Chapter 2

The sounds of the Bantu languages

Ian Maddieson

Department of Linguistics, University of California, Berkeley

1. Introduction

This chapter will describe some of the major phonetic characteristics of the (Narrow) Bantu languages based on first-hand familiarity with some of them and a reading of available literature. Since the number of languages is so large, and relatively few of them have been the object of much serious phonetic study, it will necessarily be very selective. Among phoneticians, the Bantu languages have a reputation as not having many interesting features — with the exception of the clicks and other exotic consonants introduced in the southern area largely through contact with speakers of Khoisan languages. Although it's true that many languages within the Bantu group are phonetically quite similar to each other, there is considerably more diversity in their phonetic patterns than is often believed. Some of this diversity may be disguised by the widespread use of simplifying transcriptions and orthographies which normalize away variation within and between languages, or underrepresent distinctions. Part of the aim of the chapter is therefore to draw greater attention to this diversity. Since this group of languages has received very extensive historical analysis the Bantu languages also provide a fertile field for examining inferences about the nature of phonetic sound change. Some sections of this chapter will therefore consider phonetic properties from a comparative or historical perspective.

The chapter is organized into two major sections, on vowels and consonants, with a shorter section on prosody. There are many important interactions between these three aspects of phonetic structure and some of these will be taken up at the point where it seems appropriate to do so. For example, the Bantu languages provide very striking examples of vowels affecting consonant realizations, particularly considered diachronically, and the nature of particular segments also has significant impacts on prosodic quantity and on tonal patterns. It is hoped that the brief discussions of selected issues here will encourage more attention to be paid to phonetic aspects of these languages.

2. Vowels

Vowel spacing and high vowels

The majority of Bantu languages (with some notable exceptions, particularly in the North-West) have simple-looking systems of five or seven vowels in which the expected relationships between the features of vowel height, backness and rounding hold. That is, the back non-low vowels are rounded, and the low and front vowels are unrounded. The vowels of the five-vowel systems are therefore usually transcribed as /i, e, a, o, u/ and the seven-vowel systems are most often transcribed as /i, e, ɛ, a, ɔ, o, u/ (as noted in Hyman 1999). However, these standardized transcriptions may disguise significant differences between languages, especially with respect to the nature of the vowels written /e/ and /o/.

In the five-vowel system of Xhosa (S.41), for example, /e, o/ are genuinely mid in character. The positions of vowels in an acoustic space are often shown by plotting values of the first two formants (readers unfamiliar with acoustic analysis might see Ladefoged 2000 for an introduction to the concept of a formant). The mean formant values for Xhosa vowels given by Roux and Holtzhausen (1989) are plotted in this way in figure 1. In this and following figures of the same type, the origin of the axes is in the upper right, with first formant (F1) values increasing down from the origin, and second formant (F2) values increasing to the left. The distances along the axes are scaled to reflect auditory/perceptual intervals. This kind of display closely parallels the traditional auditorily-based vowel space based on perceived 'height' and 'backness' values used, for example, in the IPA Handbook (IPA 1999) but has the advantage of being based on verifiable measurement. In figure 1 it can be seen that in Xhosa /e, o/ are located almost equidistant from the high vowels /i, u/ and the low vowel /a/. (There is a raising process in Xhosa which results in higher variants of /e, o/ when /i, u/ occur in the next syllable: the means for /e, o/ plotted here do not include tokens of these raised variants).

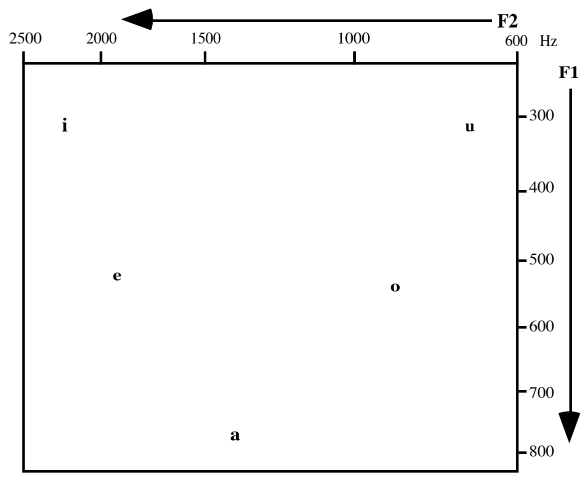

Figure 1. Xhosa vowel formant means (after Roux and Holtzhausen 1989). Each point represents the average of measurements of at least 30 tokens of the vowel from one male speaker reading a text.

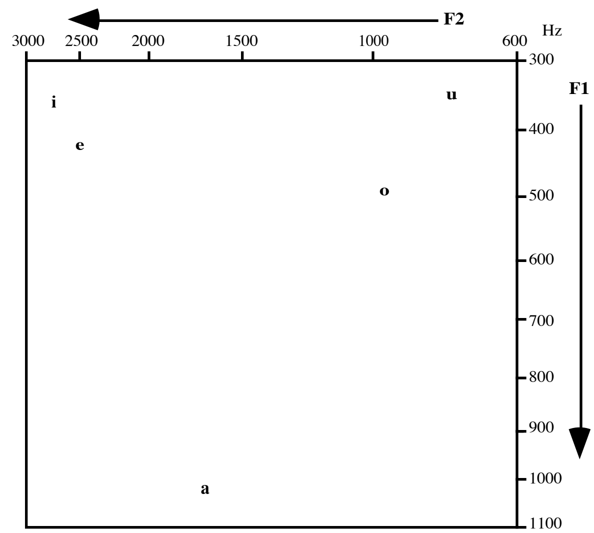

Figure 2. Ikalanga vowel formant means. Each point represents the average of at least 28 tokens of the vowel in penultimate position in a wordlist spoken by a female speaker. (Measurements by the author.).

Compare the spacing of Xhosa vowels with those of Ikalanga (S.16), shown in Figure 2 (the maxima are higher in this figure compared to figure 1, due to male/female differences in formant range.) In this case, the 'mid' vowels /e, o/ are relatively close to the high vowels /i, u/ and far from /a/. As a rough rule of thumb, vowels with a first formant lower than 400 Hz may be considered high vowels in a female voice. On this basis these particular vowels would not quite justify being considered high, but they are clearly markedly higher than those of Xhosa.

There have been relatively few acoustic studies of other Bantu five-vowel systems, but Swahili (Nchimbi 1997) has a pattern similar to Xhosa, while Ndebele (S.44) and the Zezuru dialect of Shona (S.12) (Manuel 1990) have a pattern similar to Ikalanga. The distribution seen in Xhosa or Swahili is similar to that most typically found cross-linguistically in five-vowel systems transcribed /i, e, a, o, u/, such as Spanish, Hadza or Hawaiian. This is also the pattern predicted by computational models of vowel system structure from Liljencrants and Lindblom (1972) to Schwartz et al (1997) based on the principle that vowels should be expected to be roughly equally dispersed in a space defined by the major formant resonances. In Ikalanga on the other hand the vowels are crowded into the upper part of the vowel space, with the front pair in particular being very close together.

Parallel variations in the structure of seven-vowel systems are also found. A plot of vowel distribution in Nyamwezi (F.22) is shown in Figure 3. In this language the vowels are to a large degree placed where they might be expected, given a respect for dispersion principles. This is particularly apparent for the front vowels, which are equally spaced from each other. The wordlist available for measurement included a more balanced sample of front than of back vowels, and the back vowels are probably in reality more separated than this plot indicates. These data suggest that transcription of this vowel set as /i, e, ɛ, a, ɔ, o, u/ (as in the figure) is appropriate, rather than the /i, ɪ, e, a, o, ʊ, u/ preferred by Maganga and Schadeberg (1992). For Sukuma (F.21), the more northerly dialect of the same language, Batibo (1985) also provides acoustic evidence for a relatively wide separation of the seven vowels, with /e, ɛ, ɔ, o/ all being clearly mid vowels.

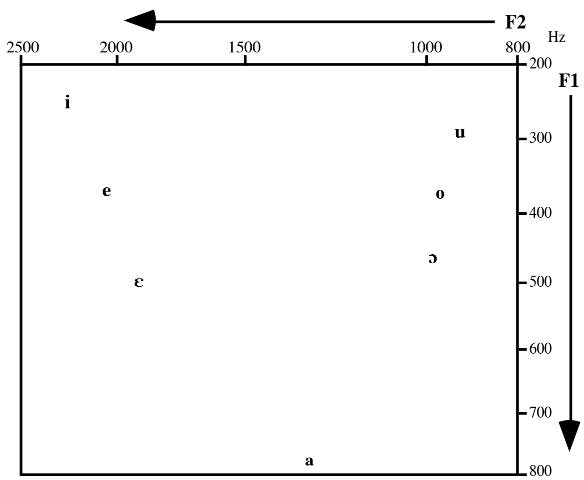

Figure 3. Nyamwezi vowel formant means. Each point represents the mean of between 9 and 23 tokens of unreduced final or penultimate vowels in a wordlist spoken by a male speaker. (Measurements by the author.)

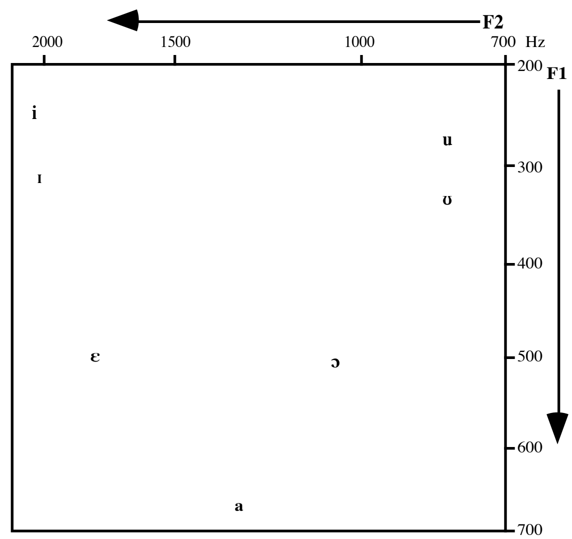

Figure 4. Bobe vowel formant means. Each point represents the mean of 20 or 30 measurements on minimal sets of words differing only in the penultimate vowel, spoken by a male speaker. (Measurements by the author on a recording made by J-M. Hombert, made available by L. Van der Veen.).

The relationship between the seven vowels of Bobe (Pouvi) (B.305) is notably different, as demonstrated in figure 4. Here a pair of vowels in the front and a pair of vowel in the back have such low values of F1 that they are all appropriately considered to be high vowels. The means are 248 Hz for /i/, 313 Hz for /ɪ/, 277 Hz for /u/, and 334 Hz for /ʊ/. The next lower vowels are markedly lower. In addition we may note that the front pair /i/ and /ɪ/, and the back pair /u/ and /ʊ/ have F2 values which are identical or nearly so, whereas Nyamwezi /e, o/ have F2 values intermediate between the higher and lower vowels in the system.

These data show that Bantu vowel inventories, both five and seven vowel systems, are split between those which are similar to global norms in their spacing and those in which the vowels are atypically crowded in the higher part of the vowel space. This fact bears directly on the issue of the appropriate reconstruction of the Proto-Bantu (PB) vowel system. According to the historical arguments of Meinhof (1899), followed by Meeussen (1969) among others, PB had a seven-vowel system with four high vowels, including two 'super-close' vowels usually notated /į, ų/ in addition to 'normal' high /i, u/. Guthrie (1967-71), in his a-historical compilation of 'Common Bantu' forms, adopts the same convention. Some recent interpretations of what the phonetic reality might have been have tended towards picturing the PB system as similar to modern West African vowel systems with harmony in which there are four high vowels, distinguished pairwise by tongue root position. This is unlikely to be correct. Under this scenario we would not expect to find reconstructed forms which mix the two high categories, such as *kúnį̀ 'firewood' *kų́pí 'be short'. We therefore prefer another hypothesis.

Among the significant facts which led Meinhof to posit two unusually high vowels are the striking diachronic effects on consonants preceding the reconstructed */į, ų/. The changes in question are often conveniently referred to as Spirantization (Schadeberg 1994/5, Mpiranya 1997, Labroussi 1999, Hyman this volume), but are more complex than this label may suggest, involving not only manner but also place, and sometimes producing unusual and complex outcomes. The results are also quite varied across languages, suggesting repeated independent developments triggered by inherited characteristics of the vowels. The patterns of the major stem-initial reflexes of PB labial and alveolar stops in Fang (dialect of Cocobeach) (A.75), Swahili, Yeyi (R.41), and Ikalanga before */į/, before */ų/ and before other vowels (_V), including 'normal' high /i, u/, are given in Table 1 (data from Medjo Mvé 1997, Nurse and Hinnebusch 1993, Baumbach 1997 supplemented by Gowlett 1997, and Mathangwane 1998, 1999, respectively).

Table 1. Spirantization in four languages

Vowel context	Fang	Swahili	Yeyi	Ikalanga
*p/_į	**f, v**	**f**	**s**	**s̫**
*b/_į	**bz**	**v**	**z**	**z̫**
*t/_į	**tʃ**	**s**	**s**	**ts^h**
*d/_į	**dz**	**z**	**z**	**dz**
*p/_ų	**f, v**	**f**	**f**	**f**
*b/_ų	**bv**	**v**	**v**	**v**
*t/_ų	**ts**	**f**	**ts^h**	**t^h**
*d/_ų	**dz**	**v**	**dz**	**d**
*p/_V	**f, v**	**p**	**p**	**p**
*b/_V	**b**	**w, ∅**	**β, w**	**β**
*t/_V	**t, l**	**t**	**t**	**t**
*d/_V	**y**	**l, ∅**	**r**	**l**

Diachronic (or allophonic) spirantization preserving consonant place in sequences such as /ti/ or /pu/ is not uncommon cross-linguistically, as in Japanese [tsi] and [ɸu], but a labial consonant outcome from inputs /tu, du/ (i.e. */tų/, */dų/) as in Swahili, or a lingual consonant outcome from input /pi, bi/ (i.e. */pį/, */bį/) as in Yeyi is much rarer. In these cases the character of the vowel has dominated over that of the preceding consonant. Particularly unusual are products such as the so-called 'whistling fricative' reflexes of */pi̧, bi̧/ in Ikalanga and some Shona varieties, which reduce the original consonant to a secondary feature. ('Whistling fricatives', notated [s̫ , z̫]), will be discussed further below.) Also the 'superclose' vowels quite frequently preserve the stop character of PB *voiced stops, either as simple stops or as part of an affricate, when sonorant realizations are found before other proto-vowels. These facts combine to suggest that the distinctive characteristic of these original vowels was indeed an unusually narrow constriction, nearly consonantal in character. This notion is supported by a particularly rare phenomenon reported in Hendo (C.73) and described in detail by Demolin et al (1999). In this language the class 5 prefix, reconstructed in PB as *di̧-, is realized as an unreleased voiced palatal stop ([ɟ˺]) before a voiceless stop or affricate, e.g. in [ɟ˺paka] 'moth'. MRI scans indicate that this segment is appropriately viewed as a hyperarticulation of the vowel /i/.

The reconstructed 'super-close' vowel pair may therefore have been similar to the 'fricative vowels' known to occur in certain modern Sino-Tibetan languages, such as Liangshang Yi (Ladefoged and Maddieson 1996) as discussed in Connell (2000). Some plausibility for this idea is provided by the comparisons provided by Connell between Proto-Bantu reconstructions and words in the Len variety of Mambila, a North Bantoid language which clearly has fricative vowels ([ʑ͡i] and [v͡ɯ] in Connell's transcription).

Tongue root position?

Co-occurrence restrictions of a harmonic nature between vowels, very typical of sub-Saharan African languages, are quite commonly found in Bantu languages (though often limited in extent, e.g. only applying in certain morphological contexts such as between verb roots and extensions). Bantu vowel harmony constraints do not seem to be a survival of an older Benue-Congo, or even Niger-Congo, harmony (Stewart 2000) but to be mostly more or less local innovations with diverse patterns of implementation (Hyman 1999). Vowel height, backness and rounding can all be factors in control of Bantu harmony. Vowel harmony in Africa often involves the independent use of pharyngeal cavity size, that is, adjustments of pharynx volume which cannot be accounted for as a function of the height and frontness of the tongue body (see Ladefoged and Maddieson 1996 for discussion). This is usually discussed as a contrast between advanced and retracted (or neutral) tongue root position, ± ATR. An interesting issue is therefore whether the Bantu languages, particularly those with seven or more vowels, make use of the ATR feature in this phonetic sense. (Phonologists often use [ATR] as a diacritic feature, even to distinguish pairs of vowels such as i/ɪ in English 'beat'/'bit' where tongue root position is not the phonetic mechanism involved.) The question of the role of ATR interacts with the question of the nature of the high vowels, as the *super-high/*high contrast might have been an expression of an ATR contrast, or transformed into one in daughter languages.

It is difficult to be certain that ATR contrasts exist in a language unless direct articulatory data on the vocal tract configuration during vowel production is available. There are very few studies of this type available so far for Bantu languages, but one data set is shown in Figure 5. These pictures are magnetic resonance images of sustained vowels produced by Pither Medjo Mvé, a speaker of the Bitam dialect of Fang (Demolin et al 1992). Figure 5 show very clearly that independent tongue root adjustment does not contribute to the distinctions between any members of the front vowel set /i, e, ɛ/, nor the back vowel set / u, o, ɔ/. The pharynx width, measured as the distance from the tongue root surface to the back wall of the pharynx at the height of the top of the epiglottis, in /e/ is intermediate between that in /i/ and /ɛ/, and that in /o/ is intermediate between /u/ and /ɔ/. It can be predicted from tongue body position — front vowels have wider pharynx than back vowels, lower vowels have narrower pharynx than higher vowels. The three front vowels and the three back vowels can therefore be distinguished one from another solely by height.

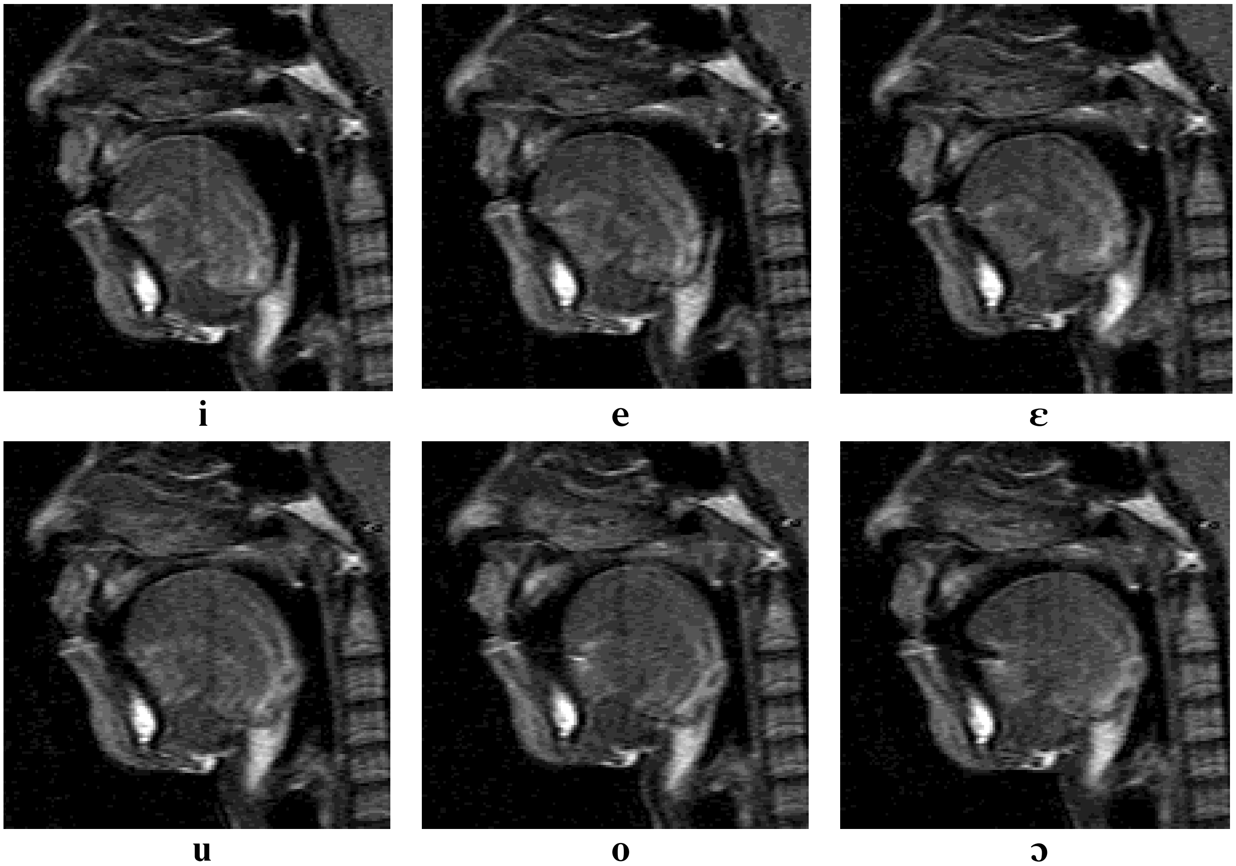

Figure 5. Articulatory positions of six of the vowels of Fang (dialect of Bitam). Top row, front vowels; bottom row, back vowels. Mid-sagittal MRI scans of isolated vowels, made available by Didier Demolin.

Bitam Fang has eight vowels, seven peripheral vowels plus mid central /ə/ (Medjo Mvé 1997). An acoustic plot of these vowels is given in Figure 6. Note particularly the slope of a line connecting the back vowels which points roughly to the position of the central vowel /a/, similar to that seen in Figures 1 and 2, and attributable to the fact that F1 and F2 frequencies co-vary in these vowels. This pattern is typical of that found in vowel systems where the back series is distinguished by degrees of height with no other factors being significantly involved. In this dialect lexical stems are marked by a strong tendency for V1 and V2 to be identical except if V2 is /a/, when /i, ə, a, o, u/ are all relatively common as V1 but /e, ɛ, ɔ/ are not (note that as many PB *CVCV items have become monosyllabic in Fang, the V2 in these cases is often not the *V2 of the reconstructed form). This pattern of co-occurrences is not one which suggests a phonological role for ATR.

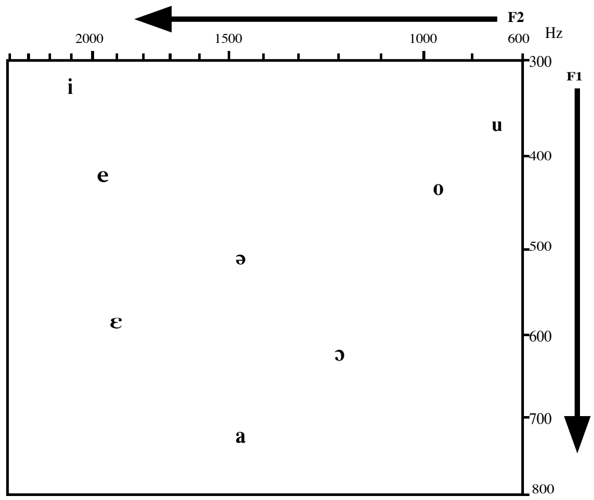

Figure 6. Fang vowel formant means. Each point represents the mean of 6 measurements, 3 of isolated vowel tokens, plus 3 tokens in final vowels in /alV/ nonsense words. Recording by Pither Medjo Mvé made available by D. Demolin; measurements by the author.

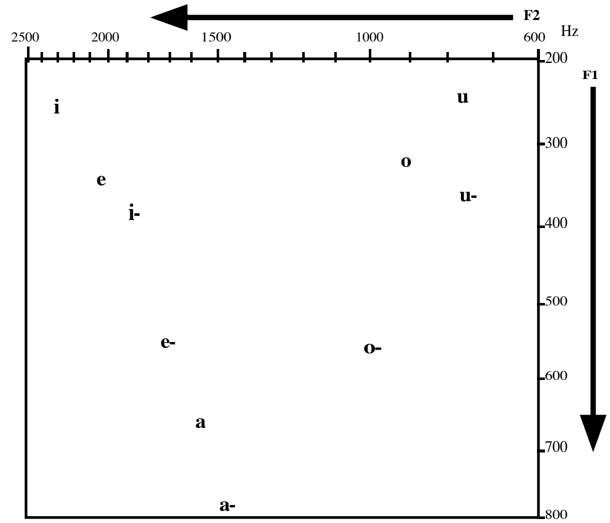

Figure 7. Nande vowel formant means. Each point represents the mean of between 6 and 21 tokens of phonetically long vowels in penultimate position in words spoken by a male speaker. Data from Ngessimo Mutaka; measurements by the author.

On the other hand the harmonic behavior and acoustic characteristics of vowels in Nande (D.42) considered together suggest that this is likely to be a genuine example of the use of ATR for phonetic distinctions. Mean formant values of the 9 phonologically distinctive vowels, plus the allophonic variations of /a/ for one speaker are plotted in Figure 7. Harmonically related pairs are noted by the use of the same symbol with and without a minus sign. In each case the putatively [-ATR] vowel has a substantially higher first formant (hence a lower position on the chart) than its harmonic counterpart. Most strikingly, the 'high' vowels /i-, u-/ are placed lower than the 'mid' vowels /e, o/. Narrowing the pharynx raises the first formant, other things being equal. The pair /u, u-/ where F2 is the same are thus quite likely (almost) solely different in pharynx width. The other back vowel pair /o, o-/ shows a smaller than expected F2 difference given the size of the difference between their first formants; a substantial pharynx width difference coupled with a degree of opening of the oral constriction may be inferred. Note that a sloping line can be fit to the vowel set /u, o, a/ and a second roughly parallel lower one to the set /u-, o-, a-/, but a straight line cannot be fit to the set /u, o, o-/ as is possible for Fang /u, o, ɔ/. The pattern for the front vowels suggests a greater interaction of the major features of vowel height and backness with pharynx width (not surprising, as retracting the tongue root is more likely to pull the tongue back and down when the tongue body position is front).

We may now revisit the Ikalanga and Bobe high vowels in Figures 2 and 4. The members of the high vowel pairs /i,ɪ/ and /u,ʊ/ in Bobe have virtually the same second formant values as each other and differ only in F1. The Ikalanga vowel pairs transcribed /i,e/ and /u,o/, which are acoustically equally as high as the Bobe pairs, differ in both F1 and F2. Plausibly, the Bobe vowel pairs differ phonetically in pharynx width, which is consistent with the auditory impression they create, while the Ikalanga pairs differ in height (and to a lesser degree in backness), which is consistent with the auditory impression *they* create. Although these acoustic measurements are suggestive, it should be borne in mind that inferences from simple formant measures concerning vowel articulation must be made with caution.

Nasalized vowels

Distinctive nasalized vowels are not particularly common in the Bantu languages but are found in certain mostly western areas, for example in Ngungwel (B.72a) in the Teke group (Paulian 1994), in UMbundu (R.11, Schadeberg 1982), in Bajele (Renaud 1976), and in a few words in the Bitam dialect of Fang (Medjo Mvé 1997). As is generally the case cross-linguistically, there are fewer distinct nasalized vowels than oral ones, at least in lexical stems. In Ngungwel (B72a), there are three oral and three nasalized vowels in prefixes [e, a, o; ẽ, ã, õ]. Lexical stems have a system of seven oral vowels but only five nasalized vowels. Nasalized vowels in the stem are reported to have the qualities [ɪ̃, ɛ̃, ã, ɔ̃, ũ] and to be invariably long (Paulian's article does include a few words with short nasalized vowels in stems; these may be misprints). There are thus seven *phonetic* qualities among the nasalized vowels, but no contrast between all seven in any environment. Examples are given in Table 2.

Table 2. Nasalized vowels in ŋgùŋgwèl

Vowel	Word	Gloss
ɪ̃ː (stem)	dzɪ̃̃ɪ̃	*tooth*
ẽ (prefix)	ẽběl	*kolanut trees*
ɛ̃ː (stem)	ntsíɛ̃ɛ̃	*horn*
ã (prefix)	ãpăb	*wing*
ãː (stem)	bã́ã̂	*children*
ɔ̃ː (stem)	ekúɔ̃ɔ̃	*broom*
õ (prefix)	ṍkáa	*woman*
ũː (stem)	ŋkṹṹ	*name*

A role for vowel nasalization in the transmission of nasal consonant harmony (Greenberg 1951, Hyman 1995) across intervening vowels seems likely in the history of Bantu.

3. Consonants

Consonant overview

Most Bantu languages are reported as having two series of plosives, voiced and voiceless, and this is the standard PB reconstruction (Meeussen 1969). Except in post-nasal environments and sometimes before the *super-high vowels, the reconstructed voiced plosives most commonly correspond to voiced continuants of one type or another, or to implosives, in the modern languages (an alternation of some kind is probably to be reconstructed to an early stage, possible even pre-Bantu). Bantu orthographies usually do not indicate these alternations, unless subsequent developments have created a contrast between, say, /b/ and /ɓ/, or /b/ and /β/. This illustrates one instance where the occurrence of cross-linguistically less common phonetic segments may be disguised by notational practices. Aspiration is a contrastive property of voiceless stops (and affricates) in some languages, principally in Eastern and Southern languages, where it is often a reflex of an earlier voiceless prenasalized stop. Engstrand and Lodhi (1985) provide one of the few phonetic discussions of aspiration in a Bantu language (Swahili).

Most Bantu languages have a full set of nasals at each place of articulation where a stop or affricate appears, but often intricate (morpho)phonological processes govern nasal/oral alternations and syllabification and other prosodic processes concerning nasals. Most of the languages have relatively limited sets of fricatives of the cross-linguistically common types, although lateral fricatives (and affricates) have developed in or been borrowed into a number of the southern languages, such as Sotho, Xhosa and Zulu. Particularly striking in this connection is the velar ejective lateral affricate [kʟ'] of Zulu, which is auditorily reminiscent of a lateral click. There is often only one liquid: /l/, /ɾ/ or /r/ (though Chaka E.62 is among those with more, see Davey et al 1982, Philippson and Motluhac, this volume). The two vocoid approximants /j, w/ occur in many languages, often alternating with high vowels /i, u/.

Although most Bantu languages use only one coronal (typically alveolar) and one dorsal (velar) place of articulation, contrasts between dental and alveolar places are found in Mijikenda and coastal dialects of Swahili (see Hayward et al 1989) and between velars and uvulars in Qhalaxarzi (S.31d, Dickens 1987). Consonant gemination has developed through internal processes in languages such as Ganda (E.15), and by contact with Cushitic languages in Ilwana (E.701).

In the rest of this section two of the particular issues of phonetic interest, the possible occurrence of articulatorily complex consonants, and the nature of the so-called "whistling fricatives" mentioned in the section on vowels will be briefly discussed. Longer sections of the chapter will be devoted to aspects of laryngeal action in consonants, to the description of clicks and their distribution in Bantu, and to some of the interesting aspects of nasality which occur in these languages.

Complex or simple consonants?

Doubly-articulated labial-velar stops (and nasals) are found almost exclusively in the languages of Africa, but they occur in only relatively few of the Bantu languages, including Londo (A.11, Kuperus 1985), 'Sawabantu' (Mutaka and Ebobissé 1996/7), Fang (Medjo Mvé 1997), and Mijikenda (E.70, Nurse and Hinnebusch 1993, Kutsch Lojenga 2001) among others. However, from the phonetic point of view, the Bantu languages have fewer articulatorily complex consonants than is sometimes suggested. An interesting process of intensification of secondary articulations into obstruents occurs, inter alia, in Nyarwanda (Jouannet 1980) and Shona. This process does not result in double articulations that are almost totally overlapped, as in labial-velars, but sequential articulations which are overlapped either not at all, or no more than is typical of sequences such as /tk/ or pk/ in English words like 'fruitcake' or 'hopkiln'. On the other hand it does produce rather unusual consonant sequences in onset positions.

Examples of the Nyarwanda strengthening of an underlying /u/ or /w/ into a velar stop after a non-homorganic nasal or stop are illustrated by the spectrograms in Figures 8-10. As these show, the first segment is released before the closure for the second is formed. When the sequence is voiced, as in /mg, bg/, a quite marked central vocoid separates the two segments. When the sequence is voiceless, as in /tk/, there is a strong oral release of the first closure. There is no overlap in the closures for the two segments, except optionally in the case of the nasal sequence /mŋ/. Somewhat similar facts have been shown for the Zezuru dialect of Shona (Maddieson 1990); however, as was observed long ago by Doke (1931a,b) the phonetic patterns vary quite considerably across the different varieties of Shona.

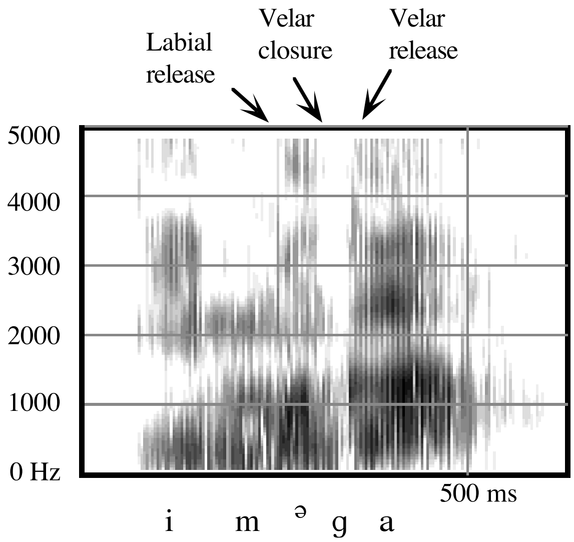

Figure 8. Spectrogram of Nyarwanda *imwa* /imga/ ‘dog’ spoken by a male speaker.

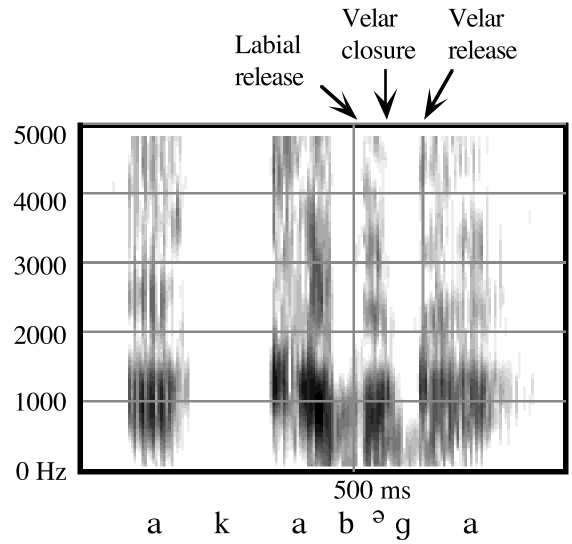

Figure 9. Spectrogram of Nyarwanda *akabwa* /akabga/ ‘dog (diminuative)’; same speaker as Figure 8.

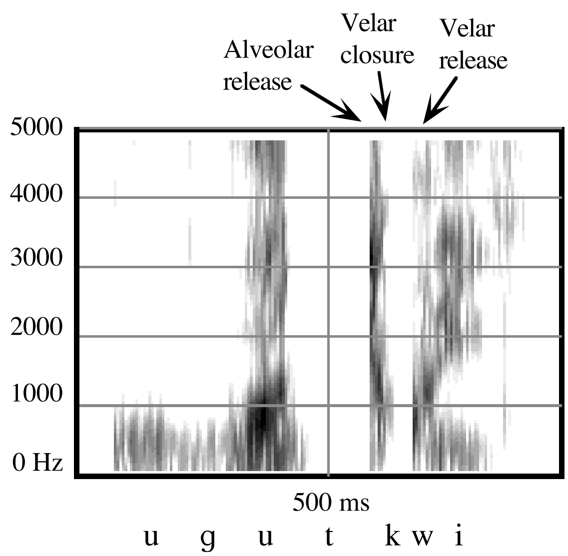

Figure 10. Spectrogram of Nyarwanda *ugutwi* /ugutkwi/ ‘ear’; same speaker as Figure 8.

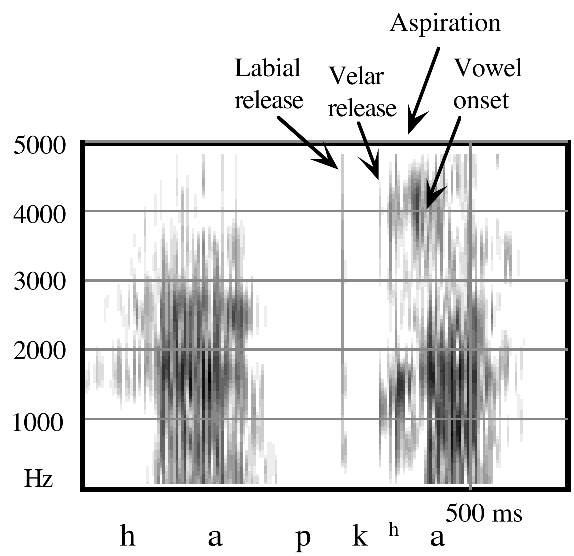

Figure 11. Spectrogram of Ikalanga /hapk^{h}a/ ‘ampit’, spoken by a female speaker from Zimbabwe.

A particularly interesting claim is made by Mathangwane (1999) concerning her pronunciation of parallel forms in Ikalanga. She suggests that elements like the /pk/ which evolves from earlier or underlying /pw/ are pronounced with almost fully overlapped closures and their duration is similar to that of simple /k/ and /p/ segments, i.e. they are [p͡k, b͡g]. She reports that the labial closure is formed first — this would therefore be an important counter-example to the normal pattern found in labial-velar doubly-articulated segments in other languages in which the labial closure is formed very slightly later (10-15 ms) than the velar one. The one spectrogram of a word containing /pk/ published in this study actually shows that the duration of the element is considerably longer than a simple stop, suggesting it contains a sequence of articulations, although no burst is visible for the /p/. Recordings made by the author of two other speakers of Ikalanga, one from Francistown in Botswana and one from Palmtree in Zimbabwe, did not replicate the pattern suggested by Mathangwane. For example, the word meaning 'armpit', transcribed by Mathangwane as [ɦap͡k^{h}a], could receive three pronunciations — [hak^{h}wa] with no labial closure, [hapxa] with a labial stop followed by a fairly long velar fricative, or [hapk^{h}a] with a sequence of stops with clearly separate releases, as illustrated in Figure 11. This third pronunciation was characterised by one of the speakers as being more typical of speakers of 50 or more years of age. Evidently more study of the phonetic and sociolinguistic variation in this area would be of great interest.

"Whistling" fricatives

Shona and Ikalanga are also marked by the occurrence of a type of labialization co-produced with alveolar fricatives which have led to these segments being named "whistling fricatives" (Doke 1931b, Bladon et al 1987). Unlike 'ordinary' labialization, which involves rounding and protrusion of the lips accompanied by a raising of the tongue back, i.e. a [w]-like articulation, this labialization involves primarily a vertical narrowing of the lips with little or no protrusion and no accompanying tongue back raising. The acoustic effect is to concentrate the frication noise in a relatively narrow frequency band which is indeed somewhat reminiscent of a whistle. The gesture is also timed differently from 'ordinary' labialization in that it covers the fricative duration rather than being primarily realized as an offglide; hence "whistling fricatives" can themselves be labialized in their release phase. Similar segments are very rare in the world's languages but do occur in the Dagestanian language Tabarasan.

4. Laryngeal action in consonants

Implosives and ejectives

Languages of the northwest, the eastern coastal area and the south-east often have at least one implosive, most frequently a bilabial, but implosives are generally absent in the languages of the Congo basin and southwest. Ejective stops and affricates are more rarely found in the Bantu languages, although they occur as variants of the unaspirated voiceless stops in languages of the south, especially in post-nasal contexts. The ejection is generally weak compared to that found in languages of the Afro-Asiatic family, except for Ilwana where the ejectives are in borrowed Cushitic vocabulary, and the ejective lateral affricate of Zulu mentioned earlier.

The segments labeled as implosives are sometimes described as if a glottal constriction is characteristic of their production. In Bantu, this is typically not the case; the vocal folds are in the normal position for voicing. Rather, what is critical is that the larynx is lowering during their production, so that the size of the supralaryngeal cavity is being enlarged while the oral closure is maintained. This may have two principal effects — firstly, it allows the amplitude of vocal fold vibration to increase during the closure, giving a particularly strong percept of voicing at the time of the release, and secondly it may mean that the intra-oral pressure is relatively low at the time when the closure is released so that at the moment of release the initial airflow is ingressive (Hardcastle and Brasington 1978). The waveform of an intervocalic bilabial implosive in GiTonga (S.62) is shown in Figure 12. Dashed vertical lines mark the onset and offset of the bilabial closure. Voicing is continuous through the closure; upper and lower lines have been constructed on the figure linking respectively the positive and negative peaks in the waveform in order to dramatize the growing amplitude of the voicing during the closure.

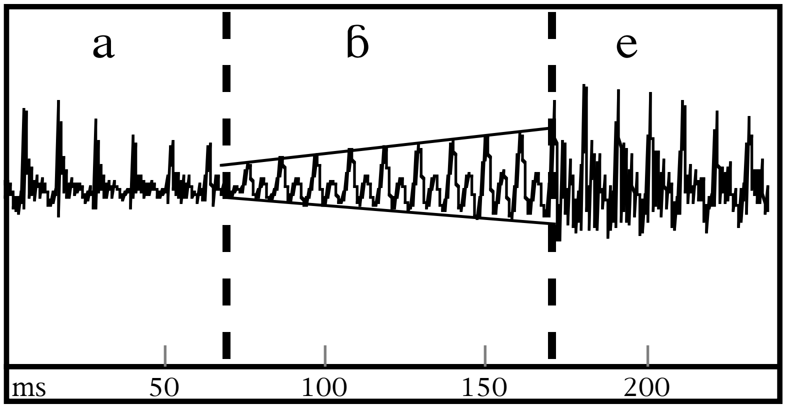

Figure 12. Waveform of the middle part of the GiTonga word /ɓàɓé/ 'father', illustrating the increasing amplitude of voicing during the implosive.

Depressor consonants

Another special laryngeal action occurs in the 'depressor' consonants which are characteristic of certain Bantu languages of the eastern and southern regions. This term was originally applied to consonants which have a particularly salient lowering effect on the pitch of the voice in their neighborhood (Lanham 1958). It has since sometimes come to be used for any consonant which has any local lowering effect on pitch (or, more accurately, on the fundamental frequency of vocal fold vibration, abbreviated Fo), such as an ordinary voiced plosive, and has even been used for those which may simply block a raising or high-tone spreading process. However, the original notion of a depressor consonant is quite different from this expanded use. The most detailed study remains that of Traill et al (1987) on depressor consonants in Zulu. This study shows that the Fo associated with depressors is lower than a low tone, and the lowest pitch is centred on the depressor consonants themselves. At vowel onset the Fo difference between High and Low tones after a set of non-depressor consonants is 22 Hz; but a High tone onset after depressor consonants is 44 Hz lower than after the non-depressors, and a Low after depressors is 23 Hz lower than after non-depressors (mean across three speakers, two male and one female). Thus a High after a depressor begins considerably lower than a Low elsewhere. Figure 13 compares the pitch contours of the Swati words /líhala/ 'aloe' and /líh̤álà/ 'harrow', where / ̤ / is a diacritic to mark the fact that the consonant is a depressor in the second word. Despite the fact that the lexical tone after the depressor is high (Rycroft 1981), the onset Fo is about 30 Hz lower than the low tone onset after the non-depressor, and a rapid pitch fall begins during the vowel which precedes the depressor.

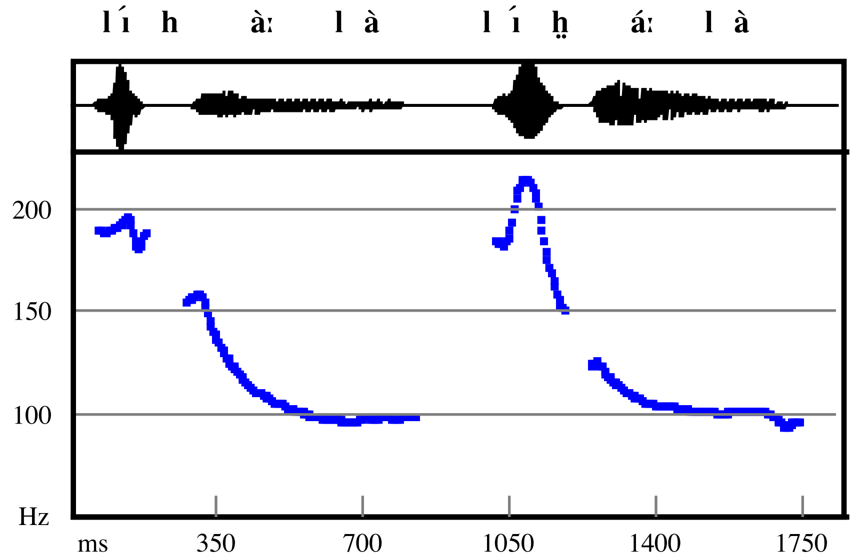

Figure 13. Pitch contours illustrating effects of non-depressor and depressor /h/ in Swati (male speaker).

Figure 13 also illustrates the fact that depression is not necessarily associated with voicing as both /h/ and /ḥ/ are voiceless. Equally, voiced segments such as nasals and approximants may contrast in depression (see also Traill and Jackson 1988, Wright and Shryock 1993, Mathangwane 1998). Since these segments make for easy tracking of Fo through the consonant, the centering of the depression on the consonant can be most easily visualized with them. Two examples from Giryama (E.72a) are illustrated in Figure 14. In these cases there is a substantial fall in Fo from the onset to the middle of the nasal, and pitch begins to rise before the consonant is released; the pitch peak on the vowel is 40Hz (left panel) or 50Hz (right panel) higher than the lowest pitch in the nasal. In these words there is noticeably breathy phonation during part of the consonant and at the vowel onset (transcribed [ɦ]), but breathiness is not an invariable accompaniment of depression as had been proposed by Rycroft (1980). Following Traill et al (1987), we understand true ‘depression’ to consist of a special laryngeal posture consistent with very low pitch co-produced with the consonant it is associated with. This gesture may become associated with any class of consonants and thus is capable of becoming itself an independent phonological entity deployed for grammatical effect as in the ‘depression without depressors’ described by Traill (1990).

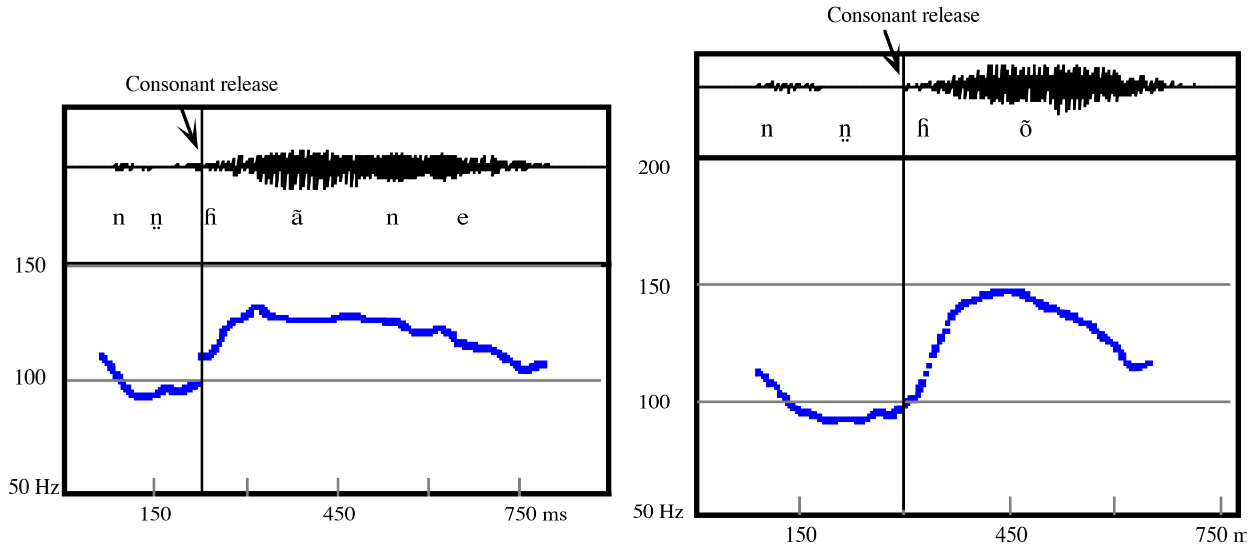

Figure 14. Pitch effects of depressor nasals in the Giryama words /nhane/ 'eight' (left panel) and ideophone /nho/ (right panel). Recording courtesy of Constance Kutsch Lojenga.

5. Clicks

Clicks are unique to Africa as speech sounds and as they are of unusually great phonetic interest they will be described in some detail. They have been incorporated into the consonant inventories of a number of languages in the southern part of the Bantu area, indisputably as a result of contact with Khoisan languages. The languages concerned are separated into two geographical clusters. In the Southeast, the core is formed by the languages of the Nguni group (S.40), especially Zulu which has 12 click consonants, and Xhosa which has fifteen, as also does Phuthi. More far-flung Nguni varieties, such as Zimbabwean Ndebele, Malawian Ngoni and Tanzanian Ngoni have (or had) fewer clicks (Doke 1954, Ziervogel 1959). Clicks also are found in Southern Sotho (Guma 1971), in some speech varieties of the Tsonga group (S.50) such as Nkuna, Dzonga and Ronga (Doke 1954, Baumbach 1974, Afido et al 1989), and, more marginally, in the Ndau dialect of Shona in Mozambique (Afido et al 1989). They do not occur in the Shona varieties of Zimbabwe, nor in Northern Sotho (Poulos and Lowrens 1994), Venda, Ikalanga, Tswa, Chopi, GiTonga and other languages neighboring or closely related to those in which clicks are found.

In the Southwest, the largest number of clicks is found in Yeyi (Gowlett 1997, Sommer and Vossen 1992), which has at least ten, and maybe as many as 20 (accounts are quite variable, only in part because of dialectal differences). Clicks are also found in Gciriku (including in the ethnonym itself), Shambiu, Kwangali, Mbukushu, Fwe, and Mbalangwe (Möhlig 1997, Baumbach 1997). They are not found in languages of the Wambo cluster (R20) such as Kwanyama, Mbalanhu and Ndonga, nor in Herero, UMbundu, Totela, Subiya, Lozi, Kgalagadi, Tswana and other languages spoken near the area where the borders of Namibia, Angola, Botswana and Zambia meet.

This distribution in shown approximately by the map of Southern Africa in Figure 15, prepared with the Bantu Mapmaker software (Lowe and Schadeberg 1997). Languages with clicks are shown stippled — heavily for those with more elaborate click inventories, and lightly for those with small click inventories. Neither the dispersed Nguni languages (Ndebele and Ngoni) nor Ndau are included here. The separation of the areas as well as the distinct nature of the vocabularies involved (Louw 1975, Sommer and Vossen 1992) indicate that borrowing of clicks happened independently in these two areas. In the southeast it is reasonably certain that clicks spread from Nguni to the neighboring languages, and in this process the number of click distinctions was reduced (in fact, click spreading is an actively ongoing process with clicks penetrating into Ndau, urban varieties of Northern Sotho and beyond; like most linguistic maps, this map represents a somewhat fictitious ethnographic idealization not corresponding precisely with any exact time or population distribution). In the southwest it is less certain that clicks were borrowed first into one language (in this case it would be Yeyi) and spread from there, but this scenario seems quite possible in view of the reduced inventories of clicks found in the other languages. It is noteworthy that none of the Bantu languages of East Africa have acquired clicks from the surviving or former languages of this area with clicks (Maddieson, Ladefoged and Sands 1999).

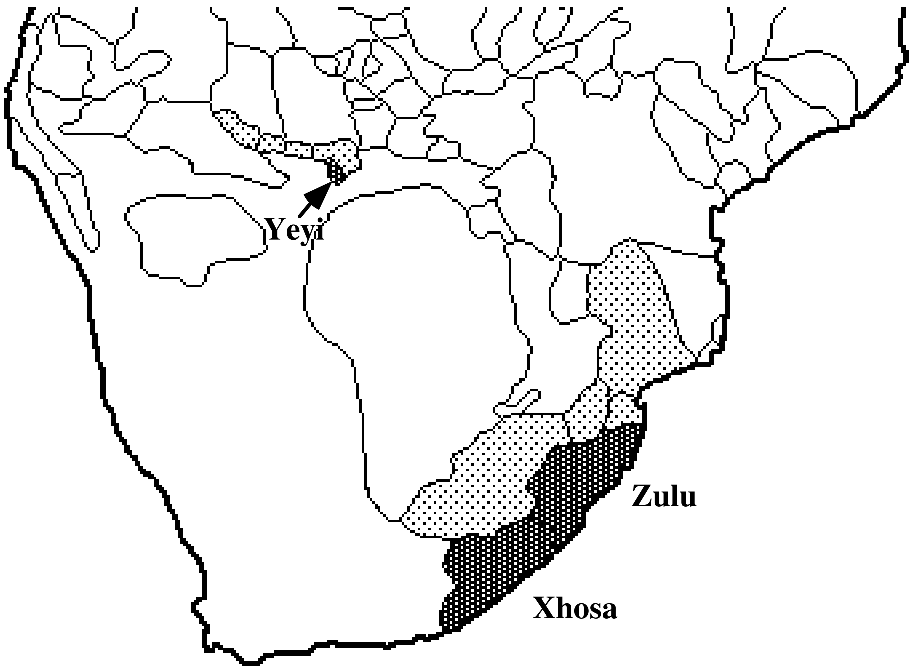

Figure 15. Bantu languages with clicks (see text for details).

The mechanism of producing clicks is now well understood. A closure in the vocal tract is formed by the back of the tongue contacting the roof of the mouth in the velar or uvular area, a second closure is then formed in front of the location of this closure by the tip or blade of the tongue or the lips. This entraps a small quantity of air between these two closures. The center portion of the tongue is then lowered while the two closures are maintained, enlarging the volume of the space between them. Consequently, the pressure in the air inside this space is reduced well below that of the air outside the mouth. Next, the closure at the front of the mouth is released and the abrupt equalization of air pressures inside and outside the mouth results in a sharp acoustic transient. Finally, the back closure is released, and this release may be separately audible. A pronunciation of a click where the velar release is clearly detectable is illustrated in figure 16.

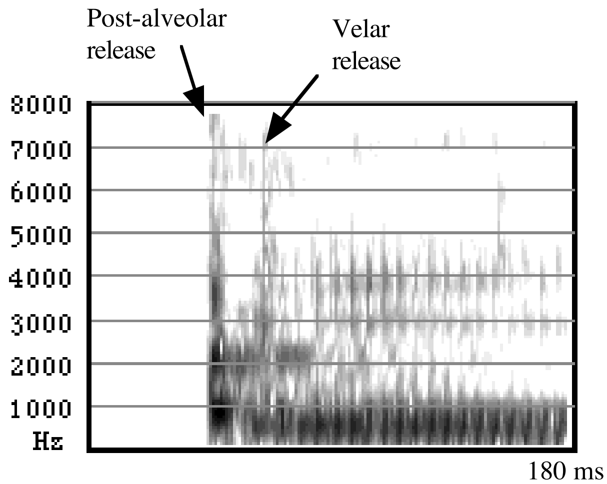

Figure 16. Spectrogram of the first part of the Ndebele word [!kòːβá] "slice!" showing an unaspirated post-alveolar click in onset position. The interval between the two releases is 20 ms.

The basic click mechanism does not determine what the larynx is doing while these movements are taking place in the oral cavity, nor whether the the velum itself is raised or lowered to block or permit air from the lungs to flow out through the nose. Thus, a click can be accompanied by simple glottal closure, by modal or breathy voicing, by open vocal folds, or by use of the ejective mechanism. Changes in larynx activity can be variously timed in relation to the action in the oral cavity, and to the timing of movements raising and lowering the velum. The possible variations are thus very numerous, and many different categories of individual clicks are found when all the languages which use them are considered (Ladefoged and Maddieson 1996).

In describing clicks it is customary to talk of the click type and the click accompaniment. The click type refers to the location of the front closure and the manner in which it is released, which may be abrupt or affricated, central or lateral. The accompaniment refers to all the other aspects of the click — laryngeal action and timing, nasal coupling, and the location (uvular or velar) and manner of release (abrupt or affricated) of the back closure. In Bantu languages, no more than four click types are used, and probably no more than seven accompaniments are found. Zulu and Xhosa have affricated dental, affricated alveolar lateral and non-affricated apical post-alveolar click types. The last of these was often described as 'palatal' in older literature. Yeyi has these three as well as a laminal post-alveolar type (variously called 'alveolar' or 'palatal' in different sources).

Thomas Vilakazi's analysis of Zulu click types (Thomas [Vilakazi] 1998, 2001), combining insights from acoustic, aerodynamic and electropalatographic techniques, is by far the most detailed study of click production in a Bantu language. Thomas Vilakazi confirms that the velar closure always precedes the front closure; this accounts for the fact that nasals preceding clicks assimilate in place to velar position, and corrects a misobservation by Doke (1926), who believed the front closure was formed first (the velar closure *must* be released after the front closure for the click mechanism to work, but it could in principle be formed later). Velar closures for all three click types in Thomas Vilakazi's data are held for about 175 milliseconds, but the front closures show some significant timing differences. The front closure for dental clicks is formed earlier and held longer (about 105 ms) than that for post-alveolar or lateral clicks (about 80 ms). The relative timing and durations of velar and front closures deduced from acoustic and aerodynamic data are graphed in Figure 16.

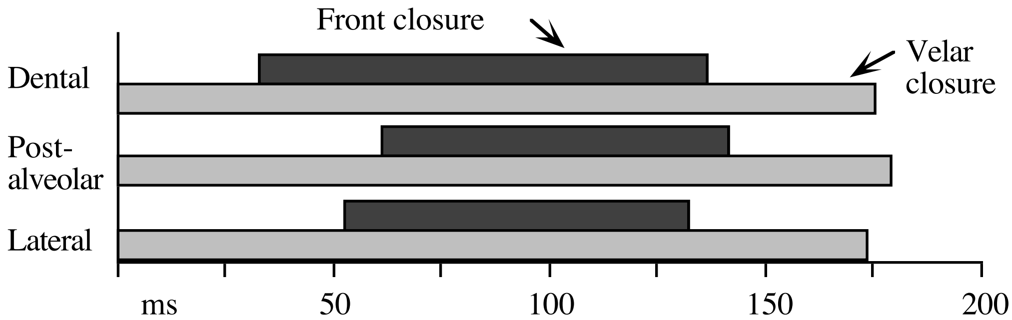

Figure 16. Closure durations and timing relations in the three click types of Zulu (means for voiceless clicks in three vowel environments spoken by three speakers, from Thomas 1999). Front closure durations are shown as heavily stippled bars.

More details on the articulations of clicks are given by electropalatography (EPG). Speakers wear a thin custom-made acrylic insert moulded to the shape of their upper teeth and hard palate in which a number of electrodes are embedded which sense contact between the tongue and the roof of the mouth. In Thomas Vilakazi's study, inserts with 96 electrodes were used, together with software allowing a sweep of the contact patterns to be made every 10 ms. The articulatory contacts can then be examined using stylized displays such as those in figures 17-19, which represents the arc of the teeth and the vault of the palate. Contacted electrodes are shown as black squares and uncontacted ones as grey dots. Figure 17 shows the production of a dental click. The first frame, numbered 0, is close to the time that velar closure is first made, as detected from the accompanying acoustic record. Because the insert does not cover the soft palate, this closure cannot be observed on the EPG record at this time. The seal around the inside of the teeth is made by 40 ms later, and as the contact area of the back of the tongue enlarges, the front edge of the velar contact is now visible as a line of contacted electrodes at the bottom of the arc. The closures overlap for 100 ms, until frame 140. During this time rarefaction is occurring. This figure makes clear that the expansion of the cavity is not due to moving the location of the back closure further back, as phoneticians have sometimes suggested. That dental clicks are produced with controlled affrication is also clear from the way the front release initially involves formation of a narrow channel, clearly visible in frame 150.

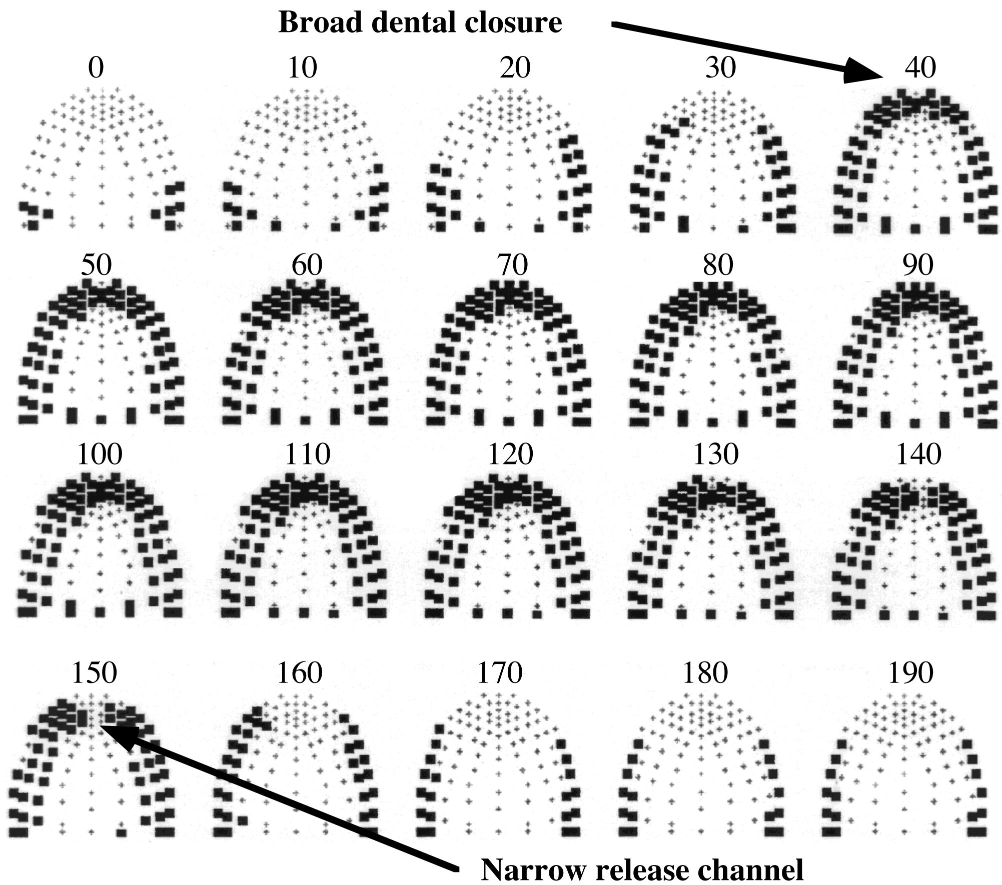

Figure 17. EPG frames showing a dental click spoken by a male Zulu speaker (from Thomas Vilakazi 1999).

Production of a lateral click is illustrated in Figure 18. In this particular token there is a long lag between the time the velar closure is made and when the front closure is sealed, about 80 ms later. The contact of the front of the tongue is asymmetrical, as the side of the tongue opposite to where the release will be made is braced contra-laterally against the palate. The release of a lateral click is also affricated, occurring initially through a narrow channel quite far back, as shown in frame 170 and continuing in frame 180. In contrast to these two affricated click types, a post-alveolar click is released without affrication. As figure 19 shows, the shift from sealed to open occurs rapidly and completely, here between the two frames numbered 170 and 180. These frames also illustrate the retraction of the tongue tip which occurs just before release of this click type. From frame 150 through to frame 170 the contacted area moves back, so that the configuration at the moment of release is clearly post-alveolar.

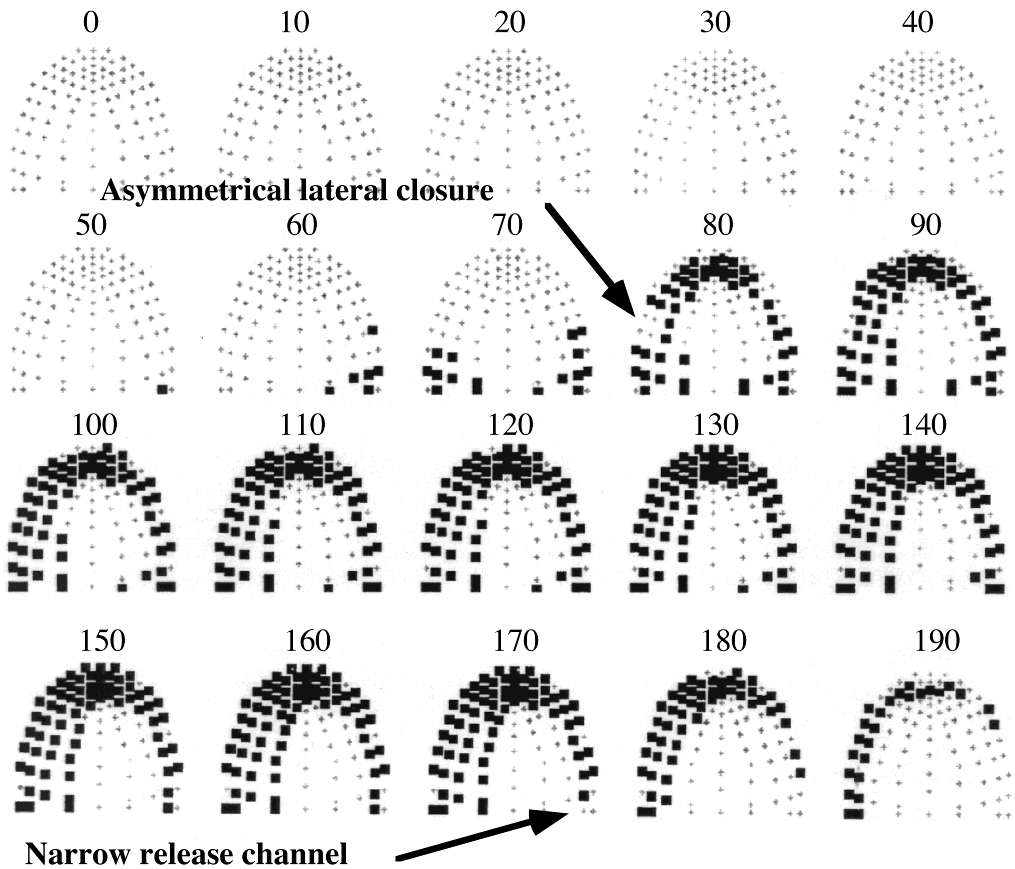

Figure 18. EPG frames of a lateral click spoken by a male Zulu speaker (from Thomas Vilakazi 1999).

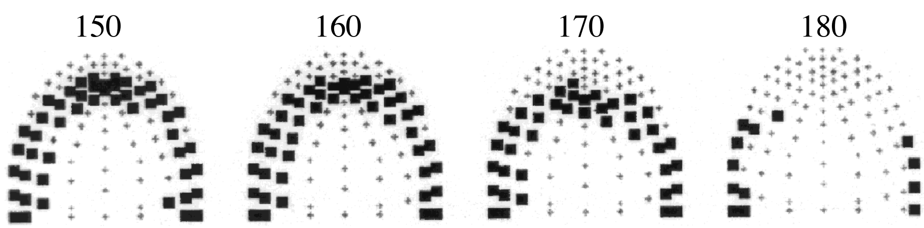

Figure 19. EPG frames showing the releasing phase of a post-alveolar click spoken by a male Zulu speaker (adapted from Thomas Vilakazi 1999).

During the time period in which the two closures of a click overlap, lowering of the center of the tongue creates a partial vacuum in the cavity between them. Thomas Vilakazi's work provides the first direct measures of how powerful the energy generated by this gesture is. Air pressure in the oral cavity is measured in relation to the ambient atmospheric pressure in hectoPascals (hPa, equivalent to the pressure required to support 1 cm of water). For an ordinary pulmonic stop, peak pressure behind the closure ranges between about 5 and 20 hPa, depending on the loudness of the voice. The peak negative pressures reached in clicks are typically -100 hPa or more and may reach over -200, as shown in Figure 20. Post-alveolar clicks have the greatest rarefaction, lateral clicks the least — perhaps because the contra-lateral bracing of the tongue in the lateral clicks may constrain the amount of tongue-center lowering that is possible. Thomas Vilakazi's aerodynamic data also reflect the different dynamics of the affricated and abrupt clicks. The equalization of internal and external pressure at release occurs much more quickly in post-alveolar clicks than for dental and lateral clicks. This can be shown by calculating the average rate of pressure change over this phase of the click, which is 14.4 hPa/ms for post-alveolars, 7.9 for dentals, and 4.2 hPa/ms for laterals. Only a small part of this difference can be accounted for by the difference in peak pressure between the click types.

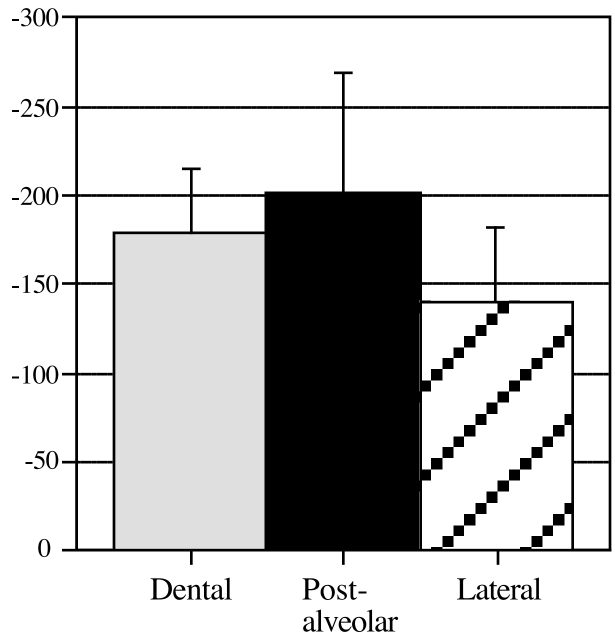

Figure 20. Peak negative pressure in the three click types of Zulu (means for voiceless clicks in three vowel environments spoken by three speakers, from Thomas Vilakazi 1999).

Zulu has four different accompaniments to its three click types, voiceless unaspirated, voiceless aspirated, voiced and voiced nasalized. The IPA recommends that these be noted in each case by writing a symbol for a dorsal consonant (velar in this case) before the click type symbol, so that click consonants are always represented by at least a digraph, and may require additional symbols or diacritics. The correspondence between IPA transcription and Zulu orthographic symbols is given in Table 3. Xhosa has five accompaniments, three of which are the same as in Zulu. The 'voiced' clicks are breathy voiced, and there is a distinct breathy voiced nasalized accompaniment; these two series are 'depressor consonants'. Yeyi has a click accompaniment described by Gowlett as 'ejective' but represented phonetically by digraphs with [ʔ]. This could indicate that there is glottal closure but no ejection.

Table 3. IPA symbols and Zulu orthography for clicks

	IPA transcription			Orthography		
	dental	post-alveolar	lateral	dental	post-alveolar	lateral
voiceless unaspirated	k†	k˜	kK	c	q	x
voiceless aspirated	k†õ	k˜õ	kKõ	ch	qh	xh
voiced	g£†	g£˜	g£K	gc	gq	gx
voiced nasalized	N†	N˜	NK	nc	nq	nx

One of the most striking things about clicks in Bantu is the lack of respect for place distinctions when few categorical contrasts exist. In Nkuna (Tsonga) Baumbach (1974) indicates that clicks are indifferently pronounced as dental or post-alveolar. In Mbukushu (K.333) the one series of clicks is reported to be pronounced “either as dental, palatal or [post-]alveolar sounds” (Fisch 1977). In Gciriku (K.332), clicks are “mostly dental, however, with a broad individual variation” (Möhlig 1997). Such context-free liberty to vary place of articulation is rarely encountered with other classes of consonants.

6. Nasals and nasality

The special phonetic interest of consonantal nasality in the Bantu languages involves principally the prenasalized segments and the realization of ‘voiceless’ nasals. In both cases aspects of timing are particularly relevant. Detailed studies of timing in prenasalized stops are included in Maddieson (1993), Maddieson and Ladefoged (1993), and Hubbard (1994, 1995). Using data from these sources, figure 21 compares the durations of nasals and voiced prenasalized stops as well as of the vowels that precede them in two languages, Ganda and Sukuma. In both languages the oral stop duration in voiced prenasalized stops is very short, so the total segment duration is not so very different from that of a simple nasal. Both languages have contrasts of vowel quantity and compensatory lengthening of vowels before prenasalized stops. But there are interesting differences between the two. Lengthened vowels are much closer in duration to underlying long vowels in Ganda than they are in Sukuma. Sukuma lengthened vowels are almost exactly intermediate between underlying short and long vowels, and the nasal portion is quite long. Nyambo (E.21) is similar to Sukuma in its pattern, and Hubbard suggests that the difference from Ganda is related to the fact that lengthened vowels count in a different way in tone assignment rules in these languages.

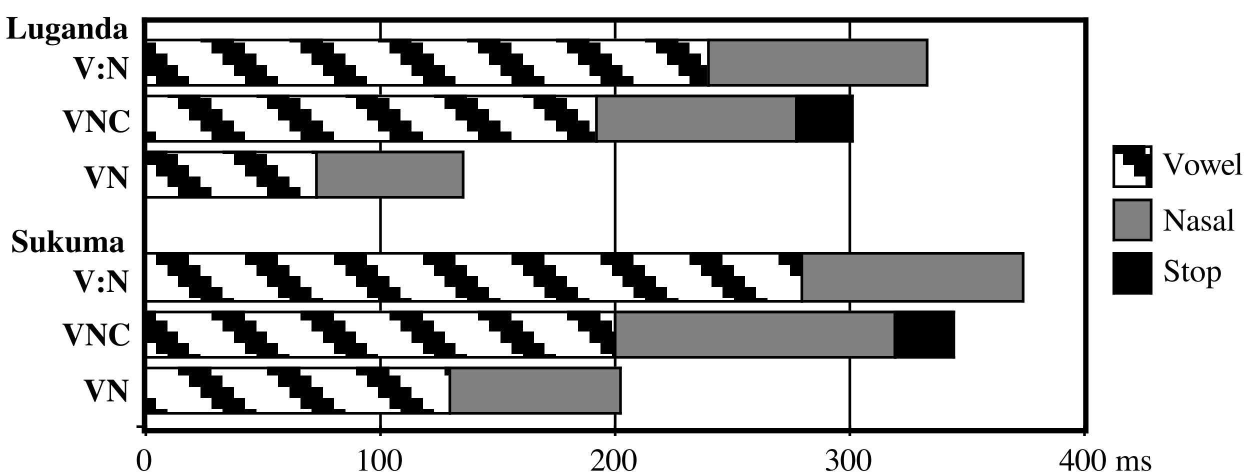

Figure 21. Comparison of selected vowel and consonants lengths in Ganda and Sukuma (see text for explanation).

Hubbard also compared the durations of vowels in three further languages with different patterns. The mean results are given in Table 4. Ndendeule has no long vowels and no lengthening. Yao (P.21) has a long/short contrast and significant compensatory lengthening so that vowels before prenasalized stops are as long as underlying long vowels and have more than double the duration of short vowels. CiTonga (M.64) has long vowels but does not show any compensatory lengthening before NC. This difference seems to be related to the different origin of long vowels; Yao maintains Proto-Bantu vowel length distinctions (and adds to them). CiTonga does not preserve Proto-Bantu vowel length, but has developed long vowels from intervocalic consonant loss.

Table 4. Comparison of mean vowel durations in three languages, one speaker per language (after Hubbard 1995).

Vowel duration in ...	Ndendeule	Yao	CiTonga
...CVC	148	61	100
...CVNC	146	130	101
...CV:C	—	132	241

In several areas earlier voiceless prenasalized stops have developed into voiceless nasals or related types of segments, including in Sukuma (Maddieson 1991), Pokomo (E.71) and Bondei (G.24) (Huffman and Hinnebusch 1998), Ikalanga (Mathangwane 1998), and Nyarwanda (Demolin and Delvaux 2001). Aspects of the original sequencing of nasal + oral and voiced + voiceless portions found in prenasalized stops are sometimes retained, and small variations in the timing and magnitude of the different component gestures create quite large variability in the acoustic pattern of these segments as critical alignments are made or missed.

This variation in the realization of 'voiceless nasals' is at least in part correlated with position in a word. Figure 22 shows a spectrogram of the Nyamwezi word /ŋ̥apo/ "basket" spoken in isolation. Dotted vertical lines separate the major phonetic components of the first syllable. The portion marked A, between the first two lines, is phonetically a voiceless velar nasal [ŋ̥]. The second line marks the time-point at which the velar closure is released. The fragment marked B has voiceless oral airflow, with resonances similar to those of the following /a/ vowel. Fragment C is the voiced portion of the vowel /a/. This type of segment might well be described as an 'aspirated voiceless nasal'. Figure 23 shows a realization of a medial instance of the same segment in the word /kɔŋ̥á/ "to suck". In this case there is no consonantal nasality; the nasal feature is realized as nasalization of the latter part of the vowel /ɔ/ in Fragment C, following an oral portion, B, and the aspiration of the initial stop, A. Fragment D, which is the consonantal part of the /ŋ̥/ is voiceless but oral, and as often in an [h]-sound, the transition of the formants of the flanking vowels can be traced through its duration.

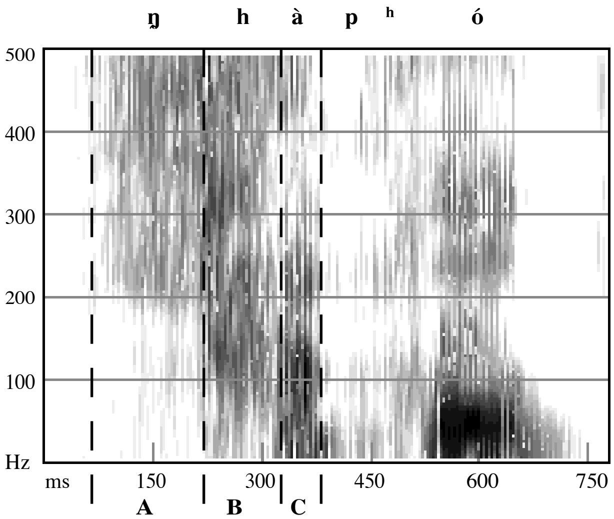

Figure 22. Spectrogram of the Nyamwezi word /ŋ̊apo/ “basket”. See text for discussion of the phonetic segmentation.

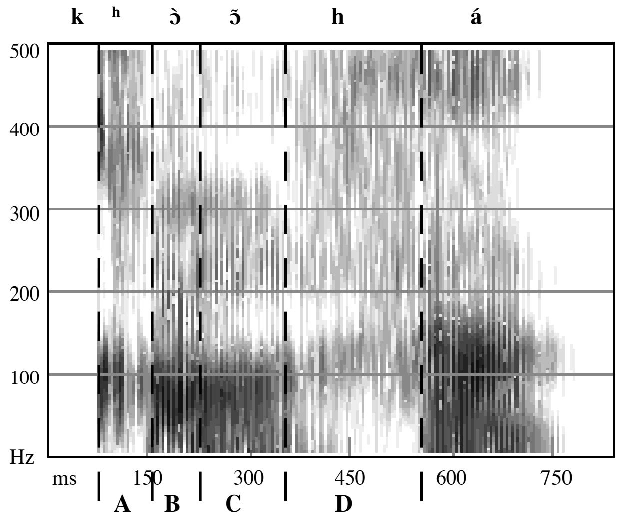

Figure 23. Spectrogram of the Nyamwezi word /kɔŋ̊á/ “to suck”. See text for discussion of the phonetic segmentation.

In Sukuma, the nasal portion of the ‘voiceless nasals’ is often at least partly voiced or breathy voiced, as described in Maddieson (1991), whereas the parallel segments in Nyarwanda are fully voiced (except after voiceless fricatives) but produced with a modified kind of voicing described by Demolin and Delvaux as whispery-voice.

7. Prosodic characteristics

A discussion of Bantu phonetics would not be complete without reference to some of the studies of the major prosodic characteristics of the languages. As is well known, the great majority of the Bantu languages have a tonal distinction of High and Low tones, which often may combine into falling contours. The phonetic shapes of tone sequences of this type can usually be modeled on the basis of the position and height of local H targets, with the Low tones treated as automatically filled valleys between these points. Certain more complex patterns, such as those noted by Hombert (1990) in Fang and Roux (1995) in Xhosa, may require a more elaborate model, and provision would have to made for the special effects of depression.

The most extensive body of work on the phonetics of tone in a single Bantu language concerns Chewa (Carleton 1996, Myers 1996, 1999a, 1999b, Myers and Carleton 1996). Detailed studies of this type not only illuminate the individual language studied but may provide insights into diachronic issues. For example, in Chewa, as is common cross-linguistically, the High pitch peak is realized at the end of the syllable to which it is associated (Myers 1999a). This pattern may form the basis for the frequent shifting of a High tone to a later syllable. These studies also address several issues in the relation between intonation and tone. For example, Myers (1999b) shows that syntactically unmarked yes/no questions are characterized by a slower rate of pitch declination than statements. Carleton (1996) demonstrated that units of paragraph length are organized by long-range patterns of tonal declination and resetting.

Some additional aspects of timing beyond those linked to prenasalization are discussed by Hubbard (1994, 1995). In both Nyambo and Ganda, she finds that the mean duration of a word is more accurately predicted from the number of moras it contains than from the number of syllables it contains. Hence these languages fit best into the typological category of mora-timed languages.

We do not know how generally true this is of Bantu languages, or how this finding relates to the phrase-penultimate syllable lengthening which is a widespread feature of the family, or if different timing relations would be found in those Bantu languages with accent or stress (with or without tone).

Compared to the quantity of work on segmental phonetics, prosodic phonetics is relatively neglected. This in part is explicable by the fact that more appropriate conceptual frameworks for dealing with such aspects have only been developed in recent years. Prosodic phonetic analysis of Bantu languages is likely to provide some of the most exciting work in the next few years.

References

Afido, Pedro, Gregório Firmino, John Heins, Samba Mbuub, Manuel Trinta. 1989. *Relatório do i Seminário sobre a Padronização da Ortografia de Línguas Moçamicanas*. Universidade Eduardo Mondlane, Maputo.

Batibo, Herman. 1985. *Le Kesukuma, Langue Bantu de Tanzanie: Phonologie, Morphologie*. Éditions Recherche sur les Civilisations, Paris.

Baumbach, E.J. M. 1994. *Introduction to the Speech Sounds and Speech Sound Changes of Tsonga*. van Schaik, Pretoria.

Baumbach, E.J. M. 1997. Languages of the Eastern Caprivi. In *Namibian Languages: Reports and Papers*, ed W. H. G. Haacke and E. D. Elderkin. Rüdiger Köppe, Köln: 307-451.

Bladon, A., C. Clark and K Mickey. 1987. Production and perception of sibilant fricatives: Shona data. *Journal of the International Phonetic Association* 17: 39-65.

Carleton, Troi C. 1996. *Phonetics, Phonology and Rhetorical Structuring of Chichewa*. Ph. D. dissertation, University of Texas, Austin.

Connell, Bruce. 2000. Mambila fricative vowels. In *Advances in African Linguistics (Trends in African Linguistics 4)*, ed. V. Carstens and F. Parkinson. Africa World Press, Trenton, 233-249.

Creissels, Denis. 1999. Remarks on the sound correspondences between Proto-Bantu and Tswana (S.31), with particular attention to problems involving *j (or *y), *į and sequences *NC. In J-M. Hombert and L.M. Hyman, eds, *Bantu Historical Linguistics: Theoretical and Empirical Perspectives,* CSLI, Stanford CA: 297-334.

Davey, Anthony, Lioba Moshi & I. Maddieson. 1982. Liquids in Chaga. *UCLA Working Papers in Phonetics* 54: 93-108.

Demolin, Didier and Véronique Delvaux. 2001. Whispery voiced nasal stops in Rwanda. Paper presented at the 32nd Annual Conference on African Linguistics, University of California, Berkeley.

Demolin, Didier, Hubert Ngonga-Ke-Mbembe, and Alain Soquet. 1999. Phonetic characteristics of an unexploded stop in Hendo. *Proceedings of the 14th International Congress of Phonetic Sciences* Vol 2: 1047-1057.

Demolin, Didier, Jean-Marie Hobert, Pierre Ondo, and Christoph Segebarth. 1992. Étude du système vocalique fang par résonance magnétique. *Pholia* 7: 41-59.

Dickens, Patrick. 1987. Qhalaxarzi consonants. *African Studies* 46: 297-305.

Doke, C. M. 1954. *The Southern Bantu Languages*. Oxford University Press for the I.A.I., London.

Doke, Clement M. 1926. *The Phonetics of the Zulu Language (Bantu Studies Special Number)*. University of the Wiwatersrand Press, Johannesburg.

Doke, Clement M. 1931a. *Report on the Unification of the Shona Dialects*. Stephen Austin and Sons, Hertford.

Doke, Clement M. 1931b. *A Comparative Study in Shona Phonetics*. University of the Witwatersrand Press, Johannesburg.

Engstrand, Olle, and Abdulaziz Y. Lodhi. 1985. On aspiration in Swahili: Hypotheses, field observations and an instrumental analysis. *Phonetica* 42: 175-1985.

Fisch, Maria. 1977 [1998]. *Einfuhrung in die Sprache der Mbukushu, Ost-Kavango, Namibia*. S.W.A. Wissenschafliche Gesellschaft, Windhoek. Translated as *Thimbukushu Grammar*, Out of Africa Publishers, Windhoek, 1998.

Gowlett, Derek F. 1997. Aspects of Yeyi diachronic phonology. In *Namibian Languages: Reports and Papers*, ed W. H. G. Haacke and E. D. Elderkin. Rüdiger Köppe, Köln: 235-263.

Greenberg, Joseph H. 1951. Vowel and nasal harmony in Bantu languages. *Zaire* 5: 813-820.

Guma, S. M. 1971. *An Outline Structure of Southern Sotho*. Shuter and Shooter, Pietermaritzburg.

Guthrie, Malcolm. 1967-71. *Comparative Bantu, Volumes 1-4*. Gregg, Farnborough.

Hardcastle, W. J. and R. W. P. Brasington. 1978. Experimental study of implosive and voiced egressive stops in Shona: an interim report. *Work in Progress (Phonetics Laboratory, University of Reading)* 2: 66-97.

Hayward, K. M., Y. A. Omar and M. Goesche. 1989. Dental and alveolar stops in Kimvita Swahili: An electropalatographic study. *African Languages and Cultures* 2: 51-72.

Hombert, Jean-Marie. 1990. Réalisations tonales et contraintes segmentales en fang. *Pholia* 5: 105-111.

Hubbard, Kathleen. 1994. *Duration in Moraic Theory*. Ph D. dissertation. University of California, Berkeley.

Hubbard, Kathleen. 1995. Toward a theory of phonological and phonetic timing: evidence from Bantu. In *Phonology and Phonetic Evidence: Papers in Laboratory Phonology 4*, ed by B. A. Connell and A. Arvaniti: 168-187.

Huffman, Marie K. and Thomas J. Hinnebusch. 1998. The phonetic nature of the “voiceless” nasals in Pokomo: Implications for sound change. *Journal of African Languages and Linguistics* 19: 1-19.

Hyman, Larry M. 1995. Nasal consonant harmony at a distance. *Studies in African Linguistics* 24: 5-30.

Hyman, Larry M. 1999. The historical interpretation of vowel harmony in Bantu. In J-M. Hombert and L.M. Hyman, eds, *Bantu Historical Linguistics: Theoretical and Empirical Perspectives,* CSLI, Stanford CA: 236-295.

IPA 1999. *Handbook of the IPA*. Cambridge University Press.

Jouannet, François. 1983. Phonétique et phonologie: le système consonantique du kinyarwanda. In *Le Kinyarwanda: Études Linguistiques*, ed. F. Jouannet. SELAF, Paris, 55-74.

Kuperus, Juliana. 1985. *The Londo Word: its Phonological and Morphological Structure*. Koninklijk Museum voor Midden-Afrika, Tervuren.

Kutsch Lojenga, Constance. 2001. The two v’s of Giryama. Paper presented at the 32nd Annual Conference on African Linguistics, University of California, Berkeley.

Labroussi, Catherine. 1999. Vowel systems and spirantization in Southwest Tanzania. In J-M. Hombert and L.M. Hyman, eds, *Bantu Historical Linguistics: Theoretical and Empirical Perspectives,* CSLI, Stanford CA: 335-377.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World’s Languages*. Blackwell Publishers, Oxford.

Ladefoged, Peter. 2000. *A Course in Phonetics, 4th Edition*. Harcourt Brace, New York.

Lanham, L. W. 1958. The tonemes of Xhosa. *African Studies* 40: 53-130.

Liljencrants, J. and B. Lindblom. 1972. Numerical simulations of vowel quality systems: the role of perceptual contrast. *Language* 48: 839-862.

Louw, J. A. 1975. Clicks as loans in Xhosa. *Bushman and Hottentot Linguistic Studies*. University of South Africa, Pretoria.

Lowe, John and Thilo C. Schadeberg. 1997. *Bantu MapMaker,* v3.1 (computer program). Berkeley and Leiden.

Maddieson, Ian and Peter Ladefoged. 1993. Phonetics of partially nasal consonants. In *Nasals, Nasalization and the Velum*, ed by M. K. Huffman and R. A. Krakow. Academic Press, San Diego: 251-301.

Maddieson, Ian, Peter Ladefoged and Bonny Sands. 1999. Clicks in East African languages. In R. Finlayson, ed., *African Mosaic: Festschrift for J. A. Louw*. University of South Africa, Pretoria: 59-91.

Maddieson, Ian. 1990. Shona velarization: complex consonants or complex onsets?. *UCLA Working Papers in Phonetics* 74: 16-34.

Maddieson, Ian. 1991. Articulatory phonology and Sukuma 'aspirated nasals'. *Proceedings of the 17th Annual Meeting of the Berkeley Linguistic Society, Special Session on African Language Structures*. 145-154.

Maddieson, Ian. 1993. Splitting the mora. *UCLA Working Papers in Phonetics* 83: 9-18.

Maganga, Clement and Thilo C. Shadeberg. 1992. *Kinyamwezi: Grammar, Texts, Vocabulary*. Rüdiger Köppe, Köln.

Manuel, Sharon. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88: 1286-1298.

Mathangwane, Joyce T. 1998. Aspirates: their development and depression in Ikalanga. *Journal of African Languages and Linguistics* 19: 113-135.

Mathangwane, Joyce T. 1999. *Ikalanga Phonetics and Phonology: A Synchronic and Diachronic Study*. CSLI, Stanford, CA.

Medjo Mvé, Pither. 1997. *Essai sur la Phonologie Panchronique des Parlers Fang du Gabon et ses Implications Historiques*. Thèse de doctorat, Université Lumière-Lyon 2.

Meeussen, A. E. 1969 [Reprinted 1980]. *Bantu lexical reconstructions*. Musée Royale de l'Afrique Centrale, Tervuren.

Meinhof, Carl. 1899. *Grundriss einer Lautlehre der Bantusprachen*. Dietrich Reimer, Berlin.

Möhlig, Wilhelm J. G. 1997. A dialectometrical analysis of the main Kavango languages. In W. H. G Haake and E. D. Elderkin (eds) *Namibian Languages: Reports and Papers (Namibian African Studies 4)*. Rüdiger Köppe, Köln: 211-234.

Mpiranya, Fidèle. 1997. Spirantization et fusion vocalique en Bantu. *Linguistique Africaine* 16: 51-77.

Mutaka, Mgessimo M and Carl Ebobissé. 1996/7. The formation of labial-velars in Sawabantu: Evidence for feature geometry. *Journal of West African Languages* 27: 3-14.

Myers, Scott. 1996. Boundary tones and the phonetic implementation of tone in Chichewa. *Studies in African Linguistics* 25: 29-60.

Myers, Scott. 1999a. Downdrift and pitch range in Chichewa intonation. *Proceedings of the 14th International Congress of Phonetic Sciences* Vol 3: 1981-4.

Myers, Scott. 1999b. Tone association and f0 timing in Chichewa. *Studies in African Linguistics* 28: 215-239.

Myers, Scott and Troi Carleton. 1996. Tonal transfer in Chichewa. *Phonology* 13: 39-72.

Nchimbi, A. S. A. 1997. Formant structure of Standard KiSwahili vowels. Paper presented at the Second World Congress of African Linguistics, Leipzig.

Nurse, Derek and Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A Linguistic History*. University of California Press, Berkeley and Los Angeles.

Paulian, Christiane. 1994. Nasales et nasalisation en ŋgùŋgwèl, langue bantu du Congo. *Linguistique Africaine* 13: 83-129.

Pongweni, A. J. C. 1990. An acoustic study of the qualitative and pitch effects of breathy-voice on Shona vowels. *Journal of Phonetics* 11: 129-138.
Poulos, George and Louis J. Louwrens. 1994. *A Linguistic Analysis of Northern Sotho*. Via Afrika, Pretoria.
Quilis, Antonio, Théophile Amdabiang, and Victoria Marrero. 1990. Phonologie et phonétique du gunu (langue bantu du Cameroun). *Travaux de Linguistique et de Philologie* 28: 343-377.
Renaud, Patrick. 1976. *Le Bajele, Vol 1: Phonologie (= Les Dossiers de l'ALCAM 1).* ONAREST, Yaoundé.
Roux, J. C. 1995. On the perception and production of tone in Xhosa. *South African Journal of African Languages* 15: 196-204.
Roux, J. C. and A. J. Holtzhausen. 1989. An acoustic and perceptual analysis of Xhosa vowels. *South African Journal of African Languages* 9: 30-34.
Rycroft, D. K. 1980. *The Depression Feature in Nguni Languages and its Interaction with Tone (= Communication No. 8, Department of African Languages).* Rhodes University, Grahamstown.
Rycroft, D. K. 1981. *Concise SiSwati Dictionary*. J.L van Schaik, Pretoria.
Schadeberg, Thilo C. 1994/5. Spirantization and the 7-to-5 vowel merger in Bantu. In *Sound Change* (= *Belgian Journal of Linguistics* 9), ed. M. Dominicy and D. Demolin, John Benjamins, Amsterdam: 73-84.
Schadeberg, Thilo. 1982. Nasalization in UMbundu. *Journal of African Languages and Linguistics* 4: 109-132.
Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée and Christian Abry. 1997. The dispersion-focalization theory of vowel systems. *Journal of Phonetics* 25: 255-286.
Sommer, Gisele and Rainer Vossen. 1992. Schnalzwörter im Yei (R.41). *Afrika und Übersee* 75: 1-42.
Stewart, John M. 2000. An explanation of Bantu vowel height harmony in terms of a Pre-Bantu nasalized vowel lowering. *Journal of African Languages and Linguistics* 21: 161-178.
Thomas Vilakazi, Kimberley. 1999. *Coproduction and Coarticulation in IsiZulu Clicks*. Ph. D. dissertation. Department of Linguistics, UCLA, Los Angeles.
Thomas [Vilakazi], Kimberly. 2000. Coproduction and coarticulation in IsiZulu clicks: aerodynamic and electropalatographic evidence. In *Advances in African Linguistics (Trends in African Linguistics 4)*, ed. V. Carstens and F. Parkinson. Africa World Press, Trenton, 265-280.
Traill, Anthony and Michel Jackson. 1988. Speaker variation and phonation type in Tsonga nasals. *Journal of Phonetics* 16: 385-400.
Traill, Anthony, J.S.M. Khumalo and P. Fridjhon. 1987. Depressing facts about Zulu. *African Studies* 46: 255-274.
Traill, Anthony. 1990. Depression without depressors. *South African Journal of African Languages* 10: 167-172.
Van Der Veen, Lolke. 1986. Notes en vue d'une description phonologique et morphologique de la langue pouvi (Gabon). Mémoire de maîtrise en linguistique africaine, Université Lumière - Lyon 2.
Van Der Veen, Lolke. 1987. De l'espace vocalique et des tons en Pouvi: Notes descriptives. Mémoire DEA, Université Lumière - Lyon 2.
Van Der Veen, Lolke. 1991. Étude comparée des parlers du groupe Okani (B30 Gabon). Thèse de doctorat, Université Lumière - Lyon 2.
Vossen, Rainer. 1997. What clicks got to do in Bantu. In B. Smieja and M. Tasch (eds.) *Human Contact through Language and Linguistics*. Peter Lang, Frankfurt: 353-66.

Wright, Richard, and Aaron Shryock. 1993. The effect of implosives on pitch in SiSwati. *Journal of the International Phonetic Association* 23: 16-23.
Ziervogel, D. 1959. *A Grammar of Northern Transvaal Ndebele*. J. L. van Schaik, Pretoria.