alphabet or a syllabary, there is another dimension to phonological competence, and that is the relationship between phonology and orthography.

A literate speaker recognizes correspondence rules of the phonemic values associated with symbols (Cook & Bassetti 2005: 6-7) and may well form judgements on the efficacy of orthography in its relationship to the alphabetic or syllabary representation of the pronunciation of words, and its efficacy in the punctuation of text. Such a speaker recognizes regularity and irregularity in the representation of sounds or syllables (see Sgall's scales of complexity and univocality above) and of words, in both reading and writing. The orthographical dimension to phonological competence enables a person to attempt to spell an unfamiliar word or name when heard and to pronounce an unfamiliar word or name when read. There is an expectation of optimal representation between phonology and orthography; this expectation lies behind what is often referred to as 'pop' (popular) or demotic spelling, such as *skool* for *school*, *hi* for *high*, etc. An expectation of optimal representation is certainly what a native speaker of an unwritten language will bring to the task of reproducing their language in writing.

In sum, a person stores in their mind the phonological components of both linguistic and communicative competence, which are

1 the phonological composition (‘pronunciation’) of every item in their mental lexicon

2 the phonological systems for the pronunciation of words:

    1 consonant systems

    2 vowel systems

    3 phonotactic systems

    4 syllable systems

    5 prosodic systems and

    6 morphophonological systems.

3 simplification systems in groups and phrases

4 rhythm in groups and phrases

5 intonation in clauses

6 phonological paragraphing

7 prosodic composition of discourse genres

8 regional and social accent variation

9 generational (‘historical’) variation

10 stylistic variation

11 developmental and interlanguage variation

12 disordered variation

13 cultural values: euphony, femininity, onomatopoeia, sound symbolism, poetic effects

14 paralinguistic variation

15 orthographic relationship (if literate)

Of all these components of competence, a native speaker of a minority language, engaged in the task of creating a new orthography for their unwritten language, will 'consult' components 1, 2, 8, 9 and 10 for the spelling of words, and components 4, 5 and 6 for punctuation and layout. Component 3 may well be involved too. Component 15 would apply if the person was literate in a relevant language of wider communication. How they 'consult' will be discussed below in terms of the methodology employed, but first, Component 15 will be presented as it relates to the case of the creation of a new orthography for Tera.

Orthography in the mind of Tera speakers

The Tera speakers involved in the 2004 workshop were four men, two older retired men (a former male nurse, and a town elder) and two younger men (one a teacher, the other a private secretary). They represented not only two generations, but also three dialects, but were also confident in being able to represent other dialect areas too. (In a subsequent workshop, three other men were involved, widening the dialect representation.) All the men were literate in Hausa and English; thus their experience of orthography reflected these two languages, which are both relevant languages of wider communication.

Hausa has a relatively 'shallow' alphabet. Its consonants can be represented by letters of the alphabet in a phonemic chart, based on Schuh & Yalwa (1999):

	bilabial	alveolar	post-alveolar	palatal	palatalized velar	Velar	labialized velar	glottal
Plosive & affricate	b	t d	c j		k g	k g	kw gw	'
Implosive & ejective	ɓ	ts ɗ		'y	ƙ	ƙ	ƙw	

Nasal	m	n						
Fricative	f	s z	sh					h
Tap Trill		r						
Approxim	w			y				
Lateral approxim		l						

It should be noted that Hausa orthography employs special 'hooked' letters for implosive and ejective consonants, that the apostrophe <'> is used for the glottal stop, that <k, g> have double functions, as does <r> for both a tap and a trill, and that a 'bigrapheme' <sh> represents /ʃ/. The five vowel letters <a, e, i, o, u> represent five vowel qualities roughly equivalent to their IPA values, but do not differentiate between long and short vowels. None of the tones (low, high and falling) are represented in any way.

English, on the other hand, has a 'deep' orthography, with a wide range of phonemic values for each letter and a wide range of letters for each phoneme. Letters from the 'English' alphabet that are immediately relevant for Tera are <p, v>.

The two older men could recall the alphabet used for the 1930 Tera Gospel of John, and all four were aware that revisions to the alphabet had been proposed since the early 1990s.

Methodology

Barnwell (2004) lists 20 steps in the procedure for creating an orthography for an unwritten language (see Appendix 1), which provided the basis for the creation of a new Tera orthography. The four participants brought all their phonological competence to bear, including their orthographic experience of Hausa and English.

A short story was recorded on audio cassette; this provides a stable basis for analysis and avoids arguments about what could and should have been said. The story was played back to ensure that all were happy with the content and then played back phrase by phrase, with the intention that each participant should independently attempt to write the words as best as they could, relying on their phonological competence in Tera and their current orthographical experience. One man was requested to write his version on a blackboard; others commented, agreed or disagreed and tentative decisions were taken. My role, as the 'facilitator', at this point was to keep a running record of the letters deployed with their phonetic values, and attempt to ensure consistency in representing sounds. My weakness was that I did not have the phonological knowledge that the participants had, and so had to check and double check on phonetic distinctions that I observed but that were ignored by them. For instance, I heard quite distinct vowel qualities [a, æ ə] for which they used only one letter <a>; later, of course, I realized that these are allophonic variations of the same vowel phoneme in Tera.

The first story comprising over 400 consecutive words yielded a mass of information. I drew up tentative vowel and consonant charts, explaining their design in terms of tongue position, place and manner of articulation, and voicing. Seeing the scheme of things and observing a certain degree of symmetry was indeed an astonishing and thrilling eye-opener for them, which added to their enormous appetite and enthusiasm for the task.

The next task was to draw on Component 1 of their phonological competence and solicit more words with the consonants discovered, in initial and final positions, and lists were drawn up. Words that began with a consonant plus [w, j] led to a search for other consonant cluster possibilities, which actually proved fruitless, as did the search for words beginning with vowels, apart from a couple of loan words from Hausa.

A second story was negotiated in the same way. Writing became ever easier as agreement was established on the letters to be used, but more consonants kept appearing until, of course, the complete inventory was reached. Their orthographic experience naturally was expanded.

One constant uncertainty was the determination of ends of words with what might become seen as inflections. This meant that attention had to be given to noun and verb morphology. What eventually emerged was that items representing number and definite and possessive reference in nouns were best regarded as inflections as they never stood alone and were always followed the noun closely; nothing else ever inserted itself between the noun stem and these morphemes. Verbs fell into classes depending on the form of their gerundive inflection, but personal pronouns were treated as separate words. Tenses were indicated also by separate words. As we drew up more and more lists of nouns and verbs with their morphological variations, confidence increased in the determination of word boundaries.

Another contentious issue was contrasting length in vowels. Some cases were relatively easy to determine, but others, particularly in final open syllables were very difficult.

Tone was a fourth contentious issue as a few otherwise identical words were distinguished solely by tone; tone appeared to be involved also with the issue of vowel length in final open syllables. We took a disyllabic word with what appeared to be the same tone on each syllable, and used that as a test mechanism for comparing the tone patterns of other words. A few words had a matching tone pattern, but many others did not. A second test word with a different pattern was taken for comparisons; some matched, others did not. Then a third, and a fourth, etc were taken until we were confident that all disyllabic words were accounted for. This procedure produced four pitch levels which were expressed as high, half high, half low, and low; this was reduced to three levels on phonological grounds, since half high could be reinterpreted as a mid level assimilating to a preceding or following high tone, and half low as mid assimilating similarly to low. This three level system was confirmed by establishing that only three levels were necessary for monosyllabic words. This analysis also resolved the case of possible long vowels in final open syllables; the relevant factor was discovered to be tone, rather than length, which meant that the contrast between long and short vowels was valid in closed syllables but was neutralized in open syllables.

The final inventory of consonants proved to be a highly complex system of 31 phonemes, with four cases of pre-nasalization, three of palatalization, and five of labialization:

	bilabial	labiodental	alveolar	postalveolar/ palatal	velar	labialvelar	glottal
Plosive	p b		t d		k g		
Affricate				tʃ dʒ			
Implosive	ɓ		ɗ		ɠ		

Prenasalized	m͡b		n͡d	ɲ͡dʒ	ŋ͡g		
Nasal	m		n	ɲ	ŋ		
Trill			r				
Fricative		f v	s z	ʃ ʒ	x ɣ		h
Lateral fricative			ɬ ɮ				
Lateral approximant			l				
Approximant				j		w	
Glottalized approximant				ʔj			

palatalization: the following labial consonants: ɓʲ, mʲ, vʲ
labialization: the following velar consonants: kʷ, gʷ, ŋ͡gʷ, xʷ, ɣʷ

The vowel system was relatively simple, consisting of six vowels, four of which had long equivalents in contrast, and four diphthongs:

i iː ɨ u uː
e eː o oː
a aː

eu oi
ai,au

Three lexical tones were identified: high, mid, low. Word divisions were also tentatively established.

Orthography decisions

The final recommendations from the workshop on the spelling of the consonants are represented in the following 'phoneme' chart, with the chosen letters in place of the IPA symbols:

	Bilabial	Labiodental	Alveolar	Postalveolar/ palatal	Velar	Labialvelar	Glottal
plosive	p b		t d		k g		
affricate				ch j			
implosive	ɓ		ɗ		q		
prenasalized	mb		nd	nj	ngg		
nasal	m		n	ny	ng		
trill			r				
fricative		f v	s z	sh zh	kh gh		h
lateral fricative			tl dl				

lateral approximant			l				
approximant				y		w	
glottalized approximant				ɗy			

For the vowels, the letter inventory was as follows:

i ii u u uu
e ee o oo
a aa

eu oi
ai,au

The criteria of maximum representation (transparency or accuracy) and maximum ease of learning (consistency) naturally led to choosing the ‘shallow’, almost phonemic, orthography of Hausa rather than the ‘deep’, lexicophonemic, orthography of English as the basis of their initial decisions. Maximum transfer (conformity) was also highly relevant as Hausa operates as the language of wider communication for a high proportion of the Tera population.

Only decisions beyond the letter values of Hausa and English need be discussed. Among the vowels, the spelling of long vowels was recommended as simply double letters; the participants decided to apply the criteria of accuracy and consistency to a level higher than that of the Hausa alphabet, which does not distinguish them. The values accorded to the letters otherwise follow Hausa more or less accurately; English values were understandably totally ignored! The letter chosen for the close central vowel was <u>, probably for two reasons: the first is historical, since the original 1930 alphabet used <ụ>, and secondly, the vowel is not distributed like front vowels. Subscript dots were abandoned in favour of underlining, as underlining was much

easier to type (consider Barnwell's principle of convenience, and Smalley's criterion of maximum ease of reproduction).

In respect of the consonants, several decisions were automatic: <b, ɓ, d, ɗ, f, g, h, j, k, l, m, n, r, s, sh, t, w, y, z> derive directly from the Hausa alphabet, and <p, v> from the English. <ch> also derives from English, but the decision was by no means automatic, since Hausa uses a simple <c>, which satisfies the principle of simplicity in the scale of complexity; to use <ch> involves a redundant letter <h>. It seems that when the participants had the opportunity of distancing their alphabet from Hausa, they took it, as a stand against a dominating culture; also, English was perceived, rightly or wrongly, as a 'progressive' language, associated with business and computing, and less of a threat to their separate identity.

Other decisions were uncontentious, if not automatic: <ny, nɡ> for /ɲ, ŋ/ respectively, and <mb, nd, nj> for prenasalized plosives. However, great debate ensued over the distinction between /ŋ/ and /ŋg/. It was clear that <ng> could be ambiguous, but a careful explanation of <ng> as a 'bigrapheme' for a single sound, and of <mb, nd, nj> as sequences representing prenasalization led to the decision to recommend <ngg> to represent prenasalized <ng>. This was declared to be the answer to what had been perceived as a very major problem. It is also a fine instance of phonological competence developing into phonological awareness.

Other decisions also required the development of phonological awareness. The very first sound of the first word of the first story was [ɣ], a sound that does not figure in either Hausa or English. <x> was suggested, but it was remembered that this letter

represented a different though related sound in the old Tera spelling system; <gh> was also suggested, from Jauro Maila's proposals and from the as yet unpublished Jideonwo (2004). [ɣ] had been spelt as <x̣> in the Tera Gospel of John (1930), but the participants showed an antipathy towards subscript dots and other diacritics, mainly because neither Hausa nor English uses them and because they were relatively cumbersome to type. Preference was given to letter combinations on the criteria of maximum transfer and maximum ease of reproduction (Barnwell's principles of conformity and convenience). The younger participants pressed for <gh>, and the older men relented. Jauro Maila later explained how he came to propose <gh>. He had noted the similar place of articulation for [g] and [ɣ] and the use that was made of <h> as a 'modifier' of a sound, particularly for a similar manner of articulation. Without knowing the technical terminology of phonetics, his phonological competence (components 2 and 15) developed a phonetic awareness as the basis of an orthography decision. The use of <h> for fricative modification explains the decisions to use <kh, zh> likewise for [x, ʒ], parallel to <sh> and eventually, <ch>. This use of <h> fulfils the criterion of maximum ease of learning (Barnwell: consistency).

Tera has two lateral fricatives, in addition to its lateral approximant. Interestingly, they were both spelt in 1930 as <ll> reflecting Welsh orthography, but without distinguishing them. Both Nyagham and Maila had suggested <tl, dl>, and the younger participants pressed for these. These choices display keen phonetic awareness: <l> is retained to express laterality; <t> is used to indicate voicelessness at the same place of articulation, <d> parallel voicing. There are no consonant sequences /t + l/ or /d + l/, even in medial position in Tera words, and so these combinations

cannot be misread. These choices indicate a superb development of phonological competence into phonological awareness.

The glottalized approximant [ʔj] was written variously as <dy, ɗy>, which puzzled me as an outsider, since there was clearly no alveolar contact implied by <d>. Hausa has a similar consonant, which is spelt as <'y>. The participants never considered this as an alternative, because there is no other need for <'> as there is in Hausa (for /ʔ/); perhaps also because this provided another opportunity to distance the spelling of their language from Hausa. But what lies behind the choice of some kind of <d>? The old Tera alphabet used <ḍy>. It seems that a historical development from a palatalized alveolar implosive [ɗʲ] to [ʔj] in Tera - losing the alveolar contact, but retaining some glottal (but 'ejective') action and a palatal tongue position – matches an identical development in Hausa (Schuh & Yalwa 1999: 92). Whereas Hausa uses <'y> because it needs <'> elsewhere in the spelling system, the Tera participants looked elsewhere in their own emerging alphabet and decided to use either <d> or <ɗ> with <y>. We experimented with <dy> as a simpler option, but <ɗy> eventually prevailed, since the <ɗ> preserved the representation of the manner of articulation.

Finally, the current implosive consonants in the language. Decisions on <ɓ, ɗ> were uncontentious as these exist in the Hausa alphabet, and local typewriters are adapted to include them. However, the representation of [ɠ] was the most contentious decision of all. Hausa does not have this implosive; its nearest equivalent, with an identical place of articulation but a different kind of glottal action ('ejective') is /ƙ/, and it is written as such. Nyagham and Maila favoured this, on the criteria of maximum transfer and maximum ease of reproduction, since Hausa typewriters are adapted for

this letter too. The old Tera *Gospel of John* had used <q> and Jideonwa used it too. Religion entered in to the argument too, as some felt that <q> was somehow an Arabic and, therefore, an Islamic letter. However, it was pointed out that <q> was, in fact, an 'English' letter, and had appeared in the (Christian) *Gospel of John*. In favour of <ƙ> were the criteria of transfer and ease of reproduction, but against it was the phonetic 'inaccuracy' of it. In favour of <q> was tradition and the opportunity to be different from Hausa, while recognizing that the Tera <q> value would not be equivalent to that of English <q>, nor even to that of Arabic <q> in transliteration. After much debating in the workshop and many debates in the community, <q> prevailed, on the strength of tradition, the issue of separate cultural identity, the 'inaccuracy' of <ƙ>, and, curiously, the design of the letter itself: a 'hook' seemed to be important to indicate implosivity, the letter <q> shared visual features with <g>, which represented the place of articulation, and the combination with a right-turning 'hook', albeit in low position, satisfied consideration of both place and manner of articulation!

Finally, in the case of tones, we decided that although tone played a significant role in the phonology of the word in Tera, there were not enough minimal pairs, and no minimal trios, to justify marking tone. In other words, the functional load of tones in minimal pairs did not justify its inclusion as a regular feature of the new orthography. A native speaker of Tera would know how to read a written word aloud in context and, in most cases, in isolation too.

Marking tone would establish maximum representation of speech but at the expense of maximum ease of learning and maximum ease of reproduction; it would have

added, in most cases, an additional, but superfluous, set of symbols to process. However, in one particular context, tone *is* recommended to be marked: to distinguish the verbal particle <á> (with high tone, 'present progressive') from <a> (with low tone, 'past tense').

Acceptability and agreement

All the recommendations were presented to two language committees, one set up by the local *Tera Forum* of Tera chiefs and their officials, and the other by local churches. Both committees were in broad agreement with the proposals, although as people have begun to write in Tera, it has become clear that one recommendation has been ignored: the diacritic <ˈ> to indicate that the preceding vowel is 'very short' has obviously been considered superfluous for reading as well as writing.

One member of the team has produced a mini wall chart of the Tera alphabet for use in homes and has published two booklets of Bible stories in the new alphabet. The beginnings of a dictionary have appeared, as has a manual for helping teachers to recognize and use the new orthography. A proposal has been submitted to a local government to introduce the development of literacy in primary schools. A hymnbook is under preparation and also a translation of the Gospel of Luke. It is also planned to revise the old 1930 'tentative' Gospel of John.

There is great enthusiasm in the local communities for the new orthography and a great desire to have more elementary literature published to enable not only the children but adults too to acquire literacy in their mother tongue. The project has raised the hopes of the people for a status of dignity within the wider political region

and has begun to fulfil their aspirations for a recognized separate social and cultural identity.

The project has demonstrated the worth of Smalley's criteria and Barnwell's principles. The new orthography enjoys maximum motivation for the community (= 'acceptability' and 'agreement') and a judicious balance of maximum representation of speech (= 'accuracy'), maximum ease of learning (= 'consistency'), maximum transfer (= 'conformity') and maximum ease of reproduction (= 'convenience'). It has been created with a methodology that

1 exploits native speaker phonological competence, particularly in word phonology for the development of the alphabet, and the phonology of grammar for punctuation;
2 develops phonological awareness;
3 exploits native speaker orthographical experience of languages of wider communication and develops it;
4 engages the native speaker community in a determinative role; and
5 engages the services of a phonologically aware facilitator.

In sum, the new Tera orthography has proved to be

(1) based on a variety of the language which is acceptable to the majority of the speech community;
(2) easy to learn;
(3) easy to write;
(4) easy to read;

(5) founded on a phonemic analysis of the language while affording access to the morpho-phonemic and lexical levels;

(6) transcending the limitations of the sign inventory of the orthography of the respective major contact language as little as possible; and

(7) in as much agreement with the available printing technology as the internal consistency of the system and the requirement of indicating the basic repertoire of phonemes will permit.

(Coulmas 1989: 238)

References

Aitchison, J (2003) *Words in the Mind* 3rd ed. Oxford: Blackwell

Barnwell, K (2004) *Developing a Writing system* (Appendix 1)

Cook, V & Bassetti, B (eds) (2005) *Second Language Writing Systems* Clevedon: Multilingual Matters

Coulmas, F (1989) *The Writing Systems of the World.* Oxford: Blackwell

Gippert, J, Himmelmann, N P & Mosel, U (eds) (2006) *Essentials of Language Documentation* Berlin: Mouton de Gruyter

Gordon, R.G. & Grimes, B.F. (eds) (2005) *Ethnologue: Languages of the World.* Dallas: SIL International

Grenoble, L A & Whaley, L J (2006) *Saving Languages* Cambridge: CUP

Igboanusi, H & Peter, L (2005) *Languages in Competition; the struggle for supremacy among Nigeria's major languages, English and Pidgin*. Frankfurt aM: Peter Lang

Jideonwo, A I (2004) *Let's Develop Nyimatli Language*. Lagos: Divine Peace

Concept Publication Ventures

Schuh, R G & Yalwa L D (1999) ‘Hausa’ in *Handbook of the International Phonetic Association.* Cambridge : Cambridge University Press

Sgall, P (1984) ‘Towards a theory of phonemic orthography’ in Luelsdorff, P A *Orthography and Phonology.* Amsterdam: Benjamins

Smalley, W A (1963) ‘How shall I write this language?’ in Smalley, W A *Orthography Studies* London: United Bible Societies

Tench, P (2007) ‘Illustrations of the IPA: Tera’ in *Journal of the International Phonetic Association 37/2*: 227-234

Winter, W (1983) ‘Tradition and innovation in alphabet making’ in Coulmas, F & Ehlich, K *Writing in Focus.* Berlin: Mouton

Appendix 1

Luke Partnership project: preparation phase
Preparation workshop, (three weeks)

Developing a writing system[1]

Plan to work in language groups, each group working with the guidance of a linguistically trained facilitator. Each group should consist of at least three mother-tongue speakers of the language concerned. In choosing participants for this process, consider those who may have potential to become the mother-tongue translators in the future. It is helpful if the facilitator has already had experience in studying related languages.

Before the workshop, identify any linguistic or cultural descriptions of the language, or of closely related languages and get copies of relevant materials.

Equipment needed:
Cassette recorder (with pause button and tape counter),
notebooks and pencils,
blackboard with chalk (some different coloured chalk is useful, or alternatively large sheets of paper (flip charts) and coloured felt-tip pens that can be fixed up on the wall with bluetac or sellotape),
computer for recording results and
a means to printout the drafts for review by the team.

Beginning to write

1. Have one of the mother tongue speakers of the language being studied tell a short story in the language. Record it on a cassette recorder.

2. All participating speakers of the language write down the text, writing it as they think is best. (The text should be replayed slowly, pausing and repeating each sentence several times.)

3. Then have one participant write the text on the blackboard. Discuss and compare the other written texts to see how much agreement there is in writing the sounds.

4. Begin to identify and list problem areas that need further investigation.

Finding the consonants and vowels and deciding how to write them

5. On the board, make a list of the consonant sounds that occur in the text. Give particular attention to sounds that do not occur in a familiar language of wider communication (e.g., Swahili, Hausa, Kriol) or which seems to be difficult.

6. Also list groups of sounds (consonants clusters, or consonant followed by w or y, or vowel clusters).

7. For each consonant or group of consonants, think of other words that contain the same sound at the beginning of the word - write the list on the board. All members of the group suggest additional words that have that sound to add to the list. If the sound never occurs at the beginning of a word, choose another position (e.g., first consonant in the word). Write lists for each of the positions in which the sound may occur. For example, as first consonant

---

[1] This summary is written as a guide for teams of mother-tongue speakers of a language working to develop a tentative writing system for their language, working with the guidance of a linguistically trained facilitator.

in the word, in the middle of a word, at the end of the word. Read through each list of words to check for the consistency of the sound and the way you have written it.

8. Members of the group write the lists in their notebooks. Write the words in columns, a new column for each sound. Leave space at the bottom of the column so that further words can be added.

9. Do the same for vowels. On the board, make a list of the vowels or cluster of vowels, on the board. Start with words that have the vowel sound at the beginning. If no vowels occur at the beginning of words, take vowels in another position (e.g., the vowel in words that have the pattern Consonant Vowel (CV) or Consonant Vowel Consonant (CVC) or Consonant Vowel Consonant Vowel (CVCV). Again, check for the consistency of the sounds and the way you have written them.

   These lists will later become the basis for a word list or mini-dictionary.

   Note examples of any words that differ only in one sound. (Examples of this are shape and sape, which differ only in the initial consonant.

   Also note examples of any words that differ only in the tone (pitch) of the voice.

10. List examples of common Consonant-Vowel word patterns (e.g., CV, CVC, CVCV, CVCVCV). Are certain patterns typical of (a) verbs or (b) nouns?

11. With the assistance of the facilitator, make a phonetic chart for the consonants and a chart of the vowels. Observe the pattern and symmetry of the sounds.

12. List any restrictions on the distribution of consonant and vowels. For example, do only certain consonants (such as p, t, k, m, n, ng, r and/or l) occur at the end of a word?

13 Keeping in mind the five principles of a good writing system (see below), discuss together and agree tentatively on how to write each sound. Again, pay particular attention to sounds that do not occur in the LWC.

Deciding where to divide words

14. Identify the pronouns of the language: (a) subject pronouns (b) object pronouns (c) possessive pronouns

Singular	Plural
1st person (I, me, my)	1st person (we, us, our) - inclusive/exclusive
2nd person (you, your singular, m/f)	2nd person (you, your plural, you dual)
3rd person (he, she, it)	3rd person (they, their)

15. Using a coloured pencil, underline any examples of these pronouns in your text. Did you write these as separate words or joined to a verb or another word? Discuss whether the pronouns should be written as separate words or joined to another word. Make a tentative decision. (This will need to be reviewed later.)

    Some guidelines that may indicate a pronoun should be written alone as a separate word: Can it be pronounced alone? Do other words come between the pronoun and the verb?

16. Explore how verb tenses are indicated.

He goes	"present simple"
He is going	"present continuous"
He will go	"future"
He went	"past"

Can you identify any verbal tense markers? Look for examples of these in your text. Discuss whether these should be written joined to the verb or as separate words. Make a VERY TENTATIVE decision.

17. Identify other questions to be explored concerning where words should be divided. Common questions are:

18. a. If there are nouns or other words with a reduplicated root, should the reduplicated parts be written (a) as one word (b) joined by a hyphen or (c) as separate words.

b. If there are compound nouns, should the parts be joined with a hyphen or written together as one word?

Deciding whether tone marks are needed

19. a. Look at any pairs of words that you have found that are different only in their tone. In many cases, the context in which these occur will help the reader to know which word is intended.

b. There may, however, be certain grammatical differences that are marked only by tone. Check carefully for the following:

1. Pronoun differences (e.g., 1st and 2nd person singular) marked only by difference in tone.
2. Singular plural differences marked only by tone
3. Verb tense differences marked only by tone
4. Negative meaning marked only by tone

In these cases, it will be necessary to mark one of the forms to indicate this difference. Discuss with you facilitator. Usually it is best to mark the least common form, leaving the common form unmarked. Usual ways to mark tone are as follows:

ò for low tone ó for high tone ô for falling tone

Writing up the tentative proposals

19. a. Review the text, making any revisions needed to fit with the tentative decisions made so far. Are all the group agreed on how to write? What differences of opinion are there?
b. If possible, keyboard the text and make printouts, so that all members of the group can read it clearly and study it carefully.
c. Record another text and have each member of the group write it out. (Have one person write it on the board.) Discuss places where different speakers have written the text differently.
d. Prepare a summary of the writing system you have agreed, with examples. Note alternative possibilities. (Booklets in other languages may serve as a model.)
e. Discuss and compare any previous writing systems used for writing the language.
f. Prepare a "back-translation" (gloss) for this text in a language of wider communication (LWC).

Prepare to share the tentative proposals you have made more widely, for discussion with others

20. Make a list of community and church leaders to whom you will want to present these tentative proposals for writing the language, for further discussion and decision. Make plans on how to present these proposals and how to process them until agreement is reached.

The five principles of a good writing system

Accuracy: The writing system should reflect the sound system of the language, so that all the important sound differences are recognized and written in a distinctive way.

Consistency: The same sound should always be written in the same way. The same symbol always represents the same sound. There should be no "silent" letters (unless they have a clearly defined function.)

Convenience: Any special symbols used should be easy to type and keyboard on a typewriter or computer

Conformity: As much as possible, follow the writing system of the language of wider communication in the area. This will make it easier for people who can already read in that language to read this language also. Also consider how other languages of the same language family or spoken in the same region are written.

Acceptability and Agreement: It is important that the proposals be presented to interested leaders and others in the area for discussion so that agreement can be reached on how to write the language. It will take time and discussion to achieve consensus.

Katherine Barnwell, Gombe, June 2004