Transcription for the Sara Languages

The transcription for the Sara Languages used in these works is derived from one developed by French linguists working at the Collège Charles Lwanga (CCL) in Sarh during the 1970's. Tone (high, medium, low) is marked above all vowels and tone-bearing consonants, and vowel nasalization is marked with a tilde beneath the vowel. There are three principal differences between the system used here and the CCL system:

1. two letters, nasalized 'r̰' and nasalized 'y̰', both marked with a tilde beneath the consonants, have been dropped. Although I have found no discussion of nasalized 'y̰', it is evident that this letter was intended to represent the palatal nasal 'ɲ̰', and Fournier (1974) specifically states that the nasalized 'y̰' is pronounced [ɲ̰] in initial position. Both these sounds occur only adjacent to nasalized vowels, whereas 'y' and 'r' occur only adjacent to oral vowels. Hence, it is clear that the y with a tilde, or [ɲ̰], is an allophone of /y/, and that their nasalized '**r̰**' is an allophone of '**r**'. There is no need to use either symbol in this work.

2. CCL uses the schwa 'ə' to indicate the neutralized vowel in bi-syllabic morphemes. The pronunciation of this sound varies greatly, but it is normally a high vowel, not a mid vowel. Jim Roberts' work in Gor has shown that 'ə' contrasts with 'ɨ̰' in Central Sara languages,[1] and for this reason it makes better sense to use the symbol 'ɨ̰' for the neutralized vowel in all the Sara Languages. This is what I have done here.

3. CCL does not mark nasalized vowels as nasalized when they immediately follow a nasal consonant. For reasons discussed below, in this work heavily nasalized vowels are marked, even when they immediately follow a nasal consonant.

This system of transcription is intended to capture all the phonological distinctions found in these languages, thus allowing us to transcribe all the words and sentences these works contain. It is not intended as a proposal for a writing system for these languages. Many Chadian linguists working on the Sara languages do not consider tone marking necessary, viewing it as cumbersome. Literate Chadians are not accustomed to writing tone, and few are able to do it with accuracy. There is also a distinct advantage to using a system which allows the language to be written using letters found on a standard keyboard. The system used here does not. Finally, the hyphens used in this work to indicate that a word is morphologically complex should be avoided in the everyday writing of the languages.

The transcription employed here is fairly phonological; for the most part only sounds which are distinctive are used as letters. However, in cases where I felt that a valid phonological analysis was excessively abstract, then I use a transcription which might be considered phonetic. It appears that it

[1] However, he does not consider either of them to be phonemes in most Sara languages.

might be possible, for example, that an abstract analysis of the phonology of vowels in morphemes would make 'ə' and 'ɨ' allophones of other vowels. To do so, however, would make the words unreadable for anyone not well versed in the details of the particular analysis.

The diphthongs listed below are found in almost all of the Sara Languages. Others exist but are rare, and will be discussed with the language in which they are found.

Three of the sounds listed below, [č][2], [ɛ], and [ɲ], are placed in square brackets to indicate that they are only used in the phonetic description of sounds, and not in the actual transcription of words and sentences.

[2] I use the symbol [č] rather than [tš] because the former makes it clear that it is a single consonant.

Consonants

b like ‘b’ in English 'bad' or in French ‘beau’.

ɓ implosive bilabial, extreme lowering of the larynx which produces the distinctive sound of the sharp intake of air as it rushes in to equalize the pressure when the stop is released.

[č] pronounced like the ‘ch’ in English ‘cheese’.

d like the 'd' in English 'die' or French ‘des’.

g similar to the 'g' in English 'go' or 'give'.

h similar to the 'h' in English 'have'.

j like the 'j' in English 'joke'.

k similar to the 'c' in English 'scoot'; relatively unaspirated.

l like the 'l' in English 'lake' or in French ‘les’.

m like the 'm' in English 'miss' or in French ‘mes’.

mb a combination of 'm' and 'b', where the 'm' is lightly pronounced, and does not constitute a separate syllable.

n like the 'n' in English 'never' or Spanish ‘nunca’

nd a combination of 'n' and 'd', where the 'n' is lightly pronounced, and does not constitute a separate syllable.

ng in initial and inter-vocalic position, similar to 'ng' in English 'single'. In final position it is pronounced like the 'ng' in 'sing'.

[ɲ] palatal nasal, similar to the ‘gn’ in French ‘onion’, or ‘ñ’ in Spanish ‘caña’. In the Sara (Proper) languages, this sounds occurs only adjacent to nasal vowels.

nj a combination of 'n' and 'j', where the 'n' is lightly pronounced, and does not constitute a separate syllable.

p My impression is that ‘p’ is not very aspirated, closer to the 'p' in English 'spit' than it is to the ‘p’ in 'pit'.

r Shortly trilled 'r', similar to the 'rr' in Spanish 'burro', but shorter; for speakers of some languages, the 'r' acquires a distinct lateral sound, sounding as they are starting to say 'l' but ending with an 'r'.

ɗ Implosive alveolar, extreme lowering of the larynx which produces the distinctive sound of the sharp intake of air as it rushes in to equalize the pressure when the stop is released.

s Like the 's' in English 'soup'.

t like the 't' in English 'stoop'; less aspirated then the ‘t’ in English ‘top’.

w like the ‘w’ in English ‘wet’. See also diphthongs.

y when beginning a syllable, like the 'y' in English 'yes'. See also diphthongs.

Vowels

a like the 'a' in Spanish ‘gato’.

ą̃ heavily nasalized 'a'; like the 'ans' in French 'sans'.

e like the 'é' in French 'arrivé'.

ę̃ heavily nasalized 'e', somewhat similar to the 'ain' in English 'train', but without pronouncing the 'n'.

[ɛ] like the ‘e’ in English ‘met’.

ə a central mid vowel pronounced between ‘e’ and ‘o’, but not rounded.

ə̃ heavily nasalized ə.

i like the ‘i’ in Spanish ‘si’.

ĩ heavily nasalized ‘i’, similar to the ‘i’ in Spanish ‘sin’ or ‘ee’ in English ‘ seen’, but without the ‘n’ and more heavily nasalized.

ɨ̃ central high vowel. Normally similar to the unaccented 'e' in English 'tenacious'. Pronunciation of this phoneme often changes in different environments. See also discussion of harmonic constraints on page 13-15.

o like the ‘o’ in Spanish ‘gota’, similar to first 'o' in English 'coconut'. Very rounded.

o̰ heavily nasalized ‘ɔ’, similar to the 'on' in French ‘bon’. In no languages is there a distinction between nasalized ‘ɔ’ and nasalized 'o'.

ɔ open 'o', pronounced similar to the ‘aw’ in English ‘law’, or 'augh' in English 'caught', as spoken in most of the Northeast Corridor or the Great Lakes area, but shorter.

u similar to the ‘u’ in Spanish ‘cuna’.

ṵ heavily nasalized ‘u’.

Diphthongs

aw similar to the ‘ou’ in English ‘couch’.

ay like the ‘ay’ in Spanish ‘hay’, or the ‘y’ in English ‘cry’.

oy very rounded ‘o’ followed by ‘y’. No English equivalent.

ɔw the ‘aw’ of ‘law’ followed by ‘**w**’ (which sounds to an English speaker like ‘u’).

ɔy very similar to the ‘oy’ in English ‘boy’.

uy the ‘u’ of Spanish ‘cuna’ immediately followed by the glide ‘**y**’.

Vowel Length

Vowel length does not appear to be significant in any of the Western Sara languages. In cases where a vowel has two tones, it is written with 2 vowels, and when it has a single tone, it is transcribed with one. (e.g. tàā̄ ‘to take’ vs. tà ‘mouth’ in Kaba). My impression is that vowels with two tones tend to be phonetically longer than vowels with a single tone.

Tone

All of the Sara Languages have three even tones: high, mid and low:

tá	'to assemble (Laka)'
ɗā	'to do' (Kaba, Laka, Ngambay)
tà	'mouth' (Kaba, Laka, Ngambay)

Tone-bearing phonemes are vowels and sonorant consonants when they are in final position. When a sequence of identical vowels occurs, or when a sonorant consonant closes a syllable, complex tones occur:

wàā	'millet (Laka)'
tàá	'up high (Kabba)'
jōó	'two (Kabba, Laka, Ngambay)'

Of the nine possible combinations, Mid-Low is extremely rare, and is found in three or four morphemes. High-Low is quite rare in most native morphemes, but common in loan words and in ideophones. The High-Low tone on the final syllable of multi-syllabic foreign words often becomes Mid tone when used in a sentences. Thus, the word for 'airplane' in Bebote is pronounced àbɨ̀sṍò in isolation, but àbɨ̀sō when in a sentence.

Alphabetical Order

a, a̰, b, ɓ, č, d, e, ḛ, ə, g, h, i, ḭ, ɨ̰,
j, k, l, m, mb, n, nd, ng, nj, o, o̰, ɔ, p, r,
ɗ, s, t, u, ṵ, w, y

Changes in pronunciation

There are a number of automatic changes in the pronunciation of certain phonemes when they occur adjacent to other phonemes that apply to all of the Sara languages.

1. Nasalization of vowels. Whenever a vowel immediately *precedes* a nasal (**m**, **mb, n, nd, ng,** or **ng**), that vowel is pronounced nasalized:

bānjɨ̄	[bã̰njɨ̄]	'leprosy'	(Mango)
rēngē	[rḛ̃ŋgē]	'be thin'	(Nar)
bìn̄g	[bḭ̀n̄]	'thigh'	(Gulay)
ɓōng	[bṍŋ]	'anger'	(Gor)

This change is not indicated in the written transcription.

2. The transcription of vowels that follow a nasal consonant merits additional comment. I will examine these occurrences in three distinct environments.

i. When a vowel *follows* a nasal consonant in a morpheme consisting of a single open syllable, it is usually heavily nasalized (e.g. mḭ́ 'five' in all Sara languages). One might be tempted to write the vowel as an oral vowel, and appeal to a rule of nasalization to explain its phonetic realization. I have not done so, for three reasons. First, nasalized vowels freely occur in this environment without nasal consonants (e.g. sḭ́ 'a little' (Mbay)). The notation I have adopted shows that sḭ́ rhymes with mḭ́, but that mḭ́ does not rhyme with lī 'snake'. Second, a vowel that follows a nasal consonant is not always heavily nasalized (see discussion of nasalization in bi-syllabic morphemes and in syllables closed by a sonorant consonant in 2ii and 2iii below). Third, it is not possible to predict the occurrence of the nasalization of vowels in all languages. In Central Sara languages such as Bebote, for example, there exist a set of morphemes which end in a vowel, but where in other Sara languages a final 'r' is found. In these cases, the vowel is not nasalized after a nasal consonant (e.g. mà 'crocodile', which contrasts with nã̰ 'to taste' in Bebote).[3]

ii. When a vowel *follows* a nasal consonant in a morpheme with a single syllable or in the first syllable of a morpheme with two or more syllables, the vowel is not heavily nasalized (e.g. nàsɨ̄ 'type of antelope'). In some languages, a phonemic distinction can occur in the former situation (e.g. màr 'crocodile' vs. mə̀r 'to insist' (Mbay) and màý 'inside out' vs. mə̀y 'resist sickness (Mbay)'). Naturally, our transcription will allow us to maintain this distinction.

iii. When a vowel occurs in word-final position in a multi-syllabic morpheme, and immediately *follows* a nasal consonant, it is always pronounced heavily nasalized (e.g. Ngambay úmã̰ [ṵ́mã̰] 'ancestor'). For consistency, I have chosen here to transcribe the vowel here as nasalized, thereby making it clear that it is the same vowel as the one found in mono-syllabic morphemes (e.g. **2-i** above). CCL has taken the opposite approach.

3. When an oral vowel occurs adjacent to a nasalized vowel, the oral vowel is pronounced nasalized:

yã̰-í	[yã̰ḭ́]	'of you (Ngam)
m̄-gḛ̀-ī	[m̄gḛ̀ḭ̄]	'I love you (Ngam)

This change is not noted in the transcription.

4. The glide y is normally pronounced [ɲ] when it is in initial position and immediately followed by a nasalized vowel. This change also occurs in non-initial position with some of the Sara languages.

---

[3] Adami et al (1981:p. 101) specifically make the same claim for Bediondo. My recordings also show that the same is true for many speakers of Mango.

5. As indicated in the guide to pronunciation above, the n in ng is always pronounced [ŋ]. In final position, the g is normally dropped.

6. Younger speakers normally pronounce the implosive dental ɗ as r. The recordings I have made demonstrate that most speakers do not distinguish between these two sounds with any consistency. [4]

Sara Morphemes

Morphemes are a set of phonemes used to build words. Morphemes in the Sara languages have strict constraints on their structure, constraints that do not apply to words. In these works, a hyphen is used between morphemes to indicate that a word is morphologically complex.[5]

Consonants clusters in morphemes

1. With the exception of the pre-nasalized stops **mb**, **nd**, **ng**, and **nj**, clusters of consonants do not occur.

2. Any consonant can occur in morpheme-initial position or between vowels.

3. In most Sara languages obstruents (b, ɓ, d, g, h, j, k, p, ɗ, s, t) do not occur in morpheme-final position. The pre-nasalized stops **mb**, **nd**, and **nj** also do not occur in morpheme-final position. In the dialect of Daba (described in the forthcoming Eastern Sara Languages volume of this series), this constraint does not apply, and final obstruents occur freely.[6]

4. When the pre-nasalized stop **ng** occurs in morpheme-final position, one of two things can occur: either the **ng** is pronounced [ŋ], without the final '**g**', or else a neutralized vowel is appended to the end of the morpheme to avoid allowing the '**g**' to be morpheme-final. (For example, in Mango and other languages wōng 'anger' can be pronounced either [wo̰ŋ] or [wo̰ŋ.gɨ̰]. In writing, I have avoided including both variants in the dictionary, but rather have selected, at times arbitrarily, the pronunciation that I have found most prevalent.

---

[4] In most Sara languages, certain morphemes are pronounced only with ɗ.

[5] Using a single symbol, the hyphen, for this purpose has the disadvantage of not distinguishing between non-productiive word-formation processes and completely productive derivational processes.

[6] The constraint also doesn't apply with ideophones, where final voiceless obstruents occur quite freely in most languages. Normally, an alternant to this is a form where a final barred-i is pronounced.

5. A similar variation occurs with morpheme-final 'r', and sometimes with 'l' and the nasal consonants as well. Thus Kaba gèr̄ can be pronounced either [gèr̄] or [gèrɨ̄].

6. As a result of these constraints, only two syllables are possible in all of the Sara Languages except for Daba: **CV** or **CVC**, where the pre-nasalized stops **mb, nd, ng** and **nj** are considered single consonants, and where any final consonant must be a sonorant consonant.

7. These constraints do not apply with morphologically complex words or with inflected words (e.g. Kaba dèr̄-dé 'their older sibling').

Vowels

There are also relatively strict harmonic constraints on the co-occurrence of vowels within a morpheme. I have discussed these constraints elsewhere (Keegan 1989 and Keegan 1997).[7] Although important differences occur among the Sara languages, there are three patterns of vowels that dominate:

1. A morpheme where the vowels in the morpheme are the same:

bə̀yə̀	'female initiation' (Daba)
dènné	'woman' (Gor)
sīrí	'seven' (Ngambay)
bòlò	'hole' (Sar)
hɔ̄pɔ̄	'sheathe (Nar)
ɓūgúlū	'to stir' (Bebote)

2. A pattern where all of vowels except the first are neutralized, typically to barred 'i' (ɨ):[8]

bàtɨ̄	'sheep' (Goulay)
kèdɨ̄	'elephant' (Mbay)
ədɨ̀	'to rain' (Mango)
ìbɨ̀	'activate fire' (Mango)
ɓōkɨ́	'pour' (Sar)
ùsɨ̀	'to scrub' (Nar)

[7] In these works, I represented the neutralized vowel using schwa (ə). Since then Jim Roberts has convinced me that they should be transcribed as barred 'ɨ', since non-neutralized schwa is required independently in many Sara Languages.

[8] Jim Roberts argues that the barred i is not phonemic.

3. A pattern where all of vowels except the last appear as barred 'i' (ɨ):

bɨ̀là	'coop for animals' (Sar)
kɨ́tē	'to tickle' (Mbay)
bɨ̀lò	'hole' (Ngam)
kīngɔ̄	'bone' (Daba)
bɨ̀lə̀	'hole' (Gor)

Of these three[9] patterns, *pattern 2* is by far the most common. *Pattern 3* is only found with non-high vowels, and more often than not, that vowel is 'a'.

In *pattern 2* morphemes, when the final neutralized barred-i is followed by a suffix beginning with a vowel, there are two possible outcomes. In many cases, the barred 'i' is dropped, and its tone is lost (e.g. in Mango, àdɨ̄ + é ---> àdé 'he.gave-him/her'). In other cases, the barred 'i' assimilates to the affixed vowel, and the tone is not lost: (e.g. in Mbay, àdɨ̄ + á --> àd āá 'he gave him/her').[10]

The actual pronunciation of the 'neutralized vowel' varies somewhat from language to language, and especially in *pattern 3* morphemes where the barred-i is not word final. And in Central and Western languages, the first vowel in *pattern 3* morphemes is not neutralized: in Gulay, for example, kīngə̄ 'bone' contrasts with kūngə̄ 'axe'. Further, in the Eastern languages where the non-final vowel becomes barred-i, its actual pronunciation depends greatly upon the nature of the adjacent consonants.

In many languages *pattern 1* morphemes containing the vowels 'u' or 'o' the second vowel is not always pronounced as barred-i (e.g. Kaba kùtù 'rump', tógō 'to wash' in 'Gulay'). Nonetheless the final '**u**' or '**o**' behave like neutralized vowels in that they are dropped when immediately followed by a suffix beginning with a vowel (kùt-é 'his rump'). In such cases, I have elected to use for the lexicon the form as it is pronounced in isolation, and noted variant pronunciations with the phonetic pronunciation (marked by square brackets right after the word).

Among the different Sara languages, there is a high degree of consistency in regards to the vowel pattern of morphemes. Viewed from a very abstract level, I believe that it is likely that there is actually only a single harmonic constraint on the co-occurrence of vowels in morphemes: the vowels of most morphemes are identical (that is, 'a...a', 'e...e', 'i...i', 'o...o', 'ɔ...ɔ', or 'u...u'). The only real difference between them is which, if any, of the vowels have become neutralized. I have not attempted to incorporate this abstract model into the writing system.

---

[9] These are not the only patterns. In some of the Sara languages, for example, *pattern 3* morphemes show up with 'o ...e' for non-low vowels. These patterns and others will be discussed in the introduction for the languages where they occur.

[10] This process also depends on the tonal sequence of the stem and the tone of the vowel sufix.

Phonological Issues

There are two phonological issues that merit greater discussion, as they are relevant to the transcription system used here.

1. The palatal nasal ɲ.

As noted above (p. 7), the palatal nasal occurs only adjacent to nasalized vowels, while y occurs only adjacent to vowels that are not nasalized. Hence, the distribution of these two consonants is complementary, and the ɲ needs to be considered an allophone of y. This was first pointed out in Keegan (1989:234), and the issue is discussed in more detail in Keegan(1997:2-7). To my knowledge, the validity of this claim has never been challenged, yet linguists studying the Sara Languages continue to treat [ɲ] as a phoneme. This goes directly against one of the most fundamental rules of phonological theory. It also adds an unnecessary phoneme to the phonemic inventory, and an additional and uncommon rule is required to explain why nasalized 'y' is found in its place in non-initial position.

2. Nasalized Vowels.

It is also common to claim that vowels following the nasalized consonants m, n, and ɲ are not nasalized at the phonemic level, but rather are only pronounced nasalized, and that therefore the vowel should be transcribed as an oral vowel. I believe this to be an error. In the first place, since nasalized vowels are independently required in all the Sara Languages, there is nothing to be gained by adding this rule, a rule which ends up being quite complicated. As noted above, there are two cases where vowels are not heavily nasalized after nasal consonant: in multi-syllabic morphemes where the first letter is a consonant (e.g. nàsɛ̃ 'antelope' in most Sara Languages), and in mono-syllabic morphemes which end in a sonorant consonant (e.g. màr 'crocodile' vs. mə̀r 'to insist' in Mbay). It might appear possible to limit the rule to one which claims that only morpheme final vowels are affected by the rule. But even this does not work in some languages. In Bebote, for example, mà 'palm tree sp.' contrasts with nə̃ 'moon', and Adami et. al (1981:101) makes it clear that the same is true for the closely related Bejondo. The same issue exists in Mango.[11]

[11] This contrast is due to the fact that in these languages final 'r' found in other Sara Languages (e.g. **màr** 'palm tree' in Sar, Mbay, etc.) is dropped (e.g. Bebote **mà**).