Dictionaries and Orthography in Modern Africa

M. Lynne Murphy, *Department of Linguistics, University of the Witwatersrand, Johannesburg, South Africa*

Abstract: This paper considers the issues and options in creating orthographies for languages without alphabetic traditions. In particular, it looks at issues that face linguists and anthropologists in creating such orthographies for preliminary descriptive work on a language. It is often the case that these orthographies by linguistic outsiders later serve as the bases for literacy-aimed orthographies. Thus, the lexicographer is constrained not only to present the language accurately, but to not undermine indigenous literacy projects for that language. This paper looks at the particular case of devising an orthography for Beng, a Southern Mande language of Côte d'Ivoire, in order to illustrate the problems inherent in alphabetising an unfamiliar language.

Keywords: AFRICAN LANGUAGES, ALPHABETISATION, BENG, CÔTE D'IVOIRE, DICTIONARY, FIELDWORK, INTERNATIONAL PHONETIC ALPHABET, LEXICOGRAPHY, LITERACY, MANDE, ORTHOGRAPHY

Opsomming: Woordeboeke en ortografie in Moderne Afrika. In hierdie artikel word die kwessies en opsies in die skepping van ortografieë vir tale sonder alfabetiese tradisies oorweeg. Daar word veral gekyk na kwessies waarmee linguiste en antropoloë te doen kry by die skepping van sulke ortografieë vir voorlopige deskriptiewe werk aan 'n taal. Dit gebeur dikwels dat hierdie ortografieë deur linguistiese buitestaanders as basis dien vir ortografieë wat op geletterdheid gerig is. Gevolglik word die leksikograaf gedwing om nie net die taal akkuraat weer te gee nie, maar ook om nie inheemse geletterdheidsprojekte vir daardie taal te ondermyn nie. Hierdie artikel ondersoek die spesifieke geval van die ontwerp van 'n ortografie vir Beng, 'n Suidelike Mande-taal van die Ivoorkus, om probleme inherent aan die alfabetisering van 'n onbekende taal te illustreer.

Sleutelwoorde: AFRIKATALE, ALFABETISERING, BENG, IVOORKUS, WOORDEBOEK, VELDWERK, INTERNASIONALE FONETIESE ALFABET, LEKSIKOGRAFIE, GELETTERDHEID, MANDE, ORTOGRAFIE

1. Introduction

This paper considers contemporary methods for assigning alphabets and orthographies to languages without an alphabetic tradition. Issues relevant to orthographic assignment are surveyed with particular attention to issues for African languages. The ultimate goal of the paper is to assess the possible ramifications of such assignments for lexicographers in the modern African context.

Lexikos 6 (AFRILEX-reeks/series 6: 1996): 44-70

These ramifications are discussed with reference to the orthography for Beng (a minority language of Côte d'Ivoire) created by American scholars for a Beng-English ethnographic dictionary.

Throughout history, conquering peoples have set the languages of their victims to print in writing systems mimicking the conqueror's own. This continues to be true in the present day, when former colonising and imperialist cultures exert orthographic influence (by means of missionaries and academics) on the same peoples they once conquered. European (including Euro-American) missionaries have frequently pursued alphabetisation of the local languages of former colonies, but academics, especially linguists and anthropologists, and governments have also played rôles in developing orthographies. In the present day, national governments and international organisations (such as UNICEF) are acting upon an increased interest in promoting literacy in developing nations. This interest presses for further development of orthographies for languages with no writing customs, and the application of existing European alphabets to these languages continues for both pragmatic and political reasons.

For outsiders involved in the development of orthographies and literacy, several ethical dilemmas present themselves. By introducing writing to a non-writing culture, members of a writing culture seem to make an implicit judgement: that the non-writing culture is inadequate and must be repaired through the introduction of writing. Thus, the introduction of writing to a culture may simply be a means of proselytising European (or Arabic or Chinese) values and culture. However, the introduction of writing by former colonial forces may instead be viewed as partial restitution for the injustices of colonialism: a) by allowing the victims of colonialism to share in the economic system that has benefited the colonial powers, and b) by helping to preserve languages whose livelihoods have been threatened by the intrusion of a colonialist language and the economic and political changes brought forth by colonialism.

Promoting literacy and non-traditional (i.e., Western-style) education produces more ethical and practical problems. One can make many good arguments for promoting literacy in local language. Vernacular literacy allows more people to attain higher levels of education and may preserve languages that would otherwise die out due to the predominance of official (often European) and national languages in the educational and economic system.[1] But literacy can also be detrimental to traditional village life. As young people become literate, especially as they become literate in national languages, they often become dissatisfied with traditional life in agricultural villages and leave to find work in cities.[2] Whether the cities of developing nations can support an influx of new literate people is definitely questionable.

As part of a literacy program, orthographic development should involve native speakers of the language in order to preserve the language users' rights of self-determination in linguistic and educational matters, as well as for practical reasons. (Literacy efforts are facilitated if the indigenous population sup-

ports the chosen writing system.) However, often some record of the language is made for the benefit of outsiders (e.g. academics) prior to mass literacy efforts. This paper aims to discern how academic orthographies affect future attempts at literacy-aimed orthographies and whether these effects may be harmful to attempts at literacy.

This paper consists of six sections. The remainder of this section explains the terms and typographical conventions used. Section 2 briefly raises questions concerning the nature of writing systems. The third section concerns orthographic design as it applies to previously unwritten languages, especially those of Africa, discussing first the selection of graph style, then orthographic choice. Section 4 outlines other orthographic issues: word division and capitalisation conventions. Section 5 concerns the treatment of the Beng language of Côte d'Ivoire in an ethnographic dictionary published in the United States. The final section discusses the rôle of dictionaries in establishing a practical orthography.

1.1 Terminology and conventions

In this paper, the term GRAPH refers to a minimal unit of writing which by itself represents some part of a language. For example, the 26 letters in English orthography are graphs, which may represent sounds. In the Chinese writing system, graphs semantically represent morphemes. ALPHABET refers to a set of graphs that represent sounds, though not necessarily the same sounds in every language that it serves. A SYLLABARY is a set of graphs that represent syllables, usually consonant-vowel combinations. DIACRITICS are symbols used in conjunction with alphabets or syllabaries in order to make more precise sound differentiations than the graphs alone. They differ from graphs in that they cannot stand alone in writing — they must affix to some graph. Accent marks are well-known examples of diacritics. An ORTHOGRAPHY is the writing system of a language, consisting of a set of graphs and conventions regarding the relation of those graphs to sounds or words, the division of words, capitalisation, *et cetera*. Although languages may share almost identical alphabets, their orthographies may differ. For example, according to the orthography of English, pan indicates the pronunciation [p^{h}æn], while in French orthography, pan ('bread') indicates the pronunciation [pã]. The term ALPHABETISATION refers to the process by which a language is assigned an alphabet and orthography. This should not be confused with the French term *alphabetisation* (used often in the materials cited here), which refers to the process of literacy education. Nor should it be confused with the term's use elsewhere to mean 'alphabetical ordering'.

In keeping with linguistic tradition, phonetic representations appear in brackets and linguistic examples are underscored. Exemplary graphs are represented in boldface.

2. The nature of writing

While the languages of most former European colonies (with the exception of some Asian colonies) have been considered by Westerners to be unwritten, some of these languages do have historical writing systems that have been ignored or deemed unsatisfactory by the colonial and post-colonial powers-that-be. Defining writing system is in itself a problem. The writing systems of Europe have almost one-to-one correspondence between spoken and printed words. Yet other systems, such as Chinese, do not practise such phonetic / orthographic consistency. In Chinese, the character represents a concept that is realised with different sounds in the different Chinese languages. Certainly, other cultures have means of representing ideas in concrete symbols. These symbols might be incorporated into the design of objects, in the patterns on clothing, in house decoration, or in paintings and sculpture. They might be used to record historic events, to revere heroes or deities, or to record debts or crop information. Are these writing systems? If we were to view such symbolic systems as writing, would our views toward the introduction of European-style writing systems have to change? In a perfect world, perhaps they would. But in a world made smaller through imperialism and subsequent world-wide economic dependencies, it may be too idealistic to suggest that orthographers should accept local symbolic systems as writing systems comparable to those of Europe and Asia. The writing systems discussed in this paper conform to Western standards; however, the reader should keep in mind the possible inadequacy of this standard.

3. Common goals of orthographic design

In designing or choosing an alphabet and developing an orthography for a language, many practical and political issues must be considered. However, these practical and political concerns may point to conflicting orthographic choices. In this section, I consider first the choice of graph style (roman, indigenous, original, etc.) and then consider the development of orthographies using the selected graph style, focusing on the roman graphs which are most often chosen for African languages.

3.1 Graph style selection

Most orthographies developed in Africa and other former European colonies have been based on roman graphs. While this choice of graphs is rarely questioned or challenged in orthographic literature, this selection raises many practical and ethical questions. Export of these western European graphs to colonies around the world smacks of cultural and linguistic imperialism. In fact, many

Christian missionaries have seen roman alphabetisation of local languages as the means by which to introduce their religion to the peoples of Africa and America. Richard Lepsius, in the advertisement for his own international orthography, insisted that, "in the case of Africa, [use of a standard system based on roman graphs is necessary, for] in this way only can we hope for Evangelization of that vast continent" (1863: v). This sentiment is echoed today in the actions of the Summer Institute of Linguistics, a Protestant missionary organisation that develops literacy in vernacular languages in order to translate and distribute the Bible and other Christian texts.

3.1.1 Indigenous graphs

In the European fervour to "evangelise" or "civilise" colonised peoples, European orthographies have sometimes supplanted indigenous orthographies. In western Africa, indigenous orthographies were developed for several languages prior to or separate from European influence. At the turn of this century in Cameroon, King Njoya of the Bamum developed an ideographic script of 466 characters which was used in writing the history of the Bamum and for royal correspondence. In Sierre Leone and Liberia, syllabaries for five Mande languages were developed by native speakers between 1830 and 1940. Each has been used for correspondence and record-keeping, and one (Vai) has been standardised and used in translations of the Bible and Koran. These systems were developed without overt influence from non-African cultures, although some similarities have been found between the Mande syllabaries and a secret writing system used to transcribe Arabic by the Hodh of Mauritania (Dalby 1986). This similarity of writing systems may indicate that a common writing system was used by the medieval Mandiŋ empire, which spread the linguistic ancestor of the Mande languages (Dalby 1986). Thus, these cultures, often described as having unwritten languages, may have centuries-old traditions of writing.

For languages with extant or historical writing systems, the introduction of a new orthography would seem redundant. Furthermore, replacement of indigenous orthographies by imported systems is blatant cultural imperialism. But despite the fact that syllabaries existed for several African languages when missionaries contacted their speakers, the westerners considered it necessary to introduce roman writing systems. This was not entirely due to western ignorance of the indigenous systems. In his advertisement for his standard alphabet, Lepsius (1863) mentions the Vai syllabary as evidence that African people are adept at teaching each other writing systems, ignoring the point that, for the Vai, his alphabet is redundant.

Dalby (1986 — one of the few orthographic histories that discusses native West African scripts) gives no indication that any of these indigenous writing systems flourish. Because mass literacy was not established with the invention

of the Bamum or Mande systems, they may be destined for extinction. Few people already know the indigenous systems, but many indigenous people know the writing system of the language used in government or church-run schools, often the official language or a regional *lingua franca*. Widespread use of the official writing system hinders the survival of the indigenous system because of the time, expense, and commitment necessary for teaching and maintaining two systems. So, while vernacular literacy programs could preserve and even promote indigenous writing systems and related cultural traditions, these efforts instead ensure that the indigenous writing systems are abandoned.

3.1.2 Entrenchment of roman graphs

The persistence of European official languages entrenches roman graphs in former European colonies. As well as preventing the revival of extant indigenous systems, this prevents the introduction of original graph systems from within the language community. For members of the language community to create their own orthography from scratch would be the utmost in self-determination in the promotion of literacy. However, such an option is barely considered because of the practical matters of printing technology and managing language policy in multilingual states.

In African nations with a European official language,[3] the use of roman graphs for indigenous languages allows for the use of the same printing equipment to produce materials in more than one language. Thus, the reasons for the choice of roman-based systems are usually economic in nature: a) printing equipment for roman graphs is readily available, and b) use of the same graphs for local and official languages avoids the economic and mental burdens of maintaining separate equipment and educational efforts.

In this age of printed communication, a primary goal of orthographic design is that the graphs be readily convertible to type. In the First World, where computer-based printing technology is very accessible, invention of new graphs poses few problems for the publishing industry and increasingly fewer problems for those who have abandoned typewriters for word processors. However, in developing nations, such technologies (and the skills for using them) are rarely available or affordable. Because new writing systems are often developed as parts of national literacy campaigns, they must be printable immediately, for use in primers and later in other educational, political, medical, and agricultural materials. For these reasons, officials often seek to utilise extant alphabets for the new orthographies, so that existing equipment may be used in printing.

However, in practice, application of the roman alphabet to non-Indo-European languages is often difficult and necessitates added expense. Because languages vary in their inventories of consonants and vowels, assigning

African significances to roman graphs is not a simple matter. Even if the sounds of the language number fewer than 26 (the number of roman graphs used in English), a basic principle of orthographic practice bars any assignment of graphs to sounds that fundamentally differs from their assignment in other languages (for example, assigning the vowel sound [ɔ] to the letter **b** is *verboten*).[4] For this reason, non-roman graphs are often introduced to a basically roman system. Often, these graphs are modified roman characters, such as **ŋ** or **ɛ**, or are borrowed from other alphabets, such as **θ** and **ð**. Once non-roman characters are introduced, the adoption of a roman-style system is no bargain. New typewriters and printing equipment must be designed to service the language.

Although new printing equipment must be manufactured, economic efficiency and political expedience have led many African nations to adopt official alphabets that include roman and non-roman graphs. While the modified roman alphabets involve an initial expense, cost-efficiency can be attained by applying the same graphs to all indigenous languages. For example, the Nigerian National Language Centre has consolidated the writing systems of over 250 indigenous languages and introduced a pan-Nigerian typewriter which contains all of the letters of the roman alphabet except x as well as nine modified roman characters and a number of diacritics (Dalby 1986: 28). While adoption of any writing system involves expense, this tactic dilutes the expense by applying the same printing equipment to many languages and thus creating a large market for the new equipment, allowing it to be mass-produced.

3.1.3 Conclusion

Modern, state-supported orthographies for African languages are most commonly based on a roman-style graph system. This holds true despite the existence in some areas of indigenous writing systems. Harmony among the writing systems of various indigenous languages and (official) European languages is usually encouraged. Proponents of this method cite economic reasons for the choice of roman graphs as well as pedagogical reasons (discussed further below). While the involvement of roman script in non-Indo-European languages is problematic, both from orthographic and political / ethical standpoints, it is now so deeply entrenched in African political and educational systems that alternatives, such as invention of new alphabets, no longer seem viable.[5]

3.2 Phoneme / graph assignment

Once the roman style of graphs has been selected for a particular language, planners must determine how roman graphs will represent the sounds of the language in question. Two methods are most popular: (1) mimicry of the ortho-

graphic system of another language, and (2) adoption of a standardised orthography which has been created in order to represent the sounds of many languages. The borrowed orthography or standard orthography is usually used as a model for a new orthography, rather than adopted without changes. Thus, the systems are tailored to fit the sounds of the languages they are intended to serve. Orthographies that adequately represent the sounds of the language are often adapted through use so that the characters are easier to produce or to read or so that less phonetic detail is represented. Standard orthographies have come to surpass the orthographies of European languages in popularity, but both are discussed below.

Orthographic theorists generally agree on several criteria for an ideal orthography (see Kemp 1981: 16*, IPA 1957: 1-2, Mann and Dalby 1986: 207, Winter 1983). Foremost is the "One Sound / One Sign" principle[6] which holds that for any phoneme (minimally distinctive sound unit), one and only one graph should consistently represent that phoneme and no other. This principle is found in every list of orthographic principles in the literature cited here and is usually the first principle mentioned, as well as the most discussed. Notably, this principle is not followed by most languages with a long history of writing, such as English or French, because natural, diachronic sound change has little regard for spelling. However, One Sound / One Sign is the goal of most current orthographies and is claimed to benefit both the language scholar, who demands consistency and precision in phonetic description, and the native speaker, who is attempting literacy (the more simple and elegant the spelling system, it is reasoned, the easier the acquisition of reading skills). The adherence to One Sound / One Sign is further discussed below with respect to various orthographic systems.

3.2.1 Borrowed orthographies

In the earliest language-recording efforts by European colonists and missionaries, the orthography of the European's native language (or some other European language) was most often used for transcription. This tendency can be attributed to the lack of international standard orthographies and to simple ethnocentrism. These early orthographies persist in some areas where the original colonial forces have remained and where continual efforts have been made to record or promote literacy in the local language. (Note, for example, the French spelling of [u] as **ou** in Ouagadougou.) Such orthographic borrowing may be instigated or perpetuated today by those who believe that borrowing holds some pedagogical value.

For many individuals, literacy in a local language is seen as a stepping stone to mastery of a European language (usually the official language of the nation) and thus to greater economic mobility. Modelling the local language's orthography after that of a European language is claimed to aid in the develop-

ment of bilingual literacy. However, such orthographic mimicry in the name of educational principles may be misguided. As Yapita Moya (1981) points out for the Bolivian language Aymara, the adoption of Spanish orthography confuses the Aymaran learner of Spanish as well as the Spanish-speaker learning Aymara, as the sound systems of the two languages are different enough to cause problems for orthographic sharing. Aymara has three vowels, while Spanish has five. Yet in three historical orthographies of Aymara modelled on Spanish orthography, all five of the Spanish vowel graphs are utilised, violating the One Sound / One Sign principle, causing confusion to native speakers who cannot discern a logic to the assignment of vowel graphs to sounds, and causing many learners of Aymara to believe that five vowels actually exist in Aymara (and that they are stupid for not being able to distinguish them).

Yapita Moya argues that the use of Spanish orthography for Aymara promotes a feeling of intellectual inferiority among the Aymara people. Because their language is so easily read by Spanish speakers (whether or not the Spanish speaker can understand the content of the text), Aymarans assume that their language is "simple" compared to the Spanish language, with which they often have difficulty. Here, we can find not only abstract philosophical reasons to oppose the culturally imperialistic imposition of European orthographies on non-European languages, but also evidence that such imperialism is damaging to educational efforts, to the emotional well-being of the individuals of a culture, and to their cultural identity.

3.2.2 Standard / international orthographies

Standard roman orthographies for non-European languages gained popularity among European scholars during the 19th century. The aim of these systems is to serve more than one dialect, if not all languages, by applying graphs to sounds universally. The One Sound / One Sign principle is intended to hold not only within one language, but across many. If a sign **b** represents a bi-labial, voiced stop in language A, then it serves the same purpose in language B. If there is no bi-labial, voiced stop in language C, then the sign **b** should not be evident in the writing of that language. The efficiency of this system is affected by the differences and similarities among the languages that are served, the consistency of the application of graphs to sounds, and the number and range of graphs and diacritics included in the system.

The colonial history of Africa has resulted in the creation of many competing orthographies. Orthographies of individual African languages were modelled on many different European orthographies because different European nations had colonies in Africa, and because the colonies of any particular European nation were not necessarily contiguous and did not honour language or language-group boundaries. Thus, the name of one language family might be spelled three different ways in three different colonial territories, for exam-

ple, Bantoe, Bantou, Bantu (Wharton 1925: 76). The shift to standardised alphabets did not appreciably alter this situation because of the number of such alphabets that have been proposed and, to some extent, applied. Wharton (1925) mentions 146 transcription schemes that had been proposed from the 17th through the 19th centuries, and he bemoans the strife caused to librarians by the variant transcriptions of a single language name, for example, chi-Suahili, kiswahili, Swaheli (p. 78).

The twentieth century has seen progress in the standardisation and harmonisation of writing systems through the introduction of the International Phonetic Alphabet (IPA) in 1888 by the International Phonetic Association (Association Phonétique Internationale). The IPA is illustrated below in Figure 1.

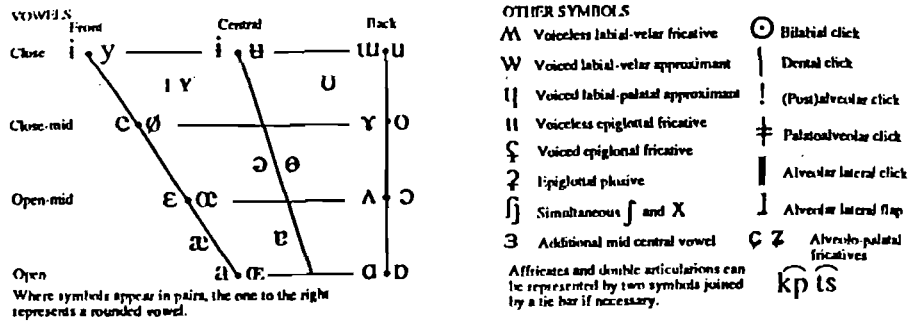
Figure 1. International Phonetic Alphabet (reprinted from Ladefoged 1990: 551)

CONSONANTS

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			
Ejective stop	pʼ			tʼ		ʈʼ	cʼ	kʼ	qʼ		
Implosive	ɓ ɓ			ɗ ɗ			ʄ ʃ	ƙ ɠ	ʠ ɢ		

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Figure 1 (*continued*)



This alphabet has had several advantages over previous alphabets such as Richard Lepsius' Standard Alphabet (1863) or Alexander Melville Bell's Visible Speech (1870). First, the IPA is a co-operative effort by an international association, unlike the previous efforts of individuals. Thus, from its inception, the IPA has had broader-based support than earlier efforts. In addition, perhaps because of the collectivity of the effort, perfection or completion has never been claimed of the IPA. Rather, it is a perpetual work-in-progress, subject to changes as new phonetic and phonological facts are learned or as problems in its practical application are discovered (see, e.g., International Phonetic Association 1989). The IPA is in this way unlike the previous works of individuals, which were often advertised as the solution to all the problems of transcription and transliteration, and which regularly fell out of fashion when they failed to adapt to new tasks.

The IPA has also benefited from the timing of its introduction. During the end of the 19th century, linguistics began to establish itself as a scientific discipline. The IPA was introduced to an academic community which had already begun to be convinced of the value of linguistic study.

The IPA allows for transcription at various levels of phonetic accuracy. This feature makes the IPA appropriate for many different orthographic tasks and audiences. Diacritics and other symbols marking non-contrastive phonetic features may be employed by linguists studying a particular language, but may be left out in the transcription of texts. Markers of features such as tone may be employed at various stages in literacy training, but may be left out by more experienced writers where the context is sufficient to make clear the relevant features. Thus, the IPA is useful as a tool of linguistic study, but may be applied such that non-linguist users are not burdened with an abundance of phonetic information.

While these features of the IPA have allowed it more success than previous alphabets, they also prevent it from being a truly One Sound / One Sign system. Among users of "the IPA", there is much variation among symbols,

such as ɨ - ɪ, ʃ - š and ʧ (tš) - č. Because most languages have few distinctive liquids, r is often used, for typographical ease, to represent whatever liquid exists in a particular language, regardless of whether the phoneme is better described by ʀ, ʁ, or ɹ. Furthermore, some phonemes present in African languages, such as prenasalised consonants ([mb], [ŋg]) and labio-velar stops ([gb], [kp]), do not have single-graph representations in the IPA.

If the IPA is used as a basis for a literacy orthography, these orthographic / phonetic inconsistencies are not problematic, despite the failure of the One Sound / One Sign principle. As a universal standard alphabet, the variation of sign-symbol correspondence is only problematic for a particular language if variant forms are used within that one language. Such random assignments of graphs are rarely included in the orthography of a single language, unless competing orthographies are designed (in which case, one form will eventually prevail). Furthermore, universality in graph assignments does not necessarily benefit the languages in which literacy is being established. For example, Winter (1983) argues that differences in orthographic convention for long vowels in two related American languages, Walapai and Havasupai, are helpful to the speakers of the languages (and perhaps to linguists as well) because they visually differentiate the two languages (which are otherwise very similar), helping to preserve the strong linguistic pride and identification which exists in both cultures despite the near-identity of the two languages in linguistic terms.

Other breaches of principle in applying the IPA, such as the use of r to represent different sounds in different languages, are beneficial (or at least, not harmful) to the language community for reasons of economy and ease in printing. The most problematic breach of One Sound / One Sign is the use of digraphs (combinations of two graphs) to represent one sound in the labio-velar stops of many African languages. The linear order of the two signs is misleading: **gb** is not a series of sounds, [g] followed by [b], but instead it is one stop with double articulation (i.e., simultaneous closure at two points of articulation). The IPA's failure to manage a single symbol for a single labio-velar phoneme may reflect the Indo-European bias of the International Phonetic Association. Having spent its efforts and graphs on the languages more familiar to its members, it fails to allow West African languages one sign for each sound. Because these African languages are not apt to have consonant clusters which would be orthographically similar to the double-articulated consonants, the use of digraphs should not be any more problematic to the native speaker of Beng than the use of digraphs such as **ch** is to native speakers of English, French, or German. However, the use of digraphs such as **gb**, **kp**, and **ŋm** is misleading to speakers of West African languages who are already familiar with the roman alphabet. Because a European alphabet is involved here, learners and speakers may tend to treat Europeans as the experts on the language, as the ones who know the "right" way to read, and this may under-

mine native speakers' confidence in their own language, language use, and intelligence, as was the case for Aymara.

3.2.3 Recent developments in Africa

In Africa, the desire for standard alphabets has grown among language and education policy-makers in order to simplify the educational needs of multilingual nations. Within languages and countries, and across the continent, standard alphabets have been proposed and legislated as official alphabets. In Tanzania, the alphabet of Kiswahili developed in pre-independence days was retained after Kiswahili was declared the official language of the nation. This system retains many orthographic anglicisms, including digraphs such as **ng** ([ŋ]) and **th** ([θ]) (Dalby 1986). In Central African Republic, an official orthography for the *lingua franca* Sango has been developed with a reduced roman alphabet, 22 letters and no digraphs.

Alphabets for multilingual use have been developed and adopted officially in several West African nations. In Ghana, a cross-lingual script has been developed which introduces non-roman graphs and avoids the need for digraphs (Dalby 1986). Other nations, such as Benin (Tchitchi and Hazoumé 1983), Côte d'Ivoire (Kokora 1979), and Mali (Dalby 1986), have adopted alphabets involving both digraphs and non-roman graphs. However, at least in the case of Côte d'Ivoire, the official alphabet has had little effect on languages for which orthographies had previously been established (except in certain official efforts to utilise the official alphabet), thus failing to harmonise the diverse orthographic systems of the nation (Alma Gottlieb, p.c.). Furthermore, the official orthography has not affected the current alphabetisation of languages, judging from the disparity between the official alphabet and the efforts of the Summer Institute of Linguistics (1984).

In spite of the establishment of so many official alphabets at the national level, a pan-African orthography has been attempted. The African Reference Alphabet (ÀRA) was developed through a series of conferences in Niamey, Nigeria with the support of the International African Institute and UNICEF. Devised in 1978 and revised in 1982, the ARA includes 60 graphs, many of them roman, roman-derived, or Greek, and many of which exist in the IPA and are given phonetic values similar to those of the IPA.[7] However, these graphs do not represent fixed phonetic values in and of themselves. Instead, phoneme-graph assignments are made on a language-by-language basis, in accordance with the following principles: (1) one sound / one sign (no digraphs); (2) graphs included in a language's orthography should be maximally distinct; (3) the same graph should represent the same sound for all the languages within a single country; (4) a language spoken in several countries should have the same alphabet in each country; (5) the same sound should be represented by the same graph within each subregion of Africa; (6) diacritics

should be avoided (Mann and Dalby 1987: 207). Figure 2 shows the corpus of graphs in the ARA.

Figure 2. African Reference Alphabet (reprinted from Mann and Dalby)

a.r.a consonants

	bilabial	labiodental		labiovelar[1]	alveolar dental[2]	alveolar	lateral	alveopalatal postalveolar/ retroflex	palatal	velar	uvular	pharyngeal	glottal
nasals[3]	m		ɱ		n̪	n		ɳ	ɲ	ŋ			
plosives	p b			ɓ ɓ	t d	t d		ʈ ɖ	c ɟ	k g	q		ʔ
implosives	ɓ				ɗ	ɗ			ʄ	ɠ			
ejectives/ aspirates[4]	p̓				t̓	t̓			c̓	k̓			
clicks[5]	ʘ				ǀ	ǃ	ʞ		ǂ				
affricates[6]								c ɟ					
fricatives	ɸ β	f v		θ ð	s z	ɕ ʑ[7]		ʃ ʒ	ç ʝ	x ɣ	χ ʁ[8]	ħ ʕ[9]	h ɦ[9]
roll/flap[10]					r		ɾ				ʀ		
continuant	w	ʋ[11]				ɹ	ɻ		ɥ				

a.r.a vowels

i	ɨ	u
e	ɘ	o
ɛ	ə	ɔ
a	ʌ	ɑ

Adoption of a pan-African alphabet is indeed practical in economic terms: typewriters and other printing equipment including these 60 characters would be marketable in any part of Africa and available to users of any African language. The ARA is also practical in its flexibility. It attempts harmonisation within languages and regions only. Thus, it does not require a separate graph for every phoneme in Africa and can easily avoid the use of diacritics for distinguishing, for example, the dental [t] of one language with the alveolar [t] of another. However, given the investments that nations have already made in their own official alphabets and other political and cultural differences amongst the African nations, widespread adoption of the ARA seems unlikely. It is likely that the ARA project was in large part an academic and bureaucratic exercise.

4. Further orthographic concerns

4.1 Division of words

One of the most basic problems in setting a language to writing is the division of speech into word units. In order to discuss such a problem, it would seem

necessary to define word in linguistic, non-orthographic terms. Yet the concept of word-hood is one that often eludes linguists. We are accustomed to thinking of words as units unto themselves. This perception is aided by our experience with words as printed units with spaces between them. In ordinary speech, however, there are no spaces between words. These spaces must be devised when we put the language to writing.

Units of meaning may be discernible in speech, but all units of meaning are not necessarily words. The minimal units of meaning, morphemes, are of two types: bound and free. Bound morphemes appear in speech only if juxtaposed with some other syntactically and semantically appropriate morpheme. Free morphemes can occur in utterances without affixing to something else. For example, the plural marker -s is a bound morpheme in English; it must attach to a count noun. House is a free morpheme. Free morphemes are not necessarily words, however. While house is a free morpheme, it is only a part of a word in houseboat. Furthermore, bound morphemes may be written as words. For example, in orthographies of Bantu languages, one sees a great deal of variety in the treatment of verbal prefixes, such as subject concord, incorporated objects, tense and aspect marking. So, subject concord markers are not treated as separate words in Lingala, but are treated as separate words (and are frequently referred to as 'pronouns') in Venda.

Wolff (1962) discusses three strategies for word division. Alternative A does not allow bound morphemes to be written as separate words — they are written as prefixes or suffixes to free morpheme words. Alternative B is to use a European orthography as a model, and to divide words on analogy with European words. Alternative C is a compromise strategy in which one tries to represent bound morphemes as parts of larger, free morphemes, while avoiding words that are excessively long. Wolff develops and argues for this last strategy, intending that it serve the structural reality of the language as much as possible while guarding the language's legibility. This seems practical and linguistically sound for most African languages (although difficult and controversial to apply in heavily incorporating languages). This, of course, requires extensive grammatical study of the language prior to alphabetisation.

4.2 Capitalisation and punctuation

Usually, capitalisation and punctuation conventions are modelled on the conventions of the colonial language or other European language. Since capitalisation is associated with importance and sacredness, this Eurocentric approach risks cultural insensitivity. If those things that are sacred to the European are capitalised, but those things that are sacred in the indigenous culture (e.g., for the Beng, the sky) are not, the orthography works a not-so-subtle devaluation of the indigenous belief system.

Aside from the ethical considerations, practical matters of graph formation create problems for capitalisation for orthographies with non-roman graphs. For example, in systems based upon the IPA, upper case characters must be designed for graphs such as ŋ. This adds to the expense of including non-roman graphs, requiring means for printing two distinct characters for each non-roman graph included. Furthermore, some IPA or pseudo-IPA graphs, such as ɪ and ʀ, resemble the upper case forms of other graphs. Thus, ɪ (the lax vowel) and i/I are likely to be confused, and an uppercase form must be invented for (small cap) ɪ that does not resemble (upper case) I.

In order to avoid these ethical and practical problems, the ARA uses no capitals and a modified European punctuation system.[8] This strategy avoids the value judgements inherent in capitalisation decisions by avoiding the decisions themselves. Thus, it is not controversial whether names of persons, sacred objects, deities, or days of the week have initial capitals, since no capitals exist.

Several national alphabets are like the ARA with respect to capitalisation. For example, the Nigerian alphabet, for which a typewriter has been produced, includes no upper-lower case distinctions among its graphs, using mostly graphs that correspond with lower case roman graphs. However, in practice, people who are literate in European languages often copy their capitalisation conventions into languages with non-capitalising alphabets. For example, the Ivoirian alphabet (Kokora 1979) includes no upper case forms. However, initial capitals at the beginnings of sentences and in proper names are used in a journal published at the University of Abidjan in the Baule language (with French translations on opposite pages), which purportedly employs the official Ivoirian orthography (Loucou 1981). However, because certain graphs included in the Baule text have no upper case forms, this capitalisation convention could not be universally applied. Furthermore, initial capitalisation is irregularly applied to the titles of articles and the journal itself, as well as to the names of ethnic groups. With such irregular observation of a capitalisation convention, it is difficult to understand the motivation for using capitals in the first place. Perhaps the authors and editors of the journal thought that a journal without capital letters might look "unacademic", or perhaps the overwhelming availability of French printed materials, in comparison to Baule printed materials, has instilled in the authors a subconscious bias toward capitalisation in the French fashion. If a system without capital letters is to survive as such, it is necessary that literacy education in that system include some sort of reinforcement of the idea that orthographies without capitalisation are as worthy of print as those with capitalisation, and that the sacred objects of cultures with capitalisation are not more sacred because they are capitalised. (These are real problems because of the use of European languages in schools. Any person with more than a couple of years of formal education will know of the capitalisation conventions in some European language.)

5. The Beng case

I now turn to a case study and critique, which concerns the following questions: Will (and should) development of an orthography for the purpose of western study of an African culture affect future attempts at alphabetisation and literacy in that language? For the Beng language of Côte d'Ivoire, what is the best orthography that can be developed for linguistic and ethnographic materials? The answers to these questions often became clear too late to influence the orthography used in the Beng-English dictionary (Gottlieb and Murphy 1995). Therefore, some of the following discussion is more rationalisation than criticism of the orthography chosen for Beng. However, close attention is paid to the usefulness of this orthography for the purpose of linguistic study and establishment of literacy.

The Beng are an ethnic group of approximately 10,000 people, divided into two kingdoms which use slightly different dialects of their language, Beng (Gottlieb 1992).[9] The language belongs to the Southern Mande group, which also includes Wan and Toura.

Information for the dictionary was first collected in 1979-80 and again in 1985 by Alma Gottlieb during anthropological field study. The result of these visits was a collection of index cards carrying over 2000 words plus hundreds of additional lexicalized phrases, transcribed in an IPA-like style. As linguistic study was not the foremost purpose of her visit, and as she is not trained in linguistics, Gottlieb's transcription of Beng words was often irregular; the sketch of Beng phonetics presented here is based on her transcriptions and a list of Beng phonemes produced by a member of the Summer Institute of Linguistics. The remainder of this section consists of subsections on the following topics: graph assignment for consonants and vowels, capitalisation, word division, and adaptation of borrowed words.

5.1 Consonant graphs

The bulk of the Beng consonants are common to West African languages. However, the written materials from Gottlieb's field research leave some questions as to the exact consonant inventory of Beng. Table 1 shows the graphs for the consonants of Beng used in the forthcoming dictionary, in contrast, where appropriate, with graphs of the IPA, the Ivoirian national alphabet, and the ARA.

Table 1. Beng consonants

	BILABIAL	LABIO-DENTAL	LABIO-VELAR	DENTAL	PALATAL	VELAR	GLOTTAL
STOP							
[-VOICE]	p		kp[1]	t	c	k	
[+VOICE]	b		gb[2]	d	j	g	
[+NASAL]	m		mgb[3]	n	ny[4]	ŋ	
FRICATIVE							
[-VOICE]		f		s			h*
[+VOICE]		v		z			
LIQUID				l/r			
TRILL				r[5]			

* used extremely rarely, mostly in onomatopoeia

[1] ARA = ƥ

[2] ARA = ɓ

[3] IPA, Ivoirian = ŋm; ARA = ~

[4] IPA, ARA = ɲ, Ivoirian = ny/ŋ

[5] Because of uncertainty as to true nature of this phoneme, equivalents cannot be given for other orthographies

[6] IPA, ARA = r; not differentiated from other central liquids in Ivoirian alphabet

The dictionary orthography has a few problems that might have been prevented by a more thorough phonological study of the language. The first of these problems involves the possible existence of a labio-velar nasal stop, which was not included in Gottlieb's original estimation of the consonantal system. Instead, this phoneme was treated as a prenasalised /gb/ and was spelled **mgb**.

However, on closer examination of the other possible prenasalised stops, i.e., those consonantal clusters spelled as a nasal followed by a stop, the existence of prenasalised consonants in Beng becomes doubtful. There are no word-initial **mb** clusters, and the recorded **nd** and **nt** clusters involve the contraction of the first-person singular possessive pronoun, ŋ, and a kinship term that begins with [d] or [t]. In these cases, the **nd** does not represent an instance of prenasalisation, but of a syllabic [n] followed by a stop, i.e., two separate phonemes (belonging to different morphemes). Yet there are several instances of initial and medial clusters recorded as **mgb** (also **ŋgb** in Gottlieb's field notes). In these cases, contraction of two morphemes is clearly not the case. SIL work on Beng (1984) also posits a labio-velar nasal consonant.

The Ivoirian alphabet and IPA represent such a phoneme with a digraph, **ŋm**, while the ARA utilises a single graph which resembles a compact, oversized tilde. The dictionary persists in using a trigraph, **mgb/ŋgb** because of the

perceived explosive nature of the consonant; at the outset, the phoneme is nasal in pronunciation, but it is released with a [b]-like sound.

The trigraphic choice is deserving of some criticism. Not only does it carry us far from the One Sound / One Sign principle, but it is misleading in its suggestion of a *pre*nasalised consonant, rather than as a nasal consonant. The explosive [b]-like ending of the phoneme is most likely a sub-phonemic, and perhaps allophonic, characteristic of the nasal phoneme. Including the **b** in the orthographic representation of the phoneme may be helpful to westerners attempting pronunciation of Beng, but it would be unnecessary for the Beng, should a popular orthography be developed from this dictionary. The Ivoirian and IPA representation, **ŋm**, would serve as a phonemically accurate representation for the sound and would be recognisable as that phoneme to any native Beng speaker who acquires literacy in Beng (a Beng speaker would not require so much phonetic detail as the **b**). By employing **ŋm**, we would also reduce the trigraph to a digraph, increasing readability. I have not lent serious consideration to the ARA single graph, as it is so unavailable in printing and typing equipment. If the ARA gains acceptance and use in western Africa before serious attempts at Beng alphabetisation for literacy are made, the ARA graphs would provide a more elegant representation of the labio-velar nasal than the suggested digraph.

The second major phonemic problem in the Beng orthography is the use of both l and r for what appears to be only one phoneme. A thorough investigation of the possible complementary distribution of [l] and [r] cannot be made from the written data; however, the written data do provide several clues which indicate that the two graphs represent the same phoneme. A first piece of evidence is the absence of words beginning with r, while many words begin with l. Furthermore, intervocalic l is much more prevalent than intervocalic r. Consonant-liquid clusters also reveal evidence of complementary distribution. The dictionary includes several words with initial **sr** but none with initial **sl**. The graphs l and r were used interchangeably in many words in Gottlieb's vocabulary cards and field notes, especially in stop-liquid clusters. Gottlieb (p.c.) confirms that [l] and [r] use seemed interchangeable to her in many words and among individual speakers.

One could rationalise that the l/r graph distinction is useful in recording allophonic distinctions. This rationalisation is actually somewhat rational for the purposes of this dictionary. As it is to be published by a linguistic organisation (Indiana University Linguistic Club), linguists will presumably be the dictionary's major audience. It is possible that our l/r distinctions will enable some linguist to determine the distribution of liquids in Beng. But should this orthography have an effect on future alphabetisation of Beng for popular literacy, the l/r distinction could be problematic. The excessive number of graphs for a single phoneme could make spelling difficult for native Beng writers, as the l/r distinction might often seem quite arbitrary.

The use of digraphs (**gb**, **kp**, **ny**) is forgivable despite the blatant disregard for the One Sound / One Sign principle. Utilising digraphs for the labio-velars simply follows the convention established by the IPA and adopted by most others, including the Ivoirian national alphabet. Though the ARA does include single-graph representations of the labio-velars, these graphs are unavailable in typing and printing equipment (except that which is especially made for the ARA), and therefore it would be burdensome for general use as well as for use in our dictionary.

The palatal nasal, which the dictionary represented with **ny**, does have a single-graph representation in the IPA and elsewhere, **ŋ**. In the Ivoirian alphabet, either **ny** or **ŋ** may be used, presumably depending upon one's typographical capabilities. Such a flexible solution seems appropriate for this sound, as the IPA graph is not only uncommon in typing equipment, it is difficult to differentiate from other nasal graphs in the handwriting of many. If the Ivoirian alphabet gains wider acceptance before Beng is alphabetised, it is reasonable to expect that both of the variant graphs will be employed by Beng-speaking individuals.

5.2 Vowel graphs

Beng has seven vowels, as illustrated in Figure 3. Two of the graphs used in the dictionary are not contained in the roman alphabet, but are used in many African orthographies (including the Ivoirian system and ARA) and in the IPA and are therefore somewhat available for typesetting and on African typewriters.

Figure 3. Beng Vowels

One controversial element of Beng phonetics is the status of nasalised vowels. The SIL missionaries researching Beng (1984) identify four nasal vowel phonemes in Beng, whereas Gottlieb and Murphy treat nasal vowels as pronunciation variations of vowel-nasal clusters (the nasal consonant of these clusters is deleted in certain contexts, leaving a nasal-coloured vowel). However, the dictionary treatment does not address these situations in a regular manner. The tendency is to use a vowel graph + nasal consonant graph in word-fin-

contexts and to indicate nasal(ised) vowels elsewhere by a diacritic tilde (~) over the vowel.

The SIL treatment of nasal vowels posits four nasal vowel phonemes: /ĩ/, /ũ/, /ɛ̃/, and /ɔ̃/. In this situation, the nasal consonant (usually an /ŋ/) that follows these vowels in many contexts is explicable as some type of *liaison*. The lack of /ẽ/ and /õ/ in the SIL system concurs with the apparent lack of distinction between medial lax and tense nasal vowels. The Beng orthography makes a distinction between the lax and tense vowels, based on distinctions in Gottlieb's field notes. This distinction is likely to be meaningless, however, considering that many languages have fewer nasal vowels than oral vowels, since fewer vowel distinctions can be made with a lowered velum.

SIL's orthography uses digraphs (vowel + n) in describing the nasal vowels, despite the phonemic status that SIL accords the nasal vowels. Considering the popular use of the tilde as an indication of the nasal feature worldwide, the SIL's use of a digraph is puzzling and misleading. By its *sequential* ordering of the graphs that usually stand for (oral) vowels and nasal consonants, the SIL suggests that these two sounds are produced sequentially.

Before mass literacy in Beng is attempted (and before further lexicographical work in Beng is done), the status of the nasal vowel must be investigated. If the nasal vowels of Beng are best interpreted as different phonemes from the oral vowels, then regular use of the tilde-marked vowel graph should be employed. In any case, regularisation of the treatment of nasal vowels is necessary, but further phonological study of Beng is needed in order to decide the best course of action.

5.3 Representation of tones

In Beng, as in many other African languages, words carry lexical tone. The style of marking of tones reflects a prevalent custom in Africa and in the romanisation of Asian languages. This custom involves iconic diacritics borrowed from European orthographies, where they serve as accent marks. The five tones of Beng are represented as the following:[10] v (mid — no diacritic), v̋ (high), v̀ (low), v̂ (falling), v̌ (rising). Other options for marking tone have been established, but they involve either marking some tones by sub-graph diacritics or placing the tonal symbol before or after the syllable instead of placement on the vowel (IPA 1957). While these practices may make symbolic or phonetic sense, popular preference for the iconic accent system did bias the choice, as did the relative ease with which the accent diacritics can be produced in writing, typewriting, or word-processing. Use of this system should not be problematic in either academic work or in a future alphabetisation of Beng for literacy purposes. In the latter situation, tone-marking might be eliminated in most contexts, save for those in which lack of tone-marking induces a pragmatically plausible ambiguity. Our dictionary should not be used as a model for tone-marking in Beng, however, because many tones may be missing or

incorrect. Because Beng tones undergo changes dependent on neighbouring tones and because many lexical items in the dictionary were recorded from experiences of them in sentences, the underlying tone of any word may have been misinterpreted by the American recorder.

5.4 Adaptation of borrowed words

Because our dictionary attempts a phonetic orthography, borrowed words have been spelled as they sound when pronounced by a Beng speaker, not as they are spelled in their original language, for example, Beng karfur for French carrefour. For personal names borrowed from neighbouring languages (such as Baule), the dictionary indicates the popular spelling of the name within the entry. While it may seem unorthodox to use a spelling other than the spelling that is used locally, the very French-style spellings of the names represent much different sounds in the Beng orthography than the names actually have (e.g., nesan replaces the spelling Nguessan and yakuba replaces Yacouba(h)). This juxtaposition of phonetic spellings with the popular spellings (which are somewhat official because of their use in identification papers) may help distinguish the dictionary's orthography as a phonetic, descriptive orthography, which is not in itself a practical orthography for popular use.

5.5 Capitalisation

The dictionary follows the ARA in its non-use of capitals in all contexts, even personal names. This strategy avoids forcing western value judgements upon the orthography of Beng. Thus, it does not allow westerners to decide which (if any) sacred objects, characters, or traditions should have capitalised names. The non-use of capitals also reinforces the idea that ours is a phonetic orthography designed for linguistic study, not for popular writing. If the ARA system or the non-capitalised Ivoirian alphabet becomes popular in Côte d'Ivoire, future orthographies may not include capitals as well. If not, choices must be made for a popular orthography as to which words deserve initial capitalisation. Ideally, the Beng people themselves will be accorded the choice of capitalisation conventions.

5.6 Word division

Decisions as to which morphemes and strings of morphemes would be treated as words depended upon both knowledge of Beng morphology and English-speaking biases, an accidental adherence to Wolff's (1962) compromise strategy C (see Section 4). Bound morphemes are treated as affixes, such as the passive

suffix -lɛ (used in the verbal citation form, as in trilɛ 'to be / make black' < tri 'black' + -lɛ) and the 'one removed' temporal suffix -ze (gbleze 'day before yesterday' < gble 'yesterday' + ze). As a matter of policy free morphemes are treated as words in any case where sound changes have not rendered the morpheme orthographically dissimilar to its pronunciation in isolation. So, for example, blonyilɛ ('to be happy' < blon 'liver' + nyilɛ 'to cool') is treated as one word because of the loss of the [ŋ], while a compound without sound change is treated as two words, such as yiru yali ('hyena' < 'night' + 'walker'). (One exception to this is the set of patriclan names, which all end in the morpheme lɛŋ 'child'. For consistency across patriclan names, all are written as compounds, regardless of whether any sound change is evident in the constituent morphemes.)

Apostrophes indicate certain contractions, most notably the address forms of many kinship terms, which contract the first-person, singular, possessive pronoun with the reference form of the kinship term. In many of these cases, some sound change does occur in the contraction, usually the deletion of (or assimilation of nasal qualities in) the initial stop of the kinship term and place-of-articulation assimilation in the possessive pronoun (e.g., ŋ + da -> n'a '(my) mother'). Use of the apostrophe differentiates the contractions from similar, uncontracted words (e.g., n'a 'mother (address)' vs. ɲa 'wife (reference)') and does not misleadingly suggest the existence of a prenasalised consonant phoneme (as in n'toma '(my) namesake'). Because the pronoun does not undergo assimilatory sound changes in other contexts of use, possessive pronouns are treated as words in other contexts.

These choices in word division seem morphologically and phonologically wellmotivated, however at least one other choice may deserve some criticism and revision. The postpositions of Beng, on analogy with the prepositions of English, were treated as separate words from their objects. It is possible, however, that these postpositions might be better treated as case-markers (or semantic *rôle* markers), and therefore (on analogy with languages with rich case-marking systems) would be better represented as suffixes, rather than words. Furthermore, many of the postpositions vary in pronunciation, depending on their phonological contexts (e.g., o/wo 'in'). Thus, in treating postpositions as words, we have broken our own rule of word division according to phonological stability. Because no serious grammatical study of Beng has been attempted, it is impossible to draw firm conclusions about the syntactic function of the so-called postpositions, but such premature assessment of the situation may bias future studies or literacy efforts.

6. The dictionary as orthographic model

Any lexicographic study of a non-alphabetic culture risks affecting that culture. If people of this culture are interested in having an orthography for their lan-

guage, but feel they lack the skills to create one, the lexicographer's orthography may be seen as a solution to their problems. For forces outside the culture that are interested in seeing the language alphabetised (e.g., government, missionaries), exposure to the lexicographer's orthography may (correctly or misleadingly) give the orthographers linguistic information with which to work. As discussed above, the lexicographer, in creating an orthography, makes many judgement calls about the language: what is a word, how is a particular sound represented, what is worthy of initial capitalisation. These decisions provide a medium by which biases toward the lexicographer's native (or preferred) tongue or his / her misunderstanding of the target language may be carried into the writing system of the language and ultimately into indigenous and external popular (and even academic) perceptions of the language.

English-speaking culture's attitudes toward dictionaries indicate the reverence we have toward their entries. Because most standard English dictionaries only record the terms of the higher social registers of English, the dictionary is used as a tool of linguistic discrimination. The grade school teacher's cliché, "Ain't isn't a word — it's not in the dictionary," demonstrates the authority that English speakers give to dictionaries (and lexicographers). Dictionaries are respected as fonts of linguistic "truth" and are expected to be very conservative. The esteem in which we hold the dictionary is evident in references to *the* dictionary (like *the* Bible), in which we ignore the fact that competing dictionaries exist and may represent different "truths" about our language.

If the same sort of authority is given to a preliminary dictionary of an unalphabetised language (such as Beng), the consequences could be disastrous if popular use of the orthography prescribed by the dictionary is attempted, since the dictionary is based on an outsider's perceptions of the language and not on native speaker intuitions (although it may be produced in consultation with native speakers). Therefore, the description of a language through a dictionary should be as accurate as possible, not only for the benefit of western science, but so as not to sabotage later efforts at establishing literacy in the language.

In the case of the *Beng-English dictionary*, thorough linguistic study could not be made prior to publication. The ethical issues surrounding the alphabetisation of Beng with only a cursory study of its linguistic structure conflict with the concern that other English-speaking academics have access to information about the Beng.[11]

The sins in the *Beng-English Dictionary*'s alphabetisation of Beng are mitigated by a number of factors. First, the text of our dictionary is in English, which is not widely used in Côte d'Ivoire. Thus, Beng speakers and Ivoirian educators will have limited access to the dictionary and will be less likely to use its orthographic model than if we had written in French (the official language of Côte d'Ivoire). The fact that the dictionary is published by a small organisation in the United States adds to the dictionary's inaccessibility as a model (although at least a dozen copies are now available in Beng villages). As discussed above, some of our orthographic conventions (non-capitalisation, phonetic spelling of

borrowed names) reinforce the perception of this orthography as a scientific endeavour, not a first step toward a Beng written literature. Disclaimers in the dictionary front matter may also serve to lessen the perception of our dictionary as authoritative.

Upon hearing of the imminent publication of this dictionary, some Beng people expressed the hope that this dictionary might help them learn English (Philip Graham, p.c.). It is interesting to note that Beng people who are literate in some language (most likely French) have assumed that they will be able to recognise Beng in print. The fact that this orthography is based upon the IPA may make this task more difficult than these Beng individuals imagine (they may be expecting the French-based type of spelling one often sees for their neighbouring languages, including Baule / Baoulé). This dictionary was not intended for such a task, and therefore would be of very limited use to anyone wishing to learn English. However, with the input of a literate Beng person, it may be possible to revise the dictionary, first through more careful linguistic study, then by the addition of an English-Beng section, so that this study can serve the Beng people.

Notes

1. Arguments that literacy aids intellectual development have been challenged in recent years. Eisenstein (1979) traces European history to the conclusion that many of the social advances thought to have been caused by the invention of writing were actually products of the invention of printing technology. Scribner and Cole (1981) argue that while the development of certain intellectual habits is frequently attributed to literacy, these habits are instead the product of the discipline of formal education. In their study, individuals who attain literacy skills in the informal village settings in Liberia and Sierre Leone did not attain the same intellectual skills characteristic of those who gained literacy through formal schooling.
2. According to Bendor-Samuel and Bendor-Samuel (1983), this problem has been encountered by missionary literacy workers in Ghana. Although village schools taught students to read and write in English, the missionaries could not find literate, adult villagers. This was because those who completed their educations quickly became unsatisfied with agricultural life and moved to the cities to find means of employment that would allow them to utilise the skills gained in school.
3. Of the 43 African nations discussed in Chrystal (1987), only seven do not have European official languages. Of these seven, six have Arabic as an official language. Only Ethiopia has only an indigenous language (Amharic) and its writing system predates European influence (Dalby 1986).
4. For further discussion of orthographic principles, see International Phonetic Association (1957), Williamson (1984), Mann and Dalby (1987: 207-211).
5. Other alternatives are apparently considered so unviable that they are not even discussed. In *Practical Orthography in Nigeria*, Kay Williamson (1984:13) takes "for granted that the Latin

alphabet, particularly as it is applied to English, is to form the basis of the representation of Nigerian languages."

6. This principle is also known as "One Phoneme/One Graph" and "One Phoneme/One Grapheme." In order to avoid the debate over the usefulness of the concept 'grapheme', I have chosen to use a more neutral, albeit less precise, label for the principle.

7. The ARA's Principle 9, with relation to existing orthographies, states "any a.r.a letter that coincides with an i.p.a letter should have as its most frequent realisation a value concordant with its i.p.a value, or other widely established usage" (Mann and Dalby 1987: 208, punctuation and capitalisation as in original).

8. See the quote in footnote 7 for an example of ARA-style punctuation and (non-)capitalisation conventions.

9. The Beng have been referred to by other sources as the Ben, the Ngan, the Gan, and the Ngen (Mundt 1987: 31). The name Beng is preferred here as it is an auto-ethnonym.

10. The v in these examples is intended to serve as a generic symbol for "vowel" and is included in order to demonstrate the placement of the diacritic above the vowel graph.

11. The desire to produce a record of Beng is not completely self-serving for the academic community. With the epidemic of language death around the world, it is vital to record minority languages now in an effort to retard their demise or at least have a cultural record in case of the language's death. (See Linguistic Society of America 1993.)

Acknowledgements

I am most grateful to Alma Gottlieb who has taught me everything I know about Beng and who has patiently put up with my incessant questions. Virtually all of the information concerning Beng in this paper is derived from my conversations with Alma, however, I bear all responsibility for the interpretations of her work here, and I sincerely hope that I have not misrepresented Alma or the Beng in any way. Parts of this work were presented at the Annual Conference on African Linguistics in East Lansing, Michigan in 1992, and I am grateful to the audience there for their comments.

References

Bell, Alexander Melville. 1870. *Explanatory Lecture on Visible Speech, the Science of Universal Alphabetics.* London: Simpkin, Marshall, and Co.

Bendor-Samuel, David H. and Margaret M. Bendor-Samuel. 1983. *Community Literacy Programmes in Northern Ghana.* Dallas: Summer Institute of Linguistics.

Coulmas, Florian and Konrad Ehlich (Eds.). 1983. *Writing in Focus.* Trends in Linguistics: Studies and Monographs 24. Berlin: Mouton.

Crystal, David. 1987. *The Cambridge Encyclopedia of Language.* Cambridge: Cambridge University Press.

Dalby, David. 1986. *L'Afrique et la lettre / Africa and the Written Word.* Paris: Fête de la Lettre.

Eisenstein, Elizabeth L. 1979. *The Printing Press as an Agent of Change.* Cambridge: Cambridge University Press.

Gottlieb, Alma. 1992. *Under the Kapok Tree: Identity and Difference in Beng Thought.* Bloomington: Indiana University Press.

Gottlieb, Alma and M. Lynne Murphy. 1995. *Beng-English Dictionary.* Bloomington: Indiana University Linguistics Club.

Hardman, M.J. (Ed.). 1981. *The Aymara Language in Its Social and Cultural Context. Gainesville: University Presses of Florida.*

International Phonetic Association. 1957. *The Principles of the International Phonetic Association being a description of the International Phonetic Alphabet and the manner of using it, illustrated by texts in 51 languages.* London: Department of Phonetics, University College.

International Phonetic Association. 1989. Report on the 1989 Kiel Convention. *Journal of the International Phonetic Association* 19: 6780.

Kemp, J. Alan. 1981. Introduction to Lepsius's 'Standard Alphabet.' Lepsius, Richard. 1863: ix-99*.

Kokora, Pascal Dago. 1979. *Une orthographe pratique des langues ivoiriennes.* Abidjan: Institut de Linguistique Appliquée.

Ladefoged, Peter. 1990. The Revised International Phonetic Alphabet. *Language* 66:(3): 550-552.

Lepsius, Richard. 1863/1981[2]. *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters.* (Ed. by J. Alan Kemp.) Amsterdam: John Benjamins.

Linguistic Society of America Committee on Endangered Languages. 1993. The need for the documentation of linguistic diversity. *LSA Bulletin* 141, Oct. 1993: 5.

Loucou, J. N. (Ed.). 1981. *kokowa: baule amannien fluwa* 2 (Université d'Abidjan).

Mann, Michael and David Dalby. 1987. *A Thesaurus of African Languages: A classified and annotated inventory of the spoken languages of Africa, with an appendix on their written representation.* London: Hans Zell.

Mundt, Robert J. 1987. *Historical Dictionary of the Ivory Coast (Côte d'Ivoire).* Metuchen, NJ: Scarecrow Press.

Scribner, Sylvia and Michael Cole. 1981. *The Psychology of Literacy.* Cambridge, MA: Harvard University Press.

Summer Institute of Linguistics. 1984. *Kwe Scwe. 1985.* Abidjan: SIL.

Tchitchi, Toussaint and Marc Laurent Hazoumé. 1983. Le développment [*sic*] d'un alphabet national dans une communauté multilingue. Coulmas, Florian and Konrad Ehlich (Eds.). 1983: 187-207.

Wharton, L.C. 1925. Transcription of foreign tongues. *Transactions of the Philological Society 1921-24*: 59-112.

Williamson, Kay. 1984. *Practical Orthography in Nigeria.* Ibadan: Heinemann Educational Books.

Winter, Werner. 1983. Tradition and Innovation in Alphabet Making. Coulmas, Florian and Konrad Ehlich 1983: 227-238.

Wolff, H. 1962. Problems in Vernacular Orthographies: Word Division. *Journal of African Languages* 1(3): 221-31.

Yapita Moya, Juan de Dios. 1981. The Aymara Alphabet: Linguistics for Indigenous Communities. (Trans. by M.J. Hardman.) Hardman, M.J. (Ed.). 1981: 262-270.