FACTORS IN DESIGNING EFFECTIVE ORTHOGRAPHIES FOR UNWRITTEN LANGUAGES

Michael Cahill
Elke Karan

SIL International
2008

SIL Electronic Working Papers 2008-001, February 2008

Contents

ABSTRACT

Recent interest in preserving endangered languages has led to a corresponding interest in orthographies for such languages. This paper, based on SIL's decades-long corporate experience as well as literature studies, summarizes the major factors which must be considered when planning an orthography. Issues of acceptability and usability must be balanced with the obvious prerequisite of linguistic soundness. The linguistic issues include the choice of graphemes ("letters of the alphabet"), decisions on word breaks, level of phonological representation, representation of suprasegmentals such as tone, etc. However, sociolinguistic and other political factors (e.g. government policies, varied dialects, language attitudes, and the influence of other orthographies) often take precedence over a linguistically ideal orthography. Finally, the practical issue of local font availability must be addressed. The development of an effective orthography is thus of greater complexity than commonly conceived.

1. Introduction[1]

An orthography is a system for representing a language in written form. It encompasses more than the symbols that represent the sounds (often called characters, letters or graphemes). An orthography also covers relative placement of these symbols, word breaks, punctuation, diacritics, capitalization, hyphenation and other aspects which might be regulated in a written standard. As Coulmas (1996:379) notes, "all orthographies are language specific."

There has been a trend of increased attention to orthographies in recent years as shown by a great increase in major publications on the topic.[2] Several factors account for this, most importantly the increased awareness of endangered languages, and the positive effects literature and literacy in these languages can have in preserving them. The attention to universal human rights (such as the *Education for All* movement) and the Unicode movement have also been influential.

Not just any orthography will do; it needs to be effective. That is, it needs to be (a) linguistically sound, (b) acceptable to all stakeholders, (c) teachable, and (d) easy to reproduce. These roughly can be thought of as scientific, political, educational, and technical aspects. These four criteria often conflict with each other.

SIL has been working for several decades in literacy and has published many works related to orthography.[3] This paper is not based on the authors' own experience and research alone, but takes into consideration SIL's extensive corporate experience, and questions raised by others on the issue.

We have severely limited the topics covered in this paper. Orthography design is only one part of language development or revitalization programs. Attention needs to be given to other aspects as well. For instance, an effective literacy program in a language community must also deal with developing motivational, pedagogical, and post-literacy materials as well as assure teacher and author training, effective instruction and distribution of literature.

This paper very briefly outlines the most crucial factors that must be considered in devising an orthography for an unwritten language. For further details, the reader is referred to the works cited in the References. In Section 2, we discuss governmental policies. In Section 3, we discuss the linguistic foundation for orthographies. In Section 4, we discuss various educational factors. In Section 5, we discuss sociolinguistic factors, especially what can be termed "local politics." In Section 6, we examine technical issues such as font availability and compatibility. In Section 7 we discuss orthography testing. We conclude in Section 8 with comments on community involvement and the benefits of literacy to marginalized groups.

---

[1] A prelliminary version of this paper has benefited from comments from audiences at Rice University and the University of Florida. We also appreciate input from Lynn Landover, Pete Unseth, and Keith Snider. The authors are listed alphabetically.

[2] This includes works such as Taylor and Olson 1995, Daniels and Bright 1996, Tabouret-Keller et al 1997, Coulmas 1996 and 2003, Rogers 2005, Anderson 2005, Joshi and Aaron 2006, as well as a new journal *Written Language and Literacy* published since 1998. See Karan 2006 for a more in-depth treatment than is in this paper.

[3] Examples of some of the published works related to orthography are Pike 1947, Powlison 1968, Gudschinsky 1973, Simons 1977, Litteral 1999, and SIL's *Notes on Literacy* 1965–2001.

We use the normal conventions of enclosing a symbol in forward slashes /a/ for phonemic representation, in brackets [a] for phonetic representation, and angled brackets <a> for graphemic representation – the actual spelling system.

2. Governmental Policies and Restrictions

Sometimes the national government has regulatory policies that one must follow. These generally restrict the options available in orthography decisions. (For example, some countries regulate script choice.) Although linguistic analysis should lay a foundation for an orthography, governmental policies, as the law of the land, trump other factors. It is no use proposing an orthography that would, in effect, be illegal.

For example, for some years in Ghana, the Bureau of Ghana Languages, the relevant regulatory agency, did not permit tone markings in orthographies. Whether one thought tone needed to be marked or not, it was not an option. There were long-term effects of this policy even though it was eventually changed. By the time the change was made, several languages had strong literacy programs with considerable amounts of literature, so a change in orthography would have been difficult.

Another example comes from Cameroon where, in 1979, the government established a unified alphabet for Roman-based writing systems (Tadadjeu and Sadembouo 1984). At present, new orthographies must conform to this standard, which regulates symbol choices. Interestingly, if a language community wants to use a non-Roman script for its orthography, such as an Arabic-based one, there are no current policies regulating symbol choices.

Sometimes the national government needs to approve individual orthographies, not just have an orthography conform to a standard. In the Central African Republic, for example, the Applied Linguistics Institute, a branch of the University, must give its approval. In Ethiopia, one of several authorized agencies must approve any new orthography. Since SIL often has a contract with governmental agencies, it must pay close attention to such policies and regulations. Sometimes policies are set in stone; sometimes they are flexible and serve only as suggested guidelines. Researching this and doing the appropriate public relations work with the agencies who make the decisions will be time well spent.

3. Linguistic Factors

3.1 Phonological analysis

As a general principle, sounds which contrast in a language should be represented with different symbols. That is, if separate *phonemes* are established, such as /r, l/ based on a contrasting word set such as ‘river, liver’, these need separate *graphemes* <r, l> in the orthography.

For decades, phonemic representation was promoted as the orthographic ideal by SIL linguists such as Pike (1947) and Gudschinsky (1973). Pike’s book *Phonemics* had as its subtitle “A Technique for Reducing Languages to Writing”. In those days, and still for the most part today, phonemic analysis and orthography design were closely linked. “One symbol per phoneme, one phoneme per symbol” was the rule. Writing systems tended to

symbolize what was "heard," i.e. represented surface sounds which had been shown to be phonemic.

But more recently, writing is regarded as more than representing speech and sounds. As we learn more about phonological systems and consider non-linguistic factors, we agree with others that "The best orthography is not necessarily the one which uses one symbol for every element which is in phonemic contrast" (Bauernschmidt 1980), as we detail below.

Phonemic analysis of a language is still foundational to orthography design, but it is only the beginning. There are at least two, and possibly more levels of representation which need to be considered. Working with multitudes of languages in all their complexities, it has seemed useful from a pragmatic point of view to consider models such as Lexical Phonology, and look at the output of the lexical level of phonology. This corresponds to the level of psychological awareness for native speakers of a language (Mohanan 1986).

Different types of phonological processes result in given surface forms. An outsider perceives the surface form (what has been called the "etic" view). However, native speakers may be aware of some phonological processes but not others, so they may only be aware of the lexical form. Clearly, orthography should be based on native-speaker perception, not on that of foreign linguists.

3.2 Levels of representation for words in context

Words are often pronounced differently in isolation than in natural speech flow. Those making orthography decisions need to consider how to symbolize these words – as they occur in isolation or as they are pronounced in context.

At one extreme is the phonetic representation, with allophones people are unaware of. An example of this would be English [p^{h}ul, spul] being spelled as <phool, spool>. This goes against the idea above of representing only the contrastive elements, the phonemes, of a language. The phones [p^{h}, p] do not contrast, so there need not be two symbols.

A more reasonable proposal is what might be called "surface phonemic," which is variants people are aware of. For example, consider the following Kɔnni data.

(1) Kɔnni:	[yí̲	bà]	[yú̲	wà]
	give	them	give	him

Both /ɪ/ and /ʊ/ are independent phonemes in Kɔnni; the alternation for *give* above is phonologically conditioned. A strictly phonemic orthography would write *give* in two different ways. However, for spelling, the better option is to keep an invariant <yɪ> in both cases. Keeping the form of a word constant when an underlying phoneme changes to another phoneme is what has been termed a "lexical" representation. Another example comes from Daai Chin of Myanmar:

(2) Daai Chin (Hartmann-So and Thomas 1981)

ʔi**m̲** 'house' + ʔsì 'floor' → [ʔi**p̲**ʔsì] 'floor of house'

Should the compound at right [ʔɨ**p**ʔsì] be *spelled* with <m> or <p>? The authors spelled it with <m>, and reported that the people read it fluently, pronouncing it with [p].

The relevant principle is that it is usually helpful to write at this morphophonemic representation level. Keeping a constant "shape" for a word facilitates reading for non-beginning readers, which is the level of readership we want to aim at.

3.3 Word divisions

Besides the issues of grapheme choice and level of representation, the question of where to insert word breaks needs to be dealt with. Are two morphemes to be attached to each other or written separately? There are a variety of cases where this type of decision must be made. These include compounds, clitics, pronouns, and prepositions, among others. A basic though overly simplistic criterion for "word-hood" is "Can that item stand alone as an utterance?" However there are other linguistic tests for determining whether a morpheme should be attached to another or not. Sometimes these tests do not all point in the same direction. For example, let us consider possessive pronouns in Kɔnni.

(3) Kɔnni possessive pronouns

bè-tígíŋ 'their house' bà-sáˡáŋ 'their porridge'

Phonologically, there is an "advanced tongue root" (ATR)[4] vowel harmony system in Kɔnni. In the above example, the ATR value of the noun controls the ATR of the possessive pronoun, thus the varying [bè/bà] pronunciations of the pronoun above. Tone rules also show the dependency of the pronoun on the noun. However, syntactically, full nouns, which definitely need to be written separately, can also fill the position filled by the pronoun, e.g. *hɔ̀wwá sáˡáŋ* 'woman's porridge.' With these potentially conflicting criteria, the determining factor was the choice of the local people themselves who preferred the pronouns to be separate: <be tigiŋ, ba saaŋ>. (See Van Dyken and Kutsch Lojenga 1993 for a detailed discussion on word divisions.)

3.4 Other issues

All the above discussion has focused on consonants and vowels as the relevant sounds that need to be considered and symbolized. However, suprasegmentals such as tone, nasality, and vowel harmony also need to be taken into consideration – they have the same types of issues. Stories abound, for example, of an orthography not marking tone, and a reader having to scan a sentence several times in order to think of a tone melody that will make the sentence make sense.

Finally, we must mention the factor of "native speaker intuition" in analysis. Native speakers of a language often have definite ideas of how their language works. There is usually some grain of truth to what they say, so the analyst should pay attention to their comments. However, caution is called for – their analysis may not be what is actually happening. For example, the first author was showing a preliminary alphabet

[4] In ATR harmony systems, the vowels are divided into two sets based on the tongue root position. A common system in many African languages is the +ATR vowels are /i, e, o, u/ and the –ATR vowels are /ɪ, ɛ, a, ɔ, ʊ/. Words usually have only vowels from one set, not mixing them within a word and sometimes even within a phrase.

book to a Kɔnni speaker who had learned to read as a soldier in World War II. The man looked at the word written <kooŋ> and insisted it was actually <kuuŋ>. But his pronunciation did *not* sound like [u]. Further investigation revealed that the phoneme in question was actually /ʊ/, which in Kɔnni sounds very much like [o] to American ears. The man did not have a mental category for /ʊ/, though. He knew it was not /o/, and his closest mental category, from English spelling, was <u>. Thus there was indeed an issue that the man identified, but his solution was not correct.

4. Educational Factors

Smalley (1964) and Malone (2004) present lists of criteria for an adequate writing system. Two of the criteria relate to education: (1) learnability and (2) ease of transfer to other languages. These constitute the pedagogical ideal. We will examine these and discuss what would enhance or undermine these ideals.

4.1 Ease of learning

Learnability relates to the inherent ease or difficulty of a system. The teaching and learning of reading and writing become more complicated when there is a mismatch between the spoken and written language. The degree of abstraction inherent in an orthography is referred to as *orthographic depth* in the literature. Shallow orthographies – those with reliable sound-symbol representation are more easily learned than deep orthographies.[5] Deep orthographies result from underrepresentation, overrepresentation (often referred to as overdifferentiation), inconsistencies in sound-symbol correspondences, silent letters, and loan words retaining unadapted spellings. Deep orthographies bring with them a cost to learners, teachers and education providers. Learners experience a higher level of frustration, resulting in possible demotivation or failure and increased learning time. For the educational system this represents increased investments of time, people resources and funds.

4.2 Underrepresentation and overrepresentation

Underrepresentation results when an orthography uses fewer graphemes than there are phonemes. For example, Akan in Ghana has nine vowel phonemes, but only represents seven in the orthography. The phonemes /e/ and /ɪ/ are both spelled with <e>. So when readers see <e>, they must decide which sound to pronounce.

Underrepresentation is never ideal from a linguistic or pedagogical point of view, but there are often sociolinguistic pressures for underrepresentation. There may be pressure from the national language, which does not provide symbols for sounds found in the local language, or perhaps the local speakers do not want additional symbols for a variety of reasons.

Underrepresentation is often a problem for reading because the reader may not be sure which sound to pronounce when he sees a given symbol. It is as if English combined the phonemes /i/ and /ɪ/ with one symbol <i>. When the reader sees a sentence "she <bit> him," there is confusion over whether she *beat* or *bit* him. Note that the difficulty is on the part of the reader. The writer merrily writes the single grapheme for both phonemes.

[5] The "shallow vs. deep" opposition is alternately referred to as "transparent vs. opaque" in recent literature (e.g. papers in Joshi and Aaron 2006).

Sometimes misanalysis is at the root of underrepresentation. An example comes from Ebrie (Côte d'Ivoire). The original orthography didn't distinguish fortis and lenis /p, t, c, k/, with the result that readers had serious problems reading. The investigators had just missed the distinction in their original analysis. More linguistic analysis revealed the pattern, and Ebrie speakers immediately approved adding <ph, th, ch, kh> to their orthography as the fortis consonant graphemes.

Tone marking deserves a special comment here. Many orthographies of tonal languages do not mark tone – a classic case of underrepresentation. Sometimes this is a result of inertia (perhaps other nearby languages do not mark tone) or a reluctance to delve into the details of analysis (Cahill 2001). Considering solely the function of tone in establishing contrast would lead to the conclusion that almost all tonal languages should indicate tone in the orthography. However, it is sometimes true that tonal contrasts can be sufficiently distinguished by context, especially by advanced readers. In these cases it may be defensible to omit tone marking from the orthography. This should be tested, not just assumed.

*Over*representation, on the other hand, is using *more* graphemes than there are phonemes. There are two different types: using different graphemes for a single phoneme, such as <f, ph> for /f/, or else, assigning allophones of a single phoneme to separate symbols. Again, from a linguistic and pedagogical point of view this is not ideal, but there are sociolinguistic reasons why this is sometimes done. Perhaps some of the dialects of the language do contrast the two sounds, while the others do not. Or perhaps a trade language or official language has contrast for the sounds and thus separate symbols, and thus the local people want both symbols for that reason.

An example is <r> in Gur languages of northern Ghana. The facts of all the languages are similar. It only occurs intervocalically, and <d> occurs word-initially. <r> is considered a phoneme in some languages but not others (Cahill 1999). However, all Gur languages we are aware of use <r> in the orthography. One main reason is its use in English, and for transferability to English it is included.

With overrepresentation, we have the opposite situation for reading and writing from underrepresentation. If a language has the extra symbols of overrepresentation, this is generally not a problem for the reader, assuming good teaching. The reader sees the symbols and learns to pronounce them the same way. However, overrepresentation *is* a problem for writers. When the writer is writing a word, he must choose between two symbols for the same psychological sound (phoneme), and the potential for choosing the wrong one is significant, unless there is a simple rule to follow.

4.3 Transfer to other languages

People often want to learn to read additional languages as well as their own. National governments are also often keen to incorporate local people into the life of the nation, which may require knowledge of a national or regional language. With this in mind, a goal for an orthography is to aim for symbols and conventions which will facilitate transfer to other languages that people are likely to want to read also. The premise is that harmonizing writing systems will minimize the effort required to transfer reading and writing skills from one language to another. Pressure toward harmonization with the language of wider communication sometimes results in under- or overrepresentation, and the value of transfer to the official language is often felt to be

worth the sacrifice. It is important to keep in mind that underrepresentation of segments is more serious a flaw than overrepresentation (Smalley 1964) and should be resisted.

Those involved in orthography decisions need to be aware of current trends. Identity and language preservation issues now often take precedence over transfer considerations, tipping the balance toward conventions which differ from prestige languages (see 5.1).

4.4 Visual appearance and the reading process

Mattingly (1992) points out that for meaning to be retrieved cognitively, some constraints in the writing system are necessary. He notes that symbols need to be "visually discriminable." This would suggest that symbols which look too much alike or crowd adjoining letters could hinder reading ease. (Consider n ɳ ɥ ɲ ŋ.) Research has shown that mirror –image symbols are not ideal for beginning readers (Pederson 2003). Overuse of a letter as part of various digraphs can also reduce reading ease (e.g. bh, dh, ph, th, sh, ph, gh, kh). Superimposing more than one diacritic is technically possible; however we do not recommend this. If information is pertinent, it ought to be marked in a salient manner. Small fonts and certain printing processes severely decrease the distinctiveness of superimposed diacritics. Baker (1997:101, 136) states that "the general view remains that no segment should bear more than one superimposed diacritic." Some languages, such as Vietnamese, make heavy use of such notation (e.g. <điển, Đấng>), but this should not be considered "best practice."

4.5 Other considerations

We have alluded to the idea that the needs of beginning and fluent readers are different when it comes to orthographies. Experienced readers read in chunks and make use of context; decoding a word may only be necessary when an unknown or uncommon word is encountered. For beginners, decoding letter by letter or symbol by symbol is much more common. An orthography that caters to beginning readers might have more of a surface representation. If the pronunciation changes due to context, the temptation might be to alter the spelling as well. This is not advisable. Two goals for reading instruction are fluency and reading for meaning. This is enhanced when words retain a constant spelling. Mattingly (1992) includes *constancy* in the representation of words as a constraint for writing systems; Venezky (2004:146) also refers to the *constancy principle*. This is not a new concept: Nida (1964:25) also referred to the "principle of unity of visual impression." This needs to be taken into consideration when deciding the spelling of words with a sound or tone change conditioned by context.

5. Sociolinguistic Factors

It has been said that all orthographies are political, and indeed the influence of the outside world, internal conflicts, as well as other social factors, can end up being the determining factors in an orthography.

5.1 Effect of other languages – attitudes and transfer/identity

Multilingualism is a natural part of the social environment for most languages of the world, and few language communities exist in isolation. Thus there are quite often influences from national or official languages, or neighboring languages. People's attitudes toward these languages may be positive or negative, and the depth of their interaction with these languages may be rare or intensive.

Local people may want an orthography to "look like" a major or national language – or *not* look like a rival one. For example, in Ghana, the Konkombas use <ln> word-finally to represent a nasalized lateral sound (e.g. <Likpakpkaln>). It is not a common sound, and the spelling is distinctively Konkomba. The neighboring and linguistically close Kombas also have this sound. However, they want it spelled <nl>, not <ln>, as a distinctive mark of identity – they are *not* Konkombas, but Kombas!

An extreme example of what can be called "hyperfragmentation" is found in Oaxaca (Mexico), where each town has its own distinctiveness, and leads to each wanting to have a different orthography. This has nothing to do with language intelligibility, but the identity of individual towns. B. Hollenbach (personal communication) tells of two Trique towns which shared a single dialect but still wanted two orthographies. In one case in East Africa, two churches with the same dialect had two differing orthographies (Keith Snider, personal communication).

5.2 Dialects

Most languages have distinct dialects, which may differ from town to town or even clan to clan. These often systematically differ in phonology. The challenge is to devise an orthography that will serve all of them. For example, where the main dialect of Kɔnni has a phoneme /h/, the village of Nangurima has /ŋ/. The orthography uses <h>, and the people of Nangurima read it as [ŋ].

It is desirable to have a single unified orthography. Whether this can be accomplished depends on the desire of the speakers of the speech varieties themselves. If they wish to be "unified" under one writing system, the most prestigious village/dialect might be chosen as a reference dialect. This is often referred to as a *unilectal approach*: one dialect serves as the basis for the written form; the others make a mental adjustment while reading and writing. Another approach is a *multi-lectal approach* where several dialects are accommodated via consideration of the various varieties and more neutral symbols are chosen. Thus one group is not clearly favored over the others.

The topic of dialect standardization involves more details than we can cover in this paper.

5.3 More politics

As mentioned above, all orthographies are political, but some are more obviously political than others. Sometimes orthographies become attached to a certain group, whether political, clan, religious, or other group. ("You're an *evangelico*, I'm a Catholic; so I support this one.") These types of affiliations must be considered when proposing or revising an orthography.

Introducing a change to an existing orthography can be even more complicated than coming up with an acceptable system for the first time (Karan 2006). One needs to

consider not only the linguistic factors, but also who designed, supports, and uses the current system(s).

6. Practical production factors (fonts)

A major challenge for orthographies in past years was the practical problem of special symbols for adapted Roman alphabets. Symbol options depended on typewriter or local printer capabilities. Rather than using a non-Roman symbol such as <ʊ>, people resorted to modifying typewritten symbols by underlining, strikethroughs, and so on. In the 1980s, both authors were thrilled to be able to order "custom keys" such as ŋ, ɛ, ɔ on their new manual typewriters, replacing some more standard symbols.

Currently, with the near-universal use of computers, and direct printing from them, there are not many limitations on what symbols can be used. Many fonts are available, including several from SIL. So there is no technical reason not to use the most linguistically suitable symbol for an orthography if it is acceptable to the local population. Local availability of printers, however, may still be a factor.

Unicode is a standard for encoding the languages of the world. Before it existed, *custom-encoded fonts* met the need for special characters in language projects and were gratefully received and used by field workers. However, these fonts have some serious limitations, and their use is now strongly discouraged.

The advent and spread of Unicode fonts is an increasing factor in the world, particularly if texts and books are to be shared between people via email or the internet. Basically, using a Unicode-compatible font instead of custom-encoded fonts is highly recommended to permit archiving and transmission of information without loss. Quite a few Unicode fonts are commercially available and come with computer operating systems. Some of these fonts include more Roman-based special symbol options than others, and some include many more non-Roman script options.

SIL has also developed several Roman-based Unicode-compliant fonts, including Doulos SIL, Charis SIL, and Andika. The first two include the entire International Phonetics Alphabet (IPA). The newest development is the Andika font, especially designed for early literacy material. Besides these fonts based on Roman scripts, there are a number of fonts that handle non-Roman scripts, including Arabic.

For free SIL fonts, see *http://www.sil.org/computing/catalog/index.asp#fonts* and *http://scripts.sil.org*.

7. Orthography Testing

We have referred to testing an orthography. This can be informal, observing qualitatively how readers respond to different orthography options. Or it can take a more formal and scientific approach. Informal testing can be used to check on acceptability and learnability of a specific system. Formal testing is usually used to discover the preferable one from among symbolization options. This type of testing might involve actual audio or videotaping of a reader, and then from the tape, taking measurements of reading speed, positions of hesitations, etc. Even in early stages, with non-fluent readers, testing can be done with a preliminary vocabulary list or elementary dictionary for individual words.

Testing the orthography to show what really works is essential. So is documenting the decision-making process and all observations and research results.

8. Discussion and further points

One of the implications of much of the above discussion is that to have an orthography that is acceptable to the local community, the involvement of that community in developing the orthography is crucial. It is true that in some remote communities with little prior education this involvement may be minimal, but the community will be much more likely to use an orthography if they have been actively involved in the decision making process.

It would be a mistake to say that a good orthography is all that is needed for effective reading. Even if a good orthography is in place, adequate pedagogical extensions, that is, good teaching materials and instruction, are also needed, for all but the most self-motivated potential readers.

A final significant topic is why people should promote literacy at all in languages and cultures that do not have a written form of the language. The accusation has been made that literacy is an imposition of Western culture, and changes local cultures in damaging ways.

To answer this, we first note that writing did not begin in the West; Asia in particular has a very long tradition of writing. China, India, and several parts of Southeast Asia had functioning writing systems in place long before Europe was literate. Also, many non-literate cultures of Southeast Asia have a tradition of having "lost" their books (e.g. Hmong, noted in Eira 1998:188). These people groups are often eager to regain what they perceive as lost ground.

Moreover, the benefits of literacy are many. Some of these are listed below.

First, it provides tools to deal with the larger world, which is unavoidably entering most of the isolated language regions. Besides reading, arithmetic is often taught, giving the mathematical tools to deal with money (often a new factor introduced by outside traders).[6] Literacy can equip local people to deal with attempted land grabs from outsiders, as in the case of a South American language group who had a city man appear with a very official-looking document. The man told them it was a deed to their land. One local who had been taught to read his own language and had then bridged into reading Spanish as well, examined it and pronounced it a very attractive Singer sewing machine warrantee! Also, computers and the internet are starting to appear in surprisingly unlikely places, and of course some degree of literacy is essential to use these.

Second, literacy provides access to both outside and local cultural materials. These can be topics related to health, such as water treatment, AIDS, avian flu, nutrition, etc. They can be government documents, such as the UN Declaration of Human Rights, which has been translated into local languages. Literacy also provides the possibility of preservation, retention, and communication of almost-forgotten folk tales and other local lore. Often elders of a group are concerned that the younger generation is not learning the old stories and oral histories and literacy offers a way to preserve them for future generations.

[6] Of outside influences, the introduction of *money* can be the biggest single disruption to a culture.

Third, a language community's own cultural identity can be strengthened and they can have a higher view of their own language once it is written. Accounts abound of how majority language speakers taunt minorities with statements like "That noise you're making isn't even a language. It's just animal noise." Literacy – a published dictionary and so forth – can help dispel this image that even the local speakers may have bought into.

For languages that are endangered, on the brink of dying out altogether, literacy can help strengthen that language. UNESCO (2003) identifies nine factors that affect language vitality, and the availability of "materials for language education and literacy" is one of these.

For women especially, literacy can provide increased self-esteem and empowerment.

Sometimes literacy can help people in ways that we in the literature-soaked West would not predict. The first author interviewed Kɔnni speakers on what they found valuable about literacy. Besides the value of reading the Bible (sometimes underestimated as a motivation for learning to read), they mentioned two things. One was that their letters could now be private. Before, if they wanted to send a letter to a relative or friend, they had to find a literate person to write it for them, and there was no privacy in such a system! Second, when they travel to other places by public transport, they do not have to embarrass themselves by asking what town they are coming to – they can now read the signs themselves.

In spite of all the benefits, literacy should not be forced on a group which does not want it. Outsiders may legitimately point out these benefits, but the choice is up to the people themselves.

To sum up, there are many factors involved in devising or adapting an orthography. Linguistic factors are basic, but many other issues make conflicting demands as well. All these must be considered and balanced for an orthography to be effective. The whole process is more complex than is commonly realized.

> Without literacy, our language was in the process of being exterminated...He who loses his mother tongue is just a slave to him who is of the lowest class...But now, even if I die today, I will die happy, because my children have a language which will endure and that they can call their own. (Josué Koné, Miniyanka speaker, Mali)

References

Anderson, Deborah. 2005. Recommendations for creating new orthographies. Unicode Technical Note 19. Unicode Inc. http://www.unicode.org/notes/tn19/ (accessed January 19, 2007).

Baker, Philip. 1997. Developing ways of writing vernaculars: Problems and solutions in a historical perspective. In: Andrée Tabouret-Keller, Robert B. Le Page, Penelope Gardner-Chloros, and Gabrielle Varro (eds.), Vernacular Literacy: A Re-Evaluation. Oxford: Clarendon Press, 93–141.

Bauernschmidt, Amy. 1980. The ideal orthography. Notes on Literacy 32:12–21.

Cahill, Michael. 1999. Aspects of the morphology and phonology of Kɔnni. Ph.D. dissertation, Ohio State University.

Cahill, Michael. 2001. Avoiding tone marks: A remnant of English education? Notes on Literacy 27(1):13–21.

Coulmas, Florian. 1996. The Blackwell encyclopedia of writing systems. Oxford: Blackwell.

Coulmas, Florian. 2003. Writing systems: An introduction to their linguistic analysis. Cambridge: Cambridge University Press.

Daniels, Peter, and William Bright (eds.). 1996. The world's writing systems. New York: Oxford University Press.

Eira, Christina. 1998. Authority and discourse: Towards a model of orthography selection. Written Language and Literacy 1(2):171–224.

Gudschinsky, Sarah. 1973. A manual of literacy for preliterate peoples. Ukarumpa, PNG: SIL.

Hartmann-So, Helga, and David Thomas. 1981. Morphemic writing in Daai Chin. Notes on Literacy 36, 30–31.

Joshi, R. Malatesha, and P.G. Aaron (eds.). 2006. Handbook of orthography and literacy. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Karan, Elke 2006. Writing system development and reform: A process. M.A. Thesis, University of North Dakota.

Litteral, Robert. 1999. Four decades of language policy in Papua New Guinea: The move towards the vernacular. SIL Electronic Working Papers 1999–001. First presented to the annual meeting of the Linguistic Society of Papua New Guinea, Madang 1995 http://www.sil.org/silewp/1999/001/silewp1999-001.html (accessed January 19, 2007)

Malone, Susan. 2004. Manual for developing literacy and adult education programs in minority language communities. Bangkok: UNESCO. http://unesdoc.unesco.org/images/0013/001351/135164e.pdf (accessed June 16, 2006).

Mattingly, Ignatius G. 1992. Linguistic awareness and orthographic form. In: Ram Frost and Leonard Katz (eds.), Orthography, phonology, morphology, and meaning. Advances in Psychology 94. Amsterdam: North-Holland, 11–26.

Mohanan, K.P. 1986. The theory of lexical phonology. Studies in natural language and linguistic theory. Dordrecht: Reidel Publishing Company.

Nida, Eugene. 1964. Practical limitations to a phonemic alphabet. Orthography studies: Articles on new writing systems. Helps for Translators 6. London: United Bible Societies, 22–30.

Pederson, Eric, 2003. Mirror-image discrimination among nonliterate, monoliterate and biliterate Tamil subjects. Written Language and Literacy 6(1):71–91

Pike, Kenneth L. 1947. Phonemics: A technique for reducing languages to writing. Ann Arbor: University of Michigan Press.

Powlison, Paul S. 1968. Bases for formulating an efficient orthography. The Bible Translator 19(2):74–91.

Rogers, Henry. 2005. Writing systems: A linguistic approach. Oxford and Maine: Blackwell Publishing.

SIL International. 1965–2001. Notes on literacy. Available in LinguaLinks Library.

SIL International. 2002. LinguaLinks Library. Version 5.0. Dallas.

Simons, Gary F. 1977. Principles of multidialectal orthography design. In: R. Loving and Gary Simons (eds.), Workpapers in Papua New Guinea Languages 21. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics, 325–342.

Smalley, William. 1964. How shall I write this language? Orthography studies: Articles on new writing systems. Helps for Translators 6. London: United Bible Societies, 31–52.

Tadadjeu, Maurice, and Etienne Sadembouo. 1984. General alphabet of Cameroon languages. Bilingual version. Propelca 1. Yaoundé: University of Yaoundé.

Tabouret-Keller, Andrée, Robert B. Le Page, Penelope Gardner-Chloros, and Gabrielle Varro (eds.), 1997. Vernacular literacy: A re-evaluation. Oxford: Clarendon Press.

Taylor, Insup and David R. Olson (eds.). 1995. Scripts and literacy: Reading and learning to read alphabets, Syllabaries and Characters. Dordrecht, Netherlands: Kluwer Academic Publishers.

UNESCO.2003. http://www.unesco.org/culture/heritage/intangible/meetings/paris_march2003.shtml (accessed Jan. 30, 2007).

Van Dyken, Julia R., and Constance Kutsch Lojenga. 1993. Word boundaries: Key factors in orthography development. In: Rhonda L. Hartell (ed.), Alphabets of Africa. Dakar: UNESCO-Dakar Regional Office and Summer Institute of Linguistics, 3–20.

Venezky, Richard L. 2004. In search of the perfect orthography. Written Language and Literacy 7(2):139–163.