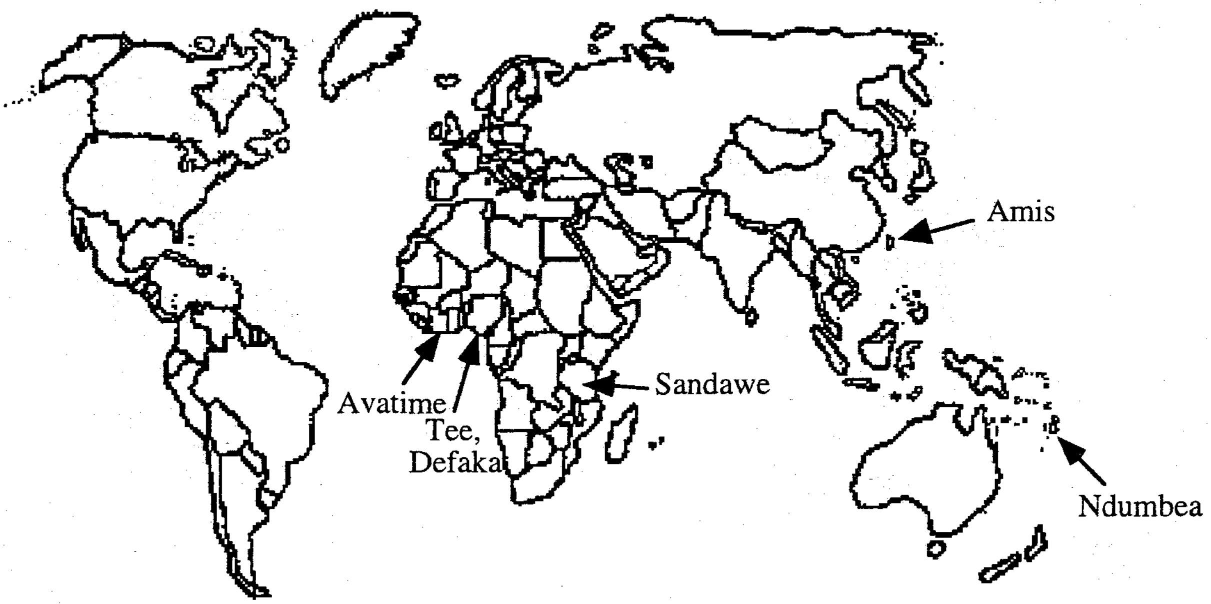
Title

WPP, No. 91: Fieldwork Studies of Targeted Languages III

Publication Date
1995-10-01

Fieldwork Studies of Targeted Languages III



UCLA WPP Volume 91

October 1995

UCLA Working Papers in Phonetics

Number 91

October 1995

Fieldwork Studies of Targeted Languages III

Table of Contents

A phonetic study of Sandawe clicks

Richard Wright, Ian Maddieson, Peter Ladefoged, and Bonny Sands

1.0 Introduction

This paper will describe in some detail the articulatory and acoustic characteristics of the clicks of Sandawe. The Sandawe language is spoken in Tanzania, in the western part of Kondoa district just to the north of Dodoma, the recently-designated capital city of the country. Figure 1 shows the location. For a phonetician, Sandawe is of particular interest because it is one of three East African languages with clicks. These languages are the only languages in the world outside of southern Africa in which clicks occur as a regular part of the phonological inventory of sounds. The other two languages in this set are Dahalo and Hadza. Dahalo, spoken around the mouth of the Tana river in Kenya, is a Southern Cushitic language. Hadza, spoken to the north and east of Lake Eyasi in northern Tanzania, cannot be persuasively classified with any others at the present time, and is best regarded as an isolate (Sands 1995). The principal phonetic characteristics of Dahalo have been described in Maddieson, Spajiç, Sands and Ladefoged (1993), and of Hadza in Sands, Maddieson and Ladefoged (1993). The three East African click languages are not currently in contact with each other, but they share certain typological similarities in the way in which the clicks pattern that are not shared by the southern African click languages.

Sandawe has no close genetic relationship to any of the other languages of East Africa, which represent all three major language families of Africa — Niger-Congo, Nilo-Saharan and Afro-Asiatic. In fact, the immediate neighbors of the Sandawe include all three of these families. Gogo to the south, Nyaturu to the west and Ilangi to the northeast are all Bantu languages in the Niger-Congo family. Datoga (Barabaig) to the north is a Nilotic language in the Nilo-Saharan family. Burunge to the east of Sandawe as well as Iraqw to the north of Datoga are Cushitic languages in the Afro-Asiatic family. The Sandawe also have some contact with Maasai cattle-herders moving out from the Maasai plateau to the northeast. Maasai, like Datoga, is a Nilotic language. The locations of the surrounding languages are shown on Figure 1. Sandawe has frequently been classified with the aboriginal languages of Southern Africa as a member of the fourth high-level language family of Africa, given the name Khoisan by Greenberg (1955). Sands (1995) provides persuasive evidence that this is likely to be correct. There is no good reason to group Sandawe specifically with the Central group of Khoisan, as has sometimes been suggested.

Based on estimates projected from the 1967 population census there may be between 70,000 and 90,000 Sandawe, most of whom are speakers of the language. From our own field observations it is obvious that the majority of young children in the Sandawe area are still learning the language, but they (and most adults) are also fluent in Swahili, the national language, and many prefer to use this language. As Swahili is the language of wider contact and is used in education, in church services, and for all government business, a relatively rapid loss of the language is a distinct possibility in the coming years.

There are no previous published instrumental studies of the phonetics of Sandawe although aspects of the phonetics and phonological system have been discussed in a number of earlier publications. Dempwolff (1916) provided the first systematic attempt at a description. Copland (1938), largely relying on Dempwolff, made some IPA transcriptions. Significant later contributions were made by Tucker, Bryan and Woodburn (1977) and Elderkin (1989, 1992). Phonetic aspects of the clicks are discussed by de Voogt (1992) in an M.A. thesis. Van de Kimmenade (1954) and Kagaya (1993) are primarily useful as sources of vocabulary. Newman (1970) is also a very valuable source of specialized terminology for cultural items and local flora and fauna.

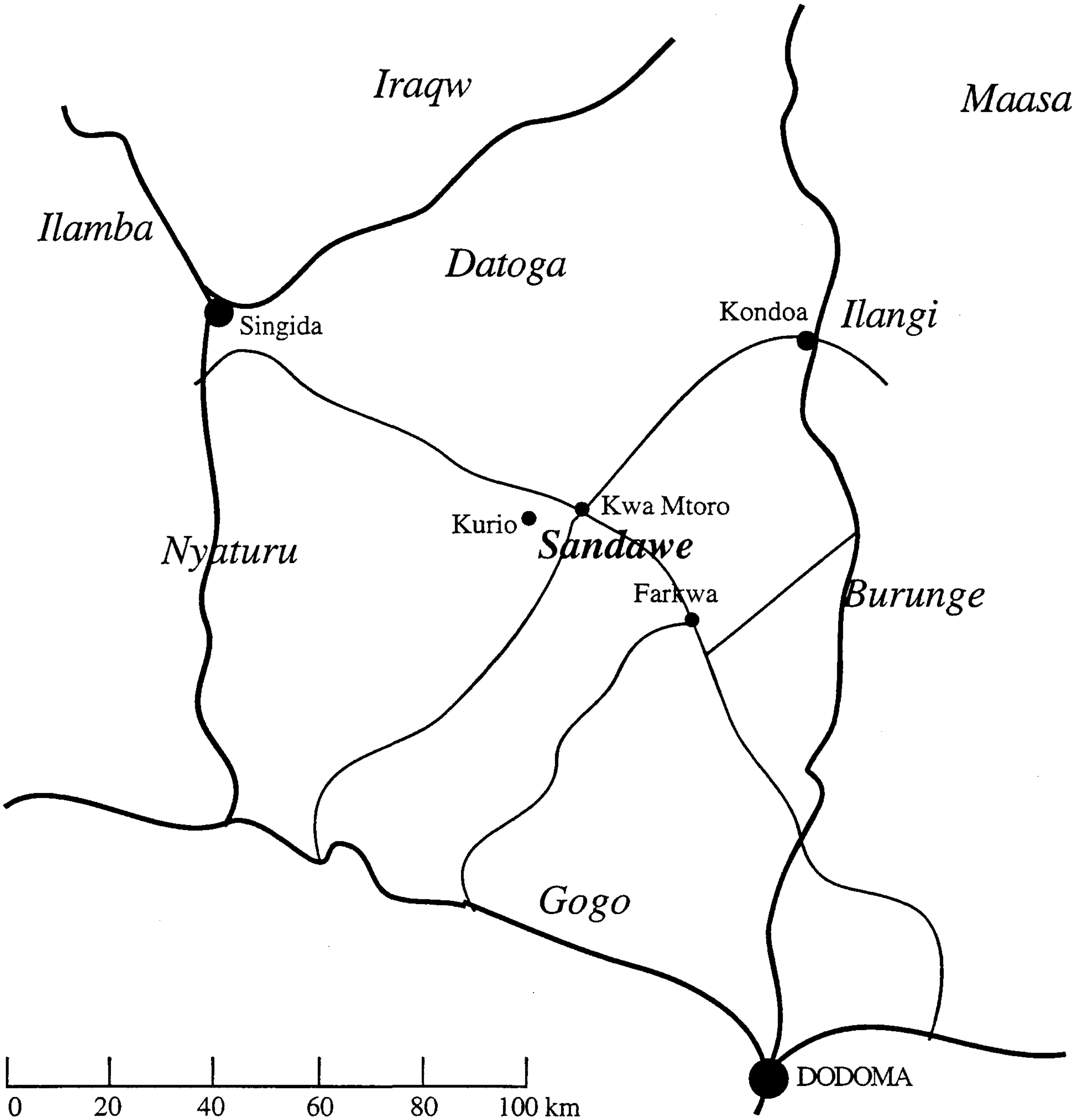


Figure 1. Map showing location of the Sandawe-speaking area, and neighboring languages. The lines represent major and minor roads in the area.

There is an important difference between Sandawe and the South African Khoisan languages whose clicks have been the focus of most recent work on the phonetics of this class of sounds. This is that clicks occur much more freely in non word-initial position in Sandawe. The study of clicks in word-medial environments provides some new insights into certain aspects of their integration into the flow of speech.

This study is based on observations and transcriptions and palatography conducted in the field, as well as on acoustic analyses carried out at the UCLA Phonetics Laboratory of audio recordings made in the field in July 1991. Most of the speakers were from Farkwa, where we observed a considerable number of people. We made systematic audio recordings of five adult male speakers whose ages ranged from 19 to 35, and of two adult females. Aerodynamic data was collected from one of the male speakers, and palatographic data was collected from three male speakers. Audio recordings of one male from Kurio, near Kwa Mtoro, were also made in order to gain some impression of the degree of variation between the two major dialects of the language.

2. Articulatory description

Clicks are complex consonants which involve two articulatory closures in the oral cavity, and the rarefaction of the air enclosed in the space between them. For those who are not familiar with the process involved we recapitulate the essentials in the sketches in Figure 2. A possible starting configuration of the articulators is shown in panel 1. A closure is formed with the back of the tongue in the velar region, as in panel 2, or in the uvular region. A closure of the front of the tongue (or, in some languages, of the lips) is made as shown in panel 3. This results in enclosing a pocket of air, indicated by the shaded area in the panel. The center part of the tongue is lowered, expanding the size of the space between the two closures, as shown in panel 4. This results in a lower pressure of air inside this space than exists in the outside environment. The front closure is then released, as shown in panel 5, resulting in an inward air flow. The sharp inrush of air creates the loud noise associated with the click release. Shortly afterward the back closure is released, allowing the normal outflow of air from the lungs through the mouth to resume, as pictured in panel 6.

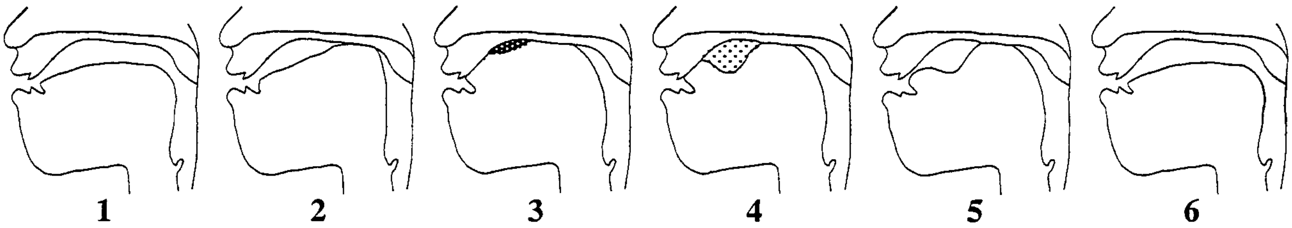


Figure 2. Sketches illustrating the production of a click.

In accord with current practice in the phonetic description of clicks, we will distinguish between click type and click accompaniment. The click type refers to the place of the articulatory closure toward the front of the mouth and to the manner of the release of this closure. The click accompaniment refers to all other aspects of the production of the click, such as the location and manner of release of the back closure, as well as the posture and movements of the larynx, the presence or absence of velum lowering (allowing nasalization), and the actions of the respiratory system. To represent the distinction between click type and accompaniment explicitly, we will transcribe all click sounds with at least two symbols, one for the type and one or more for the accompaniment. Known click types and click accompaniment in various languages are discussed in Ladefoged and Traill (1994) and Ladefoged and Maddieson (1995).

All accounts of Sandawe agree that there are three click types in the language. These may be broadly described as dental, post-alveolar, and alveolar lateral — more detailed descriptions will be provided below. The current IPA symbols used for these click types are |, !, and || respectively, corresponding to earlier IPA symbols ʇ, ʗ, ʖ. The only difference among investigators is that Kagaya prefers ǂ to !. In contrast, accounts of the click accompaniments in Sandawe vary somewhat in the number recognized and in their nature. Dempwolff distinguishes four principal accompaniments which may be interpreted as: voiceless unaspirated, voiceless aspirated, glottalized, and voiced nasalized. Tucker, Bryan and Woodburn (1977) report the same four accompaniments. Both sources note voiced clicks but consider them as occasional variants of the voiceless unaspirated. Elderkin (1989) recognizes all five accompaniments mentioned by these earlier authors, thus differentiating voiceless unaspirated and voiced. In his 1992 paper, he also notes 'predictable nasalization' accompanying glottalized clicks in non-initial position. Kagaya (1993) reports four accompaniments which he calls voiceless, aspirated, glottalized and nasalized. In a description we find confusing, de Voogt (1992) describes five contrastive accompaniments, which his summary calls 'simple', 'ejective', 'aspirated', 'voiced' and 'nasal'; the simple click is elsewhere described as a click with an accompanying glottal closure which is released immediately, in contrast with the ejective accompaniment, which has a delay before the vowel onset. In the following sections our own conclusions about the click types and accompaniments in Sandawe will be presented. Comparisons will also be made with the lateral ejective affricate because of the strong auditory similarity between it and the lateral click, already commented on by Dempwolff. The terms and symbols that we will use are shown in Table 1, which provides an overview of the Sandawe click system.

Table 1. The Sandawe click system

	DENTAL	APICAL POST-ALVEOLAR	LATERAL
VOICELESS UNASPIRATED	**k|**	**k!**	**k||**
VOICELESS ASPIRATED	**k|ʰ**	**k!ʰ**	**k||ʰ**
VOICED	**g|**	**g!**	**g||**
VOICED NASALIZED	**ŋ|**	**ŋ!**	**ŋ||**
GLOTTALIZED	**k|ʔ**	**k!ʔ**	**k||ʔ**

2.1 Click types in Sandawe

The descriptions of Sandawe click types below are based on field observations and palatographic records. Notes and casual observations were made on the entire set of speakers in the study, with particular attention being paid to the palatograms provided by three speakers. The palatographic investigation was conducted using the technique described in more detail in Ladefoged (1993). The speaker's tongue was coated with a marking mixture of equal parts powdered charcoal and olive oil. He then said the target word twice and a mirror was inserted into his mouth so that the roof of his mouth could be photographed or videotaped. The speaker's mouth was cleaned and the process repeated for each of the words listed in Table 2 below. For one speaker, the photographic record was obtained using a Polaroid camera. For the other two, a video camera was used. Both kinds of images were digitized and processed using a Macintosh computer. A cast of the speaker's palate was made using dental impression material, so that the shape of the palate can be accurately known. These casts were used in creating the sagittal views shown in some of the figures below.

Table 2. Words used in the palatographic portion of the study.

DENTAL CLICK	**k|a** 'leaf'
POST-ALVEOLAR CLICK	**k!amba** 'spleen'
LATERAL CLICK	**kǁaŋ** 'warthog'
EJECTIVE LATERAL AFFRICATE	**tɬ'a** 'to take' (Speaker 1)
	tɬ'atɬ'a 'garbage' (Speakers 2 and 3)

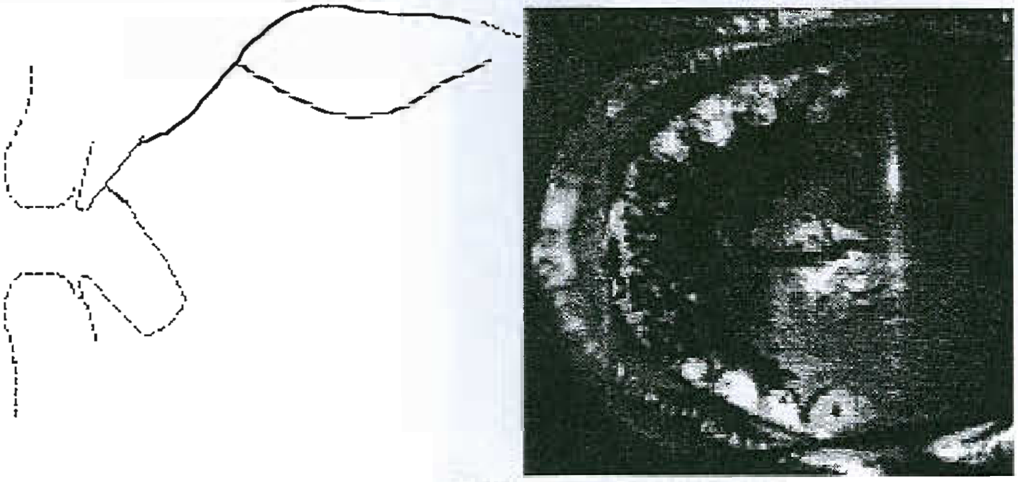


Figure 3. Palatogram (with highlights removed) and inferred sagittal view of the dental click in **k|a** 'leaf' as spoken by speaker 1.

A palatogram of one of the speakers' productions of the dental click [|] and the inferred sagittal section based on this palatogram is shown in Figure 3. We will refer to this speaker as speaker 1. The palatogram shows a contact that extends from the back of the upper teeth to behind the alveolar ridge. Note that the contact in the center is further forward than at the sides. The contact at the lateral margins is not recorded in this photograph, perhaps because it may have been on the lower edge of the molar teeth. The other speakers' palatograms show a similarly extensive contact. From direct observation, it is our impression that this palatographic pattern reflects a large simultaneous contact area, and not a moving contact of the tongue sweeping over this area. This articulation might therefore be classified as laminal denti-alveolar.

A palatogram and inferred sagittal section of the front articulation of the same speaker's post-alveolar click are shown in Figure 4. In this case the closure is made at the back of and just behind the alveolar ridge. The length of the contact from front to back is shorter than that seen in the dental click above, indicating that this closure is more likely to have been made with the tip rather than the blade of the tongue. Another speaker's palatograms showed the closure to be entirely behind the alveolar ridge. This click type has two rather distinct release patterns in Sandawe. The post-alveolar closure is often released in a way that produces a sharp inrush of air, creating the loud transient associated with canonical click sounds. However, it may also be released with a smaller prior expansion of the cavity, so that the breaking of the seal between the tongue and the palate produces only a very quiet noise. In this variant the tongue is usually allowed to strike the floor of the mouth after its separation from the roof, and it is this contact that produces the principal audible signal. We will call this a tongue slap, and where appropriate will use the symbol **¡** to transcribe the hitting of the tongue against the floor of the mouth. On some occasions both the post-alveolar release and the tongue slap create quite loud acoustic signatures.

We assume that it was the tongue slap that Tucker et al (1977) were commenting on when they mention a 'flapped click'. Similar productions are quite common in Hadza (Sands et al 1993), but are not reported as occurring in the languages with clicks in southern Africa whether of the Bantu or Khoisan families. When the speakers were pronouncing words for our palatographic data collection, they produced the canonical loud variant of this click, so we do not know if the front closures are similar in location and extent between the two variants. The acoustic characteristics of the variant with the tongue slap will be discussed below.

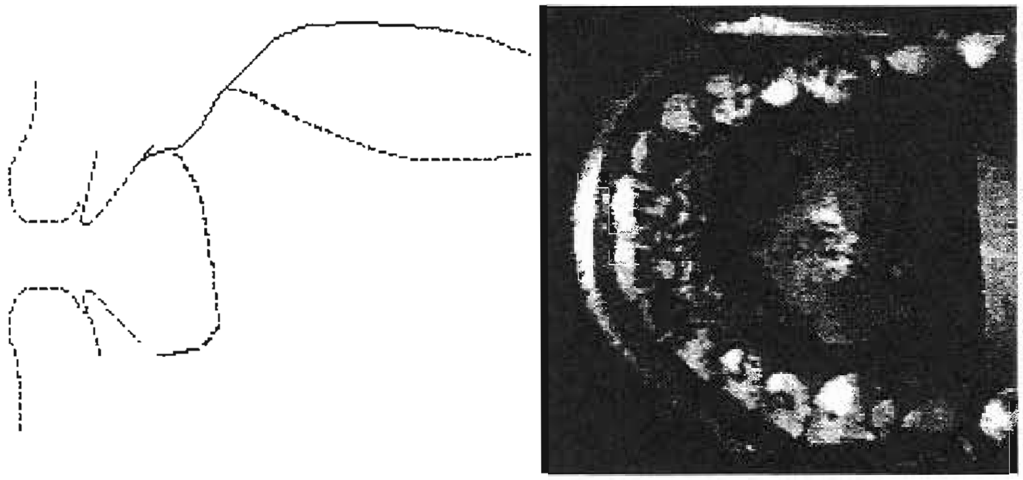


Figure 4: Palatogram and inferred sagittal view of the post-alveolar click in the word **k!amba** 'spleen' as spoken by speaker 1.

The third click type in Sandawe has a lateral release. Figure 5 shows a palatogram and inferred sagittal section of the front articulation of this click as produced by speaker 1. For this speaker, there is a broad laminal contact that covers the back of the upper teeth and extends behind the alveolar ridge, as in the dental click. The other two speakers who provided palatograms had narrower contact areas which neither included the teeth, nor reached as far back behind the alveolar ridge. (One of these productions will be illustrated in Figure 7 below.) As this articulation appears to be more typical, we consider this click type to be best described as an alveolar lateral click.

As noted earlier, Sandawe has a lateral ejective affricate which is auditorily similar to the lateral click, sufficiently similar so that care was required to avoid transcription errors. The acoustic basis of the similarity in auditory impression will be investigated in the acoustic portion of this study. The contact pattern for the lateral ejective affricate as produced by speaker 1 is shown in Figure 6. For this speaker the location and extent of the front closure is very similar for both the lateral click and the lateral ejective affricate. Both cover the back of the incisors and extend beyond the alveolar ridge. Speaker 2 also had similar front closures for these two lateral consonant types but for him they are both restricted to the alveolar region and do not involve the teeth, with the contact for the click being slightly more retracted than that for the ejective affricate. Palatograms from this speaker are shown in Figure 7. Speaker 3 showed an alveolar contact for the lateral click but his contact for the lateral affricate extended over the entire area from the incisors to the velum at some stage of its production. This speaker has a very small oral cavity and a low palatal vault, in contrast to the high vault of speaker 1. We do not know whether this speaker regularly makes such full contact as a preparatory gesture before reducing the contact area to a more specific configuration before the release. The majority of the speakers we observed typically had a closure for the lateral ejective affricate which appeared to be like that shown in Figure 7, rather than that shown in Figure 5, and it might be most appropriately labeled as an alveolar (or even post-alveolar) articulation. We also had the impression that the location of the lateral escape was relatively far back, based on the articulatory gestures that we used to produce our most successful imitations, but the correctness of this impression cannot be confirmed by static palatography.

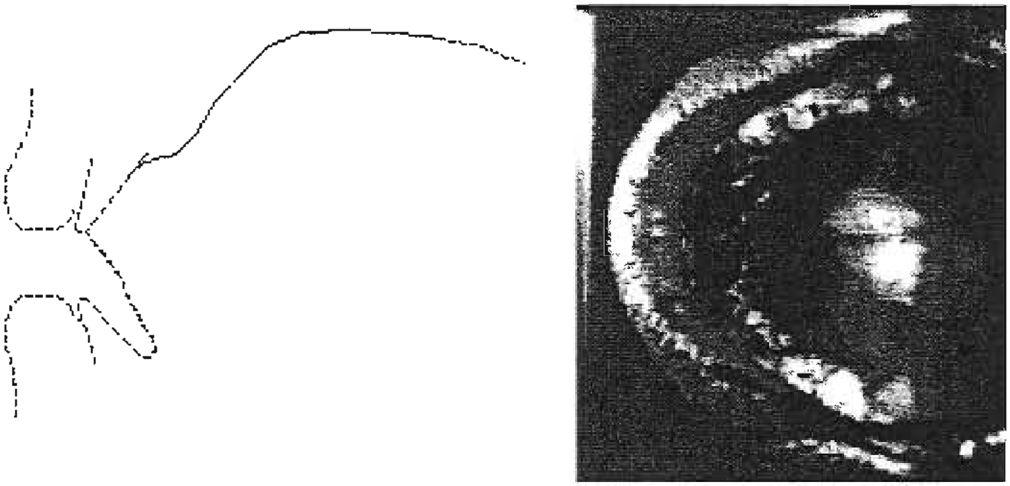


Figure 5. Palatogram and inferred sagittal view of the lateral click in ‘warthog’ **k||aŋ** as spoken by speaker 1.

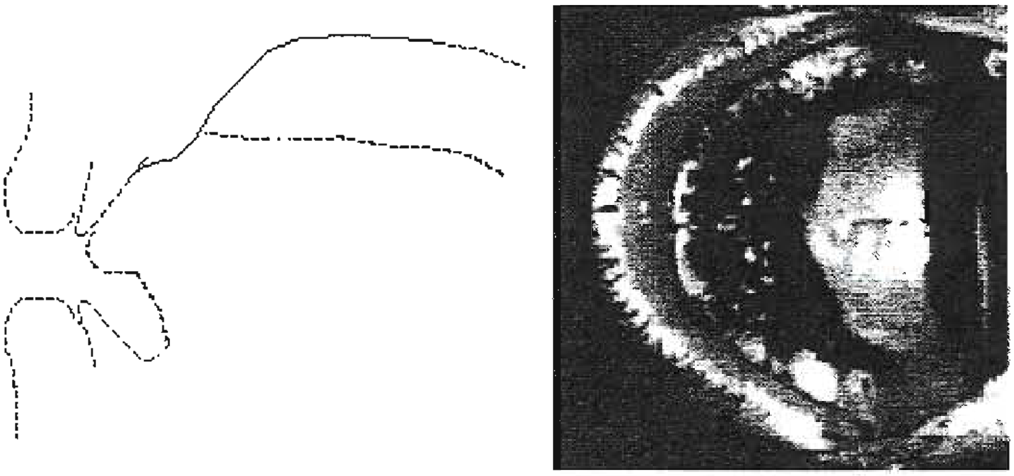
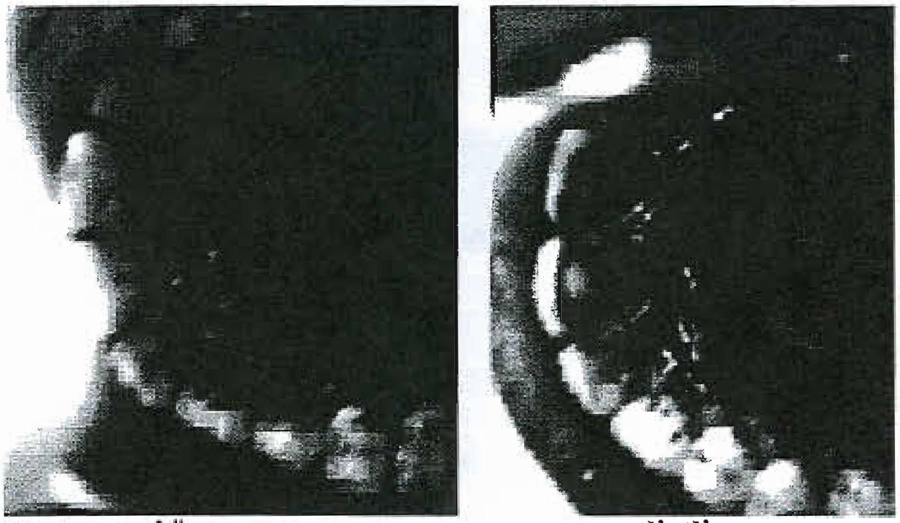


Figure 6. Palatogram and inferred sagittal view of the ejective lateral affricate in the word ‘to take’ **tɬ’a** as spoken by speaker 1.



k||aŋ **tɬ'atɬ'a**

Figure 7. Palatograms of the lateral click and lateral affricate in the words **k||aŋ** 'warthog' and **tɬ'atɬ'a** 'garbage' as spoken by speaker 2 .

2.2 Click accompaniments in Sandawe

Our articulatory investigations of the click accompaniments in Sandawe consist of aerodynamic records of the oral and nasal airflow patterns from speaker 3. Oral airflow was collected with a mask covering the mouth. Nasal airflow was recorded with a tube connected to a small foam plug with a narrow hole through it inserted in one nostril while the other was pinched closed. The aerodynamic records were produced under difficult field conditions, and flow volumes were not calibrated. However, they provide significant data of a type that is not elsewhere available, and are particularly useful for determining aspects of the timing of different actions. Two non-consecutive repetitions of each word were recorded and the aerodynamic patterns were very consistent across these repetitions, indicating that this data provides reliable, qualitative information on the production of this speaker.

Our research confirms that there are five accompaniments in Sandawe: voiceless unaspirated, voiceless aspirated, voiced nasalized, voiced and glottalized. We are also able to provide more detail on the phenomenon discussed under the name of 'predictable nasalization' by Elderkin. The acoustic signatures of the accompaniments will be discussed in greater detail below, based on a larger number of speakers. Figure 8 illustrates a token of a voiceless unaspirated post-alveolar click at the beginning of the word **k!eː** 'termitary, anthill'. The sharp inward air flow at click release is clearly shown in the oral airflow trace, and is followed by low volume egressive flow for the following vowel. The nasal airflow record shows some small perturbations at the beginning of the word but no net flow of air out through the nose. These may reflect movements of the palate during the release of the click. No significant amount of air flows out through the nose during this period or during the vowel. The speaker exhales partly through the nose at the end of the utterance, and the increase from the baseline level of nasal flow is very clear at this point. Another word-initial unaspirated post-alveolar click is illustrated in Figure 15 below.

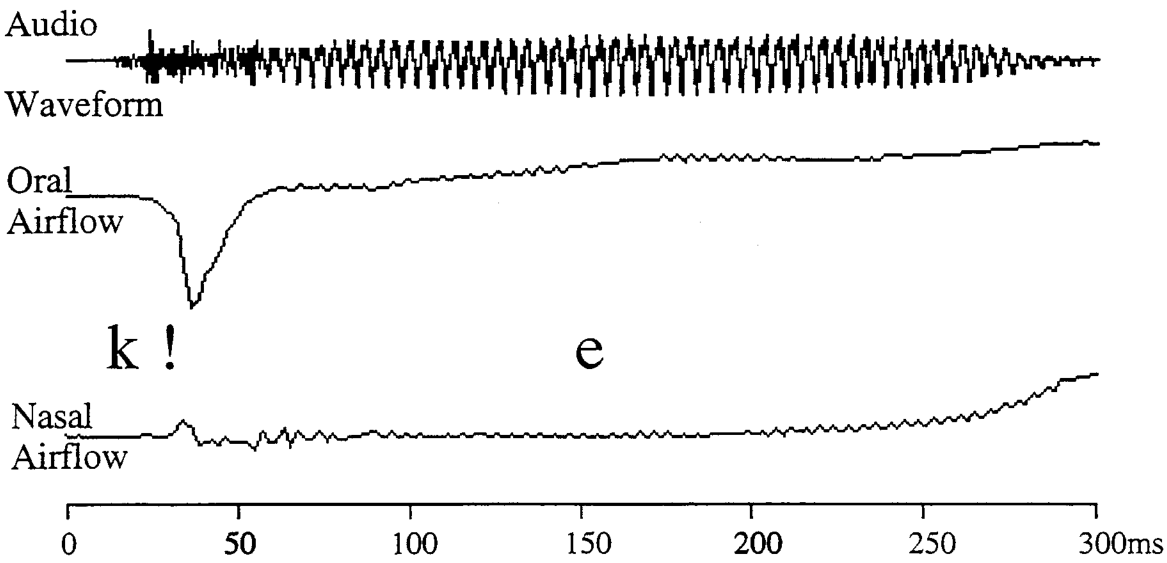


Figure 8. Aerodynamic record of **k!eː** ‘termitary, anthill’ spoken by speaker 3.

The voiceless aspirated accompaniment is shown in Figure 9, illustrating a voiceless aspirated post-alveolar click in the word **k!^{h}eŋ** ‘tongue’. Following the inward airflow due to the click release, there is a high-volume outward oral airflow, and some considerable delay before vocal fold vibration begins for the vowel. Nasal airflow is apparent for the final consonant, but not earlier. This accompaniment is found only in word-initial environments.

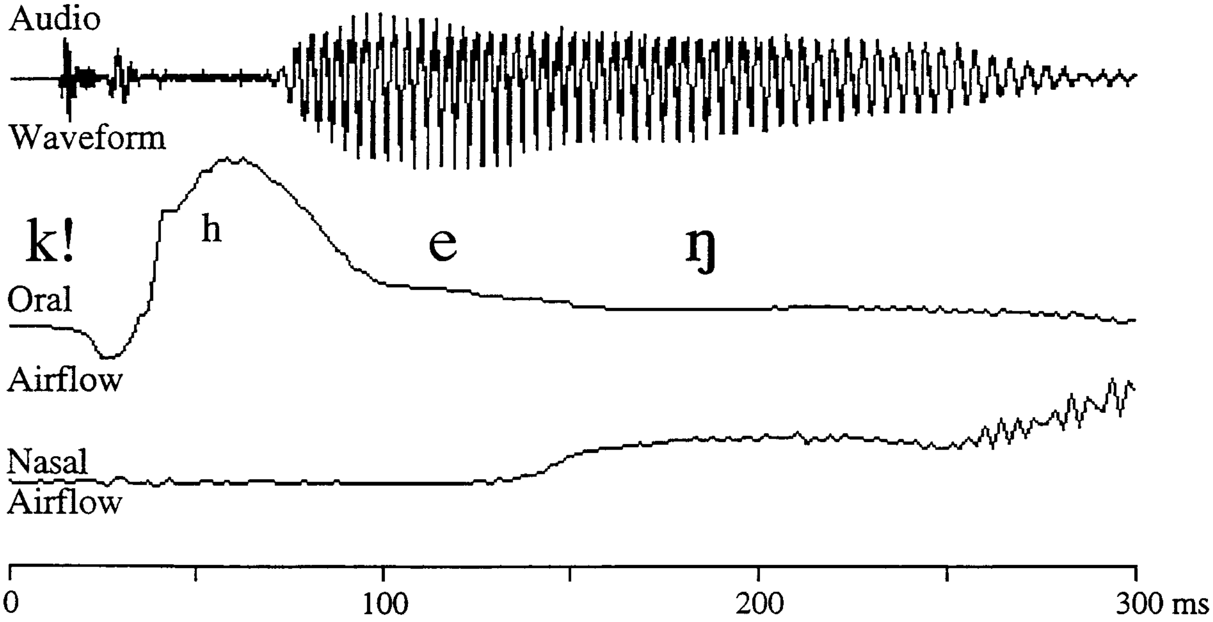


Figure 9. Aerodynamic record of **k!^{h}eŋ** ‘tongue’ spoken by speaker 3.

The voiced nasalized accompaniment is illustrated in Figure 10 which shows two voiced nasalized dental clicks in the word **ŋ|aŋ|aʔo** ‘to cut’ (reduplicated form). In the audio waveform, strong vocal fold vibration can be seen to begin well before the release of the initial click. (The oral and nasal airflow records are less reliable indicators of voicing, as they were sampled at a lower rate.) The nasal airflow record shows that at the very beginning of the record, before voicing onset, there is voiceless nasal airflow, which decreases as voicing is initiated. Continued flow through the nasal cavity for a short period after the click release is indicated by the strong vibrations in the nasal airflow trace at this time (much stronger than those in the previous two clicks), but the following vowel is primarily oral. In the medial nasalized click, the onset of the nasal component can be detected from the decrease in the oral flow and the increase in the nasal flow shortly before the time point marked by the first bold vertical bar. This bar marks the point at which the velar closure is made, and it occurs about 100 ms before the click — i.e. the dental — release occurs. The velar nasal continues to be held for about another 100 ms until the time point marked by the second vertical bar on the figure, and thus occupies a good proportion of the duration that might be ascribed to the following [a] vowel. Note that the segment transcribed as a glottal stop in the infinitive ending [ʔo] does not involve complete vocal fold closure but only a constriction of the folds resulting in reduced air flow.

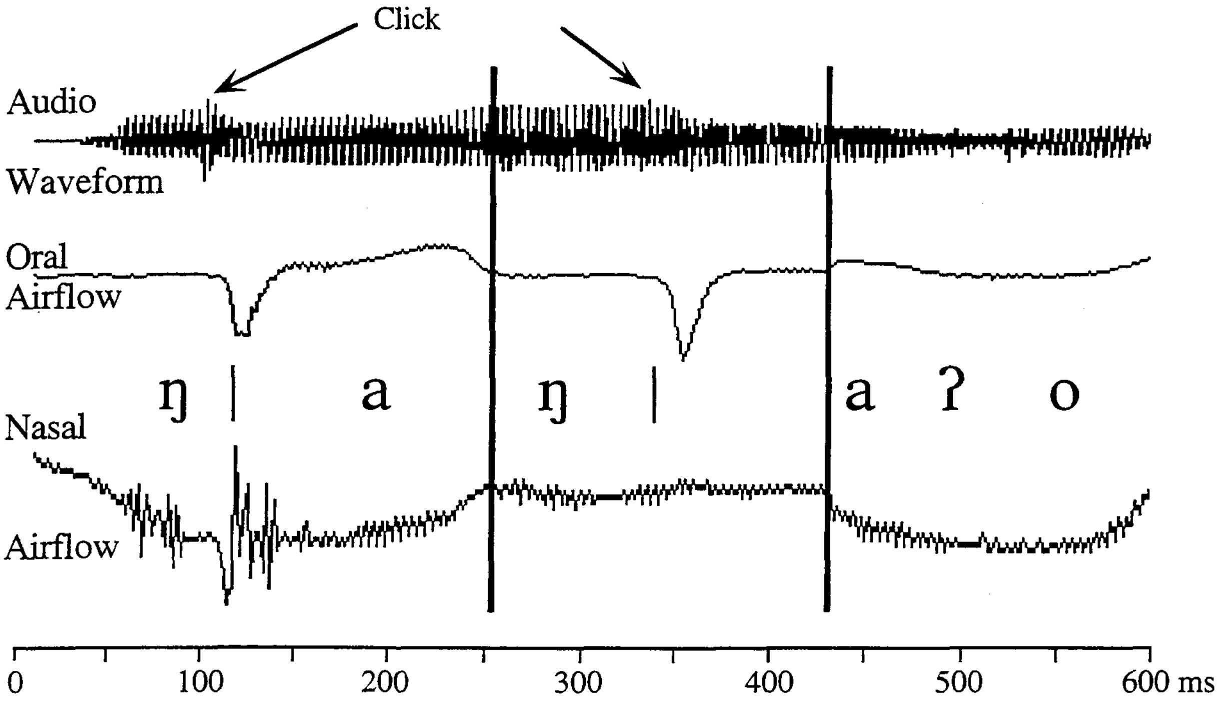


Figure 10. Aerodynamic record of **ŋ|aŋ|aʔo** ‘to cut’ spoken by speaker 3. The vertical bars mark the formation and release of the velic closure in the medial nasal click. The scale of the nasal airflow has been increased so as to show these points more clearly.

The clicks in Figure 11 in the word **g|ig|o**, the name of a species of small bird in the finch family, illustrate the voiced accompaniment with dental clicks. In initial position, the onset of voicing does not occur until closer to the click release than is the case with the initial nasalized click in Figure 10. In the audio waveform a few periods of low amplitude voicing can be observed following the click release. The vibrations produced by this voicing are observable in the nasal airflow channel, but there is no net flow of air through the nasal passage, and there is no nasal airflow preceding voice onset. We infer from this pattern that the tongue is raised to form the velar closure and that the velum is already raised to close the nasal passage before voicing is initiated. This token thus illustrates a voiced velar stop closure with about 50 ms of vocal fold vibration, which is probably close to the maximum duration that voicing can be sustained when such a configuration exists (Ohala and Riordan 1989). When this closure is released, both oral and nasal closures are broken and there is a phonetically nasalized vowel after the velar release. Because the vowel is a high vowel involving a considerable degree of constriction in the oral cavity, most of the airflow is directed through the nasal cavity.

The second voiced click in this word differs dramatically from the first in being prenasalized. The vowel in the first syllable is very short and is followed by a velar nasal which is part of this second click. There is complementary distribution such that voiced clicks in initial position occur without prenasalization and those in medial position are always prenasalized. In this token the exact time that the velar nasal begins is unclear, but might be around the 100 ms mark. Nasal airflow is shut off shortly before the click release, as shown in the nasal airflow trace and by the sharp reduction in the amplitude of the voicing vibration in the audio waveform. Both velar and velo-pharyngeal closures are maintained for about 50 ms while the dental click release is made. Then the velar closure is released and air directed solely out of the mouth. We suggest that the prenasalization of a medial voiced click is a means of retaining the relatively long lead between the formation of the velar closure and release of the front closure for a click while enabling voicing to be continuously maintained. By shortening the period during which both the oral and the nasal passages are closed to no more than 50 ms, the speaker avoids an involuntary cessation of vocal fold vibration due to pressure in the pharyngeal cavity approaching equality with subglottal pressure. In utterance-initial position, the problem is handled differently, by delaying the onset of voicing.

Another striking fact about these voiced clicks is the low oral inflow associated with the click release itself. In voiced velar plosives a forward movement of the location of the closure on the palate can be employed to assist in enlarging the pharyngeal cavity, thus enabling voicing to be sustained for a longer time than would otherwise be the case. It is possible that these voiced clicks also involve some forward movement of the velar closure to assist voicing, and that this reduces the amount that the cavity between the two closures can be expanded. The result would be a weaker inflow on release, as is shown in this record.

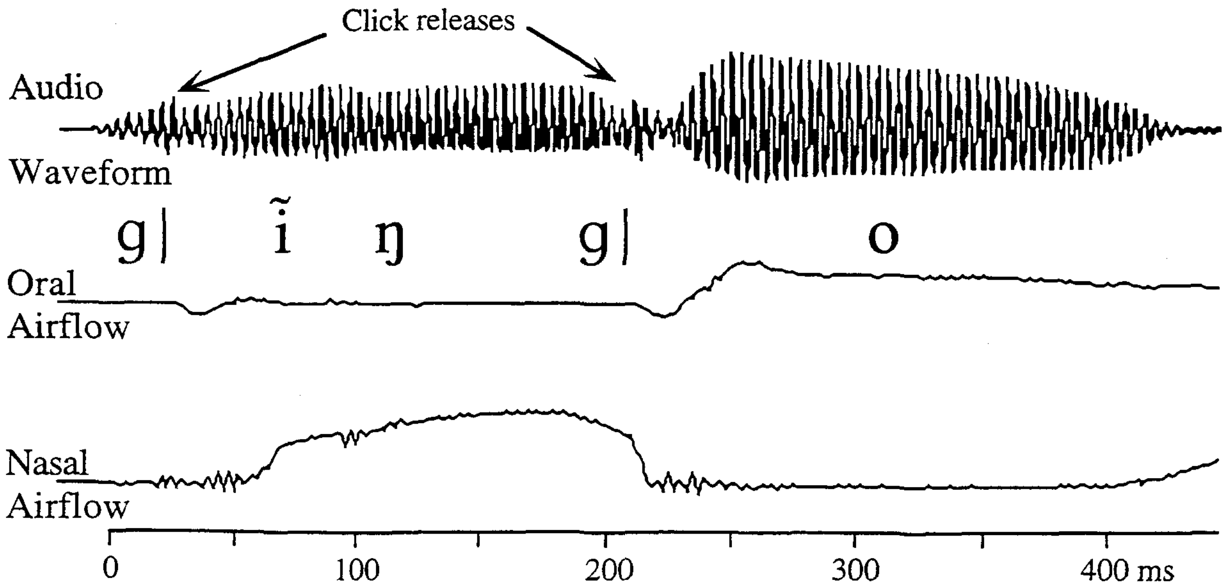


Figure 11. Aerodynamic record of **g|ig|o** (species of small bird) spoken by speaker 3. In this and subsequent figures the symbols on the figure show the phonetic elements rather than the phonological categories.

The glottalized accompaniment shows the greatest variation. Some aspects of this variation are illustrated in the remaining four aerodynamic displays in figures 12, and 14-16. A word-initial example of this accompaniment with a lateral click type is shown in Figure 12, in the word **kǁʔokǁʔa** ‘baboon’. The audio waveform shows that the lateral click release precedes the audible onset of the following vowel by some appreciable time, and the onset of phonation for the vowel is irregular. From these records and from the auditory impression in words of this type, we infer that a glottal constriction is formed and held while the click is pronounced. The timing of the velar release cannot be detected in these records, but it apparently occurs before the glottal release since the vowel onset appears glottalized. There is no suggestion of an ejective velar release, so we transcribe this accompaniment with [ʔ] following the click symbol rather than with the ejective diacritic [’] used by other writers. Note, however, that the **ʔ** symbol should be regarded as indicating a glottal constriction that is often realized as creaky voice rather than a complete glottal closure held for an appreciable period of time.

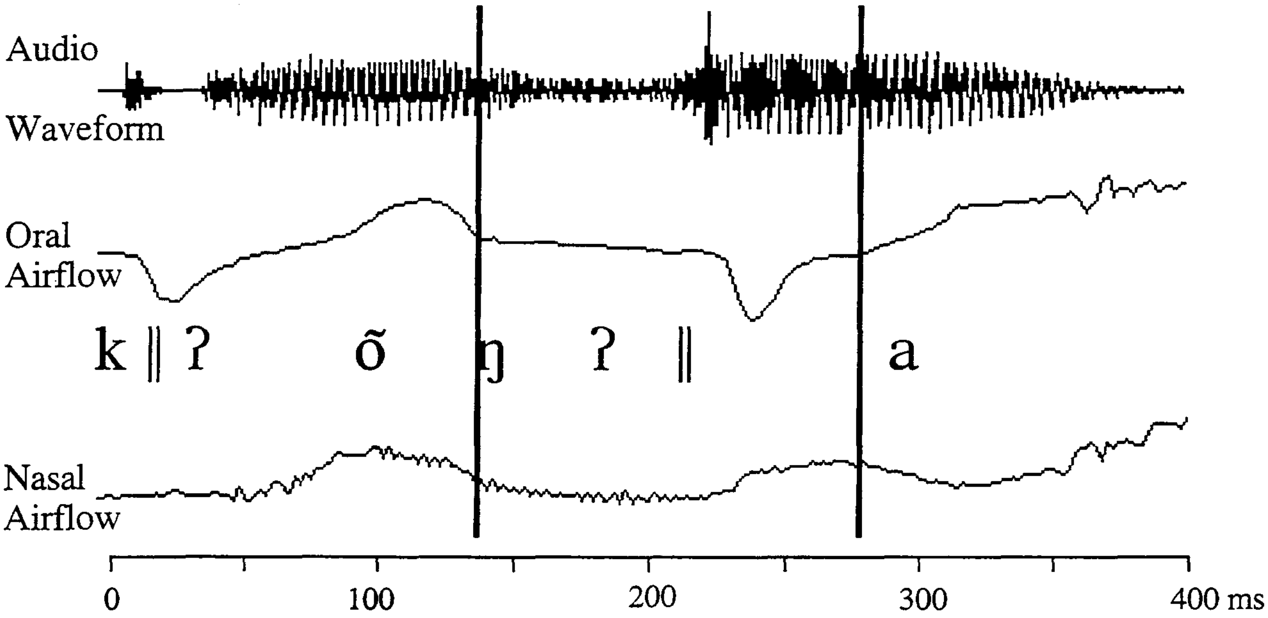


Figure 12. Aerodynamic record of **kǁʔokǁʔa** ‘baboon’ spoken by speaker 3. The first vertical bar indicates the onset of laryngealization, the second the possible time of release of the posterior (velar) click closure.

In medial position (and possibly in initial position without auditory salience) this accompaniment involves both a constriction of the glottis and a lowering of the velum, so that the nasal passage is open, as is the case with medial voiced clicks. Depending on how these movements are timed in relation to each other and to the click release, rather different outcomes result. The medial click in Figure 12 illustrates one variant. Anticipatory nasalization, indicated by upward movement of both oral and nasal airflow traces, is observed in the first vowel. As Elderkin notes, nasalization is predictable in the sense that it always observed before non-initial glottalized clicks. He also notes that it may occur before initial glottalized clicks when they follow a vowel at a word boundary. We therefore do not record it in our phonologically-based transcriptions. Both the oral and the nasal air flow volumes fall at around the time point marked by the first bold bar on the figure. However the nasal airflow does not reach zero, nor is there silence on the audio trace but rather low-amplitude irregular phonation similar to that seen in the segment symbolized by a glottal stop in figure 10. This portion of the signal might well be transcribed phonetically as [ŋ], but we have chosen to represent it as [ŋ?], indicating laryngealization as in the wor̃d initial glottalized clicks. An increase in the amplitude of the vocal fold vibration just before the click release suggests that the vocal fold constriction is relaxed shortly before the lateral release. At the click release the volume of air flowing through the nose increases; in fact there may be continuing velar closure until the time indicated by the second vertical bar, as there is at first no detectable oral airflow. The onset of the vowel is nasalized, with flow gradually shifting from the nasal passage to the mouth as it continues. (There is a final exhalation through both the nose and the mouth).

The sequence of actions we infer for this medial click is thus as follows. The velum is lowered well before the two oral closures for the click are formed, and it remains lowered through the entire click articulation and the first part of the final vowel. The back of the tongue is raised against the velum about 80 ms before the lateral click release occurs and remains raised for a total of approximately 130 ms (similar in duration to the velar nasal component in the medial nasalized click illustrated in Figure 10). At about the same time as the velar closure is made, the vocal folds begin to constrict, producing laryngealized phonation and reducing the volume of air flowing through the nose. This laryngeal constriction is relaxed before the front closure of the click is released; its presence is mainly perceptible at the offset of the first vowel. The time at which the alveolar closure is formed is unclear, but we assume this closure is shorter than the other articulations involved. The assumed temporal order of events can be approximated by the display in Figure 13.

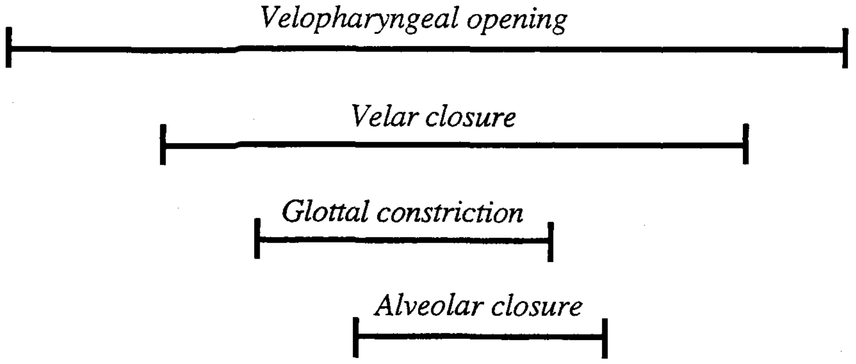


Figure 13. Schematic representation of the articulatory timing of the medial click in **k||?ok||?a** in Figure 12.

Somewhat different temporal or articulatory patterns are shown in the next three figures. Figure 14 illustrates the word **ma|?a** 'louse'. The glottal constriction is longer in this case, and lasts a similar duration to the velar closure. The velum remains down throughout the first syllable of the word, partly due to the initial nasal. In Figure 15 the glottal constriction in **k!u!?e** 'kidney' is tighter than in the preceding two figures, resulting in a particularly sharp diminution of the amplitude of the vocal fold vibrations. The final vowel is fully oral, suggesting that the velum has been raised before the click is released — possibly as early as the point indicated by the vertical bar, where nasal airflow drops. We may note that the post-alveolar clicks in this particular utterance show relatively quiet releases on the audio wave from, and show little inward airflow at release. We may infer that they have a weak degree of rarefaction. Figure 16 shows much less anticipatory nasalization before the click, so the preceding vowel is oral. The velar closure is sustained for a short while after the click release, resulting in a second peak of nasal airflow, with the valley in between these two peaks attributable to the glottalization.

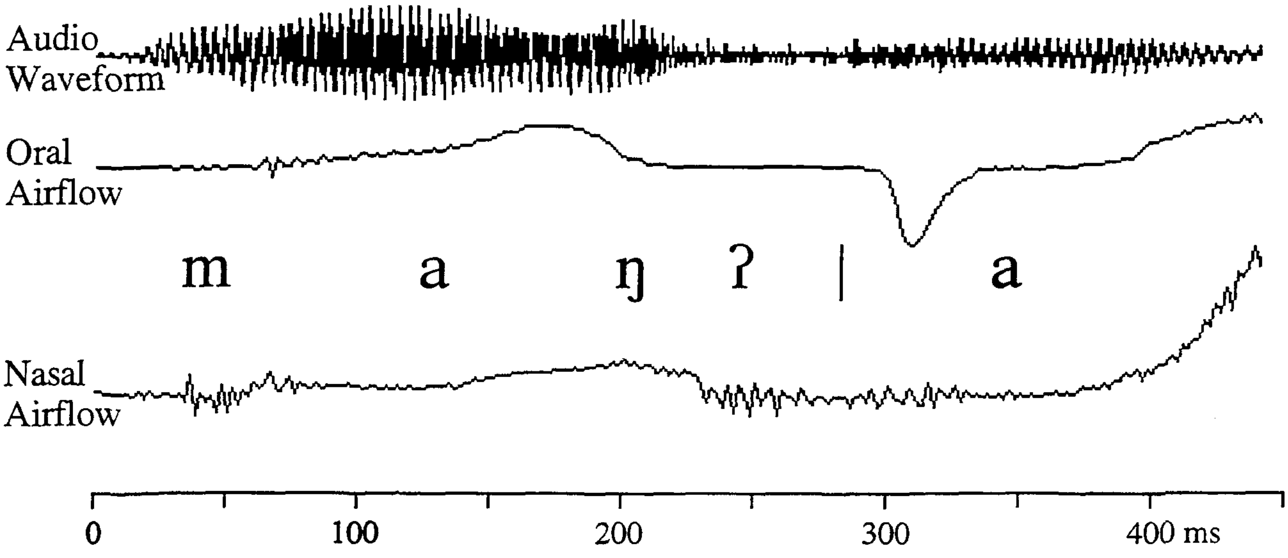


Figure 14. Aerodynamic record of **ma|ʔa** ‘louse’ spoken by speaker 3.

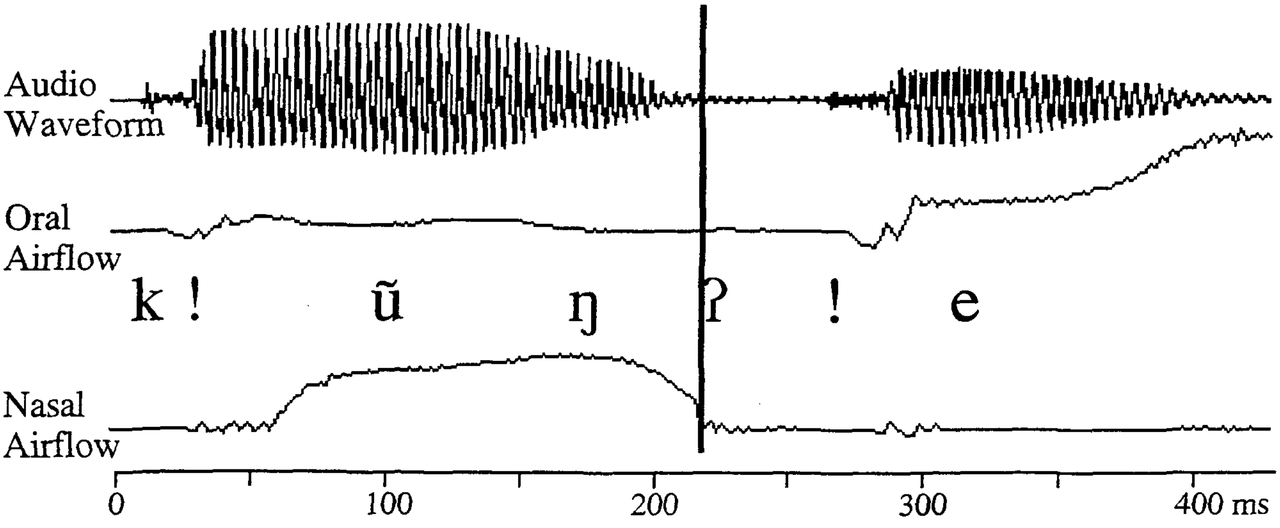


Figure 15. Aerodynamic record of **k!u!ʔe** ‘kidney’ spoken by speaker 3.

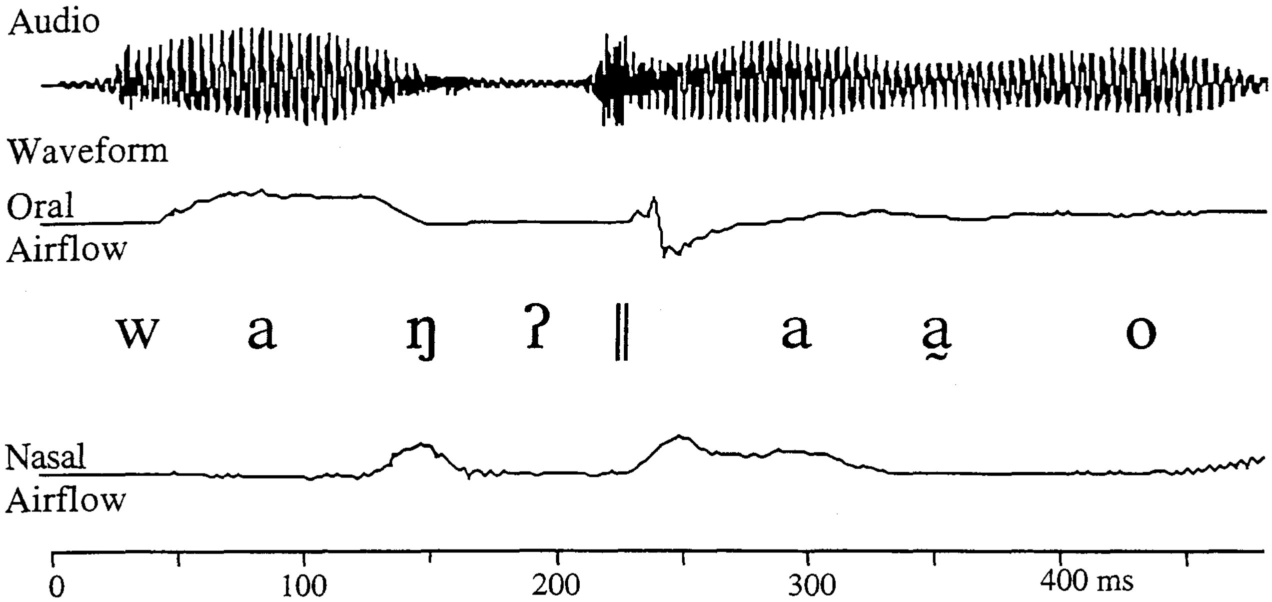


Figure 16. Aerodynamic record of **waǁʔaʔo** ‘to vomit’ spoken by speaker 3.

3. Acoustic analysis.

This section will discuss aspects of the acoustic analysis of the Sandawe clicks. As before, our presentation will be divided into two parts dealing with click types and click accompaniments respectively. The analysis of click types will focus on the amplitude and spectral characteristics of the bursts, as well as the spectrum and duration of any accompanying affrication. The analysis of click accompaniments will focus on acoustic measures of timing.

3.1 Click types

Some of the timing and amplitude differences between the click types can be seen by comparing waveforms. The three click types of Sandawe are illustrated in Figure 17 which shows waveforms of word-initial glottalized clicks on an expanded time-scale. Since the click release is followed by a period of silence when this accompaniment occurs in utterance-initial position, this environment provides a very good opportunity to examine the release itself. The particular token of a post-alveolar release illustrated here has a weak tongue-slap, which produces a second excitation, indicated on the figure by an arrow.

As noted above the dental and lateral click types are characterized by a noisy (affricated) release. In the tokens illustrated the affrication noise after the dental release has greater amplitude than the release burst, but the opposite is true for the lateral. The noise interval is also longer for the dental. The post-alveolar click type is often characterized by a higher amplitude release with little following noise, as in the figure. When a post-alveolar click has a loud release of this type it acts as an impulse that resonates in the vocal tract, creating the damped sinusoidal wave that can be seen in Figure 17 immediately following the release transient. In Sandawe, however, there is considerable variation in the amplitude of the release burst of the post-alveolar click, ranging from a release that has considerably greater amplitude than the vowel immediately following it to one with a much lower amplitude than the vowel.

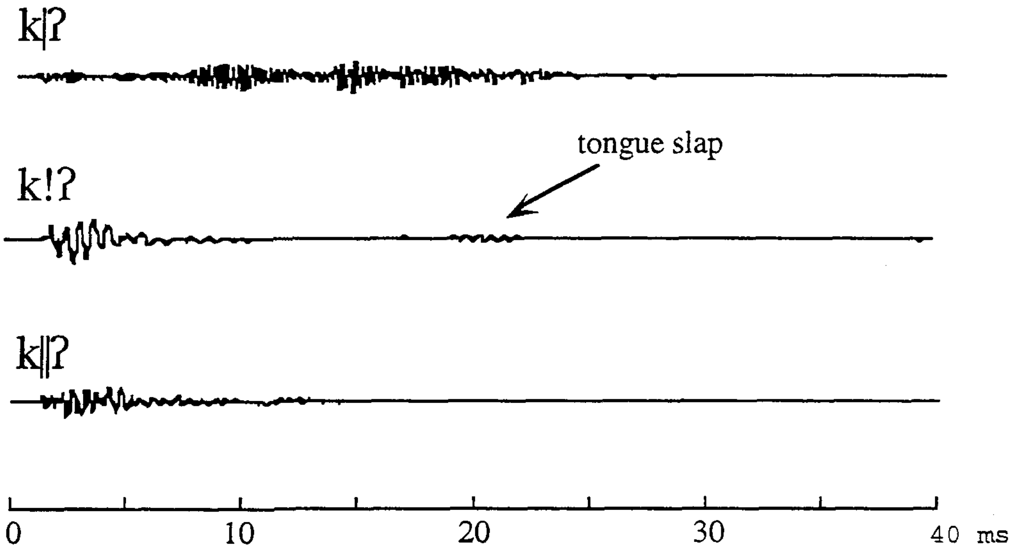


Figure 17. Example waveforms illustrating the dental, post-alveolar, and lateral click types with utterance-initial glottalized accompaniment in the words: **k|ʔa** 'bad smell', **k!ʔa** as in **k!ʔamba** 'spleen' and **k||ʔa** as in **k||ʔaŋ** 'warthog' spoken by speaker 4.

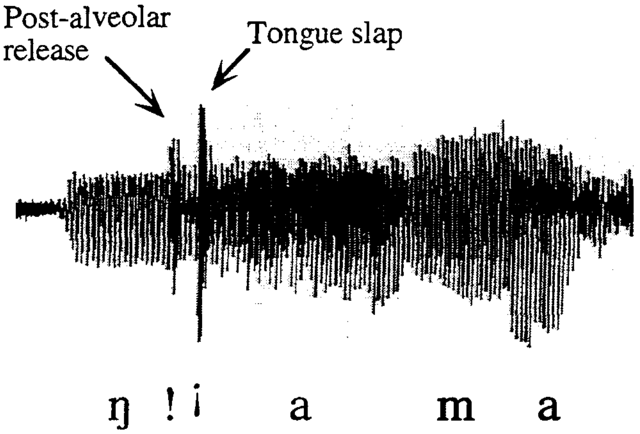


Figure 18. Waveform of the word **ŋ!ama** (tree species) spoken by speaker 3.

As described earlier, post-alveolar clicks may also contain a tongue slap, which itself varies in amplitude. A post-alveolar click with a relatively quiet tongue slap is shown in Figure 17; a token in which the tongue slap produces a louder sound than the post-alveolar release is shown in Figure 18. There are even instances where the click release is virtually inaudible while the tongue slap has the same or greater amplitude than the following vowel. Two of the five speakers recorded in this study consistently produced their post-alveolar clicks with a tongue slap, but most speakers use such variants on at least some occasions. There is about a 20 ms delay between post-alveolar release and tongue slap when both are detectable. The amplitude of the release and the presence of an audible tongue slap do not appear to be dependent on any phonological condition.

Table 3. Words used for total release duration measures

	UNASPIRATED	GLOTTALIZED	ASPIRATED
DENTAL	**k|aː** 'leaf'	**k|ʔwaː** 'wound'	**k|ʰia** 'dikdik'
LATERAL	**k||wa** 'name'	**k||ʔeŋ** 'fire'	**k||ʰoŋ** 'cave'

To compare the releases of the dental and lateral clicks in word-initial position we measured the duration of the interval from the onset of the release transient noise to the end of any affrication noise in the set of words listed in Table 3. Duration measurements were taken from waveforms displayed on the Kay CSL. Simultaneously displayed broad band spectrograms were used for reference. To maintain a balanced data set, only one word illustrating each combination of a click type and each of the three non-voiced accompaniments was included, resulting in six words altogether. However, the words are not matched for vowel quality and/or the presence of labialization. Two repetitions of each of these words from five male speakers were measured resulting in 30 measures for each of the two click types. The results are summarized in Figure 19 and Table 4. Figure 19 also includes the mean duration of the frication for the ejective lateral affricate in the word **tɬ'a** 'take' for comparison with the amount of affrication in the clicks.

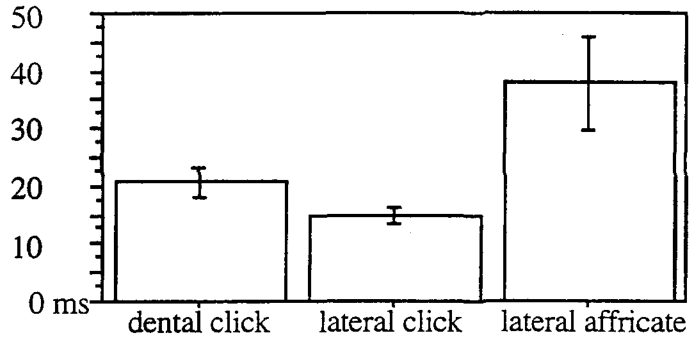


Figure 19. Mean burst plus fricative durations, pooled across accompaniment. Error bars show one standard deviation above and below the mean.

Table 4. Mean burst plus fricative durations

	MEAN DURATION	STANDARD DEVIATION	STANDARD ERROR
DENTAL	21.5	3.5	0.7
LATERAL	15.7	3.6	0.7

The three click types in Sandawe have different amounts of affrication. The post-alveolar clicks have none. Of the two affricated cases, the dental is about 7 ms longer than the lateral. The release duration measurements for the dental and lateral clicks were submitted to an analysis of variance with release duration as the dependent variable and click type and accompaniment as independent variables. The overall effect of click type on release duration was significant (F[1,48]=43.0, p<.0001), but the effect of accompaniment on release duration was not significant nor was there a significant interaction between type and accompaniment. Nonetheless, although the dental and lateral clicks have a fairly long noise component to their release, it is not nearly as long as the frication following the ejective lateral affricate despite the fact that ejective affricates have a shorter noise component than their non-ejective counterparts. Thus even though the dental and lateral clicks have a long noise component to their release they should not necessarily be grouped together with the affricates. Nevertheless, it is probable that the duration of the lateral click's noisy component contributes to its confusability with the ejective lateral affricate.

The click types differ not only in the duration of their noise components, but also in the spectral shape of the noise produced by the release. The auditory similarity between the ejective lateral affricate and the lateral click suggests that in addition to the temporal aspect of the noise there should be a spectral similarity as well. Spectral measures were taken from one example word for each of the three click types with the glottalized accompaniment and the ejective lateral affricate. Glottalized clicks were chosen for this analysis because the long VOT following the release minimized the influence of the vowel on the spectral characteristics of the release. Furthermore, this accompaniment provided the best match to the ejective nature of the lateral affricate. These consonants were all in word-initial position before /a/, as shown in Table 5. Two repetitions of each word were analyzed for each of five male subjects, equaling ten tokens for each of the four consonants.

Table 5. Words used in the spectral analysis of release noise.

ARTICULATION	EXAMPLE
dental click	**k|ʔa** 'bad smell'
post-alveolar click	**k!ʔamba** 'spleen'
lateral click	**k||ʔaŋ** 'warthog'
ejective lateral affricate	**tɬ'a** 'to take'

The spectral characteristics were analyzed using FFT spectra. The recordings were sampled at 20 kHz and the analysis was performed using a Kay CSL. The frequency range for the analysis was 0-8000 Hz. A 512 point FFT window placed over the noise was used. The results are plotted in Figures 20-23 in the form of mean spectra across all ten repetitions, displayed on a linear frequency scale. Individual differences are thus blurred, but tokens from different speakers showed generally consistent patterns.

The spectrum of the dental click, shown in figure 20, has diffuse energy that is distributed in the higher frequencies with a primary peak around 6000 Hz and a secondary peak at around 1800 Hz. This type of energy distribution is fairly typical of dental clicks although it is slightly higher than that seen in some other studies. The high frequency emphasis is due to the very small front cavity ahead of the front constriction location.

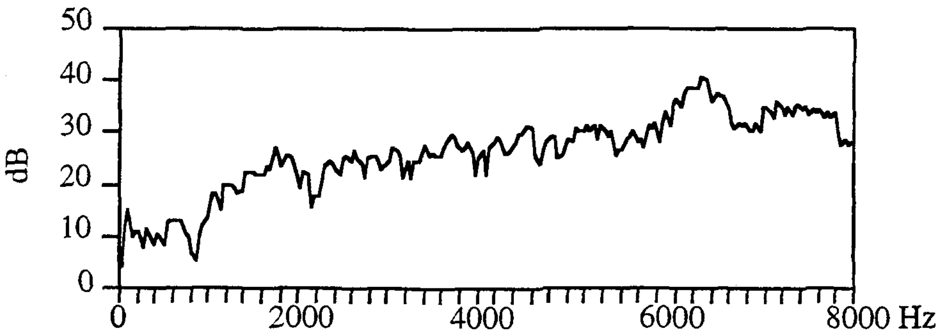


Figure 20. Mean FFT spectrum of the dental click in the word **k|ʔa** 'bad smell'.

The spectrum of the post-alveolar click, shown in Figure 21, has a high energy compact peak at around 1200 Hz, diffuse secondary peaks around 4000 Hz and 6000 Hz and a broad distribution through the rest of the spectrum. The compactness and intensity of the energy together with the very brief release noise sets the post-alveolar apart from the other two click types.

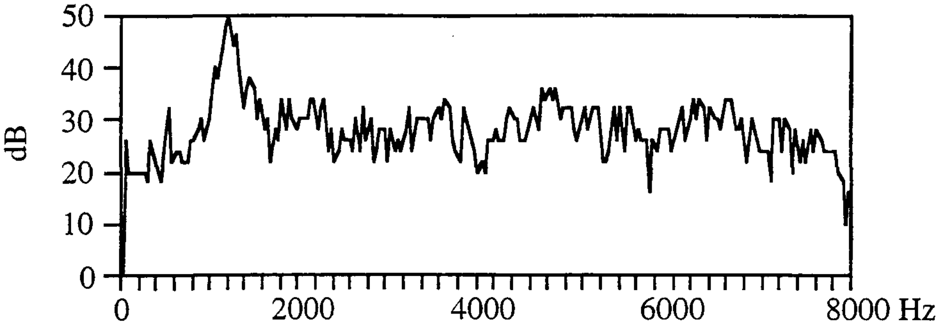


Figure 21. Mean FFT spectrum of the post-alveolar click in the word **k!ʔamba** ‘spleen’.

The spectrum of the lateral click, shown in figure 22, has peak of energy at about 2000 Hz and diffuse energy up to 6500 Hz where there is a small diffuse secondary peak. In addition to these higher frequency components, there is a lower-amplitude peak below 1000 Hz.

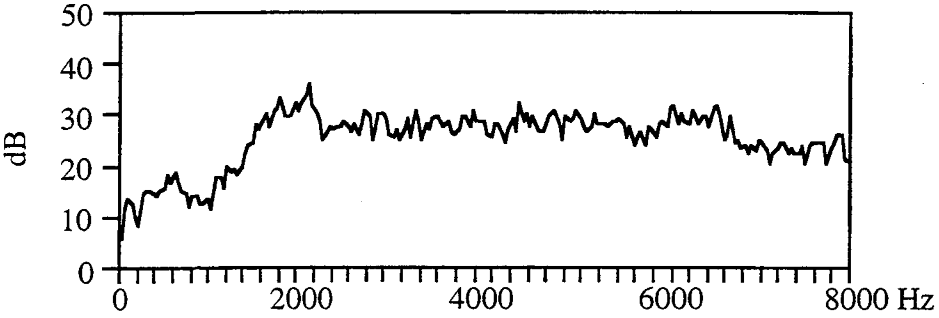


Figure 22. Mean FFT spectrum of the lateral click in the word **k||ʔaŋ** ‘warthog’.

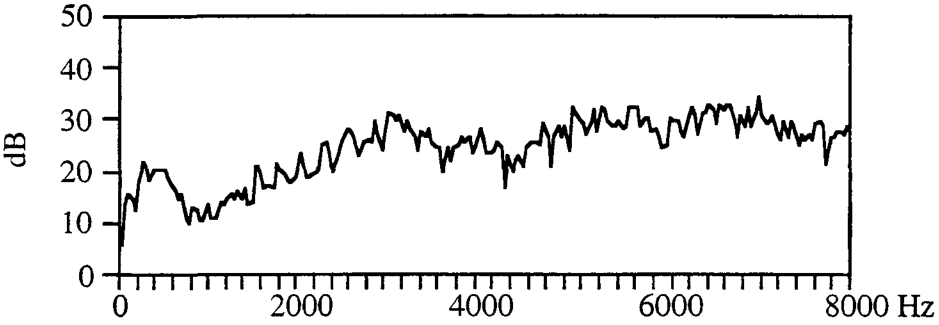


Figure 23. FFT spectrum of the ejective lateral affricate **tɬ’a** ‘to take’.

The spectrum of the ejective lateral affricate, shown in figure 23, has energy broadly distributed above 1000 Hz. There is a peak around 3000 Hz and two other diffuse peaks at 5000 and 7000 Hz. This affricate also has a low frequency peak, similar to but at a slightly lower frequency than that seen in the lateral click. This feature may be largely due to a zero at about 1000 Hz associated with the side-chamber which is created when air flows through a lateral escape channel. This would lower the energy at around this frequency, leaving a local maximum in the low frequency range.

The auditory similarity between the ejective lateral affricate and the lateral click can be partly understood in terms of spectral similarities in their noise components. Both have mid-frequency peaks between about 2000 and 3000 Hz and high-frequency peaks at around 6000 to 7000 Hz combined generally diffuse high-frequency energy. In combination with the affricate-like temporal quality of the lateral click these features may be sufficient to make the ejective affricate more confusable with the click than the spectral similarities alone might suggest.

3.2 Click accompaniments.

The acoustic patterns of the click accompaniments differ from one another in the presence or absence of periodicity and noise components and in the nature and timing of these elements. Sandawe's five click accompaniments are illustrated in Figure 23 with expanded waveforms of onsets of words beginning with dental clicks.

The three accompaniments which appear voiceless in word-initial position can be divided into two groups using the well-established measure of voice onset time (VOT). This was measured from the onset of the release transient noise to the beginning of the first identifiable glottal pulse on an expanded waveform display. The words used for the VOT measure are shown in Table 6. There were two repetitions of the three click types and the three accompaniments for all five speakers giving a total of 90 tokens. Nasalized and voiced clicks are discussed separately below. The results of the VOT measurements are summarized in Figure 24 and Table 7.

Table 6. Words used for measures of voice onset time of voiceless click accompaniments.

<table>
<tr><th></th><th colspan="2">UNASPIRATED</th><th colspan="2">GLOTTALIZED</th><th colspan="2">ASPIRATED</th></tr>
<tr><td>DENTAL</td><td>k|aː</td><td>‘leaf’</td><td>k|ʔwaː</td><td>‘wound’</td><td>k|^{h}ia</td><td>‘dikdik’</td></tr>
<tr><td>POST-ALVEOLAR</td><td>k!waʔa</td><td>‘eland’</td><td>k!ʔeː</td><td>‘anthill’</td><td>k!^{h}eŋ</td><td>‘tongue’</td></tr>
<tr><td>LATERAL</td><td>k||wa</td><td>‘name’</td><td>k||ʔeŋ</td><td>‘fire’</td><td>k||^{h}oŋ</td><td>‘cave’</td></tr>
</table>

Table 7. Mean VOT durations of the three voiceless click accompaniments.

	MEAN DURATION	STANDARD DEVIATION	STANDARD ERROR
GLOTTALIZED	61.2	22	4.0
UNASPIRATED	32.0	11.3	2.0
ASPIRATED	67.4	15.9	2.9

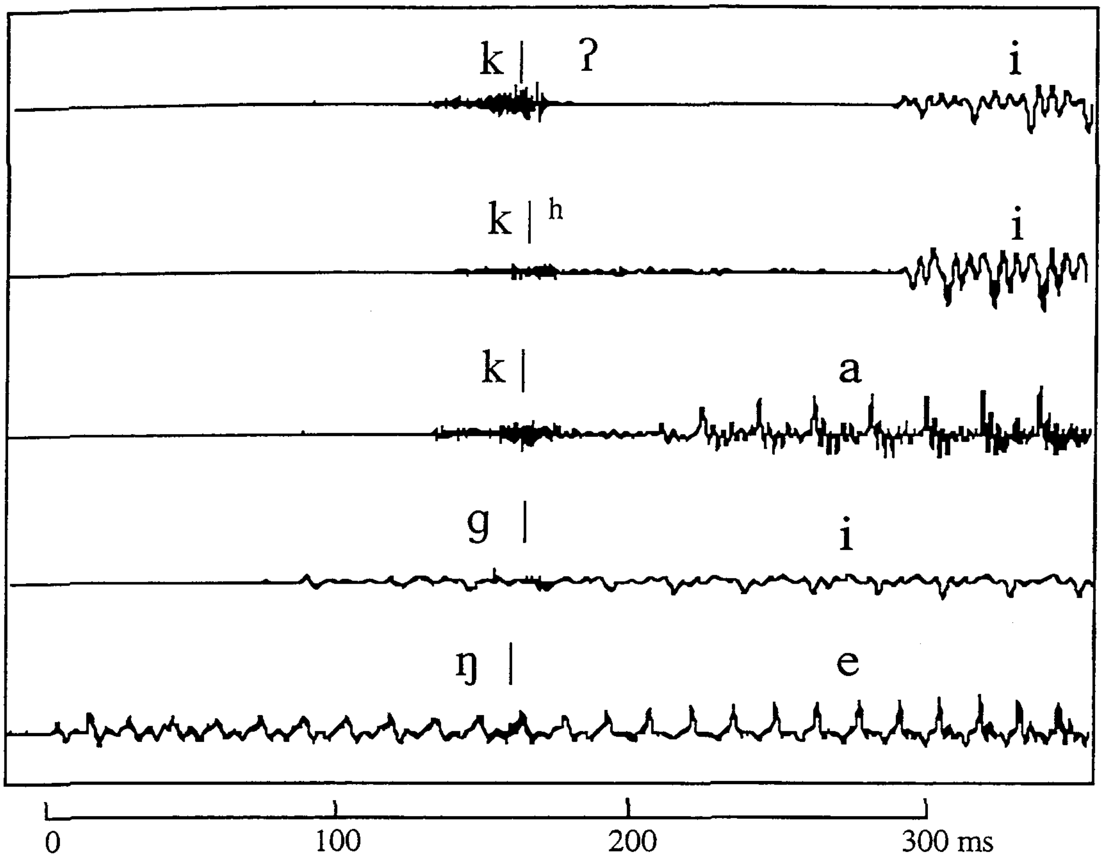


Figure 23. Example waveforms illustrating the five click accompaniments with the dental click type in the words **k|ʔiː** 'snake', **k|ʰia** 'dikdik', **k|aː** 'leaf', **g|ig|o** 'finch' and **ŋ|eʔo** 'to cut' as spoken by speaker 2.

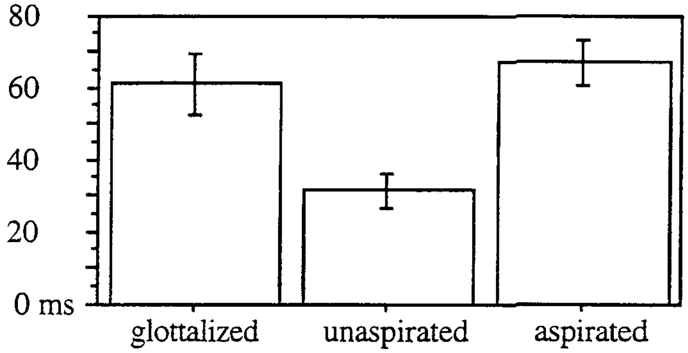


Figure 24. Mean VOT durations for three click accompaniments, pooled across click types. Error bars show one standard deviation above and below the mean.

There is very little difference between the VOT of the glottalized and the aspirated clicks both having mean durations falling between 60 and 70 ms, but the voiceless unaspirated accompaniment has a much shorter VOT of approximately 32 ms. The measurements were submitted to an analysis of variance with VOT as the dependent variable and click type and accompaniment as independent variables. The overall effect of accompaniment on VOT was significant (F[2,72]=59.8, p<.0001). The effect of click type on VOT was not significant nor was there a significant interaction between type and accompaniment.

Table 8. Words used to measure prevoicing in nasalized clicks.

DENTAL	**ŋ|weː**	'thorn'
LATERAL	**ŋ||oː**	'child'
POST-ALVEOLAR	**ŋ!iŋ**	'root'

Nasalized and voiced clicks differ from those with the other three accompaniments in that they have negative voice onset times in word-initial position. The amount of nasal prevoicing was measured from the onset of detectable voicing to the click release on the words shown in Table 7. The measurement points are illustrated in Figure 25. As we noted in connection with the aerodynamic data above, the onset of voicing does not necessarily coincide with the onset of nasal coupling or airflow. The mean duration of the prevoicing in nasalized clicks was 52 ms with a standard deviation of 25 and a standard error of 5.

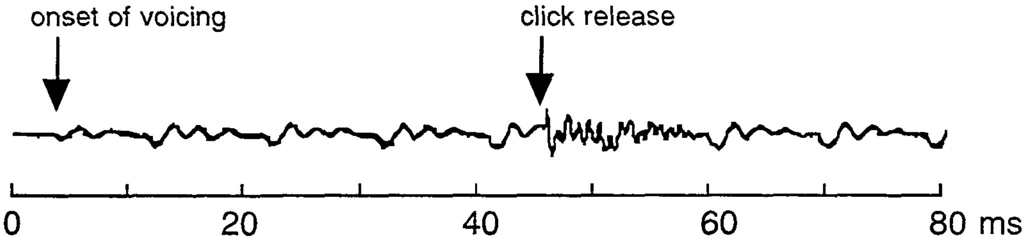


Figure 25. Expanded waveform illustrating the onset of voicing in a nasalized click and the click release in the word **ŋ||oː** 'child' spoken by speaker 2.

The data set lacked sufficient examples of the voiced accompaniment to make a statistical comparison of the duration of prevoicing in nasalized and voiced clicks. However, qualitative comparisons of the airflow and waveform displays reveals distinctly shorter prevoicing for word initial voiced clicks than for word initial nasalized clicks. There are some tokens in which the phonemically voiced clicks were produced with little or no prevoicing. We believe that tokens of this kind account for the uncertainty expressed by some earlier authors about the distinctive status of the voiced accompaniment in Sandawe since they are very similar to voiceless unaspirated clicks. However, the range of variation covered by a voiced click and a voiceless unaspirated click differ. The voiceless unaspirated clicks are never produced with prevoicing, and typically have a short VOT lag, as shown in Figure 24. Voiced clicks may have up to 30 ms of prevoicing, but may have none. Since devoicing of oral stops, particularly those with back closures, as a result of equalization of pressure across the glottis is quite common, this variation is not unexpected. The duration of voicing does not appear to be context sensitive but rather appears to be speaker and utterance dependent.

4. Summary

This paper has demonstrated several new or unusual findings concerning the phonetic possibilities for clicks, and has clarified some facts concerning the phonological patterns affecting clicks in Sandawe. The palatography illustrates the articulatory positions for Sandawe clicks, in particular confirming that the old label 'palatal' for clicks of the [!] type is a misnomer. This click type is better described as an apical post-alveolar. The common variant of this click with a tongue slap is also documented. We also confirmed the occurrence of the five distinct accompaniments in word-initial position: voiced, voiceless unaspirated, voiceless aspirated, voiced nasalized, glottalized, here transcribed **g|**, **k|**, **k|**^{h}, **ŋ|**, **k|ʔ**.

Clicks occurring intervocalically in Sandawe show anticipatory nasalization, as noted by Elderkin (1992). It is a common observation among teachers of practical phonetics that students learning to produce clicks customarily produce them with nasalization, apparently finding it easier to integrate them with a following vowel or other speech sounds if air flow can continue through the nose. Perhaps the nasalization of medial clicks in Sandawe is also a reflection of ease of production.

The acoustic study of click releases shows that there are some similarities between the spectral signature of place in clicks and in non-click consonants. Dental clicks, like dental plosives, have relatively flat spectra with a broad high-frequency emphasis. Post-alveolar clicks, like post-alveolar plosives, have a more focused frequency peak in their spectrum. The spectral properties of dental and post-alveolar plosive releases can be compared in the study of Ndumbea (Gordon and Maddieson 1995).

The demonstrable acoustic similarities between Sandawe lateral clicks and lateral ejective affricates were less substantial than might have been expected given the degree of auditory similarity noted between them.

Acknowledgments

This research was supported by a grant from the National Science Foundation for research on the phonetic structures of endangered languages. We are very grateful to the Tanzania Commission on Science and Technology and to the district and provincial authorities in Arusha and Kondoa for approving our project, and to Herman Batibo, Nicholas Blurton-Jones, the Catholic Missions at Farkwa and Kurio, Windolin Pankras and Dave and Jen Pearce for facilitating our work in various ways. We also thank D. E. Elderkin for providing us with a copy of his unpublished dissertation on Sandawe.

References

Copland, B. D. 1938. noʊts ɒn ðə fənɛtɪks əv sandawe. *lə mɛtrə fɔnetik* 16:60-64.

Dempwolff, Otto. 1916. *Die Sandawe: Linguistisches und Ethnographisches Material aus Deutsch-Ostafrika* (Abhandlungen des Hambugischen Kolonialinstituts 34, Reihe B Völkerkunde, Kulturgeschichte und Sprachen 19). L Friederichsen, Hamburg.

Elderkin, D. E. 1989. The significance and origin of the use of pitch in Sandawe. Unpublished D. Phil. dissertation. University of York, Heslington, York.

Elderkin, D. E. 1992. Predictable nasality before East African clicks. *Afrikanistische Arbeitspapiere* 29:111-119.

Gordon, Matthew and Ian Maddieson. 1995. The phonetics of Ndumbea. *UCLA Working Papers in Phonetics* 91: 25-44.

Greenberg, Joseph H. 1955. *Studies in African Linguistic Classification*. Compass Publishing Co, Branford, CT.

Kagaya, Ryohei. 1993. *A Classified Vocabulary of the Sandawe Language*. Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Tokyo.

van de Kimmenade, Martin. 1954. *Essai de grammaire et vocabulaire de la langue Sandawe. Micro-Bibliotheca Anthropos* 9: 1-118.

Ladefoged, Peter and Ian Maddieson. 1995. *Sounds of the World's Languages*. Blackwells, Oxford.

Ladefoged, Peter and Anthony Traill. 1994. Clicks and click accompaniments. *Journal of Phonetics* 22: 33-64.

Maddieson, Ian, Sinisa Spajiç, Bonny Sands and Peter Ladefoged. 1993. Phonetic structures of Dahalo. *Afrikanistische Arbeitspapiere* 36: 5-53.

Newman, James. 1970. *The Ecological Basis for Subsistence Change among the Sandawe*. (Foreign Field Research Program Publications, No 36). National Academy of Sciences, Washington, D.C.

Ohala, John J. and Carol Riordan. 1979. Passive vocal tract enlargement during voiced stops. In *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*, ed. by J. J. Wolf and D. H. Klatt. Acoustical Society of America, New York: 89-92

Sands, Bonny E. 1995. *Evaluating Claims of Distant Linguistic Relationships: The Case of Khoisan.* Ph. D. dissertation. University of California, Los Angeles.

Sands, Bonny, Ian Maddieson and Peter Ladefoged. 1993. The phonetic structures of Hadza. *UCLA Working Papers in Phonetics* 84: 67-97.

Tucker, A., and Margaret A. Bryan with a contribution by James Woodburn. 1977. The East African click languages: a phonetic comparison. In *Zur Sprachgeschichte und Ethnohistorie in Afrika: Neue Beitrage afrikanistische Forschungen*, ed. by W. Möhlig, F. Rottland and B. Heine. Dietrich Reimer, Berlin: 300-322.

de Voogt, A. J. 1992. *Some Phonetic Aspects of Hatsa and Sandawe Clicks*. Thesis for Doctorandus degree, African Linguistics, University of Leiden.

The Phonetics of Ndumbea

Matthew Gordon and Ian Maddieson

1. Introduction

The language which we refer to here as Ndumbea is one of the indigenous Austronesian languages of the southernmost part of New Caledonia, a French "Overseas Territory" in the South Pacific. The name Ndumbea more properly belongs to a people who have given their name to the present-day capital of the territory, Nouméa, as well as to Dumbéa, a smaller town to the north of Nouméa. Rivierre (1973) uses the phrase [ṇã́ã ṇḍùmbea] "language of Nouméa" for the name of the language. The spelling Drubea, using orthographic conventions based on those of Fijian, is preferred in Ozanne-Rivierre and Rivierre (1991). We prefer to represent the prenasalization overtly in the spelling of the language name, but our spelling does not note the post-alveolar place of articulation of the initial stop. In the titles of the books by Païta and Shintani (1983, 1990) and Shintani (1990) the language is referred to as Païta. This is the name of one of the clans speaking the language and it has also become the name of a modern town, Païta. The language is no longer spoken in these towns but survives in a few rural settlements to the northwest of Nouméa, and in an area reserved for the indigenous inhabitants around Unya (also spelled Ounia) on the east coast. The locations referred to are shown on the map in Figure 1. The total number of remaining speakers of Ndumbea is probably on the order of two or three hundred at most.

Previous work on Ndumbea includes a small grammatical sketch and word list by Leenhardt (1946), a detailed phonological sketch by Rivierre (1973), a grammatical sketch by Païta and Shintani (1983), an outline grammar by Païta and Shintani (1990), and a dictionary by Shintani (1990). Rivierre's work is primarily aimed at comparison of the southern New Caledonian languages, a topic that is taken further in Ozanne-Rivierre and Rivierre (1991). Ndumbea is very closely related to the Numee language, spoken in Goro and on the Île Wen, and in a slightly different form ([kweɲii]) on the Île des Pins. These locations are also shown in Figure 1.

This paper will concern itself with describing certain salient phonetic properties of Ndumbea. Ndumbea's phonological and phonetic system is of interest for several reasons. The language is one of relatively few Austronesian languages which is tonal, and it has a larger inventory of vowels than many other languages in this family. It also has a three way contrast between coronal stops, a relative rarity in the languages of the world — only 3.54% of languages in Maddieson's (1984) survey of 317 languages contrast more than two coronal stops. This rather unusual contrast will be the primary focus of this paper. Ndumbea provides a further opportunity to examine if the acoustic properties of coronals conform to a general pattern based on simple place of articulation distinctions, or whether they vary in a language-specific way and depend on differences of detail in the shape of the constriction and the nature of the release (see Ladefoged and Maddieson 1995 for an overview of this issue). However, a general description of the acoustic properties of the vowels and comments on other aspects of the consonant system will also be provided.

Materials for the present study were collected in New Caledonia in February 1993 by the second author, and include data provided by two groups of Ndumbea speakers. The principal type of data consists of audio recordings of lexical items selected from Rivierre (1973) and Shintani (1990). One group of four adults, three women and a man, was recorded at Naniouni on the west coast. Two of these subjects also provided palatographic data. A second group of two speakers, one man and one woman, was recorded in Unya.

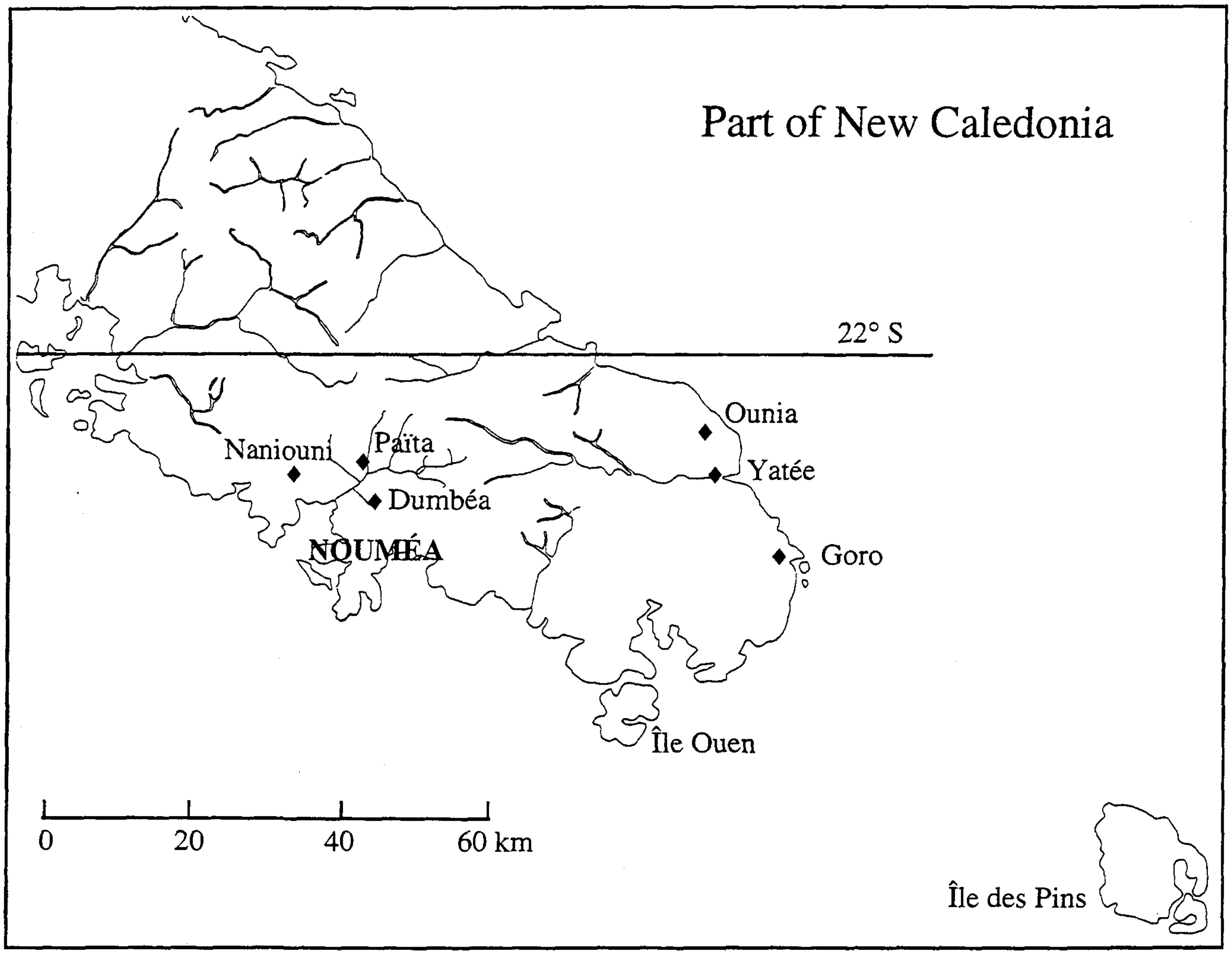


Figure 1. Map of southern New Caledonia.

2. Vowels

Interpretations of the Ndumbea vowel system differ somewhat. According to Rivierre (1973), the Unya dialect of Ndumbea has seven long and seven short oral vowels. The qualities of the long and short vowels are similar except that /ɪ/ has no short counterpart and /ɛ/ has no long counterpart. There are also five nasalized vowel qualities [ĩ, ũ, ẽ, õ, ã], which occur both short and long. This system agrees with that used in Shintani (1990), but Païta and Shintani (1983) had recognized three additional vowel qualities, [ɔ, ə, ɯ] and the nasalized vowel [ə̃].

Table 1. Ndumbea vowels (after Rivierre 1973)

	Oral		Nasalized
Short	i u ʊ e o ɛ a	Short	ĩ ũ ẽ õ ã
Long	i: u: ɪ: ʊ: e: o: a:	Long	ĩ: ũ: ẽ: õ: ã:

Long and short vowels contrast in most environments, but in Rivierre's analysis (1973: 83) only short vowels may precede the post-alveolar tap within a morpheme. The words that Rivierre and Shintani represent as containing a CV_1ɾ̠V_2 string where $V_1=V_2$ and V_1 is short seem to us to be better interpreted as containing a consonant cluster. The first vowel in such positions is usually very short and of indistinct quality. It seems primarily to serve only as onset to the tap, which, given its ballistic nature, requires an articulatory approach from a more open position to be fully realized. Back vowels do not follow the labialized bilabials and velars (/pʷ, mbʷ, mʷ, kʷ, ŋgʷ/, nor the velar nasal /ŋ/ and the velar fricative /ɣ/.

2.1 Oral vowels

In order to describe the phonetic qualities of the vowels more precisely, values for the first three formants in oral vowels in selected words spoken by the four speakers from Naniouni were measured using the Kay Elemetrics Computerized Speech Lab (CSL). The three female speakers are labeled speakers F1, F2, F3 respectively, and the male speaker is labeled speaker M1. (The two speakers from Unia are designated F4 and M2.) Data was sampled at 10 kHz. Formants were calculated for a steady state portion of the vowel at about the mid-point of its duration, as selected by eye on a spectrographic display. Superimposed LPC and FFT spectra, calculated over 30 and 25.6 ms frames, respectively, beginning with this halfway point were displayed. Formant values are usually those calculated by the LPC analysis, with the FFT spectrum being used to check the accuracy of the LPC analysis. The LPC window was calculated using 12 or 14 coefficients in most cases, though up to 20 coefficients were sometimes used in problematic cases, particularly for the back vowels.

The measured vowels were preceded by stop consonants and appeared, in virtually all cases, in monosyllabic words (although there was no observed difference in formant values according to the position of the vowel in the word). The consonants preceding the vowel were coronals before the front vowels and /a/ and velars and bilabials before the back vowels. These environments were chosen to minimize the length of consonant transitions, particularly of F2, and to provide 'prototypical' exemplars of the front and back vowels. No consistent differences were found between long and short vowels, although short vowels were sometimes sufficiently short that formants had not reached their steady state by the midpoint of the vowel, in which case measurements were taken from a slightly later point in the vowel. The number of tokens for each vowel varied, but ranged from a high of 18 tokens of /a/ for speaker M1 to a low of two tokens of /ɛ/ for speaker F1. For all speakers, /a/ and /i/ occurred in the greatest number of measurable tokens, while each of the other vowels was measured between two and six times.

Table 2. Mean values of first three formants for the vowels of four speakers from Naniouni.

Speaker & vowel		**F1**	**F2**	**F3**	**Speaker & vowel**		**F1**	**F2**	**F3**
F1	n				**F3**	n			
i	7	356	2639	3255	i	12	308	2648	3277
ɪ	2	387	2540	3003	ɪ	3	390	2580	3135
e	4	504	2445	3233	e	6	400	2348	2945
ɛ	2	583	2153	2516	ɛ	2	560	1972	2813
a	13	909	1663	3006	a	13	755	1667	2889
o	3	531	1075		o	3	554	1118	
ʊ	5	375	702		ʊ	2	406	752	
u	3	293	594		u	3	341	693	
F2					**M1**				
i	12	339	2573	3369	i	16	307	2232	2637
ɪ	3	384	2542	3030	ɪ	3	375	2134	2543
e	4	428	2275	2978	e	5	361	2026	2520
ɛ	2	461	1944	2886	ɛ	3	414	1938	2415
a	11	723	1647	2929	a	18	678	1351	2411
o	3	540	951		o	3	474	838	
ʊ	4	410	722		ʊ	4	352	603	
u	3	364	592		u	3	311	623	

The mean values of the first three formants for each vowel for each of the speakers are shown in Table 2, except for F3 values for the back vowels. These could not be reliably measured, due to background noise and faintness of energy in the higher harmonics. To provide a visual idea of the range of variation of the vowels, all measured tokens of the vowels of the three female speakers are plotted in Figure 2 in an F1 vs. F2 space. A more compact view of these distributions is provided in Figure 3, which plots the mean for each vowel of each speaker. The ellipses in Figures 2 and 3 are drawn with radii of one standard deviation along the axes of the first two principal components of the distribution. The spacing of the vowels of the male speaker in the F1 vs F2 space is similar to that of the females.

Most of the vowels are differentiated well on the basis of the first two formants in both figures, the exception being /ɪ/ in figure 2 which is completely overlapped by the ellipse for /e/ and largely overlapped by the ellipse for /i/. As shown in figure 3, however, the mean F1 and F2 values for /ɪ/ are reasonably well differentiated from both /i/ and /e/, from /i/ mainly in the F1 dimension and from /e/ on the basis of both F1 and F2. As shown in Table 2, /i/ and /ɪ/ are well differentiated by their third formant values; F3 values for /ɪ/ are lower than those for /i/ for all speakers, ranging from 339 Hz lower for speaker F2 to 142 Hz lower for speaker F3. We interpret this as indicating a more backed articulation for /ɪ/ than for /i/.

Given the vowel inventory in Table 1, one might suggest that the short /e/ and /ɛ/ are the short counterparts of the long /ɪː/ and /eː/ respectively. We have no comment to make on the phonological appropriateness of this suggestion, but our acoustic analysis indicates that the qualities are as represented in Rivierre's transcription. /eː/ and /ɛ/ sound substantially different, as borne out by the formant values, which show that /ɛ/ is more central and lower than /eː/. Although /e/ and /ɪː/ sound more similar impressionistically, their formant values are no closer to each other than /ɪː/ is to /iː/, a minimal contrast which clearly must rely on a quality difference

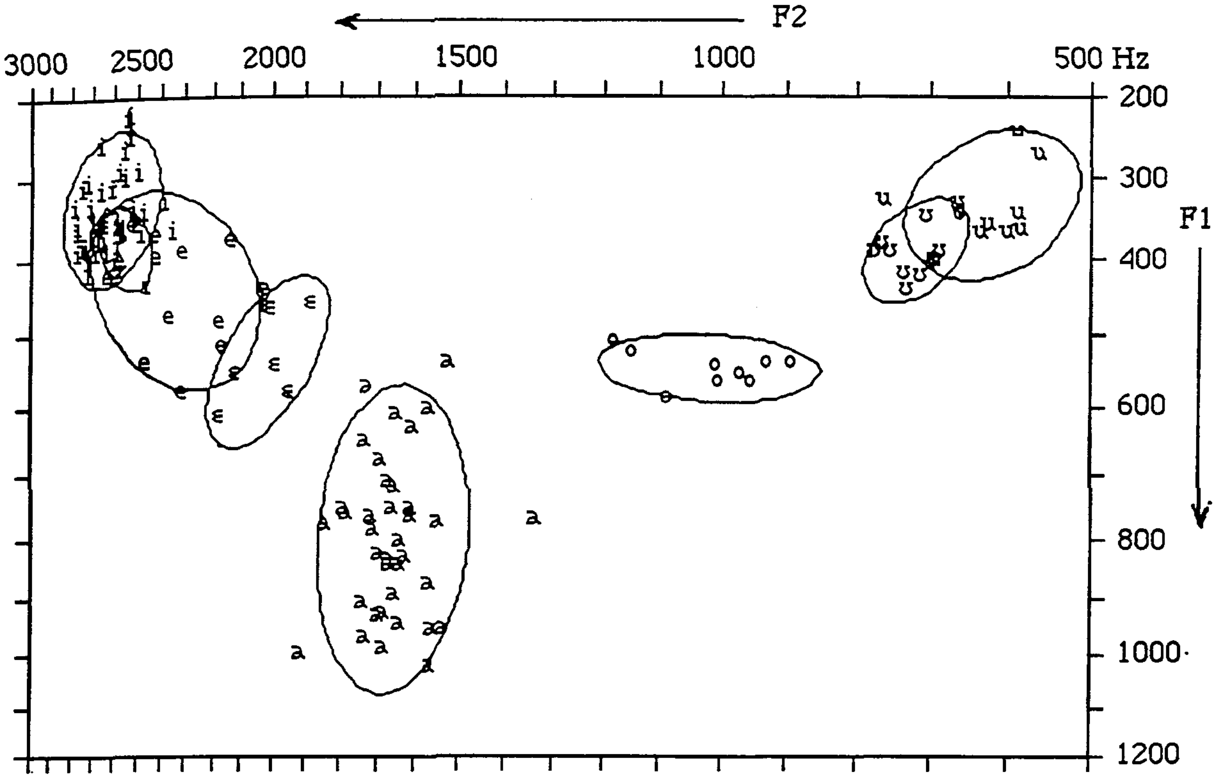


Figure 2. Plot of all measured vowel token from speakers F1, F2, F3 in F1 vs F2 space.

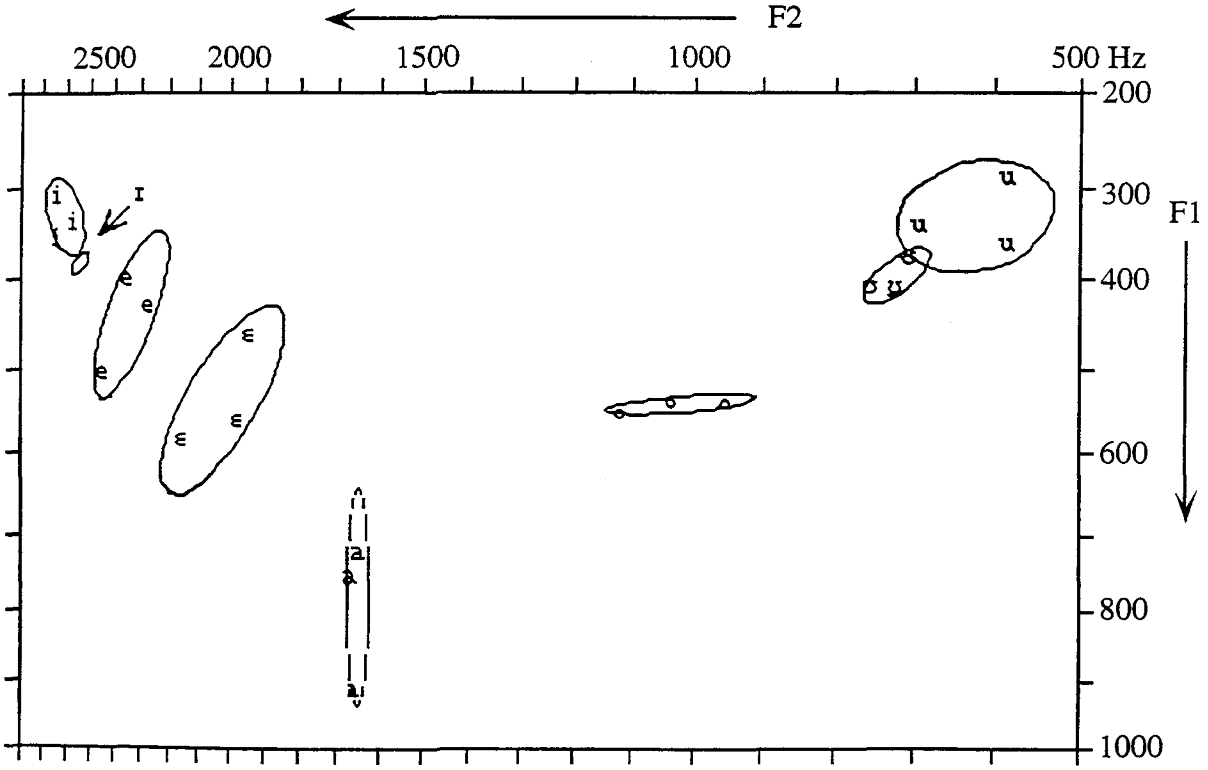


Figure 3. Plot of mean of each vowel for each of speakers F1, F2, F3 in F1 vs F2 space.

2.2. Nasalized vowels

The occurrence of nasalized vowels is largely predictable, since they are confined to environments in which nasal consonants occur, or used to occur and have left unique reflexes, as in the case of /ɣ/. (/ɣ/ is the reflex of */ŋ/ before an originally nasalized vowel; it is lost before back vowels, and the nasality of the vowel is lost.) However, nasalized vowels contrast with oral vowels after voiceless stops, the glide /w/ and the labiodental fricative /v/. Since only oral vowels follow prenasalized stops and only nasalized vowels occur after nasals, Rivierre observes that oral and nasalized vowels are in complementary distribution after these classes of consonants. This distributional pattern could also be expressed by saying that nasals before oral vowels have an oral release, thus treating the prenasalized stops as variants of the nasals. This is indeed their historical source in this language (Rivierre 1973, Ozanne-Rivierre & Rivierre 1991). Plain nasals occur in Numee where prenasalized stops are found in Ndumbea, as the language names themselves attest.

3. Consonants

Table 3 shows the consonant inventory for Ndumbea, as reported by Rivierre (1973) but generally transcribed using IPA symbols. Consonants such as /f, s, l/ which are found only in a few recent loan-words are not included in this chart. We find the same number of consonants as Rivierre, but aim to add some details of their phonetic characterization. Apart from the absence of /ŋʷ/, voiceless and prenasalized plosives and nasals occur symmetrically at the same number of places of articulation. Recall that prenasalized stops and nasals are not found in minimally contrasting positions, given their distribution with respect to vowel nasalization. The prenasalized stops are produced with a nasal portion of varying duration. A few utterance-initial tokens were observed to be produced as a voiced stop without an audible nasal portion. These are nonetheless prevoiced prior to closure release, and are thus distinct from the voiceless stops, which have a short voice onset delay.

Table 3. Consonants of Ndumbea (after Rivierre 1973)

	Bilabial	Labialized bilabial	Labio-dental	Dental	Post-alveolar	Palatal	Velar	Labialized velar
Plosive	p	pʷ		t̪	ṭ	c	k	kʷ
Prenasalized plosive	mb	mbʷ		n̪d̪	ṇḍ	ɲɟ	ŋg	ŋgʷ
Nasal	m	mʷ		n̪	ṇ	ɲ	ŋ	
Fricative			v				ɣ	
Approximant	(w)					j		(w)
Tap					ɾ̣			

The two segments listed as fricatives, /v/ and /ɣ/ (Rivierre transcribes this last sound with the symbol /x/ but notes that it is voiced), are often realized as approximants in the speech of our consultants. The approximants /w/ and /j/ are never fricated, however. /j/ is very nearly in complementary distribution with /ɲ/, as it occurs almost always before an oral vowel, whereas /ɲ/ is followed by a nasalized vowel. The production of /ɲ/ includes cases where the palatal closure in incomplete, resulting in the pronunciation of [j̃].

The places of articulation labeled dental, post-alveolar and palatal (except perhaps for the articulation of /j/) can all be considered coronal, that is, they involve the tongue tip and blade as the active articulator (Ladefoged and Maddieson 1995). The consonants labeled post-alveolar in

Table 3 are generally referred to in the literature as retroflex, but this is an ambiguous term which covers a range of degrees and modes of retracting the point of closure in a coronal consonant. Post-alveolar place is indicated by a subscript dot in preference to using the IPA symbols for retroflex consonants. The stop consonants labeled palatal are produced further forward in the mouth than prototypical palatal consonants (as exemplified in Hungarian, for example). There is no doubt that these Ndumbea consonants are appropriately classed as coronal. The characteristics of this three way coronal distinction are discussed in more detail with respect to the voiceless stops in the next section.

The post-alveolar tap /ɾ̣/ varies with a post-alveolar approximant /ɹ̣/, with the tap articulation predominating in the data set collected. In a minority of cases, this sound appeared to have been realized as an alveolar tap /ɾ/; there were also a few tokens containing an alveolar trill, involving two or three contacts between the tongue and the roof of the mouth. A following nasalized vowel typically spreads its nasality to an adjacent /ɾ̣/ with the result being either a nasalized tap, an extremely short nasal stop or a short nasalized approximant. The vowel preceding this nasal segment may also be nasalized. A spectrogram illustrating two of these variants in the word *cáɾ̣̃ã́ɾ̣ẽ* 'to run' is shown in figure 4. (Recall the suggestion that the first vowel in cases such as this might be regarded as solely a transitional element between the initial stop and the tap.) /ɾ̣/ and /ṇ/ are largely in complementary distribution, conditioned by the nasality of the following vowel and position in the word. They do not contrast in word-medial position and /ɾ̣/ does not occur initially in a word, although is very common medially and does begin some common suffixal morphemes.

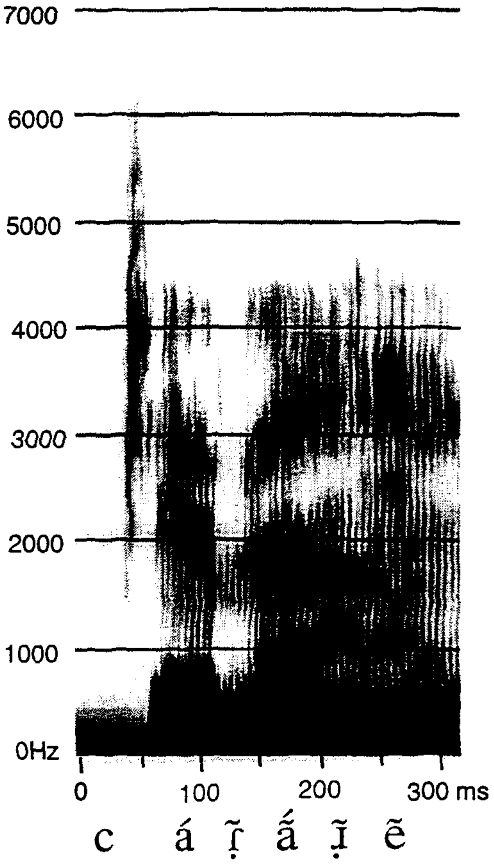


Figure 4. Spectrogram illustrating two variants of nasalized /ɾ̣/. a nasal tap and a nasalized approximant, as spoken by Speaker F1.

4. The Coronal Contrast

As mentioned above, Ndumbea is one of a comparatively small number of languages in the world which contrasts three coronal stops, i.e. stops involving contact between the tip or the blade of the tongue and tectum, or roof of the mouth. According to Rivierre, the contrast is one between an apical dental, a retroflex and a prepalatal stop, which may be followed by a period of frication following the release.

4.1 Palatography

As part of our analysis, palatograms were made of the three voiceless coronal stops as produced by two of the three Ndumbea speakers. Illustrative examples of the near-minimal triplet *t̪áa* 'reef' *ṭá* 'ground' and *cáa* 'juice' as produced by speaker F1 are shown in Figure 5. The palatograms of the other speaker show very similar articulations.

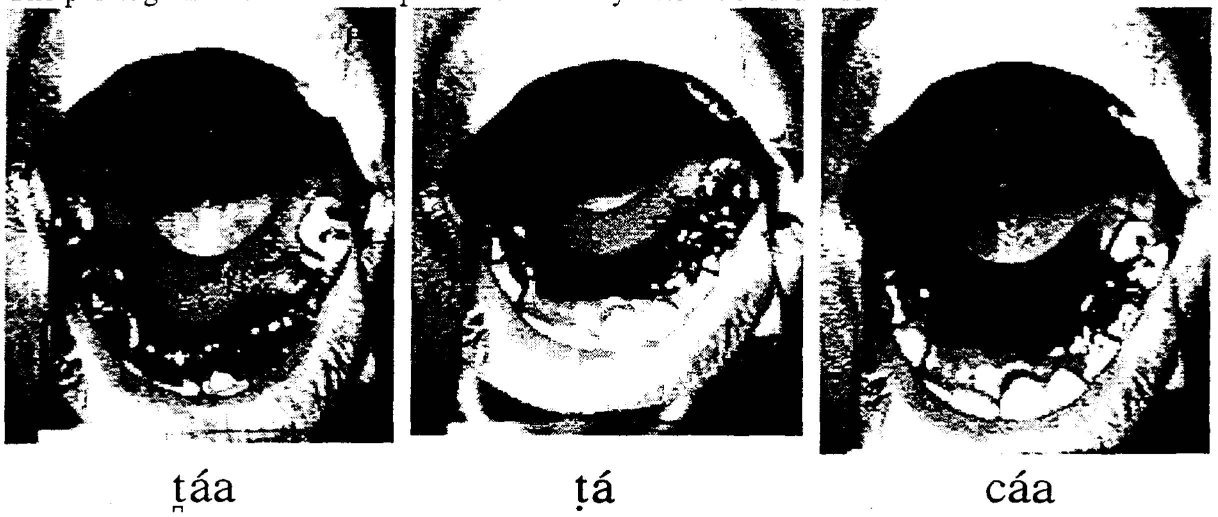


Figure 5. Palatograms of the three voiceless coronal stops of Ndumbea before /a/.

As is clear from the palatography, the contact for the dental stop is against the upper teeth and the front of the alveolar ridge. The contact is not particularly broad in the sagittal plane. Although we do not have any linguagrams to show the point of contact on the tongue directly, the palatogram for the dental stop is suggestive of an apical (or perhaps "apico-laminal" in Dart's (1991) terminology) contact between the tongue and roof of the mouth. As for the post-alveolar, its contact area is a little narrower and is markedly farther back in the mouth than that of the dental, at the back of and just behind the alveolar ridge. It does not involve any contact on the front teeth. This might also be characterized as apical. The third coronal has a very broad contact that extends a little further forward than the post-alveolar, and extends substantially further back. It is without any doubt a laminal articulation, and could be more precisely described as a laminal post-alveolar rather than as a palatal.

As in Iaai, another New Caledonian language (Maddieson and Anderson 1994), the contact area for the post-alveolar in Ndumbea is not as far back on the palate as the retroflex stops in Dravidian languages such as Tamil, Telugu (Ladefoged and Bhaskararao 1983) and Toda (Shalev, Ladefoged and Bhaskararao 1993) or the apical retroflex in the Australian language, Eastern Arrernte (Butcher, in progress, cited in Ladefoged and Maddieson 1995). In fact, although the Ndumbea post-alveolar can be classed as a type of retroflex it more closely resembles the apical alveolar stop than the retroflex of Eastern Arrernte.

4.2. Acoustic Analyses

Spectrograms of the three voiceless coronal stops in the triplet *t̪aa* 'reef', *ṭá* 'ground' and *cáa* 'juice', as produced by speaker F1 are shown in figure 6. As the spectrograms show, the stops differ in the amplitude, duration and spectral composition of their release portions. In order to describe the acoustic properties which differentiate them, as well as to draw further inferences about their articulation, a battery of acoustic analyses was performed on the three voiceless coronal stops. These analyses include the spectra of the bursts, burst amplitudes, formant transitions upon release of the stop, and durations of the closure and the noisy period following release. Closure duration and formant transitions were not examined for the palatal series for two reasons. First, the palatal stop in intervocalic position was typically characterized by an incomplete oral closure which made measurement of closure duration difficult. Secondly, the transitions are quite clearly distinct by inspection alone, with a higher origin for F2 than is seen with the other two classes of coronals.

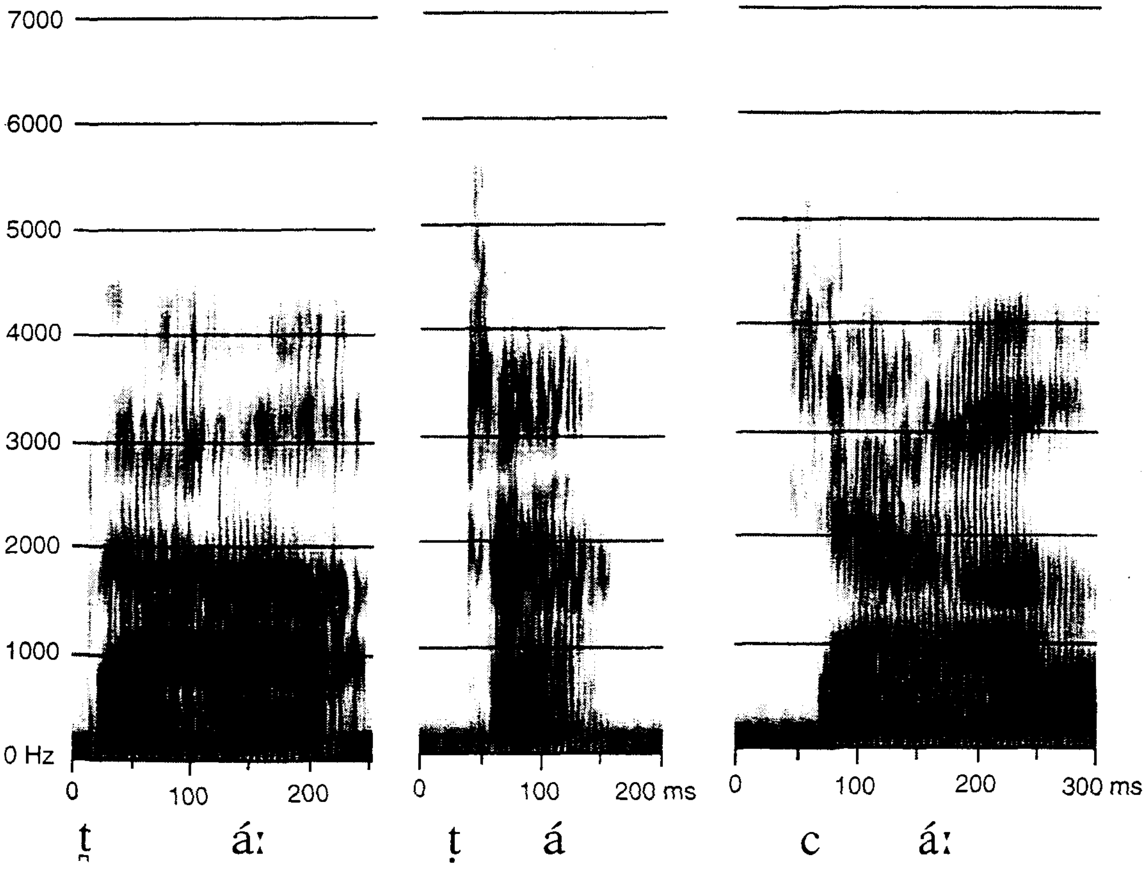


Figure 6. Spectrograms illustrating the three voiceless coronal stops of Ndumbea in word-initial position, as spoken by Speaker F1.

4.2.1. Burst Spectra

Using the Kay Computerized Speech Lab system, tokens were sampled at a rate of 20 kHz to capture the spectral properties of the burst from 0 to 10 kHz. A 256 point (12.8 ms) window was centered around the burst transient and the FFT spectrum of this window calculated. A short window was chosen since the dental stop has a short VOT and it was desirable to remain consistent in methodology for all three coronal stops. A longer window would have incorporated voiced pitch periods and formants of the vowel following the dental stop, thereby failing to represent the spectral properties of the burst itself. The words examined are listed in Table 4. Table 5 shows the number of tokens measured for each speaker by place of articulation and the following vowel. Limitations on the material recorded means that the data are not balanced across any of the three factors of consonant place, vowel context and speaker. This restricts the analysis that can be performed.

To compare the burst spectra of the different consonants, amplitude values for each point in the FFT spectrum from all tokens of a given CV sequence for a given speaker were averaged together to produce a mean spectrum. Figure 7 shows the mean burst spectra for /t̪/, /ṭ/ and /c/ in word initial position before /a/ and before /i/ from the four speakers from whom measurements in both vowel contexts were taken (Speaker F3 had no measurable tokens containing a palatal stop before /a/). The spectral displays show the frequency range from 1000-8000 Hz plotted on a logarithmic scale, thus focusing on the spectral range of greatest importance to the human listener and approximating the frequency scaling of the auditory system.

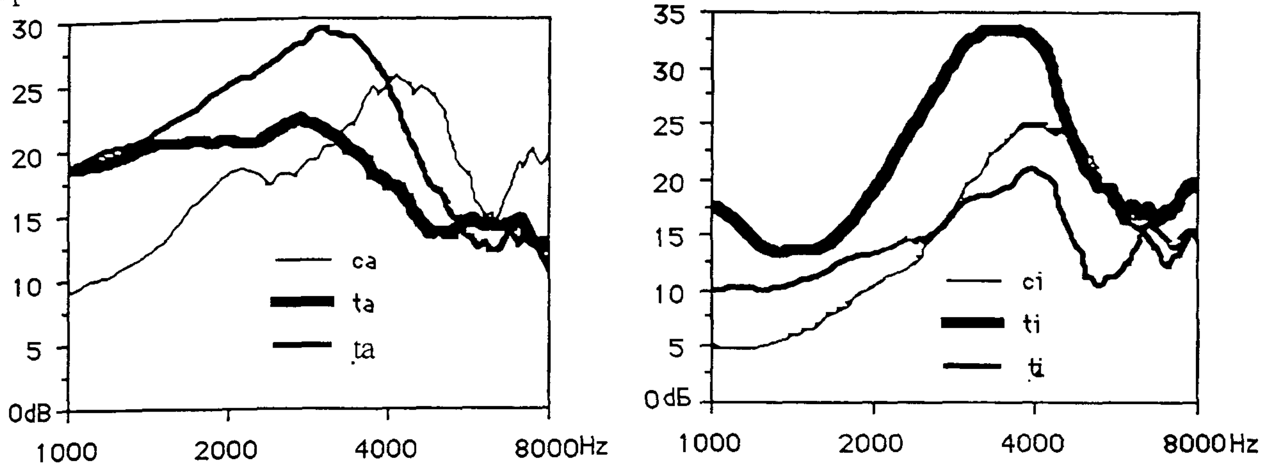
Table 4: List of words used to measure word-initial burst spectra of voiceless coronal stops.

Dental		**Post-alveolar**		**Palatal**	
Before /a/					
t̪á	'chicken'	ṭa	'ground, dirt'	cáa	'juice'
t̪áa	'reef'	ṭaúcì	'pants'	cãrã́	'run'
t̪àa	'one'				
Before /i/					
t̪íi	'blood'	ṭíi	'tea'	cící	'fishing line'
t̪í	'sugar cane'			cí	'spear'
t̪ící	'paper'			cìi	'mushroom'
t̪íi	'kidney'			cíi	'cloth ("manou")'
t̪ìi	'mute'				

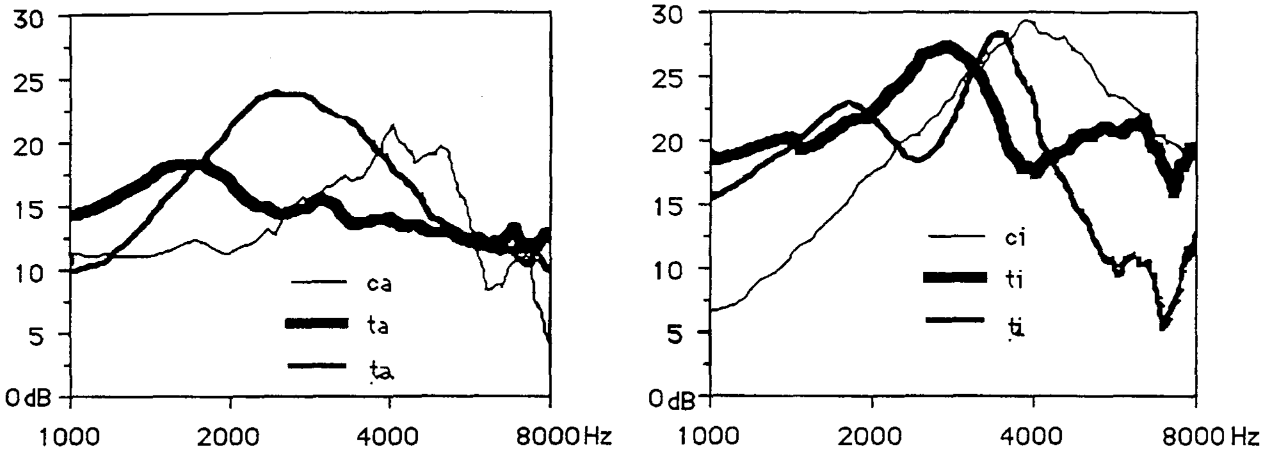
Table 5: Number of measured burst tokens by speaker, place of articulation and following vowel.

	Speaker					
	F1	F2	F3	M1	M2	F4
Before /a/						
t̪	3	4	2	2	6	5
ṭ	3	2	2	2	4	4
c	1	1	0	2	4	2
Before /i/						
t̪	3	1	3	3	0	0
ṭ	1	1	1	1	0	0
c	5	4	5	5	0	0

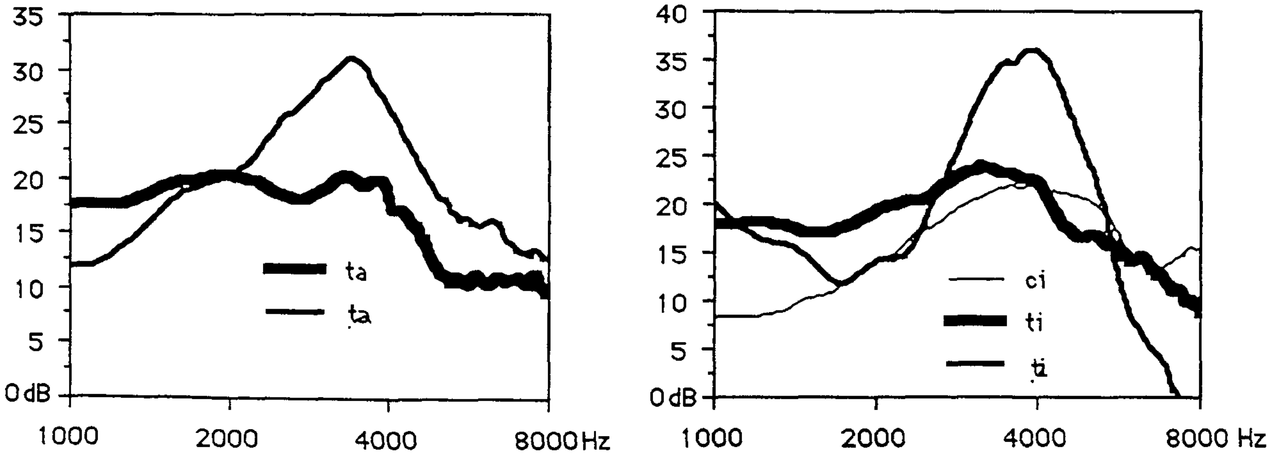
Speaker F1



Speaker F2



Speaker F3



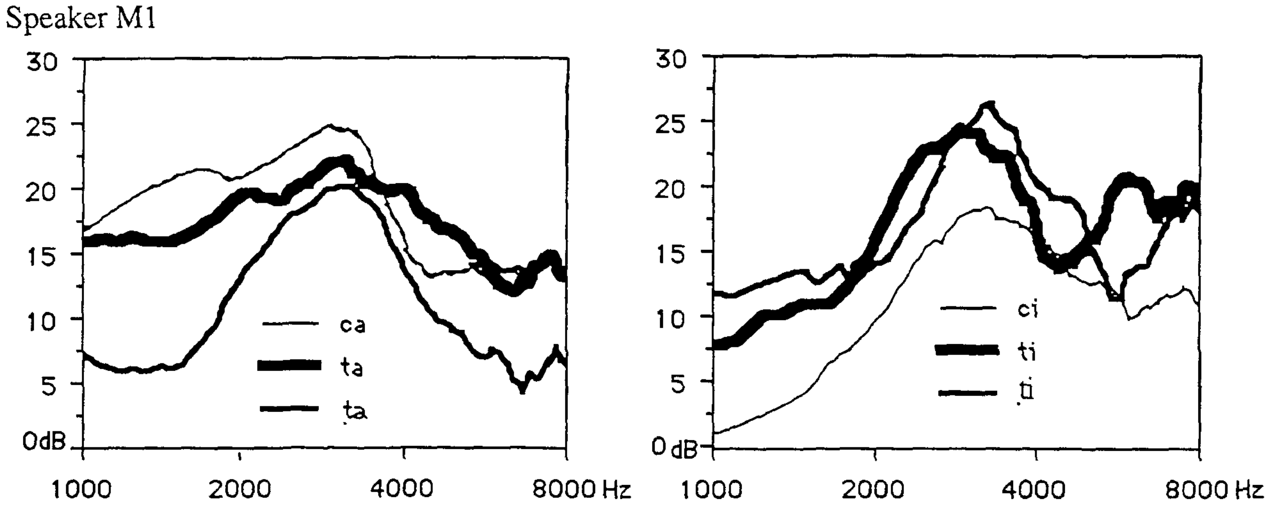


Figure 7. Mean burst spectra of word-initial voiceless coronal stops before /a(:)/ and /i(:)/ from the four speakers from Naniouni.

Spectra of the voiceless coronal stop bursts from the second group of speakers from Unia, are shown in Figure 8. This figure shows only the context before /a/, but includes the burst spectrum of the dental and post-alveolar stops intervocalically in the minimal pair /wet̪a/ 'age' (6 tokens) versus /weṭa/ 'enemy' (3 tokens) as produced by speaker M2

There are many detailed differences in such spectra, but major properties can be described in terms of three parameters, overall spectral slope, relative degree of flatness (also called 'diffuseness'), and the relative frequencies and bandwidths of principal spectral peaks. For these spectra, overall slope is not usually a good descriptor. This may reflect the fact that this measure has usually been used to characterize that part of the spectrum below 5 kHz (e.g. by Stevens and Blumstein 1978). The other parameters distinguish the consonant categories well. The spectra for the dentals, shown by the boldest lines in the figure panels, are generally flatter than those of the other coronals. A more peaked spectrum occurs before /i(:)/, particularly for speakers M1, F1 and F2. In this environment the peak is lower in frequency and broader in bandwidth than is seen with /ṭ/ or /c/. Except for the three cases of M1, F1 and F2 before /i(:)/, the overall spectral shape for /t̪/ can be described as falling. The post-alveolar stop burst, shown by the intermediate thickness line in the figures, typically has the most peaked spectrum of the three coronals. There is generally a relatively narrow bandwidth spectral peak between 3000 and 4000 Hz. The palatal stops, represented by the thin lines, are also less flat than the dental, being characterized by a peak above 4000 Hz for the female speakers, and having the lowest relative amplitudes in the lower part of the spectrum.

Some individual differences stand out in this data. Speaker M1 has a similar frequency for the spectral peak of both the post-alveolar and palatal stops, but the post-alveolar peak is narrower. Speaker F1 appears to show a greater degree of coarticulation before /i(:)/ than other speakers, since the spectral shapes of all three coronals are rather similar in this environment. In particular, there is a strong spectral peak in the region of the second and third formants of the following vowel for the dental burst. The comparison of intervocalic dental and post-alveolar stops of speaker M2, on the right of Figure 8, also seem to show the influence of a front vowel environment on the dental spectrum, in this case a preceding one.

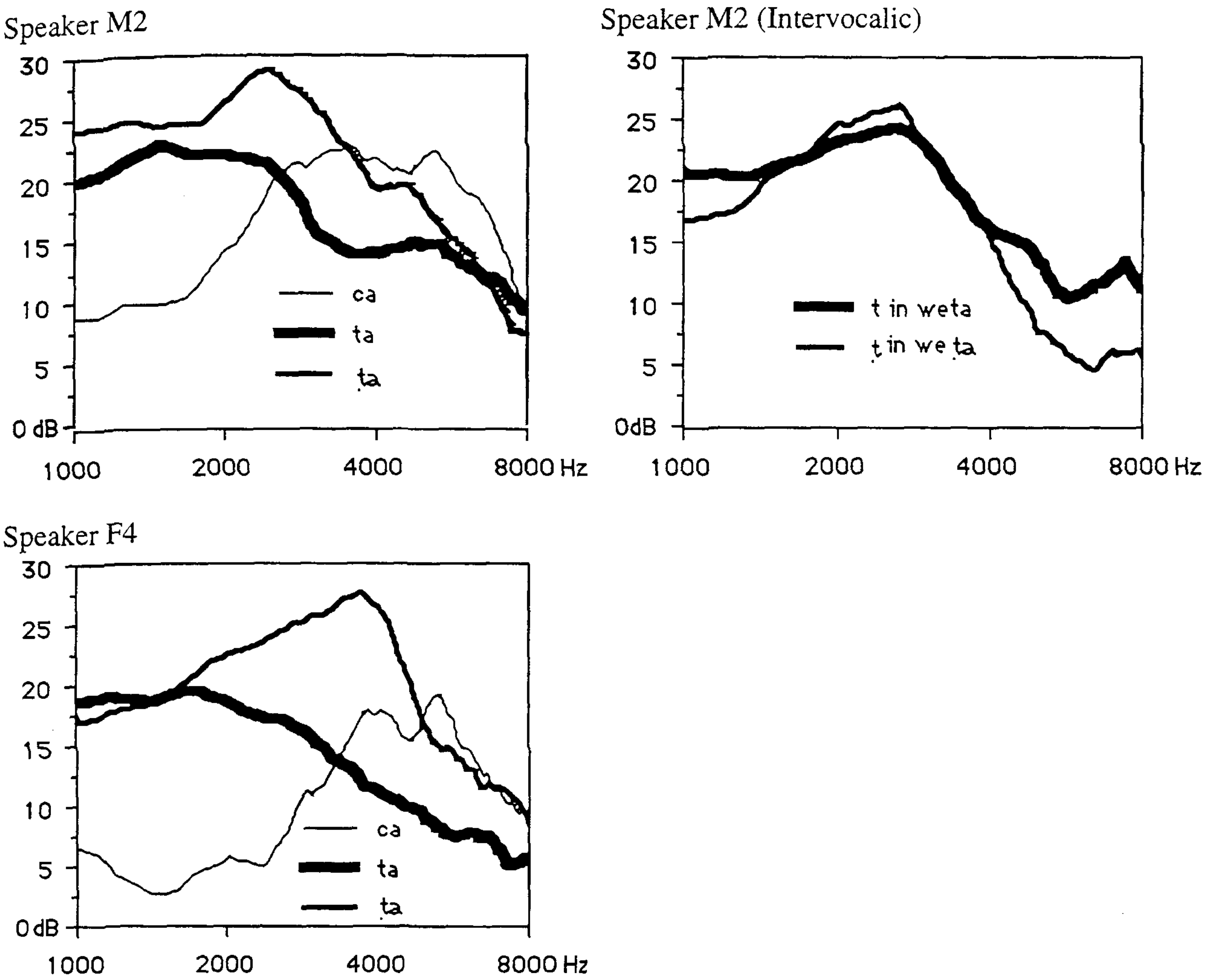


Figure 8. Mean burst spectra of word-initial voiceless coronal stops before /a(:)/ from the two speakers from Unia, and intervocalic /t̪/ and /ṭ/ from speaker M2.

4.2.2 Noisy post-release interval (Voice onset time)

The coronals are also differentiated by the amplitude and duration of the interval of noise following the release of the closure. The duration of this noisy period was measured from the moment of the release of the stop to the onset of the first formant of the following vowel. (This is essentially the same measurement as voice onset time, but we wish to clarify that the interval concerned may contain portions of both affrication and aspiration noise). Measurements were taken from the same tokens used for the burst spectra for all six speakers. Results for each speaker separately are shown in figure 9. The coronals were measured before /i/ and /a/ for the speakers from Naniouni but only before /a/ for the two speakers from Unia. For speaker M2 a word-medial comparison of /t̪/ and /ṭ/ is included.

The histograms in figure 9 suggest two principal findings. The three coronals appear to be differentiated by the duration of the noisy period, and these durations are affected by the quality of the following vowel. The palatal and post-alveolar stops have longer duration than the dental, which is markedly shorter. A tendency for the palatal to be longer than the post-alveolar is also apparent. Noise duration is longer before /i(:)/ than before /a(:)/ for every comparison that can be made. Speaker F2 has an anomalously short duration of the frication period for the palatal stop, resulting in it having a shorter duration than the post-alveolar in her case.

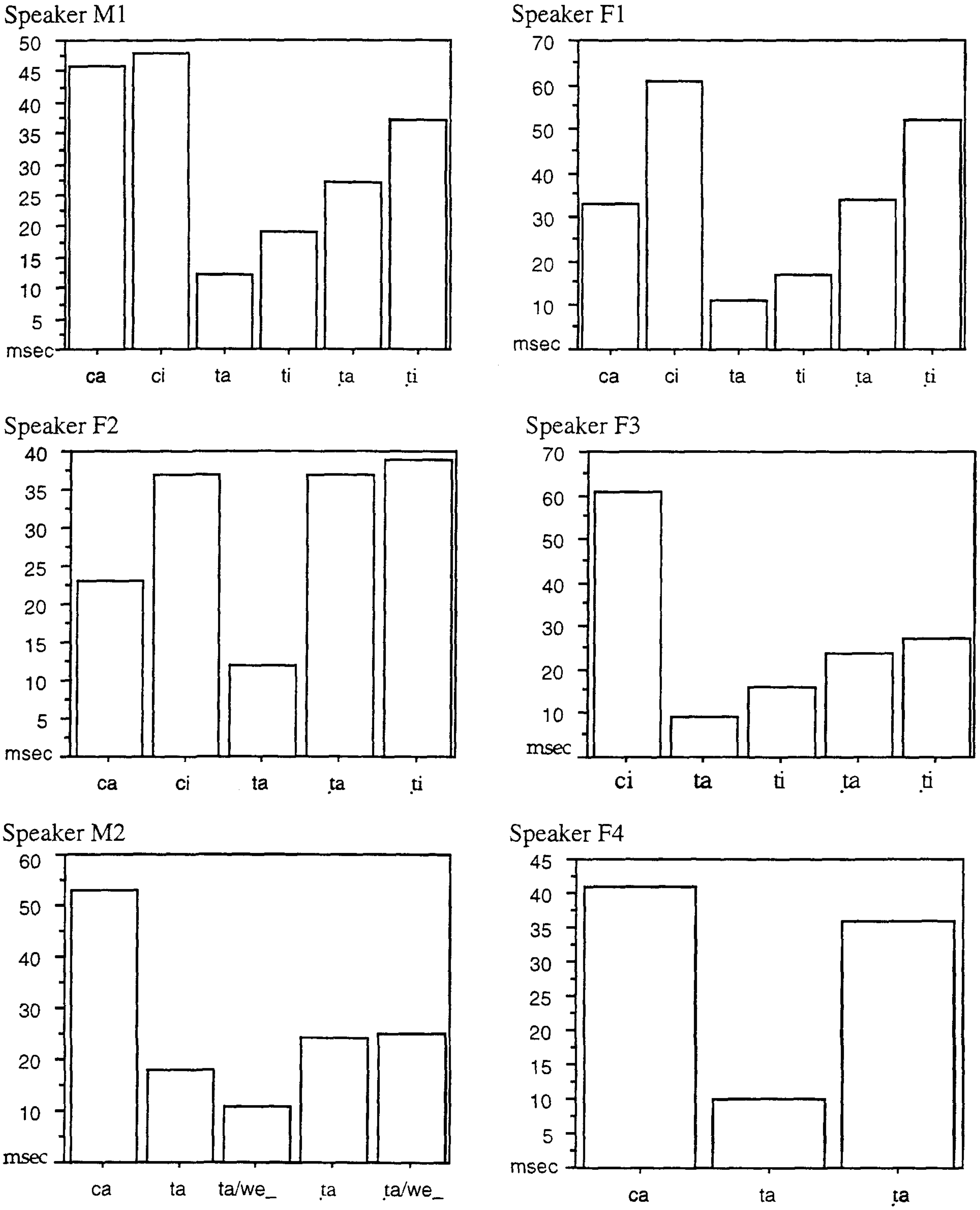


Figure 9. Histograms showing mean duration of the noisy post-release interval from six speakers.

In order to test the effects of vowel quality and place of articulation, a two factor ANOVA was performed. Because of gaps in the data set, this meant splitting the analysis into two parts. Post-alveolars and palatals were examined for speakers M1, F1 and F2 as a group, with vowel quality as a second independent variable. Results showed no significant effect of place of articulation on duration of the noisy period (F(1, 8) = .436, p = .5276). Nor did vowel quality have a significant effect, although results revealed a trend toward significance (F(1, 8) = 4.933, p = .0571). In order to test whether the failure of the effect of place of articulation on the noise duration to reach significance was due to the aberrant values of speaker F2, the test was redone using only speakers F1 and M1. Once again, neither place of articulation (F(1, 4) = .2.359, p = .1993) nor vowel quality (F(1, 4) = 5.497, p = .0790) had a significant effect on the duration of the noisy post-release period.

Dentals and post-alveolars were compared grouping together speakers M1, F1 and F3. For the dental and post-alveolar set there was a significant effect of place on noise duration (F(1, 9) = 7.039, p = .0263); the noise duration following the post-alveolar was significantly longer than following the dental. The effect of vowel quality was again insignificant (F(1, 9) = .189, p = .6737). From these two analyses we may conclude that the dental is significantly shorter than both the post-alveolar and the palatal in the duration of the following noise.

4.2.3. Burst Amplitude

Burst amplitudes of the three coronals in Ndumbea were examined using the same tokens. In order to normalize for variation in overall speech amplitude and recording conditions, measurements of the peak burst amplitude were made relative to the peak amplitude of the following vowel. All measurements were taken from a continuous energy curve using 5 ms windows calculated continuously through the time domain. The peak amplitude of the burst was subtracted from the peak amplitude of the vowel. Given the relatively low signal to noise ratio of the recordings, taking the difference of the two amplitude values resulted in a greater scatter of values and, hence, better differentiation of the stops, than is obtained by dividing the vowel amplitude by the burst amplitude. The results of the burst amplitude analysis are presented in Figure 10. A higher value indicates a burst with lower relative amplitude.

The general pattern is for the palatal to have the weakest release burst — i.e. the largest vowel-burst difference. The dental and post-alveolar release bursts are more intense, with some tendency for the post-alveolar to be stronger than the dental. This result therefore seems to be grouping the stops according to their degree of laminality. In general, the bursts of all three coronals have greater relative intensity in comparison with a following /i/ than with a following /a/, as expected given the greater intensity of /a/.

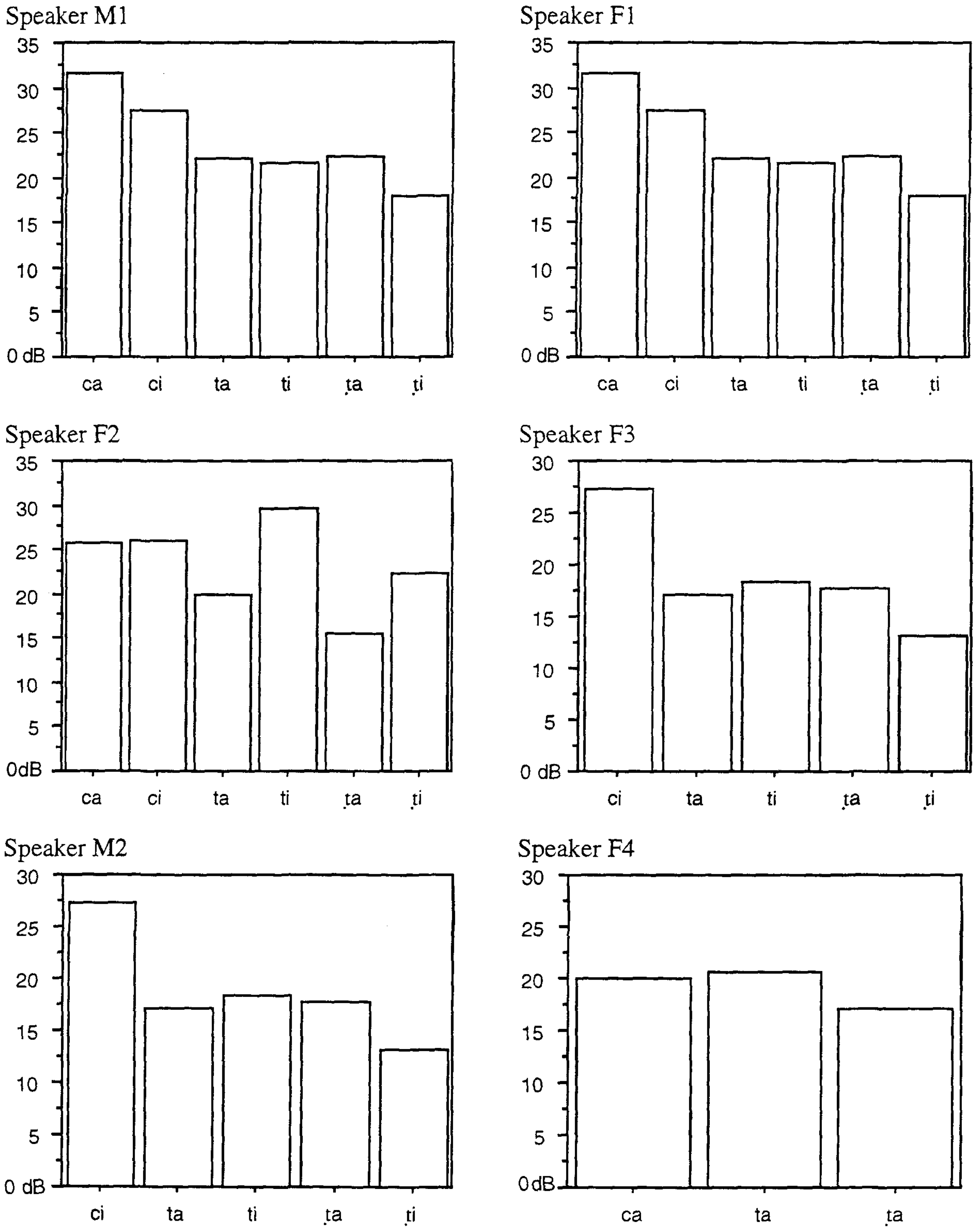


Figure 10. Histograms showing differences between peak burst amplitude and peak amplitude in the following vowel.

4.2.4. Formant transitions

The onset F2 and F3 values of /a(:)/ were measured to compare the formant transitions after word-initial voiceless dental and post-alveolar stops. A 14-coefficient LPC analysis was calculated over a 20 ms window beginning 10 ms following the release. This point represents a reasonable compromise between the times at which clear formant structure begins following these two categories of consonants. In some problematic cases, an FFT power spectrum was consulted to resolve F3 where the LPC analysis failed. Speaker M1 was not analyzed due to the presence of noisy excitation in the F2-F3 region which made formants impossible to resolve at the vowel onset.

Because there are only a few words in the data set containing short /a/ after a dental and none containing long /a:/ after a post-alveolar, the effect of the preceding consonant is confounded with a potential effect of vowel length on the formants of the vowel, which might then also extend to the onset transition. For this reason, separate analyses were performed to tease apart the effect of vowel duration and place of articulation of the preceding consonant on formant values. Words containing long and short /a/ after dentals only was examined to test the effect of vowel duration alone. This data set contained one long/short pair from each of five speakers. Formants were measured at two points using an LPC analysis: at the mid-point of the vowel over a 30 ms window (the same window used to measure formants discussed in section 2), and at the vowel onset over a 20ms window beginning 10 ms after release. F2 and F3 values for long and short vowels at the onset and the mid-point of the vowel generally differed very little from each other, as table 6 shows. None of these differences is statistically significant, but the number of data points is small.

Table 6: Formant values for long and short /a/ at onset and middle of the vowel (after dentals).

	Onset		**Middle**	
	F2	F3	F2	F3
Short	1797	2936	1671	2981
Long	1764	2938	1602	2879

The effect of the preceding consonant on formant values was examined over a set of 30 tokens balanced for speakers and place of articulation (15 following the dental, 15 following the post-alveolar). Only tokens in which the consonant was word-initial were chosen in order to exclude the possibility of coarticulation with a preceding vowel. Because of the limitations on the data set noted earlier, all 15 post-alveolar tokens contained short vowels, while 13 of the 15 dental tokens contained long vowels. The mean onset F2 was higher following the post-alveolar than after the dental, 1921 Hz to 1753 Hz, a difference of 168 Hz. In an analysis of variance, the effect of the preceding consonant on the second formant was found to be highly significant ($F(1, 28) = 14.912$, $p = .0006$). The effect of place of the preceding consonant on F3 was not significant ($F(1, 28) = .0082$, $p = .986$).

An analysis of variance in which *length* was the relevant factor confirmed that the place of the preceding consonant had a larger effect on F2 values than the duration of the vowel did. A further analysis of 34 tokens with equal numbers of long and short vowels, balanced by speakers but not by place was conducted (the set contained 17 long vowels, all after dentals and 17 short vowels, 12 after post-alveolars and 5 after dentals). This showed a difference of only 122 Hz in the F2 mean for short vowels vs. long vowels (1867 Hz vs 1745 Hz), and a lower level of significance for the comparison. Because the difference was smaller when the data set was sorted by vowel duration than by preceding consonant, one may conclude that the effect of the preceding consonant is at least greater than that of vowel duration if in fact both are present.

4.2.5. Closure duration

In our data few comparisons are available for the stops in intervocalic positions, hence little information can be given about closure duration. However, the duration of the closure was measured for the voiceless dental and post-alveolar stops in intervocalic position in the minimal pair /wet̪a/ (8 tokens) versus /weṭa/ (6 tokens) as spoken by speakers M2 and F4. Of the 14 tokens measured only two came from speaker F4, one example containing a dental and one with a post-alveolar. The measurement began with the cessation of F2 of /e/ and ended at the release of the stop closure. The mean duration for the post-alveolar was 154 ms while the mean for the dental was only 115 ms. A one factor ANOVA computed over a balanced set of 6 dental and 6 post-alveolar tokens showed this difference to be statistically significant (F(1, 10) = 6.530, p = .0286), but the limited number of tokens means that this result should be treated with caution.

4.3. Discussion

The three voiceless coronal stops of Ndumbea are quite easily differentiated in both the articulatory and acoustic domains. The dental is articulated further forward than the other two. The palatal is laminal in opposition to the other two. The dental has a short period of noise after its release, whereas the other two have longer and more intense noise. The spectra of these noisy periods show that the palatal has the highest frequency concentration of energy, and the post-alveolar has the narrowest bandwidth of noise. The onset of F2 is highest after the palatal, and significantly higher after the post-alveolar than after the dental. Closure duration appears longer for post-alveolar than for dental stops.

In many languages, dental consonants are typically laminal, e.g. in Toda (Shalev et al. 1993), and Dahalo (Maddieson et al. 1993), and they are often accompanied by considerable noisiness or affrication on their release, although this is not the case for Dahalo. Retroflex or post-alveolar stops often have a short post-release noise period, as in Tiwi and Arrernte (Anderson and Maddieson 1994, Ladefoged and Maddieson 1995), and this has sometimes been explained as a consequence of their apical articulation. The tongue-tip can move rapidly and the tongue body is lower in the mouth than for a laminal. This creates a short frication duration since the aperture increases rapidly. However, Ndumbea reverses this pattern. Although the post-alveolar is apical, it has a noisy release.

The Ndumbea post-alveolar stops shows other somewhat unexpected features. First, F3 values immediately following the post-alveolar do not differ consistently from F3 values after the dental. Post-alveolar consonants are often associated with lowered F3 values in transition from an adjacent vowel (Stevens and Blumstein 1975, Jongman, Blumstein and Lahiri 1985).. Furthermore, F2 values for the post-alveolar are higher than those for the dental. Retroflex consonants often have shorter closure durations than other coronals, but not in Ndumbea. These formant measures suggest that there is no significant sublingual cavity in Ndumbea post-alveolars, and that the tongue body and/or the jaw may be higher for post-alveolars than for dentals.

We may posit that the dental/post-alveolar contrast in Ndumbea crucially involves a more forceful articulatory contact for the post-alveolar. This is inferred from our duration measurements, and may in turn account for some of the acoustic properties, such as the apparently slower release and high F2. Further, the post-alveolar is not produced in Ndumbea with the quasi-ballistic gesture that is seen in languages with more retracted retroflex consonants. These observations reinforce the point that, despite many overall similarities, languages demonstrate many individual patterns in the relationship between broad acoustic and articulatory categories.

5. Summary

In summary, our work on the phonetics of Ndumbea has shown the following. The eight vowels of Ndumbea are for the most part well differentiated on the basis of mean F1 and F2 values, although individual tokens of /ɪ/ show considerable overlap with both /i/ and /e/. The third formant distinguishes /i/ from /ɪ/. The three types of coronal stops are easily distinguished by both articulatory and acoustic measures. However, the way that these properties are related to each other provides further evidence that there are no simple language-independent associations between acoustic realization and a given place of articulation for a class of consonants.

Much work remains to be done. Future research should include, at the very least, the analysis of the properties of the coronal nasals and prenasalized stops, the phonetic characteristics of the nasalized vowels and an analysis of the phonetics of Ndumbea tone.

Acknowledgments

Grateful thanks are extended to all the speakers who shared their linguistic knowledge with us. Stephen Schooling provided invaluable assistance in New Caledonia, and this project could not have been completed without his support. This work was funded by a grant from the National Science Foundation to Peter Ladefoged and Ian Maddieson for research on the phonetic structures of endangered languages.

References

Anderson, V. B. 1993. Acoustic Characteristics of Tiwi Coronal Stops. UCLA M.A. thesis.

Anderson, V. B. and Ian Maddieson. 1994. Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics* 87: 131-162.

Butcher, A. (in progress) *Phonetics of Australian Languages,* ms.

Dart, Sarah N. (1991) *Articulatory and Acoustic Properties of Apical and Laminal Articulations (UCLA Working Papers in Phonetics 79).* Ph. D. dissertation, University of California, Los Angeles.

Jongman, Allard, Sheila E. Blumstein and Aditi Lahiri. 1985. Acoustic properties for dental and alveolar stop consonants: a cross-language study. *Journal of Phonetics* 13: 235-251.

Ladefoged, Peter and Bhaskararao, P. 1983. Non-quantal aspects of consonant production: a study of retroflex consonants. *Journal of Phonetics*, 11: 291-302.

Ladefoged, Peter and Ian Maddieson. 1995. *Sounds of the World's Languages*. Blackwells, Oxford.

Leenhardt, M. 1946. *Langues et dialectes de l'Austro-Mélanésie*. Institut d'Ethnologie, Musée de l'Homme, Paris.

Maddieson, Ian. 1984. *Patterns of Sounds.* Cambridge University Press, New York.

Maddieson, Ian and Victoria B. Anderson. 1994. Phonetic structures of Iaai. *UCLA Working Papers in Phonetics* 87: 163-182.

Maddieson, Ian, Siniša Spajić, Bonny Sands and Peter Ladefoged. 1993. Phonetic structures of Dahalo. *UCLA Working Papers in Phonetics* 84: 25-65.

Ozanne-Rivierre, Françoise and Jean-Claude Rivierre. 1989. Nasalization/oralization: Nasal vowel developments and consonant shifts in New Caledonian languages. In *VICAL 1: Oceanic Languages (Papers from the Fifth International Conference on Austronesian Linguistics)., Part 1,* ed by R Harlow and R. Hooper. Linguistic Society of New Zealand, Auckland: 413-432.

Paita, Yvonne and Tadahiko L. A. Shintani. 1983. *Esquisse de la Langue de Paita (Nouvelle-Calédonie).* SETOM, Nouméa.

Rivierre, Jean-Claude. 1973. *Phonologie Comparée des dialectes de l'extrême-sud de la Nouvelle Calédonie.* SELAF, Paris.

Shalev, Michael, Peter Ladefoged and Peri Bhaskararao. 1993. Phonetics of Toda. *UCLA Working Papers in Phonetics* 84: 89-125.

Shintani, Tadahiko L. A. 1990. *Dictionnaire de la Langue de Paĩta.* Société d'Etudes Historiques de Nouvelle-Calédonie, Nouméa.

Stevens, K. N. and Sheila E. Blumstein. 1975. Quantal aspects of consonant production and perception: a study of retroflex stop consonants. *Journal of Phonetics* 3: 215-233.

Stevens, K. N. and Sheila E. Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64: 1358-1368.

The vowels and consonants of Amis - a preliminary phonetic report

Ian Maddieson and Richard Wright

1. Introduction

The Amis language is one of the indigenous Austronesian languages of Taiwan, spoken on the island for hundreds of years before the first arrival of the Chinese. These languages are generally called 'Formosan' based on the name given to the island by the Portuguese. Most comparative Austronesianists regard all the non-Formosan (or 'extra-Formosan') Austronesian languages as forming a large unit, the Malayo-Polynesian branch of Austronesian, in opposition to the Formosan languages (Dahl 1976). In the view of many, the Formosan languages do not themselves form a single classificatory unit, but rather fall into three groups — Atayalic, Tsouic and Paiwanic — each of which is coordinate with the Malayo-Polynesian branch (Blust 1977, 1980). In this scheme, Amis may be a Malayo-Polynesian language, as it has some features shared with the languages of the Philippines. Another Taiwanese Austronesian language, Yami, clearly belongs with Philippine languages.

Amis is spoken on the east coast of Taiwan between the large predominantly Chinese city of Hualien in the north and the city of Taitung in the south. There is also an isolated group to the south of Taitung separated from the remainder of the Amis. These locations are shown on the map in Figure 1, together with locations of some other Formosan languages. The speakers' name for themselves is Pangcah [pantsaʜ]. 'Amis' appears to be a word meaning 'northerners' but has become the standard usage in the literature (together with the variants Ami and Amitsu, accommodating to Chinese and Japanese syllable structure constraints). There are a number of distinct dialects of which northern varieties from the Hualien region are the best described (Edmondson 1986, Chen 1987). Tsuchida (1982) distinguishes twenty varieties of the language, but the dictionary edited by Fey (1986) suggests that these can be grouped into four main clusters, labeled by her Southern, Central, Kwangfu and Hualien. The isolated southern group may form a fifth cluster. The data in the present report reflects material from two locations in the Central dialect area, Kangku and Fengpin. These varieties are referred to as Makota?ay and Kakacawan by Tsuchida.

According to recent estimates there are about 130,000 Amis people (Bareigts 1976 and personal communication), but it is not known what percentage of this number is fluent in the language. In Fengpin, where our data was collected, we encountered few people under about 20 years of age who spoke Amis. Young people prefer Mandarin Chinese, which was until recently the mandatory language in schools. Parents spoke to their young children in Mandarin.

The language has several features of phonetic interest. It has a smaller than average number of distinct vowels. Lexical contrasts mainly involve only three principal vowel categories, which we will write with /i, a, u/. It is commonly assumed that languages with smaller numbers of distinctive vowels will allow greater variation in the realization of a given vowel than languages with larger numbers of vowels (for some specific discussion of this idea, see Manuel 1987). Amis provides an opportunity to test the expectation that a small vowel inventory will show great variability. The opportunity is enhanced by the presence of consonants at a wide range of places, from bilabial to epiglottal. The epiglottal consonants themselves are also of considerable interest since they represent a phonetic rarity.

No previous instrumental studies have been made of the phonetics of Amis, and the data presented here include the first technical phonetic descriptions and quantified data on the sounds of the language.

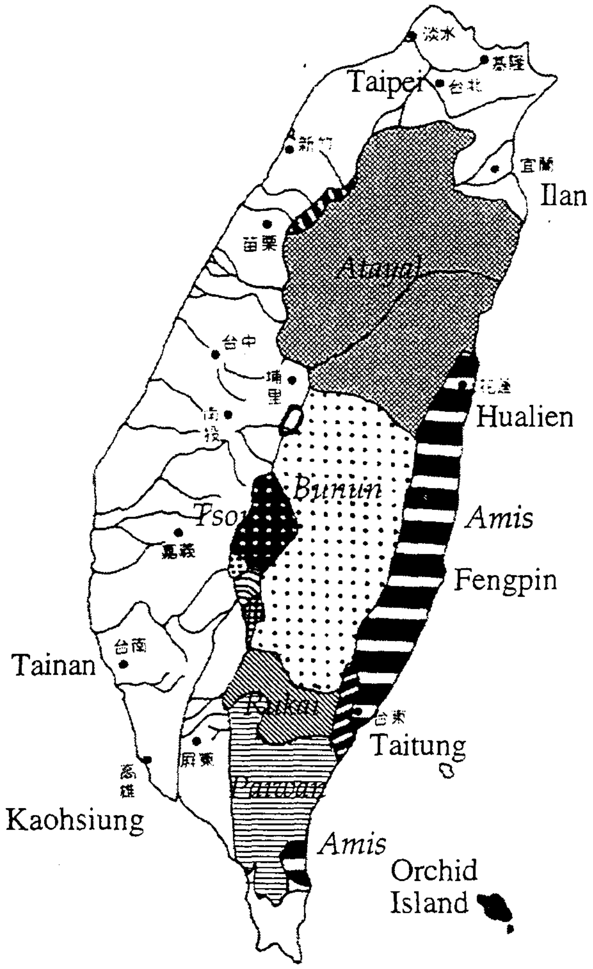


Figure 1. Map of Taiwan (after Fey 1986) showing the areas in which Amis and other principal surviving Austronesian languages of Taiwan were spoken in recent historic times. Place-names are transliterated in the form most often seen in Taiwan.

2. Data and methods

In order to study both the vowels and the consonants of the language, as well as their interactions, a wordlist was compiled illustrating the full range of consonants in intervocalic position with like vowels on each side. The wordlist therefore included words such as /siˈriʔ/ 'goat', /maˈtaʔ/ 'eye', and /vuˈŋuʜ/ 'head'. The great majority of the examples are disyllables, but a few trisyllabic forms were used where appropriate disyllables could not be found, and some speakers added affixes to disyllabic stems to make what they considered more appropriate utterances. A few other words, such as the numbers from one to ten and some examples to illustrate consonant clusters, were also included. Lexical items were selected principally from Fey (1986), but Duris (1969) was also consulted, and some additional words were suggested by André Bareigts and our language consultants. Not every consonant-vowel combination could be included in the list because of systematic gaps of various types as well as limits on available lexical data, but the great majority of possible VCV sequences are represented. Consonants before and after the VCV of interest are not controlled. Stress regularly falls on the final syllable, moving from a stem to a suffix if one is added.

This wordlist was recorded on February 23 1995 by two small groups of speakers representing the varieties of the language spoken at Kangko (Makuta in Amis) and Fengpin. Kangko is about 15 km south of Fengpin. Each subject was recorded separately onto digital audio tape using a head-mounted microphone positioned about 5 cm from the mouth. The resulting recordings have a very good signal-to-noise ratio and a wide frequency range. There were two repetitions of each token by each speaker, resulting in eight examples of each word, apart from a few cases of errors and substitutions.

The consonantal articulations have been studied qualitatively by direct observation and listening, as well as by examination of waveforms and spectrograms. For the vowels, LPC-based estimates of the frequencies of the first three formants of the steady state portion (or nearest approximation to a steady state) before and after the intervocalic consonants were obtained using the Kay CSL system. A window of 20 msec was used on signals sampled at 20 kHz. The same formants were measured at the onset and release of the consonants in order to show the magnitude of transitions. To date, only the vowels of the four Kangko speakers has been analyzed in this way; similar results are, however, anticipated from the Fengpin group.

3. Consonants

The consonants of the Central dialect of Amis as spoken in Kangko can be represented by the chart below.

	Bilabial	Labio-dental	Dental and Alveolar	Palatal	Velar	Epiglottal	Glottal
Plosives	p		t		k	ʡ	ʔ
Affricates			ts				
Nasals	m		n		ŋ		
Fricatives		v	s		(ɣ)	ʜ	
Trills			r				
Lateral fricatives			ɮ				
Lateral flaps			ɺ				
Approximants	w			j			

A number of interpretative comments can be made on the chart. The voiceless oral plosives and the affricate /ts/ are audibly released in all positions. Thus *ccay* 'one' is [tstsaj]. A very brief voiced vowel may intervene between clusters of obstruents such as this, so 'one' could be transcribed as [ts^{ə}tsaj] and *spat* 'four' as [s^{ə}pat]. An audible voiceless release occurs in final position; Chen (1987) notes the same phenomenon in northern Amis, but misleadingly describes it as aspiration. A final glottal stop often has no audible release. In our recordings the voiced fricatives /v, ɮ, ɣ/ are regularly devoiced in word-final position, and occasionally also in word-initial position. As the words were produced in isolation, this devoicing process may well be an utterance level effect. The lateral fricative /ɮ/ is pronounced with a post-dental or interdental articulation. The velar fricative /ɣ/ is parenthesized in the chart since it occurs only in a few loan words, such as /ri'ɣiʔ/ 'ridge between sections of a rice field'. The sibilant alveolar affricate and fricative, /ts/ and /s/, may optionally be pronounced as palato-alveolars (laminal post-alveolars) before /i/. The approximant /j/ does not occur initially in native words, but does occur in word-medial and final positions. The sequences /wu/ and /ji/ are systematically absent.

Despite their geographical proximity, the Kangko and Fengpin varieties have one salient difference in their consonantal systems. Where Kangko speakers use a lateral fricative, Fengpin speakers use a nonsibilant central dental fricative, [ð̪]. In northern Amis varieties, this corresponds to a voiced stop, [d], which can be laxed to [ð] in intervocalic position. Thus *ada* 'enemy' is [ʔaɮaʔ] in Kangko, but [ʔað̪aʔ] in Fengpin, and optionally also [ʔað̪aʔ] in northern Amis (Chen 1987).

Amis has two sounds that might be described as rhotics — an alveolar trill and a lateral flap which is often post-alveolar. Examples of these segments medially in the environment /u__u/ are illustrated in the spectrograms in Figures 2 and 3. Both are characterized by a sharp lowering of the third formant, which reaches a level very close to the second formant in these examples. The trill typically has two or three contacts; the example shown has three in both the medial and final positions in the word. The flap has an exceedingly short closure; 10 ms or less is not unusual. This segment requires a relatively open vocal tract configuration on either side; hence when it is not preceded or followed by a full vowel its production entails a brief vowel-like approach or release portion. This property can be observed in word-final position in the spectrogram of /ʡuʡuɺ/ "fog" in Figure 4 below, and preceding another consonant in a consonant cluster in Figure 13.

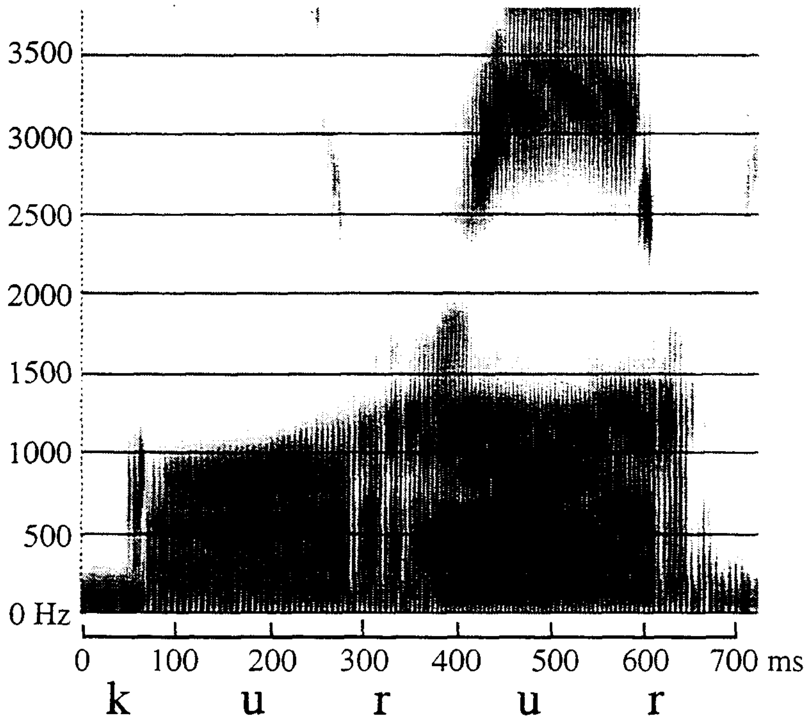


Figure 2. Spectrogram illustrating the alveolar trill in the word /kurur/ 'upper part of the back' as spoken by speaker S2 (female).

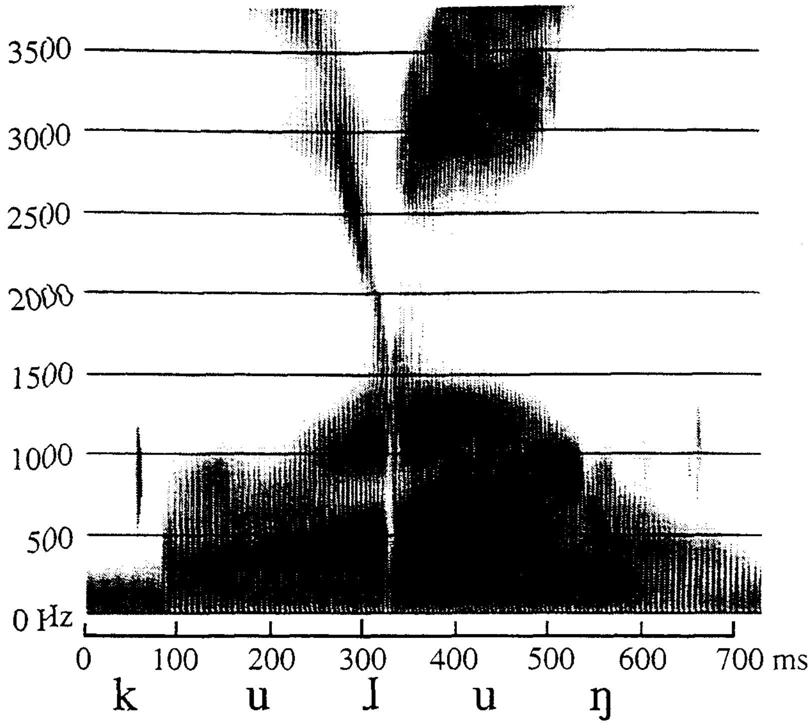


Figure 3. Spectrogram illustrating the lateral flap in the word /kuɺuŋ/ 'cow' as spoken by speaker S2 (female).

The most unusual consonants in this inventory are the two sounds which we analyze as the epiglottals /ʡ/ and /ʜ/. These are distinct from the glottal stop. In most other Formosan languages uvulars are found corresponding to the epiglottals, but similar segments occur in some dialects of Atayal (Li 1981). These unusual sounds have proved difficult to describe for several of those who have written about the language previously. Fey (1986) talks of Amis having "heavy and light glottal stops", in reference to the distinction between /ʡ/ and /ʔ/. Duris (1969), a native French speaker, perceives a similarity with French [ʀ] and describes what he writes with 'h' as "un R guttural" or "un R dur". Our perception is that /ʡ/ and /ʜ/ are a stop and a voiceless fricative made in the low pharyngeal region. The closest constriction, or a closure, in this part of the vocal tract is likely to be at the level of the epiglottis, and these sounds may involve active lowering of the epiglottis, as suggested by Laufer and Condax (1979, 1981), or be formed by strong retraction of the lowest portion of the root of the tongue. Chen (1987) does note pharyngeal articulations, writing [ħ] as the syllable-final allophone of /h/, and [ʔħ] as the word-final realization of the phoneme /q/, which is a glottal stop in all other positions. We are not sure what articulation is intended by the transcription [ʔħ], but it may represent a glottal stop with a secondary pharyngeal constriction. Chen's discussion suggests that the northern varieties of Amis may have a different distribution of sounds in the pharyngeal and glottal regions from Central dialects, but we are not sure if this is so.

In Central Amis, /ʔ/, /ʡ/ and /ʜ/ all occur in comparable environments. We are therefore inclined to treat them all as contrastive. /ʡ/ and /ʜ/ have quite marked transitions to and from adjoining vowels, whereas /ʔ/ does not. /ʜ/ patterns like /s/ in that it remains voiceless in all positions in the word. /ʡ/ often shows a number of impulse-like energy spikes as the closure is made and as it is released. We think that these are trill-like vibrations of the epiglottis.

Spectrograms illustrating /ʡ/ and /ʜ/ in comparable intervocalic environments are shown in Figures 4 and 5 for /u/, Figures 6 and 7 for /i/ and Figure 8 and 9 for /a/. Note the very strong raising of the second formant, and lowering of the third formant at the margins of the consonant in the /u__u/ context for both /ʡ/ and /ʜ/. The transitions are not as extensive in /i__i/ but both the second and third formant show some lowering. The first formant shows some raising in both these vowel contexts, but F1 is raised the most in the context of /a/. The vowel /a/ also shows a raised second formant and a lowered third formant in the environment of an epiglottal.

Because these transitions do not seem to show a clear locus, that is, an apparent point of origin in the consonant that is common across environments, it may be inferred that there is also some coarticulation of the epiglottal consonants with their vowel context.

The spectrograms in Figures 4 and 7 also illustrate an initial /ʡ/ and a final /ʔ/ in /ʡuʡuɹ/ and /tiʜiʔ/ respectively. A systematic comparison of the three consonants /ʔ/, /ʡ/ and /ʜ/ in final position is provided by the spectrograms in Figures 10-12 which illustrate them after the same vowel, /u/. Whereas the final glottal stop causes no substantial transition of the formants of the preceding vowel in Figure 10, the two epiglottal consonants show the second and third formants converging towards each other. This convergence is almost complete by the time that the closure for /ʡ/ is formed in /ŋusuʡ/ in Figure 11, but the formants continue to converge during the frication period of /ʜ/ in /vuŋuʜ/ in Figure 12. The final release of the stop constriction is also very clear in Figure 10.

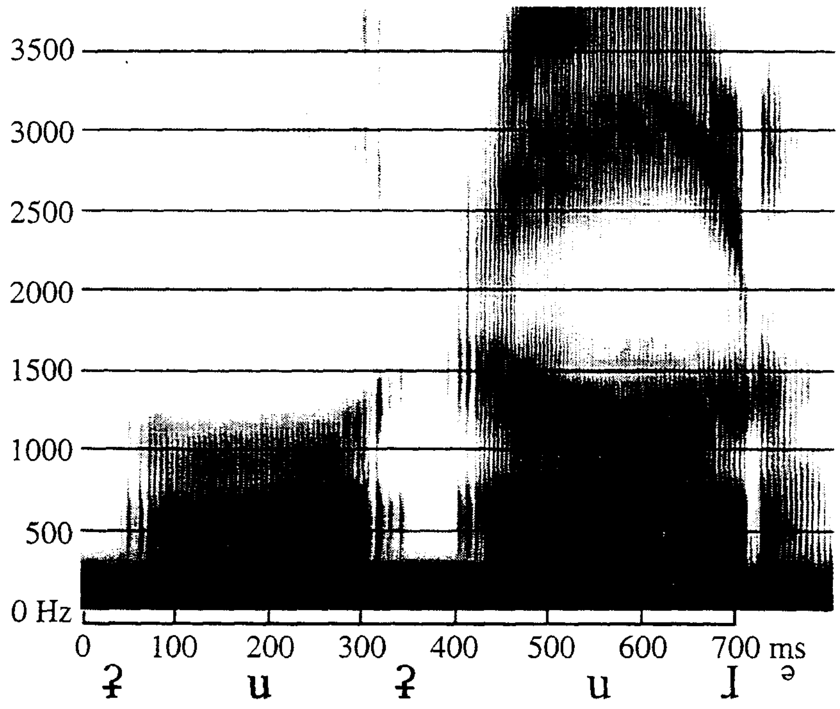


Figure 4. Spectrogram of the word /ʡuʡuɹ/ "fog" as spoken by Speaker S2.

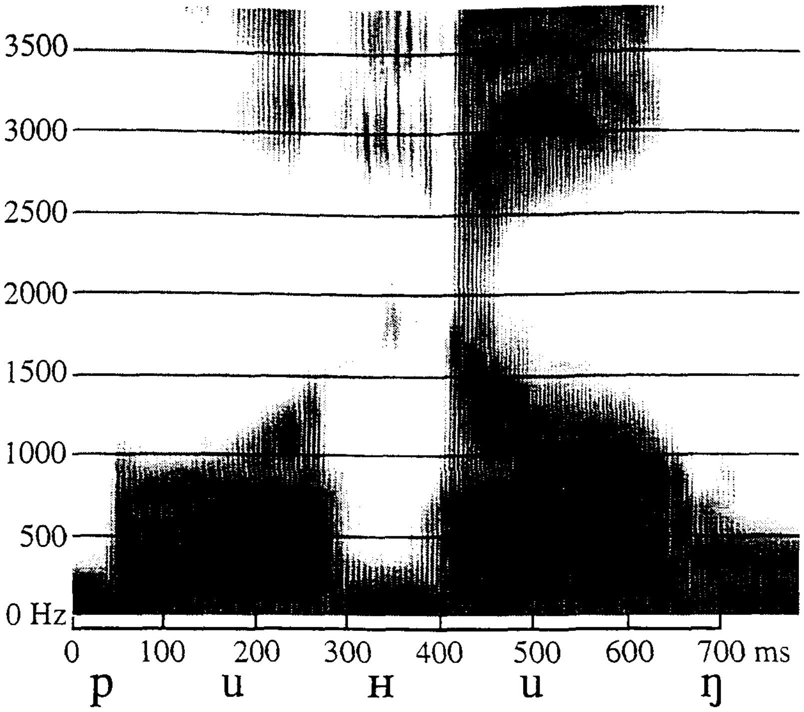


Figure 5. Spectrogram of the word /puʜuŋ/ "horn" as spoken by Speaker S2.

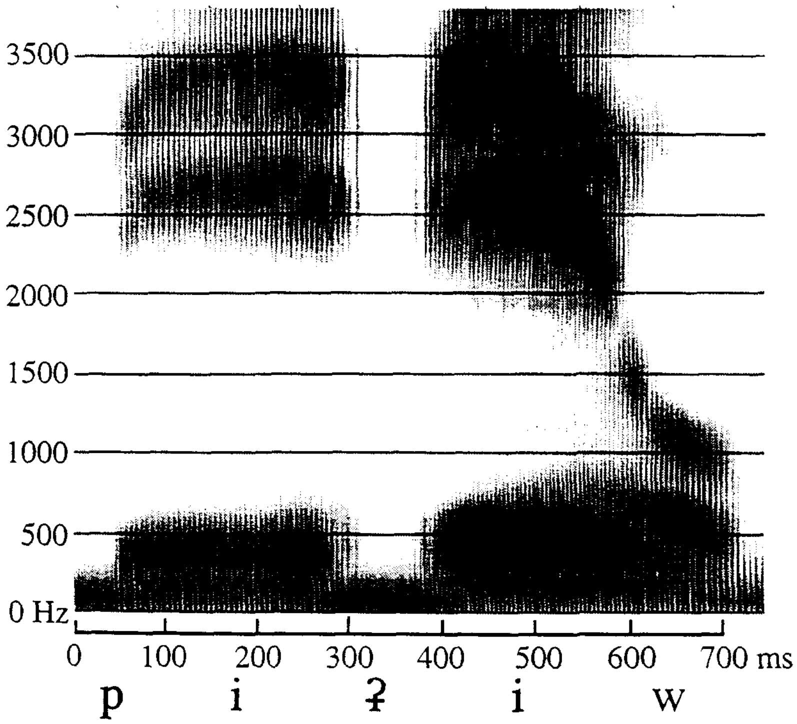


Figure 6. Spectrogram of the word /piʡiw/ "cripple" as spoken by Speaker S2.

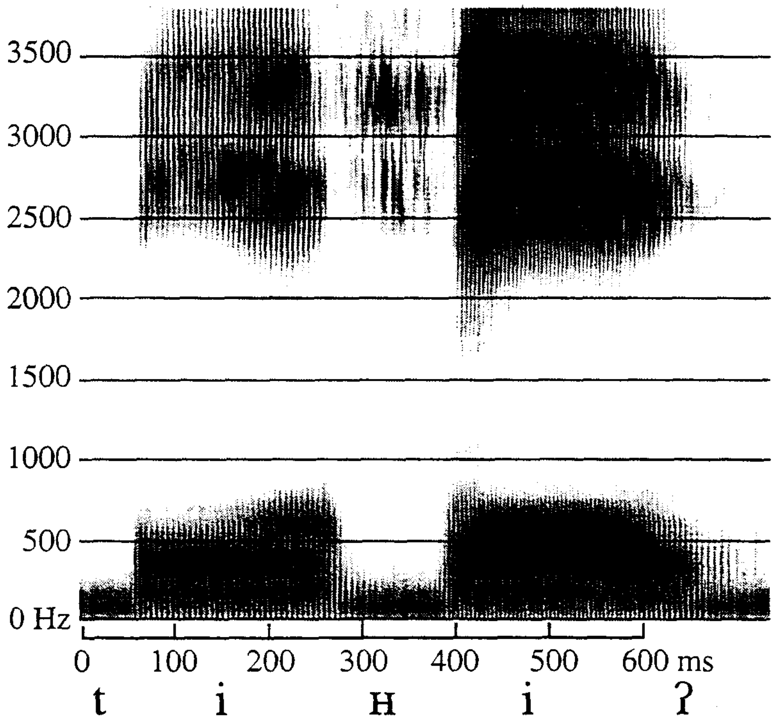


Figure 7. Spectrogram of the word /tiʜiʔ/ "spouse" as spoken by Speaker S2.

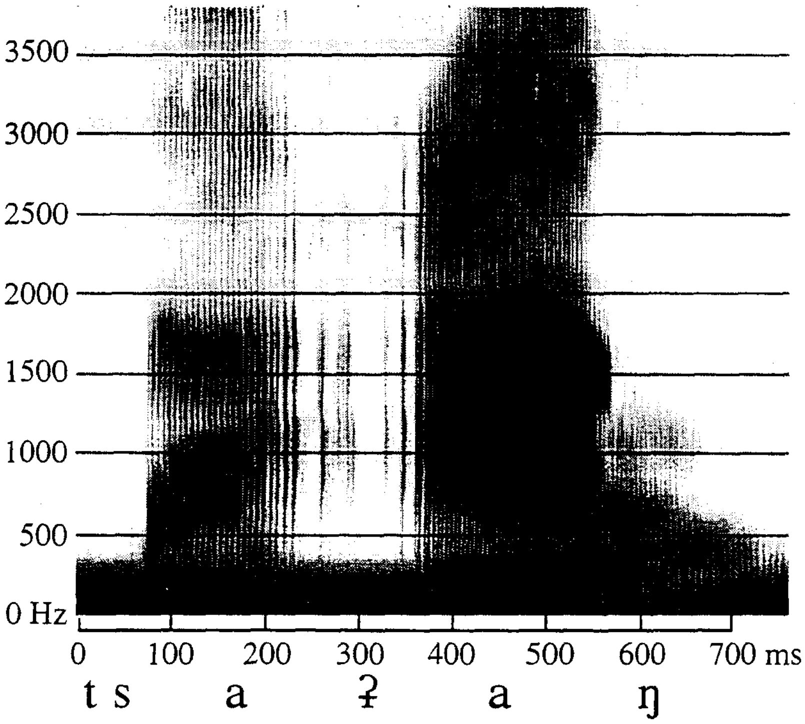


Figure 8. Spectrogram of the word /tsaʡaŋ/ "branch" as spoken by Speaker S2.

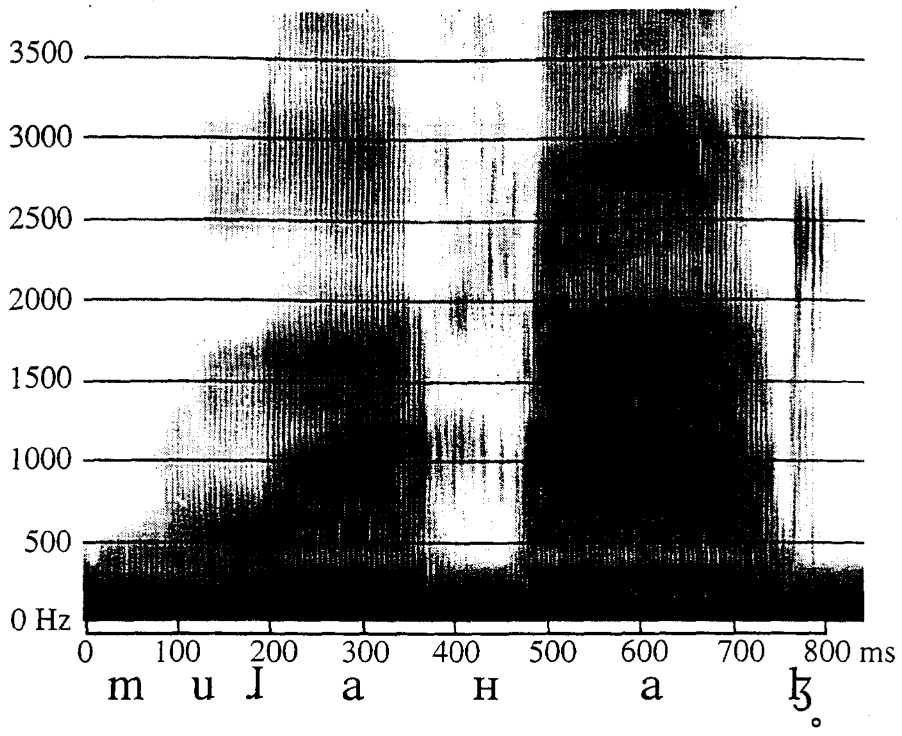


Figure 9. Spectrogram of the word /muɹaʜaɮ/ "growing" as spoken by Speaker S2.

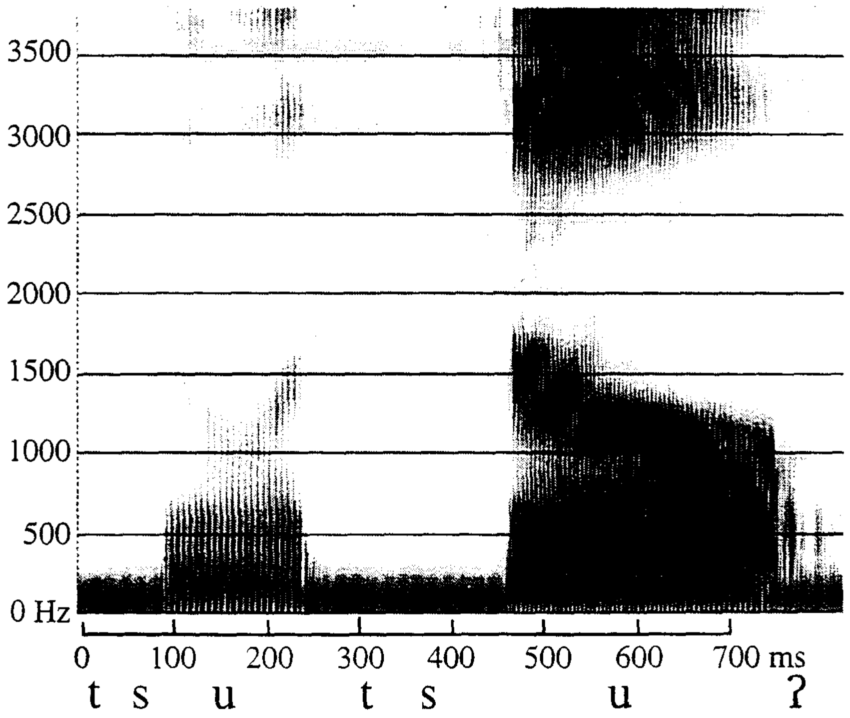


Figure 10. Spectrogram of the word /tsutsuʔ/ "breast" as spoken by Speaker S2.

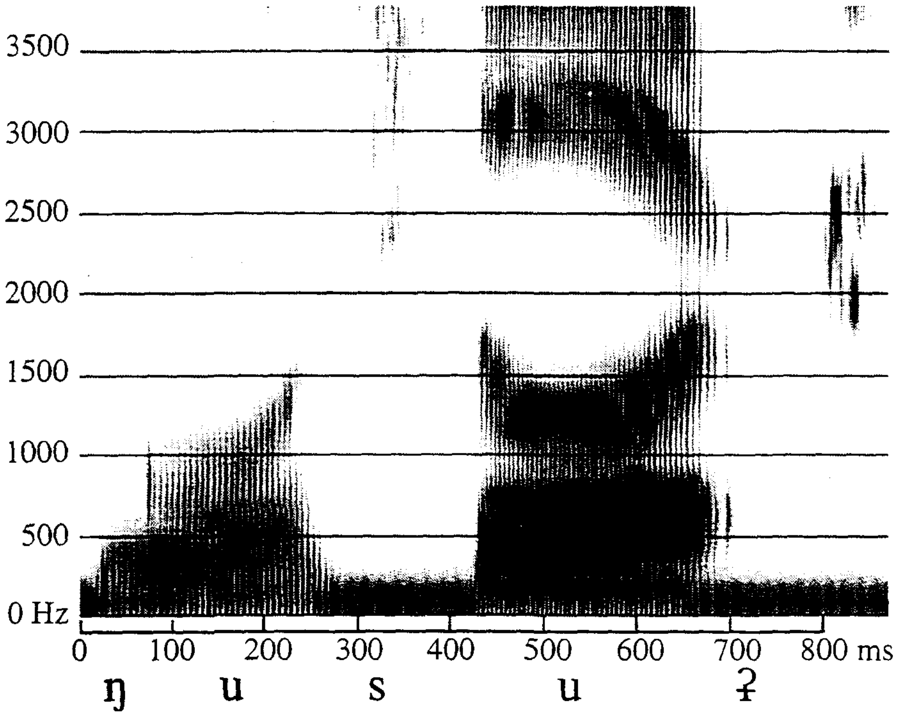


Figure 11. Spectrogram of the word /ŋusuʡ/ "nose" as spoken by Speaker S2.

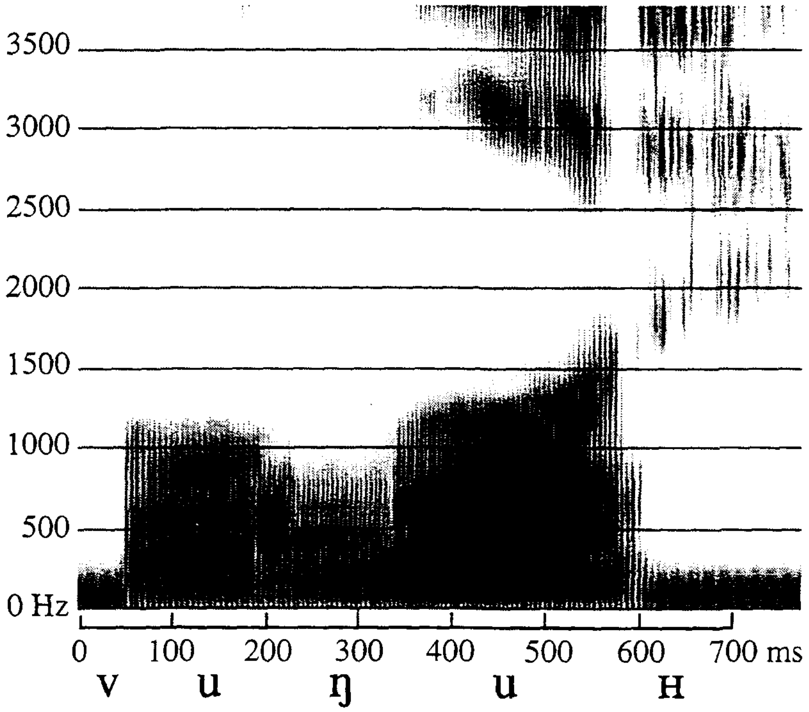


Figure 12. Spectrogram of the word /vuŋuʜ/ "head" as spoken by Speaker S2.

The mean formant values of /u/ in the window placed over the transition from the vowel to these consonants in final position in a set of words are given in Table 1. These measurements — which because they are averaged over a window are not the most extreme values — show markedly higher F1 and F2 before the epiglottals than before the glottal, and markedly lower F3 before the two epiglottals than before the glottal. In each comparison, the epiglottal/glottal difference is significant to at least the .0005 level (Fischer's PLSD). The closer approach of F2 and F3 to each other before /ʡ/ than before /ʜ/ is also apparent in these measures.

Table 1. Mean formants in analysis window for final VC transition between /u/ and /ʔ/, /ʡ/ and /ʜ/.

Final C	*n*	F1	F2	F3
ʔ	16	522	946	3020
ʡ	24	700	1563	2621
ʜ	36	676	1390	2740

4. Vowels

Amis has only three full vowels, /i, u, a/, whose characteristics will be discussed more fully below. There may be a fourth phonologically distinct vowel, a short schwa-like central vowel written in most sources with the letter 'e'. In many words where this letter has been written it is being used as a way of signifying that the first consonant in a CC cluster is released, as noted earlier. Fey (1986) argues against this usage, but it is defended by Duris (1968, 1969). Both authors fail to see that what they treat as a special variation between initial 'e' and glottal stop is part of this same phenomenon, involving variation in the perceptibility of the release of /ʡ/ when it is C1 in a CC cluster.

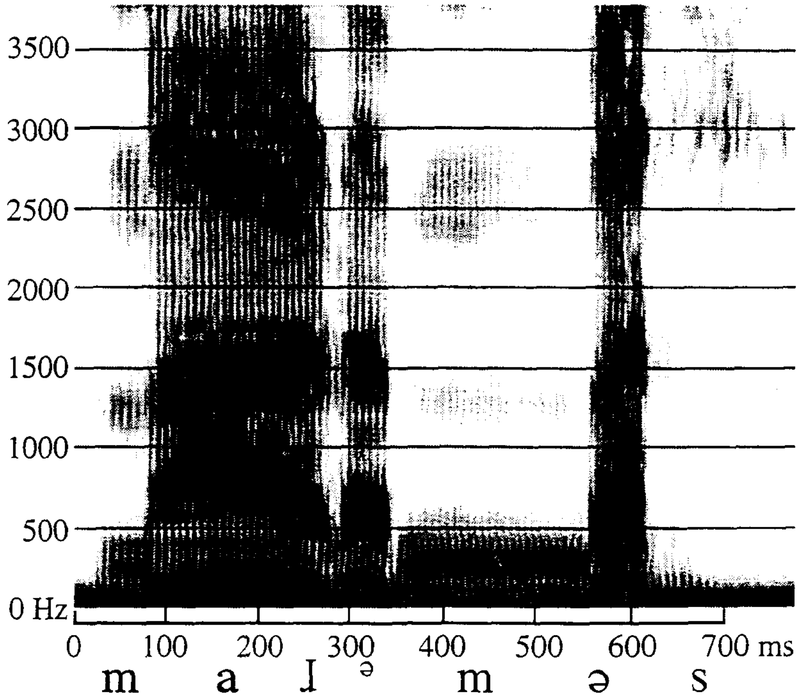


Figure 13. Spectrogram illustrating occurrence of /ə/ in a stressed syllable in the word /maɺməs/ "sad" as spoken by speaker S1 (male).

There are however, some cases where a short central vowel seems to fall in the stressed syllable of a word. An example of this is illustrated in the spectrogram of the word /maɺməs/ "sad" in Figure 13. The inherent shortness of this vowel is well-demonstrated in this spectrogram: it is of about the same length as the transitional vowel which occurs as part of the production of the flap in the cluster /ɺm/. Part of the realization of the stress in the final syllable of this word would seem to involve the marked lengthening of the initial nasal. A possible phonological interpretation would be that the vowel /ə/ here is also a transitional feature, separating the two consonants /ms/ that would otherwise adjoin. This interpretation would entail recognizing words that contained no vowel, such as /ʡnm/ 'six', pronounced [ʡnəm] or [ʡᵊnəm]. Examples of the 'e' vowel are notably rarer in Fey's dictionary than the other three vowels, and we have chosen to omit it from our primary analyses.

The three full vowels are a high front unrounded vowel, a back rounded vowel with a somewhat lower position than its front counterpart, resulting in vacillation between 'u' and 'o' in Amis orthographic texts, and a low central unrounded vowel. The distribution of their variants in the plane defined by the first two formants is shown in Figure 14. This plots the value of each individual vowel measured in both V1 and V2 position for the three subjects S1-S3. V1 is the first of the two phonemically identical vowels across the medial consonant, V2 the second. Recall that stress regularly falls on the ultimate, hence the stress relations between the two vowels when the sequence being analyzed constitutes the first two syllables of a trisyllable differs from when the vowel pair is in a disyllable. However, in most cases V1 is an unstressed vowel and V2 is a stressed one. Although S1 is a male and the other two speakers are females, the range of formant values is quite similar for these three speakers and their data can be reasonably plotted together. Speaker S4, the other male, has formants that are overall lower, as may be seen in Figure 15, and hence do not fit well on the same scaled plot. In the figure, vowels in the V1 position are plotted with open symbols, those in V2 position with filled symbols. Not all points are visible due to overlap of values. There are 96 tokens of /i/, 102 tokens of /a/ and 123 tokens of /u/ in the analysis. The list of words used as the basis of these formant measurements is given in the appendix.

Several observations can be made from this figure. Most striking is the relatively compact distribution of values for the vowel /i/ despite the contextual and speaker differences included in the data. There is very little lowering or centralization of this vowel away from the high front area. The other two vowels show more variation. It is apparent that the rounded back vowel /u/ shows a tendency to be more centralized in V2 position than in V1 position, a pattern that will be discussed further in connection with Figure 15. The low vowel /a/ varies in an area toward the back of the vowel space, such that the distributions of /u/ and /a/ F1/F2 values overlap a little at the margins.

The perceived height difference in the high vowels is reflected in the measured values of F1. For /i/ the mean F1 of /i/ of speakers S1-3 is 348 in V1 position, and 398 in V2 position. For /u/ it is 465 in V1 position and 542 in V2 position. Separate one-way analyses of variance showed that these /i - u/ differences are highly significant. In V1 position the mean F1 of /u/ is 117 Hz higher than the mean F1 of /i/, $F(1,217)=161$, $p < .0001$, and in V2 position it is 144 Hz higher, $F(1,217)=193$, $p < .0001$. All other comparisons between the three full vowels with respect to F1, F2 and F3 also show highly significant differences, except for the comparison of F3 of /u/ and /a/, which are essentially identical.

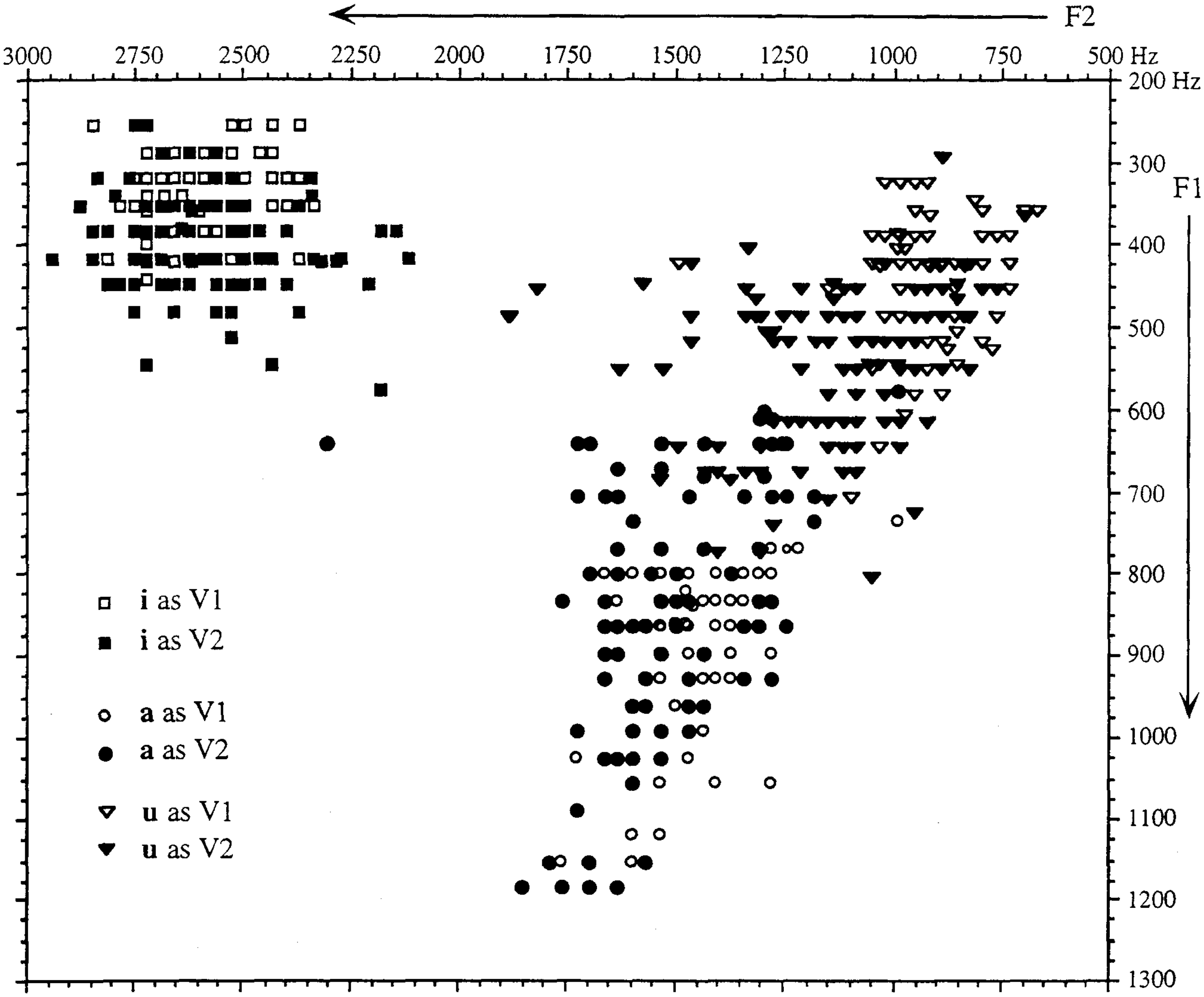


Figure 14. Plot of vowel measurements of single tokens in the plane of the first two formants. Data from subjects S1-S3.

These patterns can also be seen in Figure 15, which shows the mean values of the first two formants of each vowel for each of the four speakers plotted separately . As in Figure 14, vowels in the V1 position are plotted with open symbols, those in V2 position with filled symbols. The more compact formant space defined by the mean values for speaker S4 is obvious. The speaker differences are largest for /a/, which contributes to the greater dispersion of values seen for this vowel in Figure 14. For all four speakers, mean V1 values are invariably a little more extreme than mean V2 values, slightly higher (lower F1) for /i/, slightly higher and backer for /u/, and slightly backer (lower F2) and sometimes lower for /a/. The effect is strongest with /u/, being apparent for this vowel in the raw values plotted in Figure 14. As it is V1 which is always an unstressed vowel, this is not a case of stressed vowels being more peripheral. The position may in fact not be the relevant factor: these differences might as likely be attributable to the fact that the consonant environments preceding and following the measured vowels do not match.

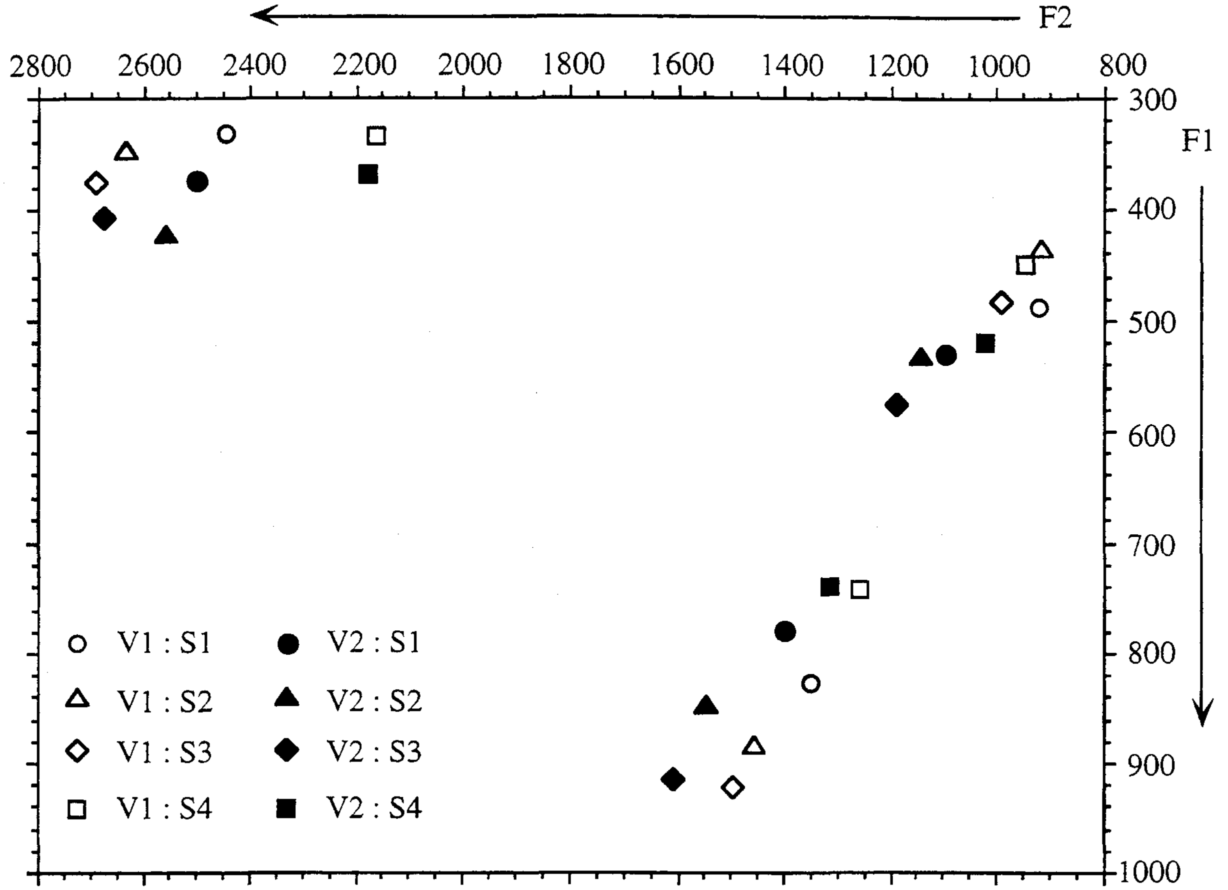


Figure 15. Plot of individual mean vowel measurements in the plane of the first two formants. Data from subjects S1-S4, separating vowels in V1 and V2 positions.

5. Vowel variation

In order to study the variation of the vowels according to the intervocalic consonantal context the formant values were normalized. This was done by subtracting the mean for a given speaker and vowel combination in V1 or V2 position from the values for that speaker's vowel tokens in a given consonant context. For example, speaker S1 has a mean value of 331 Hz for the steady state of /i/ in V1 position. In his two repetitions of /tipiɮ/ 'bowl' the measured F1 of the first vowel was 320 Hz. This speaker thus has a normalized value of -11 Hz for the F1 of V1 in each of these two utterances. These values can then be averaged with other speakers' deviations from their respective means in each consonant context. By removing the major component of the individual speaker differences from the measurements of individual tokens, this procedure enables the data from different speakers to be compared on similar scales. The results are plotted in the next three figures.

Figure 16 shows the mean deviation of F1, F2 and F3 of /i/ in V1 and V2 position in each of the medial consonant contexts averaged across the four speakers. For example, this figure shows that F1 of /i/ is lower than the speaker mean values when it precedes or follows /p/ in the words measured. Similar data for /a/ are shown in Figure 17, and for /u/ in Figure 18. Scales of these three figures are the same. The error bars show one standard deviation.

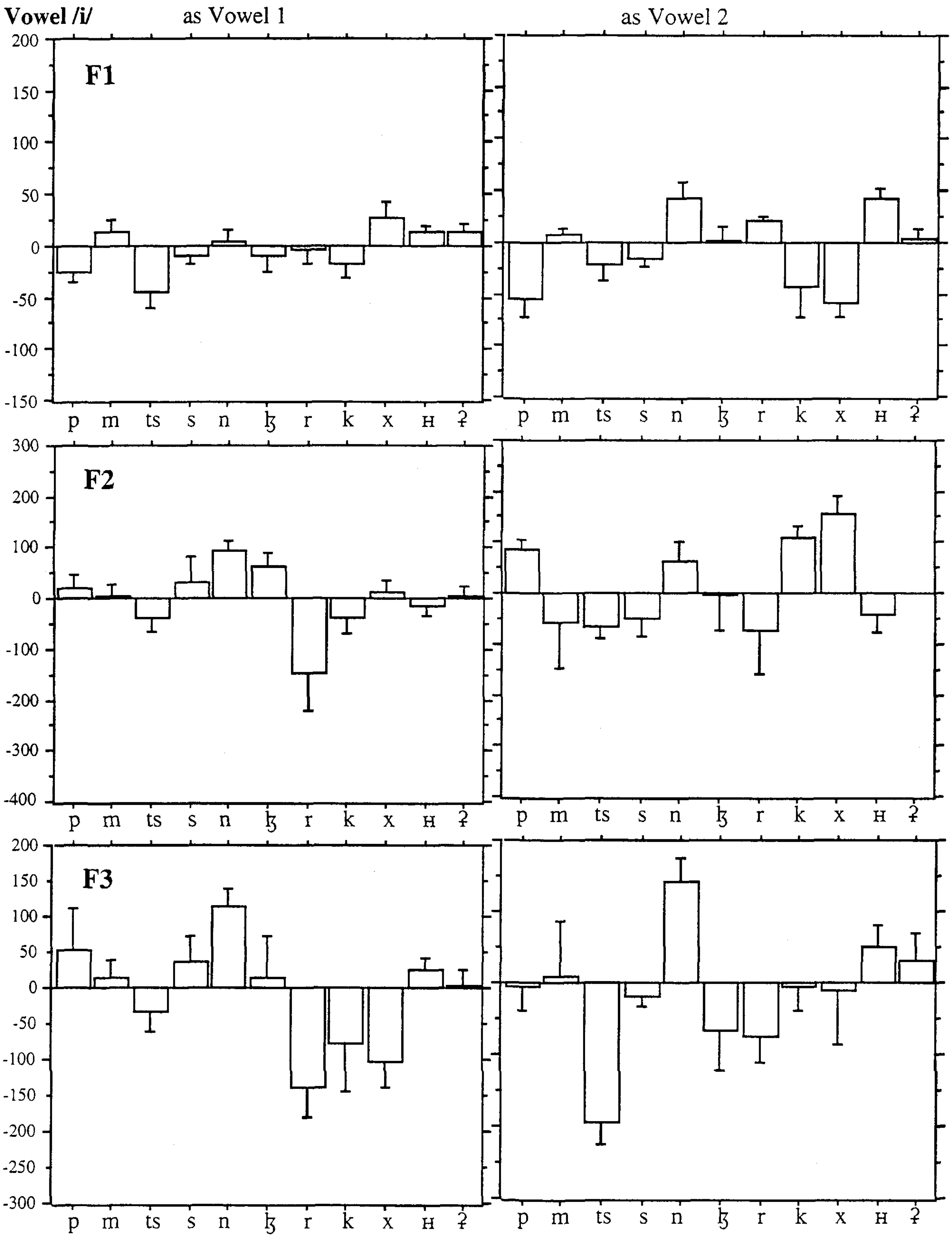


Figure 16. Mean deviations in formants of /i/ in different consonant environments.

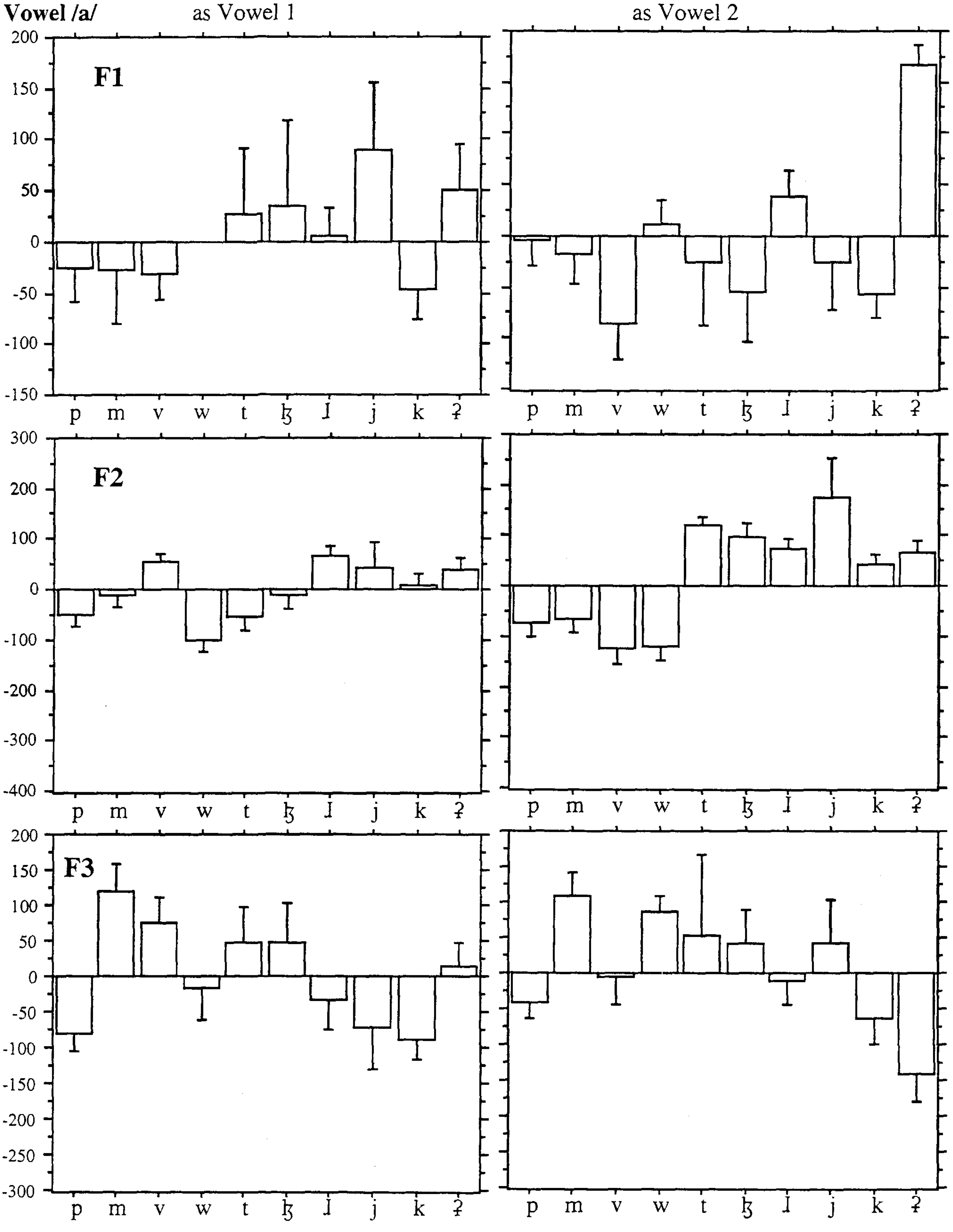


Figure 17. Mean deviations in formants of /a/ in different consonant environments.

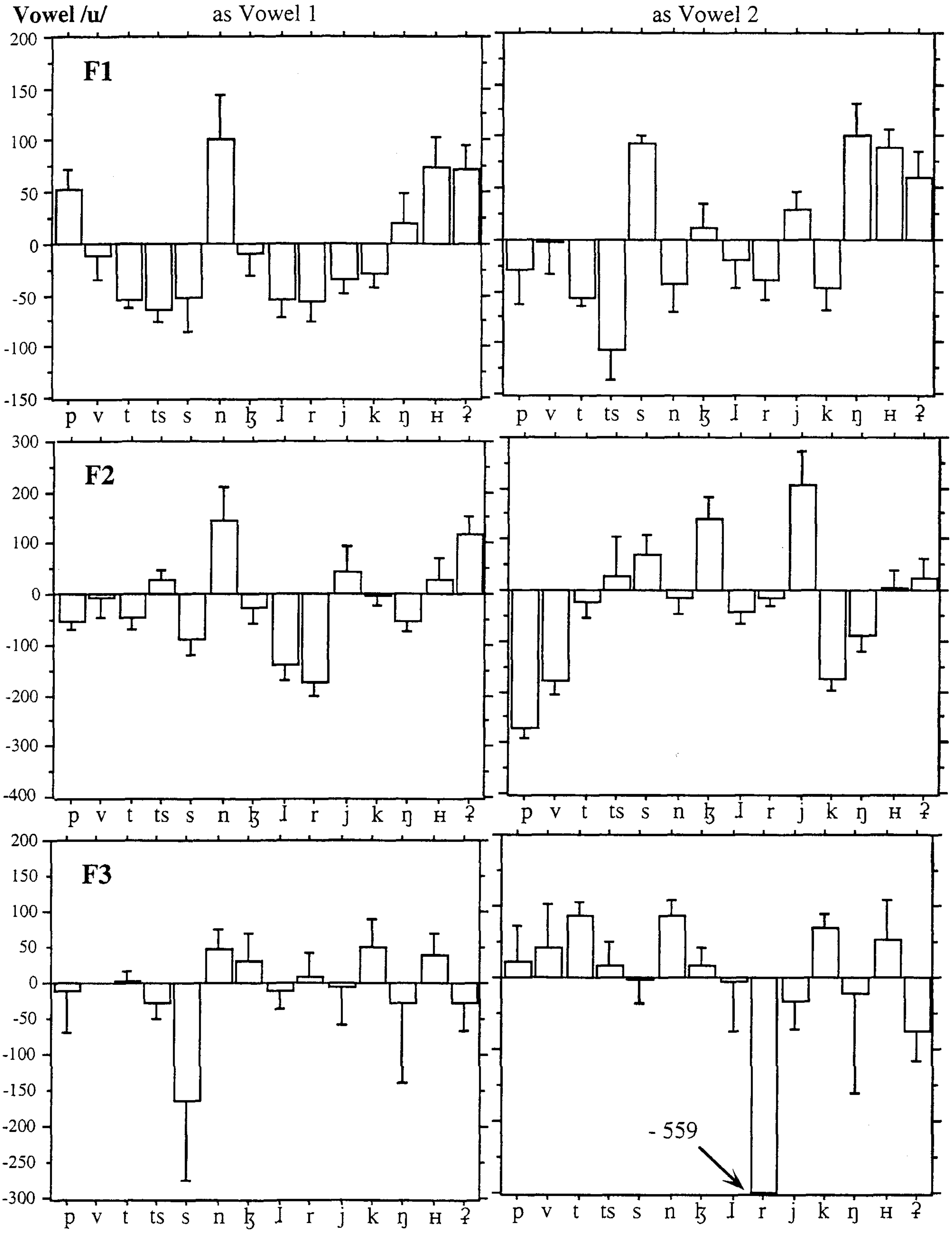


Figure 18. Mean deviations in formants of /u/ in different consonant environments.

Figures 16-18 provide a general impression of the amount of the steady-state variation that is associated with the given consonants, but they need to be interpreted with caution in light of the limitations on the wordlist (see the appendix). As an example we may consider the final panel of Figure 18. In the list the largest number of different consonants appears in the /u__u/ environment and this context provides by far the greatest deviation of the formants, the F3 of /u/ after /r/. The analyzed list includes two words containing /u/ following /r/, /muruʜ/ "bow" and /kurur/ "upper back". In both these words the final consonant is also one that has a transition with a very low F3, so both consonants flanking /u/ in V2 position exert a strong lowering influence on this formant. Note that /u/ preceding /r/ undergoes no comparable lowering of the F3 of its steady state even though the initial consonants /m/ and /k/ would both be expected to have a low F3 transition.

Despite a few extremes, the deviations are in general of relatively modest proportions. To some extent they are in expected directions: The first and second formants are most often somewhat lower in labial environments; coronals have varied effects on the F2 of different vowels; /j/ raises F2 of all vowels; the epiglottals raise the first formant. But a good part of the deviations reflect more random variability. In order to examine to what extent the patterns summarized in Figures 16-18 reflect systematic coarticulation between flanking consonants and the vowels, a series of regression analyses were performed. These compare the magnitude and direction of the transition to or from a flanking consonant with the deviation of the vowel steady state from the speaker*vowel mean. The transition magnitude is the value of a given formant in the window placed over the transition minus the value of the vowel steady state. Thus, for example, one repetition of the word /ŋujus/ 'mouth' by speaker S3 has a steady state value of 928 Hz for F2 of the first vowel. The F2 at the transition to the medial consonant is 1408 Hz. Hence the transition magnitude of this token is +480 Hz. The mean F2 of /u/ as V1 for this speaker is 997, and therefore the deviation of this token from the speaker*vowel mean is -69 Hz. The analysis looks at whether there is covariation between these two types of measures.

The regression analyses shows that the magnitude of a transition predicts very little of the deviation of a given vowel from its mean value. The values of the correlation coefficients of three relationships for each of the first three formants are shown in Table 2. The relationships are those between V1 and the following VC transition, V2 and the preceding CV transition, and V2 and following VC transition. In both of the relationships between a vowel and its following consonant little of the vowel variation is correlated with the consonant transitions. In other words, anticipatory coarticulation between vowel and following consonant is negligible across the data set as a whole. There is some appreciable effect of the preceding consonant on the following vowel with respect to F1 and F3, but F2, which carries the main burden of the distinction between /i/ and /u/, is only very weakly affected.

Table 2. Correlation coefficients (R^2) between vowel steady state deviations and magnitudes of consonant transitions.

	F1	**F2**	**F3**
V1 predicted by VC transition	.044	.01	.157
V2 predicted by CV transition	.315	.167	.348
V2 predicted by final C transition	.008	.059	.144

A visual impression of the strongest of the relationships in Table 2, between the medial consonant and V2 F3, is shown in Figure 19. Note that this result is affected by the special case mentioned above in connection with Figure 18. If the three points clustering in the upper right corner of the graph, which represent three of the tokens of /kurur/ "upper back", are excluded R^2 drops over a hundred points to .243. If all tokens of /kurur/ and /muruʜ/ are excluded, R^2 falls to only .212. We may therefore conclude that, despite possessing a small vowel inventory, Amis characteristically displays rather weak coarticulatory effects of consonants on the steady states of adjoining vowels.

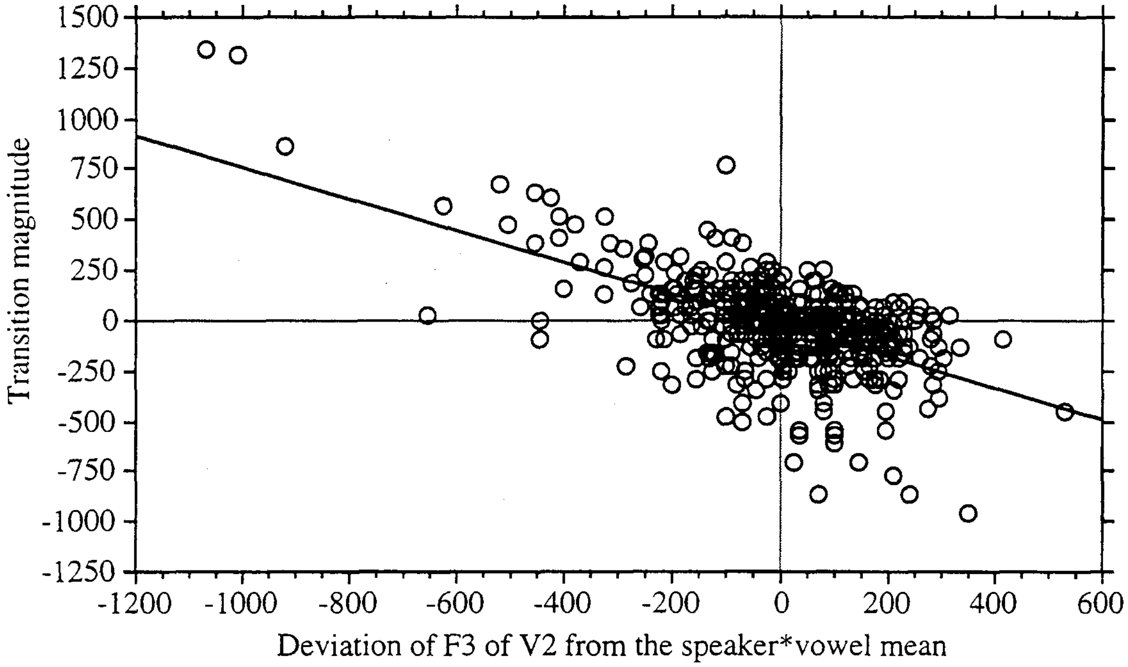


Figure 19. Regression plot showing relationship between CV transition magnitude and V2 F3 deviation. Regression equation is Y = -24.766 - 0.788 * X. n = 425.

6. Final Comment

Amis vowels are relatively stable in their phonetic shape even though few phonologically contrastive categories are involved. This language therefore stands as a counterexample to any general claim that vowels in small vowel systems will necessarily display high variability. This is not to say that there are not some languages with small vowel inventories and great variation. We believe that such vowel variation is likely to be greater in languages in which the vowel contrasts are in the height dimension alone, such as Kabardian and Marshallese, or which have secondary articulation contrasts among their consonants (Choi 1992). Languages like Amis restrain consonantal coarticulation and hence better preserve distinctions in the front-back dimension.

Acknowledgments

We gratefully acknowledge the welcome to Taiwan extended by Professors Chui-yu Tseng and Paul Jen-Kuei Li of Academia Sinica in Taipei and the generous assistance they provided for this project. Father André Bareigts of the Paris Mission was a charming host in Fengpin, and shared his deep knowledge of Amis with us. Monica Wright provided invaluable aid as translator. Funding was provided by a grant from the National Science Foundation for research into the phonetic structures of endangered languages. Despite these debts, our greatest gratitude must go to the speakers of Amis who took time to provide us with the essential material for this research - their voices. We were honored to be awarded Amis names by our friends in Fengpin.

References

Bareigts, André. 1976. *Les Principaux Mythes de l'Ethnie Amis, Volume 1: Traduction et Notes*. Privately published, Hualien and Fengpin.

Blust, Robert. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping. *University of Hawaii Working Papers in Linguistics* 9: 1-15.

Blust, Robert. 1980. Austronesian etymologies. *Oceanic Linguistics* 19:1-181.

Chen, Teresa M. 1987. *Verbal Constructions and Verbal Classification in Nataoran-Amis* (Pacific Linguistics, Series C, No. 85.) Research School of Pacific Studies, Australian National University, Canberra.

Choi, John D. 1992. *Phonetic Underspecification and Target Interpolation: an Acoustic Study of Marshallese Vowel Allophony (UCLA Working Papers in Phonetics 82).* Ph D. dissertation. University of California, Los Angeles.

Dahl, Otto Christian. 1976. *Proto-Austronesian* (Scandinavian Institute of Asian Studies Monograph Series, No. 15). Curzon Press, London.

Duris, Antoine. 1968. *Dictionnaire Français-Amitsu*. Typescript prepared for Kuangshi Press. Reproduced by André Bareigts at Fengpin.

Duris, Antoine. 1969. *Dictionnaire Amitsu-Français*. Typescript. Reproduced by André Bareigts at Fengpin.

Edmondson, Jerrold. 1986. *A Grammar of Hsiukuluan Amis*. Summer Institute of Linguistics, Dallas.

Fey, Virginia. 1986. *Amis Dictionary*. The Bible Society in the Republic of China, Taipei.

Laufer, Asher and Iovanna Condax. 1979. The epiglottis as an articulator. *Journal of the International Phonetic Association* 9: 50-56.

Laufer, Asher and Iovanna Condax. 1981. The function of the epiglottis in speech. *Language and Speech* 24: 39-62.

Li, Paul Jen-Kuei. 1981. Reconstruction of Proto-Atayalic phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* 52: 235-300.

Manuel, Sharon. 1987. *Acoustic and Perceptual Consequences of Vowel-to-Vowel Coarticulation in three Bantu Languages*. Ph. D. dissertation, Yale University, New Haven CT.

Tsuchida, Shigeru. 1982. Subclassification of Ami dialects. Manuscript. Tokyo University.

Appendix

List of words used as the basis of the measurements reported in this paper.

	/i/		/a/		/u/	
p	tipiɮ	bowl	kapaɺ	palm	hupur	silt
			papaʜ	leaf		
m	mimiŋ	little	namaɺ	fire		
			mamaʔ	elder		
v			savaʔ	younger sibling	tuvur	calf (animal)
			ɮavak	morning		
w			hawaʔ	groin		
			wawaʔ	child		
t			mataʔ	eye	tutuj	puppy
					kutuk	louse
					ɺutuk	mountain
					tutuʡ	dipper (from gourd)
n	(i)tini	here			punuʡ	brain
	isiʡ	urine			ŋusuʡ	nose
	pisiŋ	face				
ts	tsitsiw	piglet			tsutsuʔ	breast
ɮ	wiɮiʔ	leeches	aɮaʔ	enemy	tuɮuʜ	burn, roast
					(mi)muɮuʜ	sand, grind
ɺ			kaɺaŋ	crab	kuɺuŋ	cow
	siriʔ	goat			muruʜ	bow (bend from waist)
r					kurur	upper back
j			ʡajam	bird	tsujuʜ	angel
					ŋujus	mouth
k	(ma)pikiŋ	shriveled limb	pakaʜ	buttocks	tukus	peak
			kakaʔ	elder sibling		
					vuŋuʜ	head
ɣ	riɣiʔ	ridge in rice field				
ʡ	piʡiw	cripple	tsaʡaŋ	branch	ʡuʡuɺ	fog
	tsiʡim	plant sucker	waʡaj	leg	tsuʡuŋ	bud
	ɮiʜiv	cave	muɺaʜaɮ	growing		
ʜ	tiʜiʔ	spouse			puʜuŋ	antler
	tatiʜiʔ	next				
ʔ	tatiʔiʜ	bad				

Collapsing vowel harmony and doubly-articulated fricatives: two myths about the phonology of Avatime

Ian Maddieson

1. Introduction

The language generally known as Avatime is spoken by about 18,000 people in the Volta Region of Ghana (Ring 1987, 1995). It is called Si-ya(-se) by its speakers, who refer to themselves by the name /kedone/ (singular masculine form). Although Avatime is an outsiders' name (from Ewe) it is the most widely familiar designation for both the language and the people and it will be used in this paper for the sake of its familiarity. Avatime is primarily spoken in the seven small towns and villages of Vane, Amedzofe, Gbadzeme, Dzokpe, Biakpa, Dzogbefeme, and Fume as well as by members of a small diaspora scattered in various regions of Ghana and other parts of the world.

Avatime has been classed as a member of the Central Togo, Togo Mountain or "Togo Remnant" group of languages (Struck 1912, Heine 1968, Kropp Dakubu 1988). This a socio-cultural and geographic rather than a genetic group. Although all the languages are undoubtedly related and belong in the Niger-Congo family, there are at least two major subgroups represented, and these groups may each be more closely related to different non-Togo Mountain language groups than they are to each other (Stewart 1989).

Avatime was studied by the missionary linguist Emil Funke (1909, 1910) in the early years of the twentieth century following the foundation of a North German Mission station at Amedzofe. It was subsequently largely ignored in the linguistic literature until a flurry of work in the late 1960's and early 1970's, including that of Kropp (1967), Heine (1968) and Ford (1971, 1973). More recently it has been the target of fieldwork by the author of this paper and Russell Schuh (see Schuh 1995a, b).

This new work suggests that it is time to revise the received wisdom on some aspects of Avatime phonology. This paper will address two issues in particular. The first concerns the phonological structure of the vowel system, including the phonetic basis of the vowel harmony system found in the language. The second is the claim that certain Avatime fricatives have double articulations, which, if true, would conflict with the suggestion of Ladefoged and Maddieson (1995) that doubly-articulated fricatives are not used as regular linguistic sounds.

2. Vowels and vowel harmony

Funke had noted nine vowel qualities in Avatime, but chose to write only seven. Kropp (1967), Heine (1968) and Ford (1971, 1973, 1988) all agree that Avatime has a 'triangular' seven-vowel system, which they transcribe with the symbols in (1):

(1) **i** **u**
e **o**
ɛ **ɔ**
a

As in many of the Niger-Congo languages of West Africa, vowel harmony constraints apply to limit the vowel combinations which may occur within stems and to determine the selection of variant forms of many affixes, for example, the class prefixes that attach to nouns and the subject markers preceding verbal elements. These constraints are especially apparent with the mid vowels. Using the symbols in (1), we may illustrate these affixal harmony processes with nouns such as **o-ze** "thief" and **ɔ-dʒɛ** "woman", and in verbal phrases such as **wo to** "you sg. pounded" and **wɔ tɔ** "you sg. cooked". In these cases the alternate forms of the noun class prefix **o-/ɔ-** (here, the human singular class) and the second singular subject (and past tense) marker **wo/wɔ** are determined by whether the stem vowel is drawn from the set {**e,o**} or the set {**ɛ,ɔ**}. For convenience, but without prejudging the phonetic facts, the vowel set including {**e,o**} will be referred to as Advanced Tongue Root or [+ATR] vowels, and the set including {**ɛ,ɔ**} will be referred to as Retracted Tongue Root or [-ATR] vowels. The low vowel /**a**/ belongs to the [-ATR] set. A stem with /**a**/ takes affixes from the [-ATR] set, and an affix whose [-ATR] variant has the vowel /**a**/ has /**e**/ in its [+ATR] variant. Thus, the human plural noun class prefix has the variants **be-/ba-**, as we see in **o-ze/be-ze** "thief/thieves", but **ɔ-ka/ba-ka** "father/fathers". There are a variety of complexities introduced by factors such as processes of vowel coalescence and the occurrence of affixes with invariant forms but these will not be discussed here. There are also a few words with nasalized vowels. More information is given in Schuh (1995b).

At issue here is what role high vowels play in Avatime vowel harmony. In the analysis proposed by Ford, high vowel stems occur with both harmonic variants in the affixes and clitics attached to them, but there is no +/- ATR distinction among high vowels. He takes the apparent harmony among prefixes in words such as those in (2), as he would write them, as evidence of a former distinction which is now non-phonological.

(2)	**o-bu**	"bee"
	ɔ-bu	"god"
	o-kusi	"chief"
	ɔ-si	"stirring stick"

Whereas Ford (1973) would reconstruct an earlier opposition between [+ATR] and [-ATR] high vowels, Heine (1968) does not even consider that the parent language had such a contrast. However, Avatime actually maintains this distinction in its present-day phonology, and has a nine-vowel system. +/- ATR high vowels have not fallen together. This might have been suspected from the fact that Kropp (1967) wrote the vowels in [-ATR] stems and affixes with high vowels as 'e' and 'o', rather than as 'i' and 'u' as Ford does. (This results in the need for her to use diacritic symbols to show which of the verb stems written with 'e' and 'o' will take [+ATR] and which will take [-ATR] affixes.) It is the fact that particular words have been heard by different listeners as having different vowel qualities whereas other words are invariably heard with agreed-on high vowels or mid vowels that suggests that the apparently variable vowels might actually be distinct from both the sets with which they have been conflated.

2.1 Data

The high vowel distinction will be demonstrated by a phonetic analysis of two data sets. One is drawn from an extensive word list recorded in Amedzofe by 12 speakers (8 male, 4 female). This recording consists of single repetitions by the 12 speakers of items designed to illustrate all the consonants and vowels of the language. Nouns were usually elicited in both singular and plural forms, which typically involves an alternation of prefixes, and sometimes were also elicited with the numeral 'one' following. As a result there may be more than one token of a given word or its stem syllable from each speaker. The second data set consists of verbal paradigms from a single male speaker giving past, future and continuous forms for six person-number categories. This speaker is not a member of the group of 12. Two pairs of minimally contrastive verb stems with high vowel were used, giving 18 repetitions of each of the four stems.

For both these data sets, the first three formants of the vowels were measured on combined displays of wide-band spectrograms and LPC 'formant histories' produced by the Kay Elemetrics CSL system. The analog recordings were sampled at 20 kHz. A 16 order LPC was used for most tokens, but adjustments were made where this resulted in a poor match between formants chosen by eye in the spectrogram and the LPC-derived values. Each measured value was obtained by averaging the value of each formant over a steady state portion in the center of the vowel. The third formant was often difficult to determine, especially for back vowels, and few measurements of F3 can be reported with confidence. The results will be discussed as 'Experiment 1' and 'Experiment 2'.

2.2 Experiment 1

The selected subset of words from the 12-speaker data set used in this analysis is given in Table 1. In this table and elsewhere in the paper where it is important to place emphasis on the phonological facts of vowel harmony, high and mid vowels in the two harmony sets are transcribed with the IPA diacritics for the ATR distinction. Vowels in the [+ATR] set are noted as [i̘, e̘, o̘, u̘] and those in the [-ATR] set are noted as [i̙, e̙, o̙, u̙]. Consonant environments are matched as well as the structure of the wordlist permits. Between two and four words are used to exemplify each of these eight vowels. /a/ is omitted from consideration as it is so sharply distinct from all non-low vowels.

Table 1. Words used for formant measurements in the 12-speaker data set.

Vowel(s)	Word	Gloss
i̘	o̘ku̘si̘	chief
i̘, u̘	ki̘bu̘	honey
i̙, u̙	ki̙fu̙	fire
i̙, u̙	ki̙gu̙	war
e̘, u̘	be̘bu̘	bees
e̙	bi̙de̙	mortars
e̙	ki̙de̙	mortar
o̘, e̘	o̘se̘	tree
o̘, e̘	o̘ze̘	thief
o̘, u̘	o̘bu̘	bee (2 repetitions)
o̙, i̙	o̙si̙	stirring stick
o̙, u̙	o̙bu̙	god

The measurements show clear evidence of an acoustic distinction between [+ATR] and [-ATR] high vowels. The mean values of the first formant for the eight vowels are shown in Figure 1. Since equal numbers of tokens of any given vowel are provided by each speaker, any speaker-dependent variation in these values is controlled. Figure 1 shows that, for each +/-ATR pair the [-ATR] vowel has a higher F1 than its [+ATR] counterpart. Importantly, the F1 of [i̙] is higher than that of [i̘] and the F1 of [u̙] is higher than that of [u̘], but in neither case does F1 of the [-ATR] member of the pair reach the same level as that of the nearest mid vowel, [e̘] or [o̘]. The +/-ATR high vowels are distinct from each other and, moreover, have not merged with mid vowels. The significance levels of comparisons between adjacent vowel pairs are shown in Table 2. All the comparisons of interest are significantly different at least the .0005 level, as indicated by a post-hoc comparison of means using Fisher's PLSD (adjusted for unequal cell sizes) following a one-way analysis of variance which had showed that there was a highly significant main effect of the vowel category.

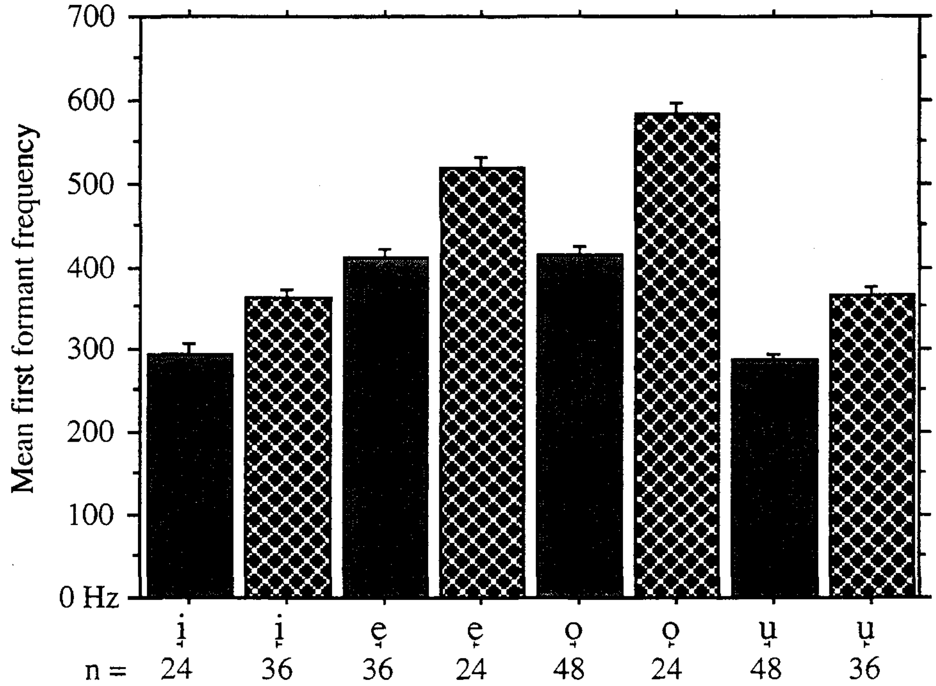


Figure 1. Mean first formant in words in Table 1 spoken by 12 speakers. Cross-hatched columns are [-ATR] vowels. Error bars show one standard error.

Table 2. Significance levels of F1 difference between vowel pairs.

Vowel pair	P-Value
i̘, i̙	<.0001
i̙, e̘	=.0005
e̘, e̙	<.0001
o̘, o̙	<.0001
o̙, u̘	=.0001
u̘, u̙	<.0001

The mean second formant values from the twelve-speaker group are shown in Figure 2. Here the pattern is not consistent across +/-ATR pairs. For three of them the F2 of the -ATR vowel is slightly higher than that of its +ATR counterpart. But for the high front pair i̘ , i̙ it is the +ATR member of the pair which has the higher F2. As is shown in Table 3, this pairwise distinction is significant, but all of the other comparisons between vowels of the same height category and degree of backness do not reach significance. However, the F2 of both the -ATR high vowels is significantly distinct from the F2 of the nearest +ATR mid vowels, providing further confirmation that these vowels have not merged with vowels in the mid category. The third formant is also significantly different between /i̘/ and /i̙/ (mean F3 for /i̘ / is 2881 vs 1679 for /i̙/, n = 23 vs 36; p = .002). Note that for one token of /i̘/ F2 and F3 could not be discerned, so there are only 23 values of these formants for this vowel.

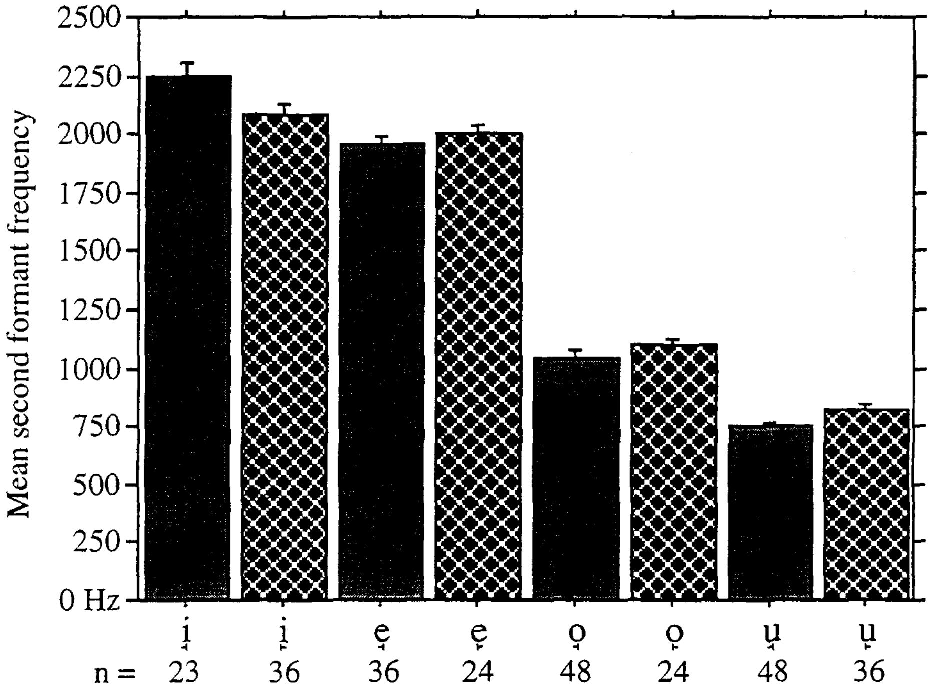


Figure 2. Mean second formant in words in Table 1 spoken by 12 speakers. Error bars show one standard error.

Table 3. Significance levels of F2 difference between vowel pairs.

Vowel pair	P-Value
i̘, i̙	=.0017
i̙, e̘	=.0043
e̘, e̙	n.s.
o̘, o̙	n.s.
o̘, u̙	<.0001
u̘, u̙	n.s.

2.3 *Experiment 2*

The second data set analyzed focused solely on the high vowels. Multiple repetitions of the minimal verb pairs cited in (4) were obtained from a single speaker using a paradigm of six person-number forms in each of three tense-aspect forms. The elements that differentiate the terms in the paradigm are listed in Table 4. Because of the critical role of tone in these contrasts, high tones are marked (although up to this point tonal distinctions have been ignored as not relevant for this study). Note the invariant form of the future marker /tá/, which consequently requires -ATR vowels in the preceding person-number markers. The other two tense-aspect forms are marked by variations in the shape of the basic person-number markers. In each case these markers must also harmonize with the ATR category of the verb stem they precede. The verbs in (4) also require this harmonization of prefixes.

(4)

gu̘	"pluck"
gu̙	"sniff"
tʃi̘	"grow old"
tʃi̙	"tear"

Table 4. Person-number and tense-aspect markers used in Experiment 2.

	Past		Future		Continuous	
	+ATR stem	- ATR stem	+ATR stem	- ATR stem	+ATR stem	- ATR stem
1 sg	me̘	**ma**	ma tá	ma tá	me̘é̘	me̙é̙
2 sg	wo̘	**wo̙**	wo̙ tá	wo̙ tá	we̘é̘	we̙é̙
3 sg (class 1)	e̘	**a**	a tá	a tá	e̘é̘	e̙é̙
1 pl	kwi̘	**kwi̙**	kwi tá̙	kwi tá̙	kwí̘	kwí̙
2 pl	mle̘	**mle̙**	mle̙ tá	mle̙ tá	mlé̘	mlé̙
3 pl (class 2)	be̘	be̙	be̙ tá	be̙ tá	be̘é̘	be̙é̙

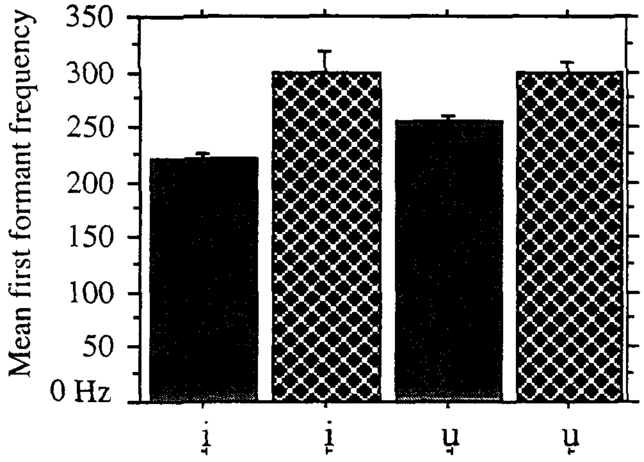


Figure 3. Mean first formant of the high vowels in the verb stems in (4).

The mean first formant frequencies of the high vowels in the verb stems in this data set are shown Figure 3. As in the group data illustrated in Figure 1, the first formant of the -ATR member of the high vowel pairs has a higher F1 than the corresponding +ATR member. Both pairwise distinctions are highly significant, as is shown by the numbers in Table 5. The comparison of the second formant in this data set is also consistent with that seen in the group data in Figure 2. F2 in i̘ is higher than in i̙ but F2 in u̘ is lower than in u̙. However, in neither of these instances is the difference very substantial. The third formant means for the high front pair in this data set are essentially equal, and the difference is self-evidently non-significant.

Table 5. Mean F1 in the verbs in (4).

Vowel	mean F1	P-Value
i̘	222	<.0001
i̙	299	
u̘	254	<.0001
u̙	299	

Table 6. Mean F2 in the verbs in (4).

Vowel	mean F2	P-Value
i̘	2148	<.0184
i̙	2091	
u̘	611	<.0748
u̙	654	

2.4 *The phonetic basis of vowel harmony*

The results of Experiments 1 and 2 show clearly that the high vowels are just as involved in whatever phonetic parameters underlie the vowel harmony system in Avatime as the mid vowels, and the distinction between what we have transcribed as i̘ and i̙ and u̘ and u̙ is a real and acoustically demonstrable one. It is now time to ask if the acoustic measurements indicate what the basis of the vowel harmony system is in Avatime. A two-dimensional view of the vowels may help to address this issue.

The distribution in the plane of the first two formants of the vowels measured for Experiment 1 is shown in Figures 4 and 5. Figure 4 plots the values for the 8 males in the group and Figure 5 those for the 4 females. Ellipses with a radius of one standard deviation along each of the first two principal components of the distribution are plotted for each vowel. The symbols used are the traditional height-related ones, so that the vowels of the [+ATR] harmony set are plotted with [**i, e, o, u**], and those of the [-ATR] harmony set with [**ɪ, ɛ, ɔ, ʊ**]. (The substantial overlap in the distribution of the high vowels and the [+ATR] mid vowels in these figures illustrates one source of the difficulty that non-native speakers have in differentiating them.)

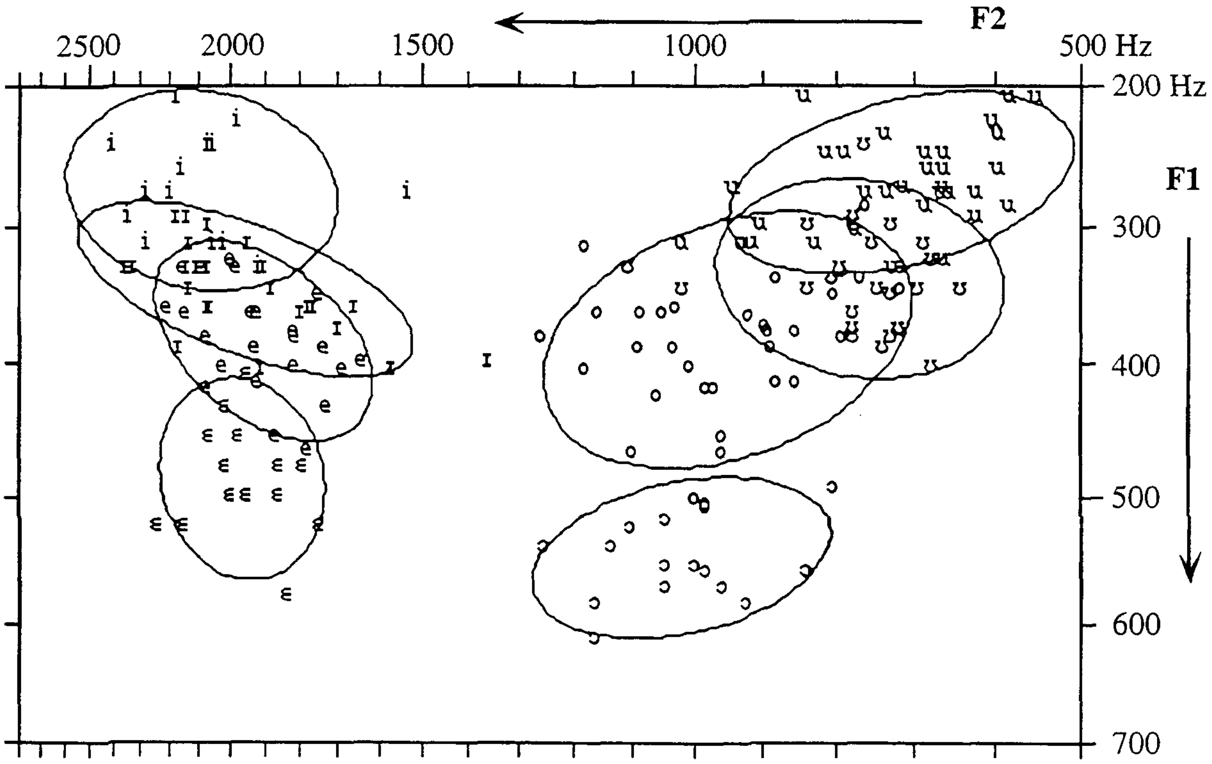


Figure 4. F1 and F2 of the 8 male subjects in Experiment 1.

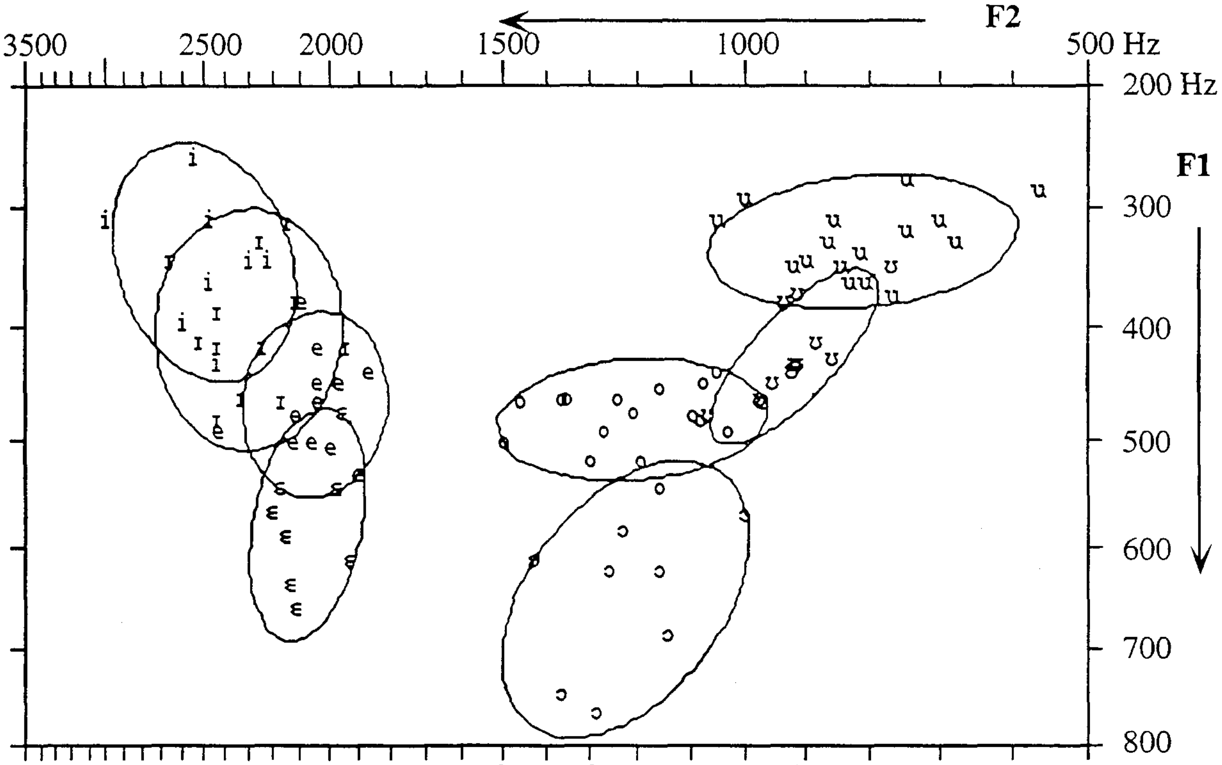


Figure 5. F1 and F2 of the 4 female subjects in Experiment 1.

In their discussion of the acoustic characteristics of the articulatory gesture of tongue root advancement, and its usual concomitant of larynx lowering, Ladefoged and Maddieson (1990, 1995) suggest that a useful diagnostic for its presence can be found in the relationships among back vowels. Discussing data from six languages from West and East Africa, most of which are definitively known from x-ray studies to have tongue root advancement as the basis of their vowel harmony, they note that in addition to an F1 difference "the high back retracted tongue root vowel is always further back" — that is, it has a lower F2 than its advanced counterpart. A similar relationship typically obtains between the back mid vowels in this data, and this parallels the relationships seen between the vowels in the front high and front mid pairs. A prototypical vowel system using tongue root advancement to distinguish four vowel pairs might therefore be expected to have an acoustic vowel space similar to that shown schematically in Figure 6.

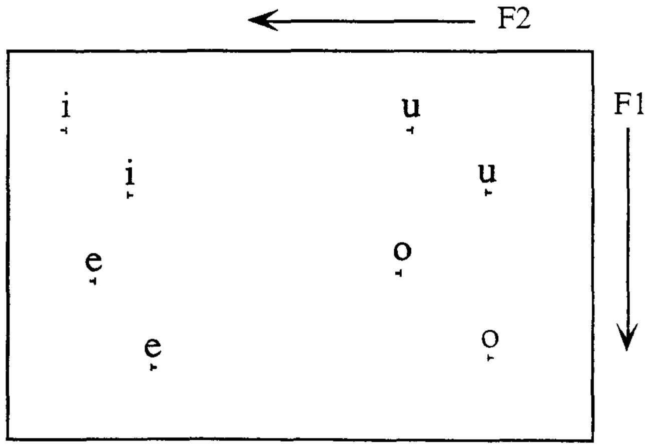


Figure 6. Idealized scheme of +/- ATR acoustic pattern.

Avatime obviously does not fit this idealization. Only the high front pair shows a lower F2 for the [-ATR] member. For the other pairs, the F2 values are quite similar, and even show means that are slightly higher for the [-ATR] member.

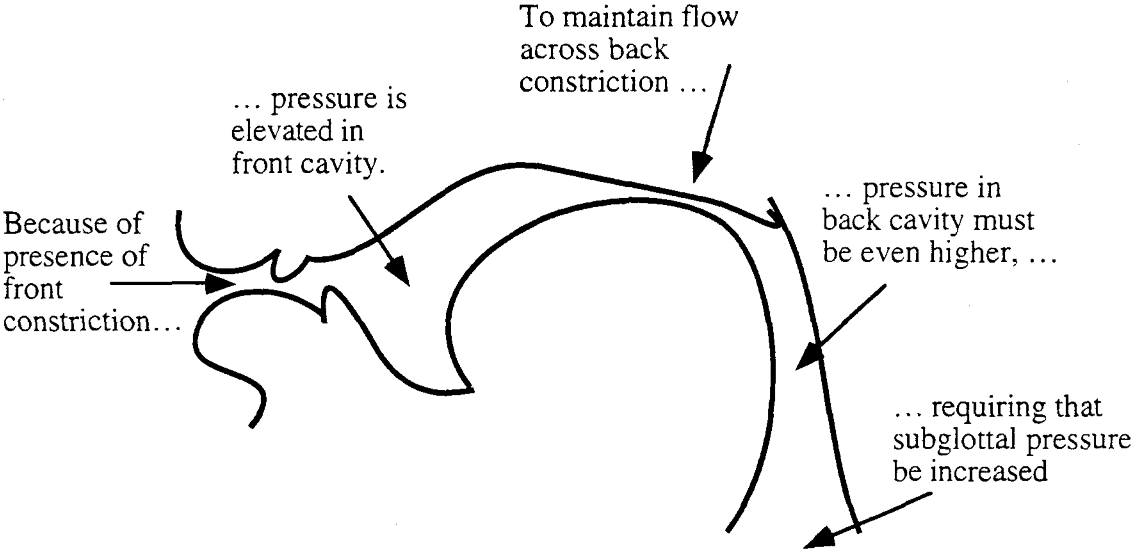
This might call into question whether the Avatime phonological contrast we have been referring to as [+/-ATR] truly involves the tongue root position. Given the strong and consistent role that F1 plays in distinguishing the members of +/-ATR pairs, it might be suspected that the distinction is only one of height, and there actually is — to revive an old term — "cross-height vowel harmony" in Avatime. Unfortunately, the situation is not so simple. Data from Akan, the prototype for tongue-root based vowel harmony, show that contrasting tongue root positions correlate less uniformly with values of F2 than Ladefoged and Maddieson may have implied. The speaker studied by Hess (1992) showed a considerably higher F2 for [u̘] than for [u̙], and the subject studied by Tiede (1993) showed no effect of the [+/-ATR] feature on F2, despite substantial differences in pharyngeal volume.

Tiede's articulatory data, obtained using magnetic resonance imaging, also replicate in a richer dimensionality results obtained by Lindau (1979), Jackson (1988) and Hess (forthcoming) from sagittal x-ray studies. In Akan, unlike in English, tongue dorsum height does not co-vary with tongue root advancement or pharyngeal volume. The F1 distinctions that consistently distinguish between +/-ATR pairs in Akan are not simply the result of height variations. They are more likely due in large part to variation in the height of the larynx, and hence of the effective vocal tract length, which does co-occur with tongue root variation. A lower larynx position for [+ATR] vowels generates a lower F1. Hess (1992) also shows that [+ATR] vowels have a narrower bandwidth of the first formant, attributed to greater tension of the cavity walls associated with the tongue fronting gesture. A study of formant bandwidths in Avatime might help to clarify whether the vowel harmony in this language has tongue root position as its phonetic basis.

3. Doubly-articulated fricatives?

The other principal issue of phonetic interest in Avatime are the purported labial-velar fricatives, transcribed by Ford (1988) as /x͡ɸ/, /ɣ͡β/. Ladefoged and Maddieson (1995) propose that doubly-articulated fricatives should not be expected to occur as regular speech sounds in the world's languages. If labial-velar fricatives occur in Avatime, this would be significant evidence that this claim is false.

The reasoning behind Ladefoged and Maddieson's claim is as follows. A supraglottal constriction narrow enough to produce friction on typical pulmonic air-flow creates elevated pressure behind it. To create sufficient velocity airflow to generate frication through a second constriction further back would require even higher pressure behind that constriction, so that a sufficient pressure differential could be sustained. To generate this pressure differential, elevated subglottal pressure would be required. Elevating subglottal pressure is extra work, and hence is disfavored on grounds of articulatory effort. This articulatory pattern is shown by the diagram in Figure 7. At the same time doubly-articulated fricatives also suffer an acoustic disadvantage. The radiated sound from the back constriction will be substantially attenuated by the effects of the one in the front, resulting in poor ability to recover the presence of the back constriction from the frication noise. So although it is humanly possible to produce doubly-articulated fricatives, they are linguistically undesirable for both productive and perceptual reasons. Note that there are no similar problems for fricatives with a secondary constriction of a lesser degree, such as /x^w/.



To maintain flow across back constriction ...

... pressure is elevated in front cavity.

Because of presence of front constriction...

... pressure in back cavity must be even higher, ...

... requiring that subglottal pressure be increased

Figure 7. Articulatory disadvantage of doubly-articulated fricatives (Drawing based on a figure in Lindblad (1980)).

Before discussing the Avatime facts, the production of the Swedish fricative transcribed "[ɧ]" will be briefly reviewed. This has sometimes been described as having a labio-dental + velar double articulation, which would be a counterexample to Ladefoged and Maddieson. Lindblad (1980) describes several variants of the phoneme of which this is one realization, but the sagittal tracing that he shows of the particular variant that has been described as doubly-articulated shows only one constriction narrow enough to cause friction. There is some narrowing in the velar region, as illustrated in Figure 8, but it is not comparable to a velar fricative articulation.

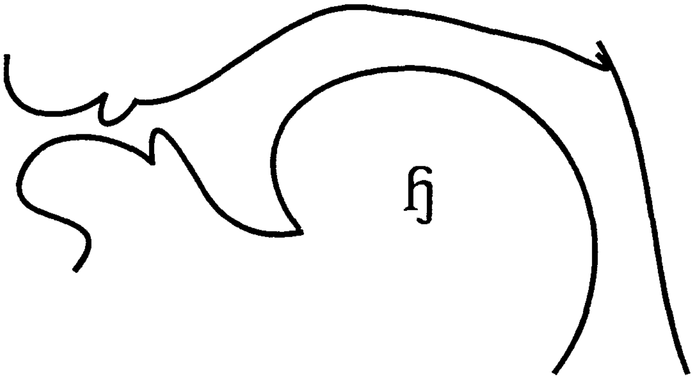


Figure 8. Sagittal section of the articulation of [ɧ], based on an x-ray tracing published by Lindblad (1980).

As noted above, Avatime has been claimed to have bilabial + velar fricatives. This might be considered plausible on grounds of symmetry. Avatime has bilabial and velar stops and doubly-articulated labial-velar stops. Since it also has both bilabial and velar fricatives (although the voiceless bilabial only occurs in loanwords, mostly from Ewe), it might be considered plausible that the combination of these articulations should also occur. Ford, who has worked extensively and first-hand with Avatime speakers, explicitly claims that this is so, and provides a consonant chart that includes the segments shown in Table 7 (Ford 1988).

Table 7. Partial consonant chart of Avatime, after Ford (1988).

	Bilabial		Velar		Labial-Velar	
Plosives	p	b	k	g	k͡p	g͡b
Fricatives	(ϕ)	β	x	ɣ	x͡ϕ	ɣ͡β

Since more information-rich techniques, such as x-ray microbeam or magnetometer systems, are at present infeasible in a fieldwork setting in Ghana, the following research strategy was adopted. The labial aperture in the native bilabial fricative [β] and the loan segment [ϕ], and the labial aperture in the alleged doubly-articulated fricatives was documented using videotape. The degree of labial constriction in these segments can then be compared. If the lips are clearly more open in the claimed double-articulations [x͡ϕ, ɣ͡β] than in [ϕ, β], it demonstrates that the front closure in the former is not narrow enough to be a fricative constriction. This technique will provide no information on the degree of the velar constriction, although acoustic and auditory evidence can be considered in conjunction with the articulatory data.

Five speakers were recorded, all male, four in Amedzofe and a fifth in Accra. Adhesive paper dots were placed on the lips to provide clear measurement points. To provide a scale for measurements, a ruler was placed in the plane of the lips and filmed at the outset of filming each subject. Detailed data from only one subject is reported here, but careful review of the tapes indicates that the results are similar for all five subjects. The analysis proceeded as follows. The videotapes were viewed frame-by-frame using a time-code superimposed on the tape to identify frames. The frame containing the maximal constriction of the lips captured during the consonants of interest was identified. Video frames of interest were then digitized and measurements made using the Image program from Warren Rasband of NIH. Since the video format used has a frame rate very close to 30 frames a second, successive frames are about 33 ms apart. A frame containing the ruler was digitized to provide the scale.

In the following discussion, fairly extensive exemplification of the results will be provided. Such careful documentation is crucial, as reports based on observation alone can be challenged. By publishing the labial views, we enable other linguists to form their own conclusions about the production of these sounds.

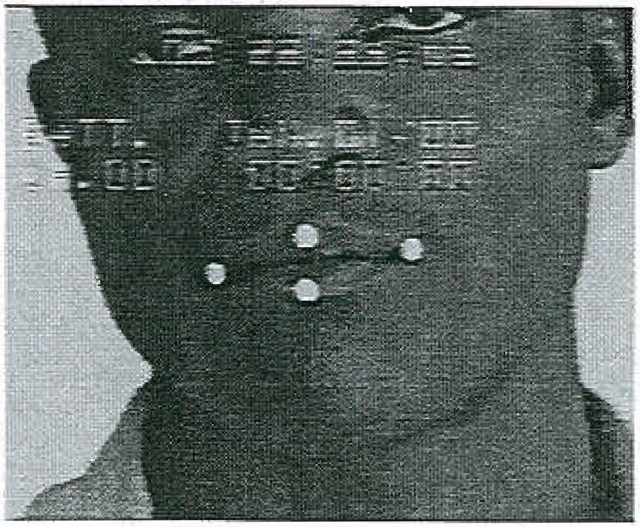


Figure 9. Bilabial closure for [m] in [ma].

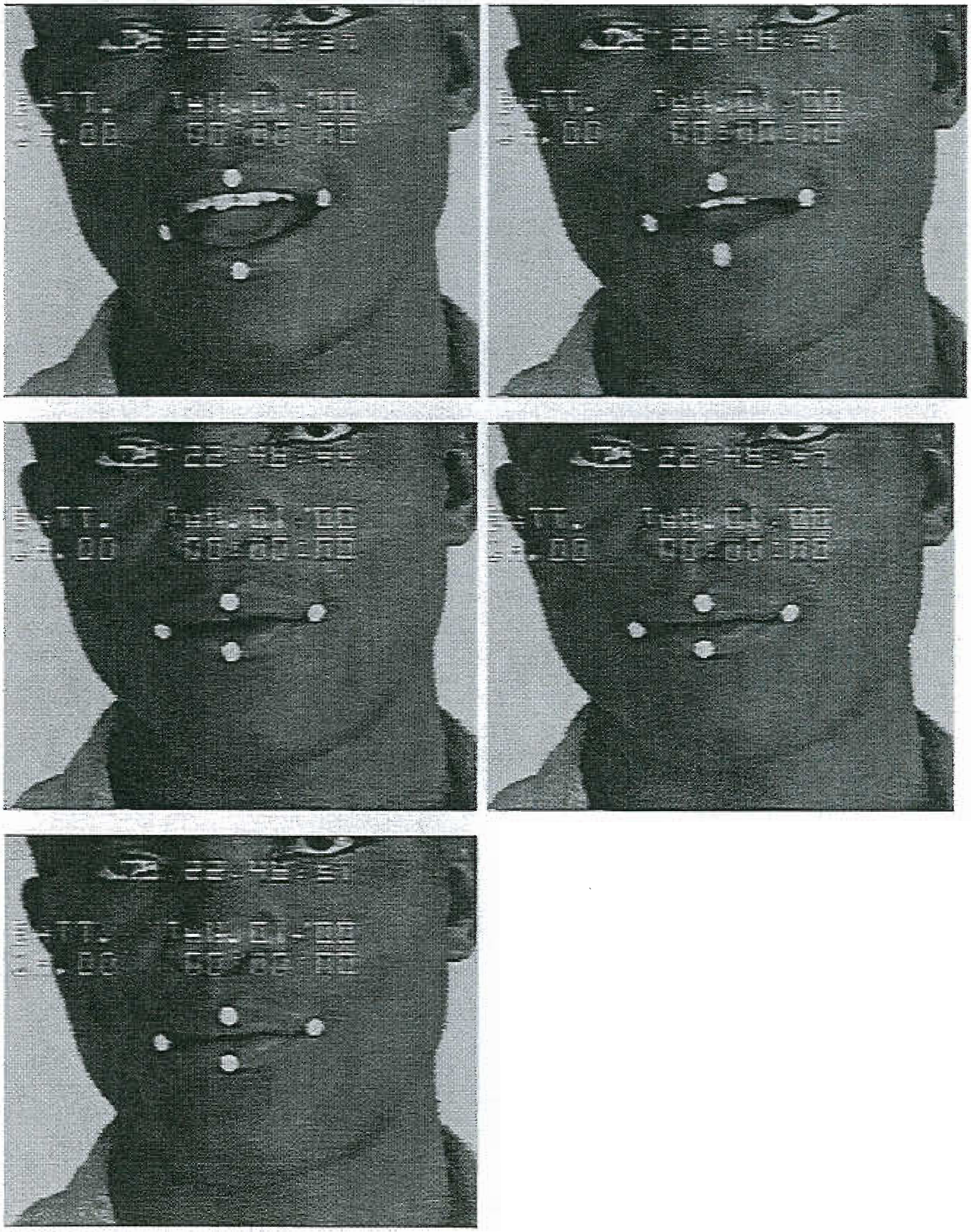
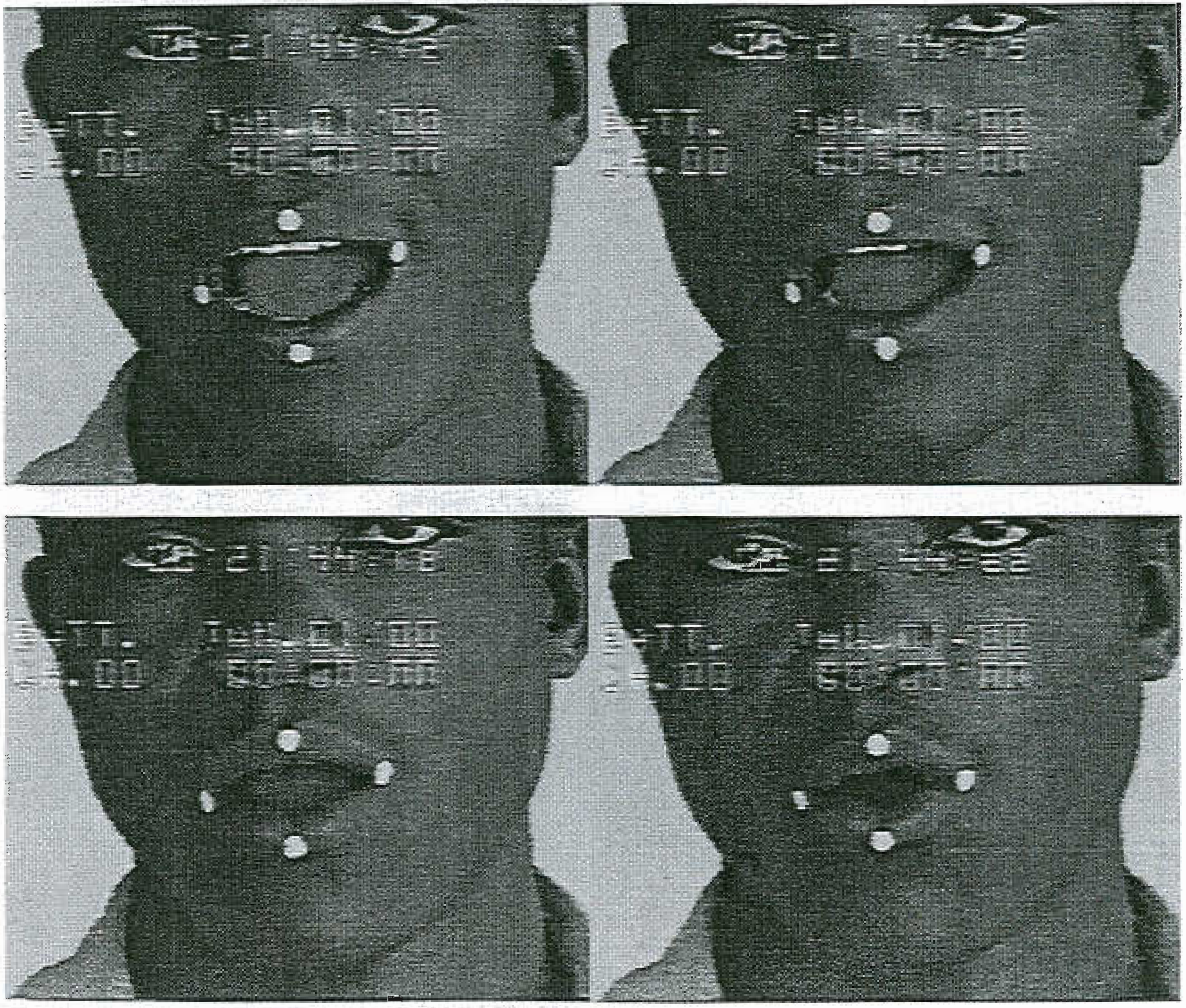


Figure 10. Five successive frames in the production of [aβa].

Figure 9 illustrates a complete bilabial closure for [m] in [ma], the first person singular past tense marker. Compression of the lips against each other brings the marker dots on the upper and lower lips very close together. A sequence of five successive frames covering the transition from the vowel [a] to the bilabial fricative [β] in [aβa] "beans" is shown in Figure 10. This sequence starts with the maximal opening for the initial [a] vowel and finishes with the maximal constriction for the fricative. The sequence follows from left to right in each row, then down to the next row. A small slit aperture remains at the center of the lips during [β], but the lips are very close together in the vertical dimension. There is no noticeable approximation of the corners of the lips.

Figure 11 illustrates the putative bilabial-velar fricative [x͡ɸ] which occurs intervocalically between [a]'s in the word meaning "charcoal (plural)". The sequence of ten frames shows the sequence from the initial vowel [a] into the consonant and up to the onset of the second [a]. In the production of this consonant the vertical separation of the lips is much greater than in [m] or [β] and the corners of the lips are drawn toward the center, resulting in an aperture shape that is rounded rather than slit.



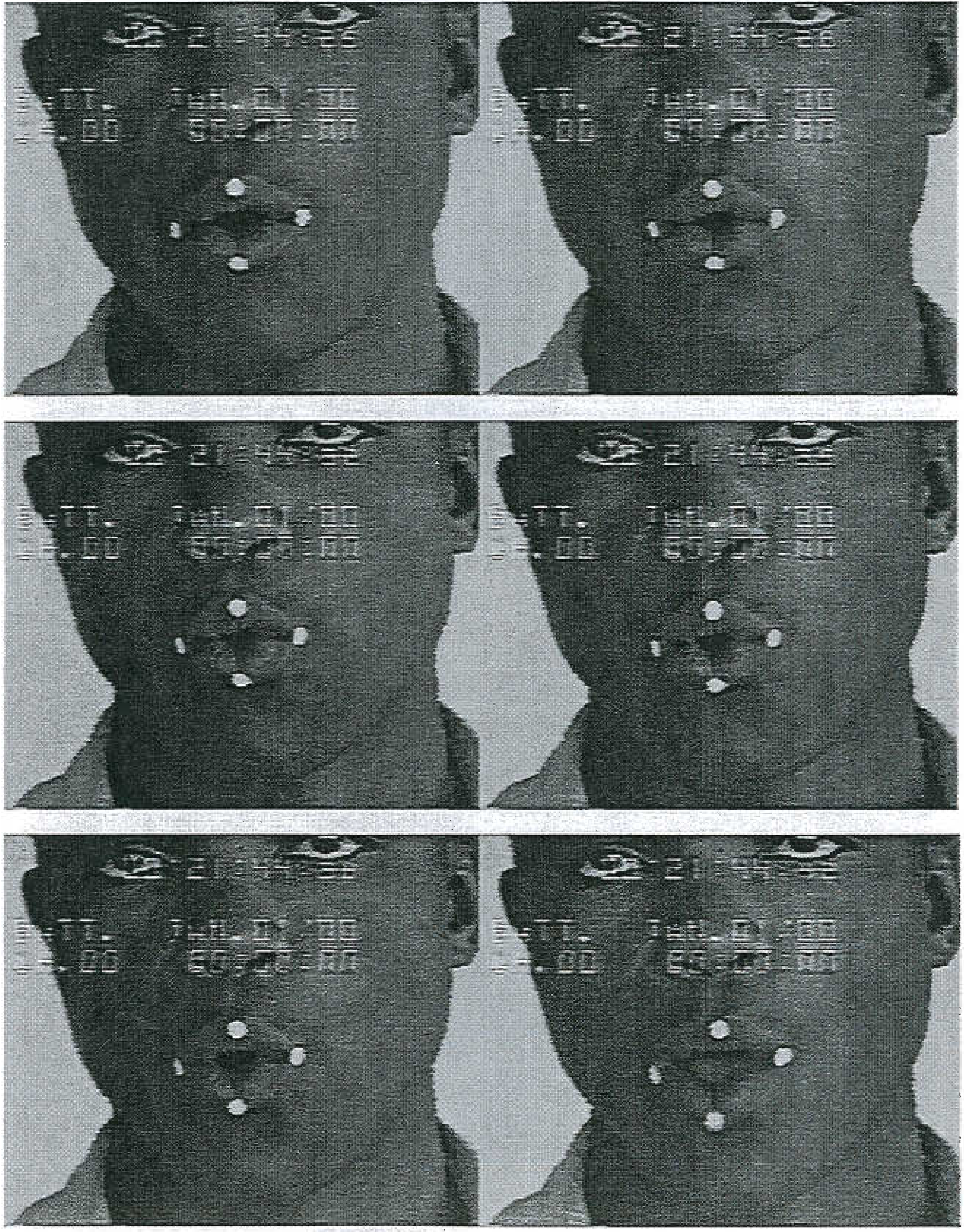


Figure 11. Ten successive frames from the /aCa/ sequence in the word meaning "charcoal (plural)", where C is the putative voiceless labial-velar fricative.

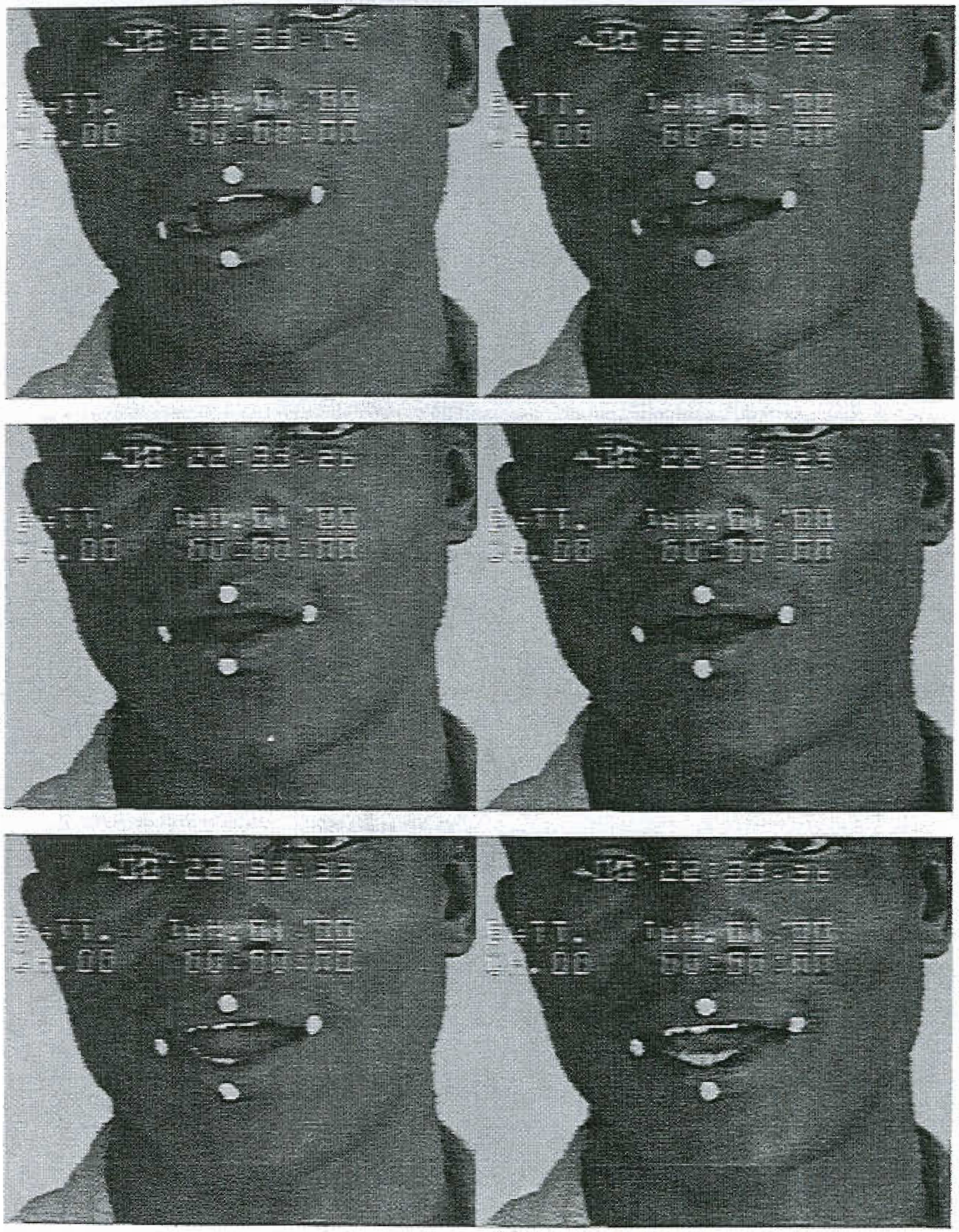


Figure 12. Six frames in the /aCa/ sequence in the phrase meaning “I pushed”, where C is the putative voiced labial-velar fricative.

Figure 12 shows that the vertical lip separation is even greater in the putative [ɣ͡β] in an [a__a] context in the phrase meaning “I pushed”. These six frames show the movement from the preceding [a] vowel to the maximal consonant constriction and on to the following [a] vowel. The lip corners are drawn in towards each other a little, but somewhat less vigorously than is seen in Figure 11.

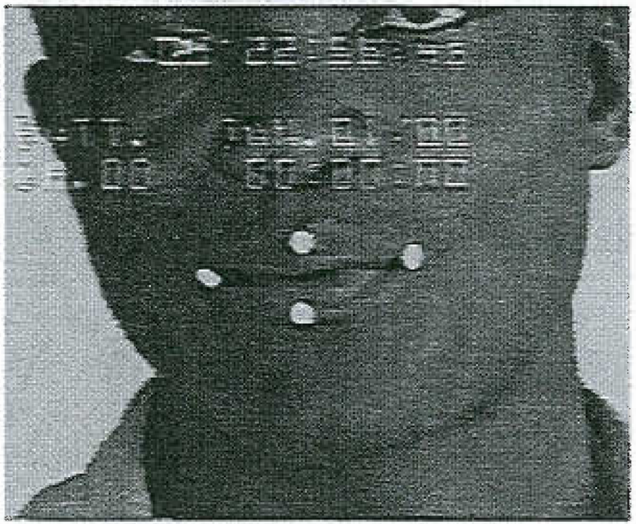


Figure 13. Maximal constriction recorded during the bilabial fricative in the word [koϕe] “farm village”.

Finally, Figure 13 shows the voiceless bilabial fricative in the loanword [koϕe] “farm village”. Only the frame containing the maximal constriction observed is shown. The narrow slit between the lips in this consonant is very similar to that seen in [β] in Figure 10.

It is very clear from Figures 9-13 that the labial articulations in the putative doubly-articulated fricatives are quite different in their nature and in their constriction degree from those found in the bilabial fricatives [β] and [ϕ]. These differences are quantified in Table 8, which gives the measured horizontal and vertical distances between the centers of the dots placed on the lips, and the horizontal and vertical distances across the aperture between the lips measured as accurately as is possible from the digitized frames containing the maximal constriction. Rather than writing [x͡ϕ, ɣ͡β], these segments would appropriately be transcribed as [xʷ, ɣʷ] since the labial articulation is approximant-like and clearly involves rounding. This transcription agrees well with the auditory impression of these sounds, which is very similar to the labialized velar fricatives found in other languages of the world.

Table 8. Minimum distances between dots and across aperture in mm.

	Between dots		Across aperture	
Segment	Vertical	Horizontal	Vertical	Horizontal
β in aβa	17.0	56.3	1.2	10.6
ϕ in koϕe	19.4	56.9	1.2	12.0
xʷ in axʷa	26.7	43.8	5.3	13.8
ɣʷ in ma ɣʷani	27.3	54.0	10.0	34.0

Without a field analysis of the type reported in this paper, the existence of [x͡ɸ, ɣ͡β] might be taken as established. This is an issue that has not only theoretical but also practical implications, since how to write the sounds [x͡ɸ, ɣ͡β] was considered at the 1992 conference on a Unified Orthography for Ghanaian Languages. If these were indeed unique sounds, the principles of that conference would have required that a distinct grapheme be found for them. This is not necessary. However, a satisfactory orthography for Avatime should distinguish among [+ATR] and [-ATR] high vowels, since these do both occur in the language.

Acknowledgments

The field research reported in this paper was carried out in November and December 1994 as part of the UCLA-University of Ghana faculty exchange program, funded by the US Information Agency through the James S. Coleman Center for African Studies. Much gratitude is due to Edmond Keller, Merrick Posnansky and their staffs for making my participation possible, and to my colleague on this trip, Russell Schuh for drawing my attention to this program. A debt is also due to Professor Florence Dolphyne of the University of Ghana, who arranged our welcome in Legon. Warm thanks are also due to Okusie Akyem Foli V, Chief of Amedzofe, and the elders of the town for giving their support to this project during the visit by Russell Schuh and myself to Amedzofe. Wilson Adinyira, Joseph Mawuena, Christian Dedume, Martin Osika and, especially, Chris Bubuama served as principal subjects and assisted in making contacts with others. Andy Ring and other members of the Ghana Institute of Linguistics, Literacy and Bible Translation (GILLBT), provided both essential material and logistical support, and an interested audience for talking about linguistics and life. The National Science Foundation supported the analysis and write-up phase through a grant for research on the phonetic structures of endangered languages. Finally, thanks are due to Russell Schuh for companionship in the research enterprise, as well as on the footpaths of the Achimota Forest Reserve.

References

Ford, Kevin C. 1971. *Aspects of Avatime Syntax.* Ph.D. dissertation, Department of Linguistics, University of Ghana, Legon.

Ford, Kevin C. 1973. On the loss of cross-height vowel harmony. *Research Review (Institute of African Studies, University of Ghana)* 4: 50-80.

Ford, Kevin C. 1988. Structural features of the Central-Togo languages. In Kropp Dakubu (1988): 126-154.

Funke, Emil. 1909. Deutsch-Avatime Wörterverzeichnis. *Mitteilungen des Seminars für orientalische Sprachen* 13: 1-38.

Funke, Emil. 1909. Versuch einer Grammatik der Avatimesprache. *Mitteilungen des Seminars für orientalische Sprachen* 12: 287-336.

Heine, Bernd. 1968. *Die Verbreitung und Gliederung der Togorestsprachen*. Dietrich Reimer, Berlin.

Hess, Susan. 1992. Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20: 475-492.

Hess, Susan. Forthcoming. *Pharyngeal Articulations.* Ph. D. dissertation, University of California, Los Angeles.

Jackson, Michel. 1988. *Phonetic theory and cross-linguistic variation in vowel articulation (UCLA Working Papers in Phonetics 69)*. Ph.D. dissertation, University of California, Los Angeles.

Kropp Dakubu, Mary Esther. 1988. *The Languages of Ghana*. Kegan Paul International, London.

Kropp, Mary Esther. 1967. *Lefana, Akpafu and Avatime with English Gloss (Collected African Wordlists No 1)*. Institute of African Studies, University of Ghana, Legon.

Ladefoged, Peter and Ian Maddieson. 1990. Vowels of the world's languages. *Journal of Phonetics* 18: 93-122.

Ladefoged, Peter and Ian Maddieson. 1995. *Sounds of the World's Languages.* Blackwells, Oxford.

Lindau, Mona. 1979. The feature expanded. *Journal of Phonetics* 7: 163-176.

Lindblad, Per. 1980. *Svenskans sje- och tje-ljud i ett allmänfonetisk perspektiv (Travaux de l'Institut de Linguistique de Lund 16).* C. W. K. Gleerup, Lund.

Ring, J. Andrew. 1987. *Planning for Literacy: a Sociolinguistic Survey of Multilingualism in Ghana.* Ph. D. dissertation, Georgetown University, Washington, DC.

Ring, J. Andrew. 1995. Revisiting the Central Volta Region: Avatime/Santrokofi/Bowiri. Manuscript. Ghana Institute for Linguistics, Literacy and Bible Translation, Tamale.

Schuh, Russell G. 1995a. Avatime noun classes and concord. Paper presented at the 26th Annual Conference on African Linguistics, UCLA, March 24-26, 1995.

Schuh, Russell G. 1995b. Aspects of Avatime phonology. *Studies in African Linguistics* 24: 31-67.

Stewart, John. 1989. Kwa. In J. Bendor-Samuel, ed., *The Niger-Congo Languages*. Summer Institute of Linguistics and University Press of America, Lanham MD: 217-245.

Struck, Bernhard. 1912. Einige Sudan-Wortstämme. *Zeitschrift für Kolonialsprachen* 2: 233-253, 309-323.

Tiede, Mark K. 1993. An MRI-based study of pharyngeal volume contrasts in Akan. *Haskins Laboratories Status Reports on Speech Research* 113: 107-130.

Voiceless approximants in Tee*

Peter Ladefoged

Most Ogoni (Kegboid) languages do not have voiceless nasals and laterals. But, as Kay Williamson pointed out in a personal communication, they do occur in Tee, (tɛ̀ɛ̀) a form of Kana, spoken by people of the Tai clan in the Tai district of the Tai-Eleme Local Government Area in Rivers State, Nigeria. Tee and Kana are closely related, with some degree of mutual intelligibility, but speakers of Tee regard the two as distinct languages, often noting, for example, that a particular form is a Kana word that is not used in Tee. In addition to the lexical differences, there are notable differences in the phonological inventories. Tee has four sounds that do not occur in Kana, a voiceless alveolar nasal n̥, a voiceless lateral l̥, and voiceless approximants or fricatives ç and ʍ. This paper will describe the characteristics of these sounds.

Tee has been discussed by Ikoro (1989), who used a Tee word list in his reconstruction of proto-Kegboid. He did not, however, record examples of the voiceless sounds mentioned above. The only other linguistic work is by Nwi-Bari (MS, 1993), a native speaker of the language, who compiled a dictionary and a short commentary on the language after attending a course taught by Kay Williamson and her colleagues. Nwi-Bari's work has been especially helpful in the preparation of the current report, which may therefore be taken as one more example of the fruits of Kay Williamson's work in Nigeria.

Tee phonological inventory

The consonants of Tee are shown in Table 1. As may be seen, the chart is typical of an Ogoni language, apart from the presence of the additional voiceless consonants. The vowels are also typical of this group. There are 7 oral and 5 nasalized vowels, as illustrated by the words in Table 2. There is no contrast between the mid vowel pairs **e** and **ɛ**, and **o** and **ɔ** when nasalized, even if the nasalization is simply allophonic; all vowels have nasalized allophones when adjacent to nasal consonants. There are three contrasting tones, symbolized as shown in Table 3, with the mid tone being left unmarked.

Table 1. Chart of Tee consonants

	Bilabial	Alveolar	Palatal	Velar	Labialized Velar	Labial velar
Plosive	**p b**	**t d**		**k g**	**kʷ gʷ**	**kp gb**
Nasal	**m**	**n̥ n**	**ɲ**		**ŋʷ**	
Fricative		**s z**				
Lateral		**l̥ l**				
Approximant	**ʍ w**	**ɹ**	**ç j**			

There is also a glottal stop which may not be contrastive, in that it occurs predictably before all vowel initial stems, but it can be considered to be phonologically required if all verb stems are said to be CV.

---

* Contributed to a Festschrift in honor of Professor Kay Williamson, University of Port Harcourt, Nigeria.

Table 2. Words illustrating the Tee vowels

Oral vowels			Nasalized vowels		
i	**bî**	faeces	ĩ	**bĩ́ĩni**	cooked, done
e	**bé**	fight,war	ẽ	**bẽẽ**	bend
ɛ	**bɛ̀ɛ̀**	rule (vb)			
a	**ba**	yam porridge	ã	**bã́**	pot
ɔ	**bɔ́**	to be fit			
o	**bo**	deformed person	õ	**bõ̀m**	puff adder
u	**bu**	belly	ṹ	**bṹ**	bush baby

Table 3. Contrasting tones in Tee.

High	**bé**	fight, war
Mid	**be**	home
Low	**bè**	enclose, make a fence around

The purpose of this paper is to demonstrate the voiceless sounds, including the voiceless nasals, which do not occur in other Ogoni languages, nor, indeed, in any other Nigerian language, as far as is known. Voiceless nasals are comparatively rare in the world's languages, although they occur in a number of languages spoken in South-East Asia, for example, Burmese (Maran 1971, Dantsuji 1984) and Angami (Bhaskarao and Ladefoged 1991), and in some Bantu languages spoken in Southern Africa, such as Kwangali. The other voiceless sounds in the set we are considering here occur in other Nigerian languages as well as elsewhere in the world. Voiceless lateral approximants are fairly unusual; but they occur in, for example, Burmese, Tibetan, and Klamath (Ladefoged and Maddieson, in press), as well as in some Nigerian languages, such as Bura (Hoffman 1957). The voiceless counterparts of the central approximants **j** and **w** are found in some forms of English, and in other languages, such as Yao (Purnell 1965), Klamath (Barker 1964) and Aleut (Bergsland 1956). The voiceless palatal **ç** is found in Nigerian languages, such as Bura, and the voiceless labial velar **ʍ** occurs in other Ogoni languages. However, I do not know of any other language that has the set of four voiceless sounds, **n̥**, **l̥**, **ç** and **ʍ**, which are found in Tee. Words showing these Tee contrasts are given in Table 4. We should also note that, in addition to the voiceless alveolar nasal **n̥**, there may be a voiceless bilabial nasal **m̥**, but this sound has been observed in only one word, **àm̥èè** 'a gland in the abdomen' (sic), and then in the speech of only one of the seven speakers observed for this paper, and occasionally in the speech of one other speaker.

Table 4. Some contrasting consonants in Tee.

	Voiceless		Voiced	
Alveolar nasal	**n̥à** **n̥ɔ**	hornbill war	**náa** **nɔ**	gun study
Alveolar lateral	**l̥a** **l̥éégа**	yawning to dislocate	**la** **lee**	palm frond bitter
Palatal approximant	**çáà** **çéé**	yam roll into a ball	**jaa** **jèègà**	thank you terrified
Labial velar approximant	**ʍàá** **ʍéé**	canoe, boat calabash	**wa** **wee**	wife flute

In order to investigate these sounds, the words in Table 4 were recorded by seven speakers of Tee, and spectrograms were made on a Macintosh computer using SoundEdit. The contrast between the voiceless and voiced alveolar nasal is shown in Figure 1. The voiceless nasal on the left consists of a voiceless portion (which may have some breathy voicing), followed by a voiced portion of approximately the same length. It might seem appropriate to regard this as a sequence of sounds which could be symbolized **hn**. But when this sound is compared with the voiced alveolar nasal in a minimal pair, shown on the right of Figure 1, it is apparent that it would be inappropriate to regard the voiceless sound as composed of two segments. It is considerably shorter than the voiced alveolar nasal, which there is no reason to consider as anything other than a single segment.

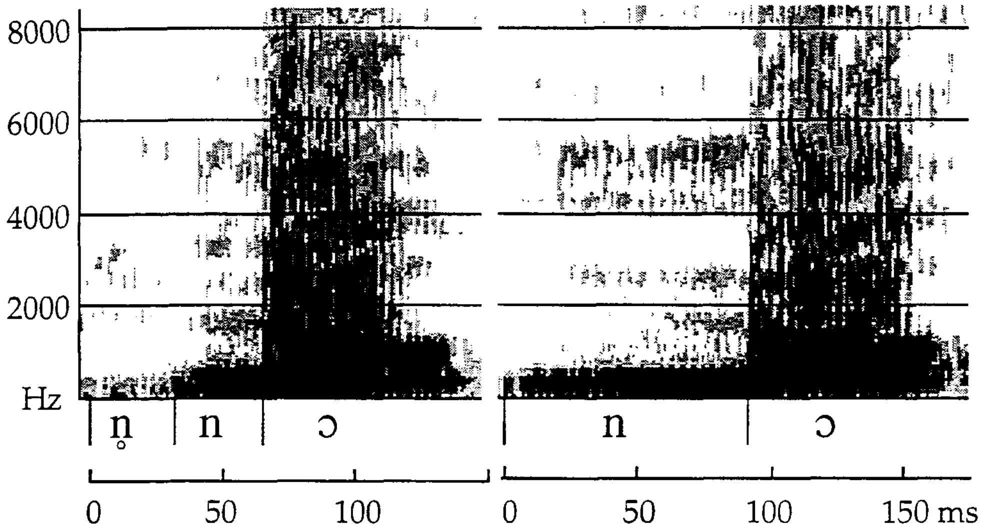


Figure 1. Spectrograms of the Tee words **n̥ɔ** 'war' and **nɔ** 'study'.

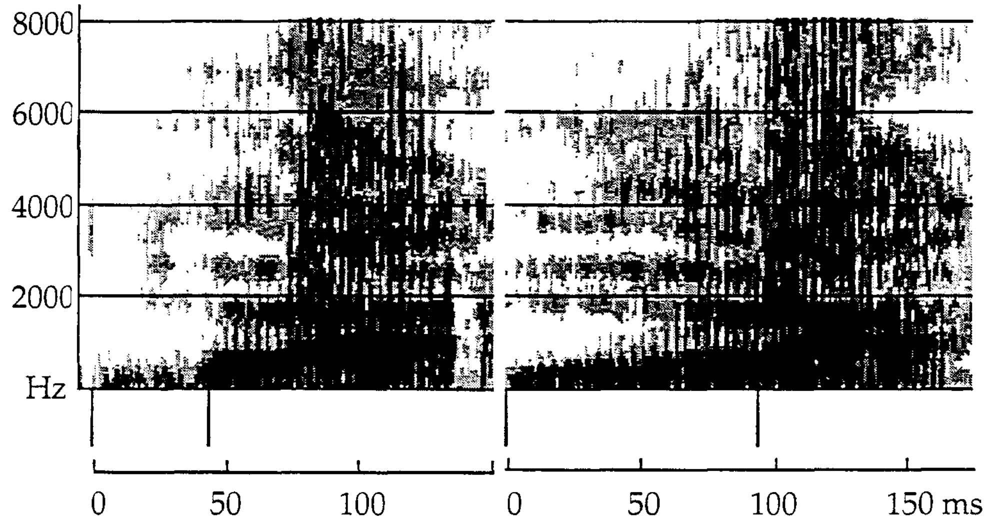


Figure 2. Spectrograms of the Tee words **l̥a** 'yawning' and **la** 'palm frond'.

The contrast between the voiceless and voiced alveolar laterals is somewhat similar, as is demonstrated in Figure 2. Once again, half the voiceless sound is actually voiced; and this sound, taken as a whole, is considerably shorter than the corresponding voiced sound. Maddieson and Emmorey (1984) have shown that there are two types of voiceless laterals: those that are frictionless, and can be classified as voiceless approximants, as in Burmese, and those that have a considerable strident component, making them lateral fricatives. It is apparent that the Tee voiceless laterals are of the Burmese type, in that they are accompanied by very little friction. Like the other alveolar sounds in Tee, most speakers who could be observed pronounced them with the tip of the tongue raised, making them apical alveolars.

The remaining sounds to be noted in this paper are illustrated in Figure 3. In the case of these glides, it is impossible to segment them into parts, but it is interesting to note that there is a tendency in these sounds (particulary the **ʍ**) for part of the sound to be voiceless, and part voiced, as was the case for the sounds previously considered.

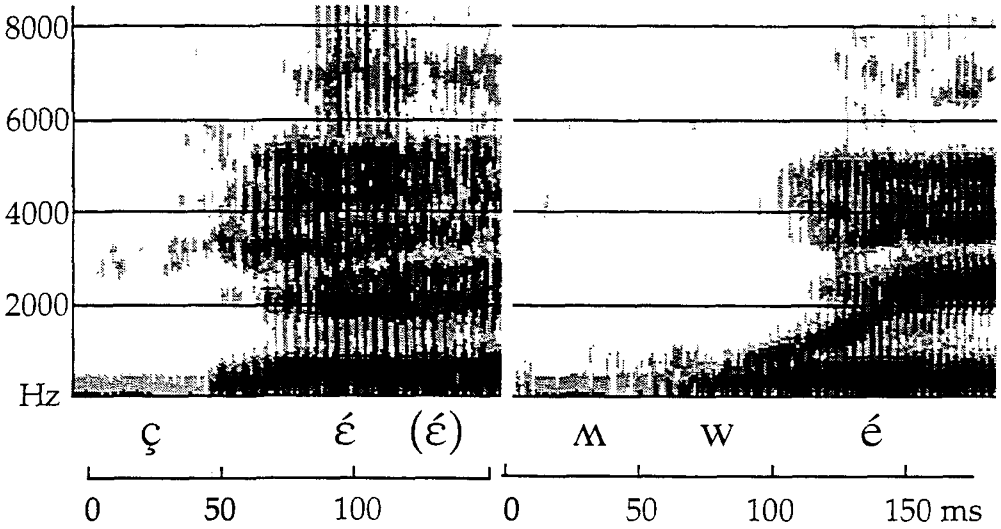


Figure 3. Spectrograms of the Tee words çɛ́ɛ́ 'roll into a ball' and **ʍ**éé 'calabash'. Part of the long vowel in each of these words has been cut off.

The origin of these voiceless sounds is not known. No doubt Kay Williamson, with her extensive knowledge of the history of the languages of Nigeria, will be able to suggest relevant cognates and phonological rules for their derivation.

References

Barker, M. A. R. (1964). *Klamath Grammar.* Berkeley and Los Angeles, University of California Press.

Bergsland, K. (1956). Some problems of Aleut phonology. *For Roman Jakobson.* The Hague, Mouton. 38-43.

Bhaskararao, P. and P. Ladefoged (1991). "Two types of voiceless nasals." *Journal of the International Phonetic Association* : 80-88.

Dantsuji, M. (1984). "A study on voiceless nasals in Burmese." *Studia Phonologica*: 1-14.

Hoffman, C. (1957). *A grammar of the Bura language.* Hamburg.

Ikoro, S. M. (1989). *Segmental phonology and lexicon of Proto-Kegboid.* University of Port Harcourt.

Maddieson, I. and K. Emmorey (1984). "Is there a valid distinction between voiceless lateral approximants and fricatives?" *Journal of Phonetics* : 181-190.

Maran, L. R. (1971). *Burmese and Jingpho: A Study of Tonal Linguistic Processes.* Urbana, University of Illinois.

Nwi-Bari, W. K. (1993). *Tee - English Dictionary.* Unpublished MS.

Purnell, H. C. (1965). *Phonology of a Yao Dialect.* Hartford, Hartford Seminary Foundation.

The phonetic structures of Defaka

Aaron Shryock, Peter Ladefoged and Kay Williamson

1. Introduction

Defaka (**défàkà**) is an Ijoid language, belonging to the Atlantic-Congo group of the Niger-Congo family. It is spoken by a small community, most of whom live in a single ward of Nkọrọọ town in the Bonny Local Government Area, Rivers State, Nigeria. Defaka is the only known language which shows a considerable similarity to Ijọ without actually being Ijọ (Williamson, 1989). The language was first described by Jenewari (1983, 1989), who was primarily concerned with its linguistic affiliation. The present paper incorporates much of the only other linguistic work on Defaka, a long essay for the B.A. (Hons) degree by Oseinte Daba Bob-Manuel (1990), now unfortunately deceased. As so much of her work has been incorporated here, she deserves to have been listed as a co-author, but we have departed somewhat from her analysis, and she should not be considered responsible for the present form of this paper.

The Defaka live very much among the Nkọrọọ, whose language is a dialect of Ijọ. Both our principal informants came from mixed marriages, in one case with an Nkọrọọ mother and a Defaka father, and in the other case with a Defaka mother and an Nkọrọọ father. Nearly all the Defaka are bilingual in Defaka and Nkọrọọ (often speaking other languages as well), and most of the Defaka under 30 do not speak the language. In these circumstances we must acknowledge that the phonetics and phonology of the language we are describing are heavily influenced by the Nkọrọọ dialect of Ijọ, and the two languages may be indistinguishable at this level. They are, however, quite distinct languages lexically, and not mutually intelligible. There are probably no more than a few hundred speakers of Defaka. The 1963 census reported that the Defaka and the Nkọrọọ together totaled 5,468 people (Jenewari 1983:6).

This study is based on data collected in August 1994 at the University of Port Harcourt and in the Defaka ward of the village of Nkọrọọ. The data consists of recordings of 12 male speakers of Defaka.

2. Vowels

Proto-Ijọ has been reconstructed as having 9 vowels in two harmony sets, **i**, **e**, **a**, **o**, **u** and **ɪ**, **ɛ**, **a**, **ɔ**, **ʊ**, with **a** being common to both sets (Williamson forthcoming). Defaka and the Nkọrọọ dialect of Ijọ are the easternmost of the Ijoid languages, and are entirely surrounded by languages with considerably smaller vowel systems. Williamson (1973) has shown that Nkọrọọ has been strongly influenced by its neighbors: **ɪ** has fallen together with either **i** or **e**, and **ʊ** with **u** or **o**. In addition **ɛ** has at least partially merged with **e**, and **ɔ** with **o**. Williamson notes that the contrast between **e** and **ɛ** is rather tenuous in Nkọrọọ, and that there are comparatively few sequences of vowels in words involving **ɔ**. Nkọrọọ is thus best regarded as a 6 or 7 vowel language, which, if the language survives, may shortly be reduced to five vowels. The same may be true of Defaka.

In addition to the 7 oral vowels **i, e, ɛ, a, ɔ, o, u** , Defaka has five contrastively nasalized vowels (all the oral vowels are allophonically nasalized when adjacent to a nasal consonant). The complete set of Defaka vowels is illustrated by the words in Table 1.

Table 1. Words illustrating the Defaka vowels.

i	**ǐbǐ**	'sleep'	ī	**sì̄ì̄**	'sink'
e	**èbè**	'pig'	ē	**dḗḗ**	'chief'
ɛ	**ɛ̀wɛ̀rɛ̀**	'children'			
a	**ǐbà**	'it is done'	ā	**sà̄à̄**	'urine'
ɔ	**ǐtɔ̀**	'stomach'			
o	**ǐdò**	'farm'	ō	**óɓṓṓ**	'400'
u	**ób↓ú**	'door'	ū	**sù̄ù̄**	'stretch'

In order to examine the acoustic properties of the vowels of Defaka, the collected data was digitized at a sampling rate of 20 kHz using the Kay Elemetrics Computer Speech Laboratory (CSL) package. The first four speakers were recorded under reasonably good condition at the University of Port Harcourt. The other 8 speakers were recorded in one crowded session in a small house in the Defaka ward of the village of Nkọrọọ. Two tokens of each word were recorded for each of the four speakers at Port Harcourt; only one token was recorded for the remaining 8 speakers.

To examine the formant frequencies and amplitudes of the vowels of Defaka, we selected words which illustrate the 7 contrasting short and long vowels. For the short vowels, words were selected in which the vowel occurred word-initially preceding a labial consonant. However, it was necessary to measure vowels occurring in other environments in order to have data for the vowels **ɛ, ɔ, u**. For the long vowels, there were so few words in the data set illustrating these vowels that it was not possible to control for the environment in which the vowels occurred. The words used in the examination of formant frequencies and amplitudes are presented below in Table 2.

Table 2. Words used in formant frequency and amplitude measures for the vowels.

i	**ǐbǐ**	'sleep'	**ii**	**ǐǐmà**	'heavy'
e	**èbè**	'pig'	**ee**	**tèè**	'play'
ɛ	**ɛ̀dɛ̀lɛ̀**	'vulture'	**ɛɛ**	**jɛ̀ɛ̀**	'bird'
a	**ápá**	'wing'	**aa**	**fàà**	'say'
ɔ	**ɔ́sɔ́rɔ́**	'navel'	**ɔɔ**	**ɔ̀ɔ̀**	'roast'
o	**òbò**	'back'	**oo**	**óó**	'salt'
u	**ób↓ú**	'door'	**uu**	**túúnɔ̃**	'five'

The frequency of the first three vowel formants (F1, F2, and F3, respectively) were measured during a steady-state portion of the vowel. The measurements were taken from a wideband spectrogram supplemented by reference to FFT spectra, averaged over an 80 ms window positioned where the formant measurements were taken. In the case that the steady-state portion of the vowel was less than 80 ms in length, a 52 ms window was used. In the case of the **ii** of **ǐǐmà** 'five', the vowel formants were measured with a 52 ms window position near the onset of the vowel to minimize the effect of allophonic nasalization resulting from the adjacent nasal. Finally, in addition to the frequency measurements, the amplitudes of the formants in the FFT analysis were measured by taking the greatest amplitude of the harmonics comprising each formant .

The formant frequency and amplitude values for the mid short vowels **e, ɛ, o, ɔ** presented in Table 3 represent only the vowels of the four speakers recorded in Port Harcourt. In the case of the 8 speakers recorded in Nkọrọọ, the word **ɛ̀dɛ̀lɛ̀** ‘vulture’ was not collected. In addition, the speakers in Nkọrọọ preferred to pronounce **ɔ́sɔ́rɔ́** ‘navel’ as **ísɔ́rɔ́**. As a result, we excluded the formant frequency and amplitude values for the corresponding mid vowel **e** and **o** of these 8 speakers because including these measurements without the measures for ɛ and ɔ for the same speakers would produce an unbalanced data set for these vowels which could misrepresent the nature of the acoustic differences between them.

The mean formant frequencies and amplitudes for the short and long vowels are presented in Table 3. The means of the short and long vowels are not presented separately because a two factor analysis of variance including the factors vowel quality and vowel length indicated that vowel length does not have a significant effect on vowel quality: for F1, F[1,159]=0.39, p<.5334; F2, F[1,159]=0.469, p<.4944; and F3, F[1,159]=3.445, p<.0653.

Table 3. Mean formant frequencies (Hz) and amplitudes (dB) of the vowels.

Vowel	F1	F2	F3	A1	A2	A3
i	363	2113	2912	63	41	40
e	418	1917	2634	61	40	38
ɛ	446	1866	2675	58	42	35
a	751	1446	2451	57	51	40
ɔ	542	1144	2417	58	53	35
o	423	1064	2137	65	51	33
u	344	1022	2338	62	45	30

The mean formant frequencies are plotted in Figure 1 with F1 plotted against F2' - F1. F2' is a weighted average of F1, F2 and F3 calculated as following: F2' = F2 + (F3 - F2) (F2 - F1) / 2 (F3 - F1) (Fant 1973:52).

The first point to note about these formant frequencies is the considerable overlap between the mid vowels **e** and ɛ and, as far as F1 is concerned, between **o** and ɔ. Despite the overlap, the members of these pairs of vowels do have significantly different F1 values, showing that they do differ in vowel height. In the case of **e** and ɛ, the mean F1 of ɛ is 38 Hz greater than that of **e** (F[1,37]=6.48, p<.0152). The mean F1 of ɔ is significantly greater than that of **o** by 119 Hz (F[1,45]=34.327, p<.0001). Jenewari (1983, 1989) and Bob-Manuel (1990) suggest that the mid vowels **e**, ɛ and **o**, ɔ are also distinguished by contrasting Advanced Tongue Root (ATR). Many West African languages are known to have vowels that differ in the position of the tongue root, but there is no direct evidence of this in Defaka. However, the difference may be not simply in the tongue root gesture, but in the enlargement of the whole pharyngeal cavity, partly by the movement of the tongue root, but also by the lowering of the larynx. The lowering of the larynx sometimes results in vowels having a slightly breathy quality, which means that more harmonic energy will be present in the lower part of the spectrum. With this in mind we considered the difference in the amplitudes of the first two formants (Ladefoged, Maddieson, and Jackson 1988, Fulop et al. 1995). An ATR vowel with a breathy voice quality will have relatively less amplitude in the second formant, assuming no changes in lip rounding (which may not be valid for our data) and assuming that the formant frequencies do not come significantly closer together (which is the case in our data, where the first and second formants are far apart in **e** and ɛ and about 600 Hz apart in both **o** and ɔ). We found that **e** has a greater A1-A2 measure than ɛ; however, the difference is not statistically significant (F[1,37]=2.768, p<.1046). We also found that o has a 9 dB greater difference than ɔ, so that, relatively, the second formant in **o** has less intensity than that in ɔ. This difference is significant (F[1,45]=32.393, p<.0001), but it may not be a valid indication of a difference in ATR, as an increase in lip rounding would also lower the relative amplitude of the second formant. There is therefore very little evidence for a difference in voice quality (and hence in ATR) among Defaka vowels.

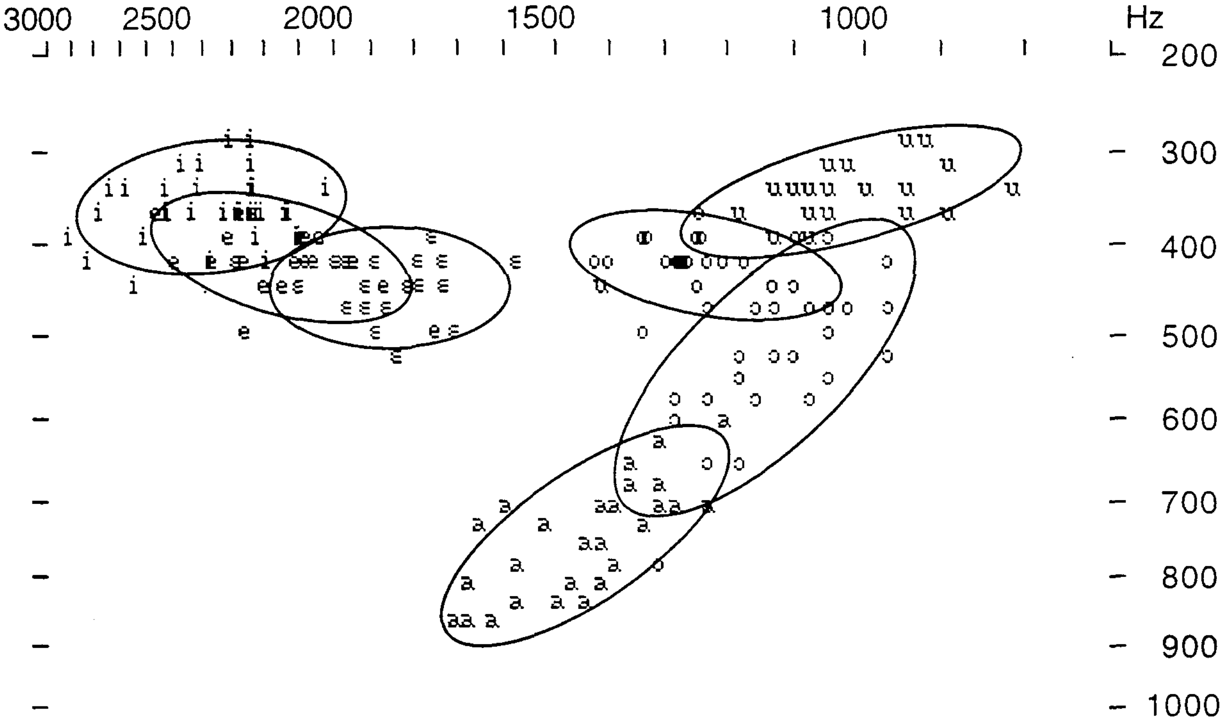


Figure 1. Means and standard deviations (Hz) of the formant frequencies of the oral vowels.

Defaka has five contrastively nasalized vowels. In order to examine the acoustic properties of the nasalized vowels, we measured the formant frequencies and amplitudes of these vowels. In addition, we measured the frequency and amplitude of the principal nasal formant. In the case of **a** it was not possible to distinguish the nasal formant because of its proximity to F2. Compare the formants of **i** and **ĩ** in Figure 2. Note the additional nasal formant in **ĩ** at approximately 1200 Hz. In the case of **a** and **ã**, though, the nasal formant of **ã** is not distinguishable from the F2 of the vowel. Thus, the vowels **a** and **ã** have similar formant values.

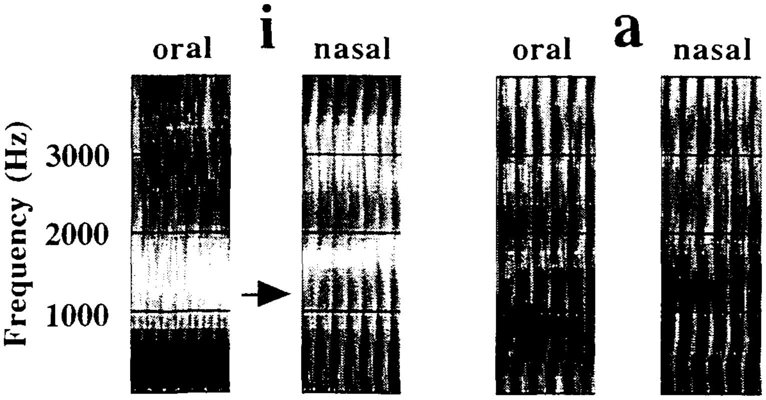


Figure 2. Spectrograms of oral and nasal vowels. The arrow indicates the nasal formant in **ĩ**.

The words used in the examination of formant frequencies and amplitudes are presented below in Table 4 and the mean formant frequencies and amplitudes of the nasalized vowels in Table 5.

Table 4. Words used in formant frequency and amplitude measures for the nasal vowels.

ĩ	**sĩ̀ĩ̀**	'sink'	**õ**	**óɓṍṍ**	'400'
ẽ	**dẽ́ẽ́**	'chief'	**ũ**	**sũ̀ũ̀**	'stretch'
ã	**sã̀ã̀**	'urine'			

Table 5. Mean formant frequencies (Hz) and amplitudes (dB) for the nasal vowels.

Vowels	F1	F2	F3	N	A1	A2	A3	AN
ĩ	356	2163	2882	1256	60	34	34	32
ẽ	541	1923	2793	1231	56	46	40	40
ã	753	1483	2533	-	53	51	35	-
õ	432	930	2446	1510	58	44	30	28
ũ	358	958	2345	1509	56	41	31	33

In Figure 3 the mean formant frequencies of the nasalized vowels are plotted with F1 against F2' - F1. As noted above, F2' is a weighted average of F1, F2 and F3 (Fant 1973:52).

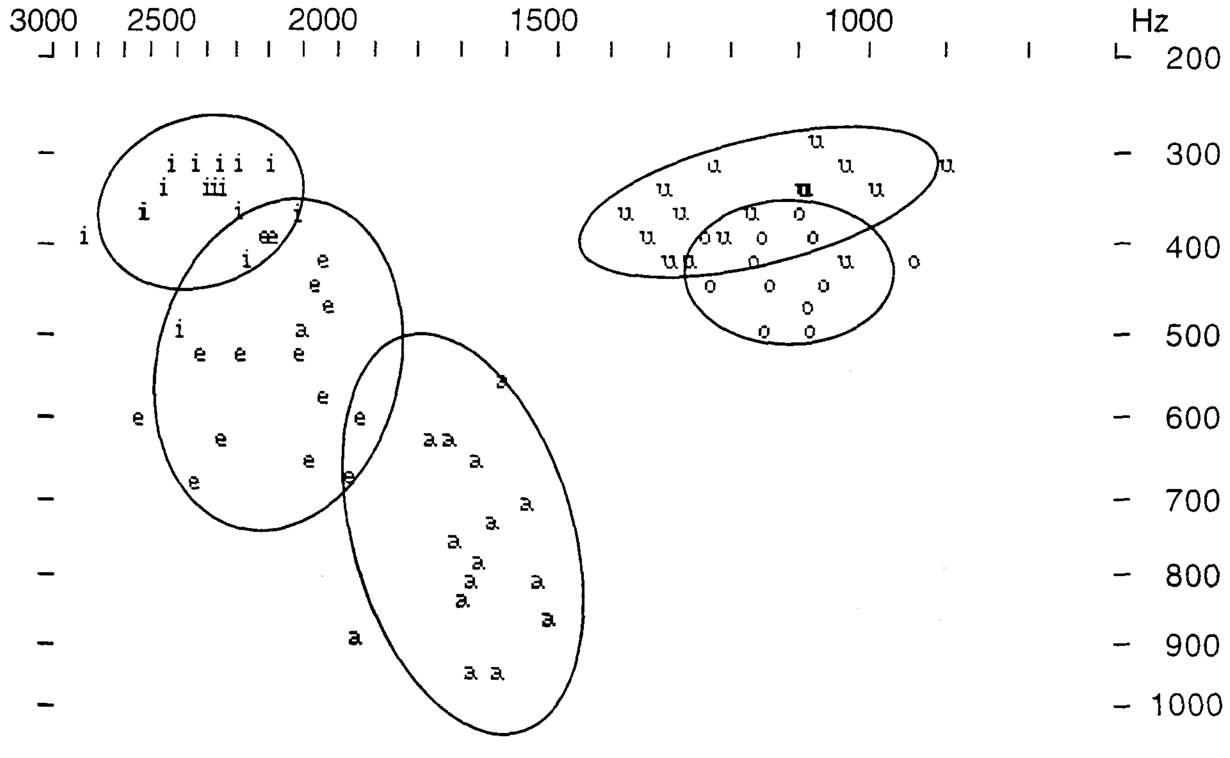


Figure 3. Means and standard deviations (Hz) of the formant frequencies of the nasal vowels.

The intrinsic pitch of the vowels was also investigated. The fundamental frequency (F0) of the low-toned vowels **i, e, ɛ, a, ɔ, o** was measured from a narrow band spectrogram display. A low-toned **u** does not occur in the collected data. F0 was calculated from the frequency of the 10th harmonic or the highest distinguishable harmonic in the steady-state portion of the vowel.

For the words of the shape VCV, the F0 of the first and second vowels was measured and averaged. The words used in the investigation of inherent pitch are presented below in Table 6 and the mean frequencies in Figure 4.

Table 6. Words used in intrinsic pitch measures.

i	**ìbì**	'sleep'	a	**fàà**	'say'
e	**tèè**	'play'	ɔ	**ɔ̀ɔ̀**	'roast'
ɛ	**jɛ̀ɛ̀**	'bird'	o	**òbò**	'back'

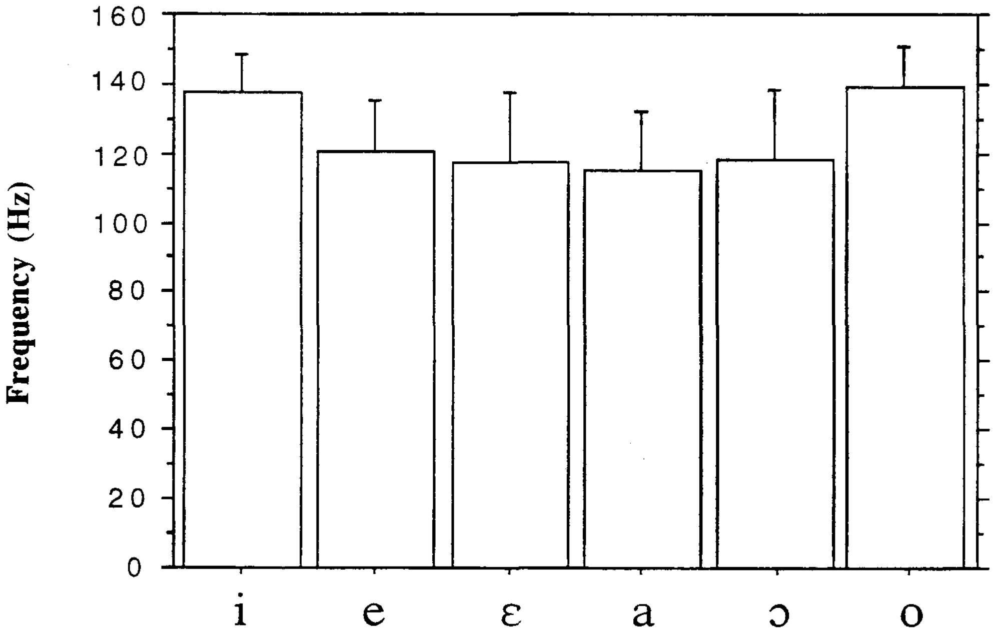


Figure 4. Means and standard deviations (Hz) of the low-toned vowels.

It has been reported that there is a correlation between vowel height and intrinsic pitch such that high vowels have an intrinsic pitch which is higher than low vowels. The intrinsic pitch of the low-toned vowels of Defaka show a tendency to conform to this correlation between vowel height and intrinsic pitch. However, the differences in pitch between these vowels are not significantly except in the case of **o** and **a**. Moreover, the pitch of **o** is higher than that of **i**, an unexpected result.

Finally, vowel length is contrastive in Defaka. The duration of the short and long vowels were measured. In the case of words of the shape VCV, the duration of the first vowel was measured. In the case of **u**, however, the vowel only occurs in the word **óbꜜú** 'door'. The vowels **ɛ** and **ɛɛ** were not measured because of the difficulty in determining the duration of **ɛɛ** in the word **jɛ̀ɛ̀** 'bird', the only word in the corpus with this vowel. The words used in the investigation of vowel duration appear in Table 7. The mean durations of the short and long vowels are presented in Figure 5.

Table 7. Words used in duration measures for short and long vowels.

i	**ìbì**	'sleep'	ii	**ììmà**	'heavy'
e	**èbè**	'pig'	ee	**tèè**	'play'
a	**ápá**	'wing'	aa	**fàà**	'say'
ɔ	**ɔ́sɔ́ɾɔ́**	'navel'	ɔɔ	**ɔ̀ɔ̀**	'roast'
o	**òbò**	'back'	oo	**óó**	'salt'
u	**ób↓ú**	'door'	uu	**túúnɔ̀**	'five'

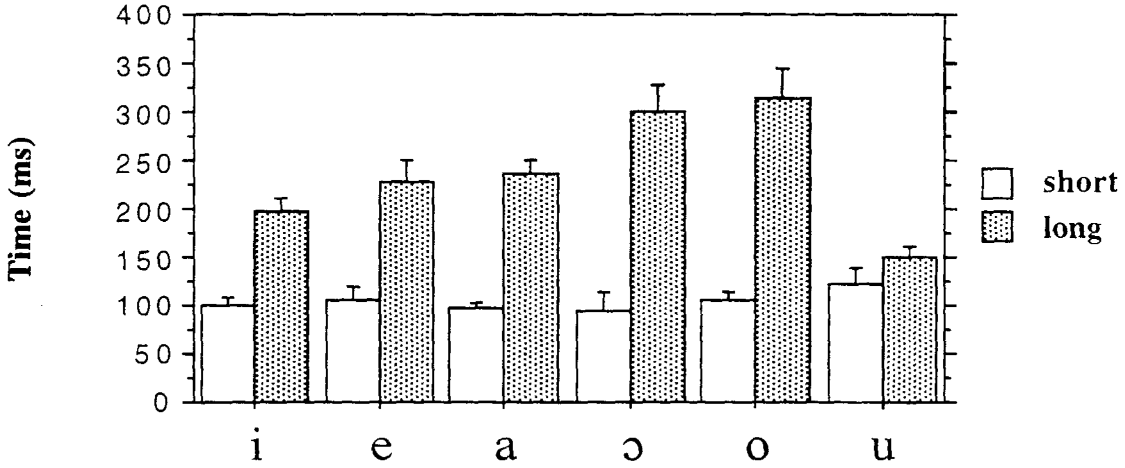


Figure 5. Means and standard deviations (ms) of the short and long vowels of Defaka.

The long vowels are at least twice the length of the corresponding short vowels except in the case of **u** and **uu**. As noted above, the vowel **u** occurs in final position as opposed to initial position as with the other short vowels. In order to determine the effect of final position on the duration of **u**, its duration was compared to the durations of the initial and final vowels of **òbò** 'back'. The initial **o** of **òbò** 'back' has a mean duration of 106 ms; the final **o** 130 ms. Assuming a comparable relation between the duration of **u** in initial and final position, an initial **u** would have a mean duration of approximately 100 ms given the mean duration of 123 ms for the final **u**. Even with the inferred duration of 100 ms for **u**, the difference in duration between **u** and **uu** is not as great as seen with the other vowel pairs. The relatively shorter duration of **uu** may result, in part, from its position in a penultimate syllable with a consonant in the onset position of the syllable.

3. Consonants

The consonants of Defaka are shown in Table 8. Notable allophones of these consonants are shown in parentheses in the table below.

3.1. Plosives

Defaka contrasts voiced and voiceless unaspirated stops at three places of articulation: bilabial, alveolar, and velar. In addition, there are two labial velar stops that might be considered to be voiced as opposed to voiceless, but which do not contrast in the same way as the other voiced and voiceless stops. There is also an implosive produced at the bilabial place of articulation. The plosives and implosive of Defaka are illustrated by the words in Table 9.

Table 8. Chart of Defaka consonants.

	BILABIAL	LABIO-DENTAL	ALVEOLAR	PALATAL	VELAR	LABIAL VELAR
PLOSIVE	**p b**		**t d**		**k g**	**kp gb**
AFFRICATE				**dʒ**		
IMPLOSIVE	**ɓ**					
NASAL	**m**		**(n)**		**(ŋ)**	
TAP			**ɾ**			
FRICATIVE		**f v**	**s (z)**			
LATERAL			**l**			
APPROXIMANT	**w**		**(ɹ)**	**j**		

Table 9. Words illustrating the Defaka plosives and implosive.

p	**páɾà**	'leg'	**ápá**	'wing'
b	**bìɔ̀**	'river'	**èbè**	'pig'
t	**tèè**	'play'	**ìtɔ̀**	'stomach'
d	**dẽ́ẽ**	'chief'	**ìdò**	'farm'
k	**káà**	'hand'	**ɓéké**	'fall'
g	**gògó**	'namesake'	**àgáɾà**	'lizard'
kp	**kpá↓á**	'wind around'	**kpìkpìkì**	'owl'
gb			**àgbà**	'bite'
ɓ	**ɓéké**	'fall'	**óɓṍõ**	'400'

The voiced plosives, **b**, **d**, **g**, are fully voiced throughout the stop closure, contrasting with the voiceless plosives **p**, **t**, **k** which have no voicing during the closure and are followed by a short Voice Onset Time (VOT) before the following vowel. The so-called voiceless labial velar **kp** is phonetically distinct from both these sets of stops in that it is characterized by the onset of voicing during the closure, prior to the release, a property of voiceless labial velars previously observed (Ladefoged 1964, Garnes 1975, Connell 1994; Pettorino and Giannini 1995). Figure 6 presents waveform displays illustrating these differences in the timing of the onset of voicing for **b**, **p**, and **kp**. The release of the stop closure is indicated by an arrow. It is difficult to be precise about this moment, but it is taken to be when the first cycle indicating vocalic quality occurs. In the case of **b** there are voicing vibrations throughout the period prior to the release. For **p**, there is a short interval after the release before voicing begins; and for **kp** the voicing commences slightly earlier than the release. This small difference between **p** and **kp** is consistent across all our speakers. The degree of voicing in the so-called voiceless labial velar **kp** is similar to that in the so-called voiced stop **b** in English.

We measured the VOT for the three voiceless unaspirated stops, **p, t, k**. The words used appear in Table 10. The mean VOT durations are presented in Figure 7. As has been found for many other languages, the VOT for the velar stop is significantly longer than that for the alveolar and bilabial stops, which, in Defaka, are not significantly different from one another.

Table 10. Words used in duration measures for Voice Onset Time (VOT).

p	**páɾà**	'leg'
t	**táátú**	'three'
k	**kàɾì**	'curse'

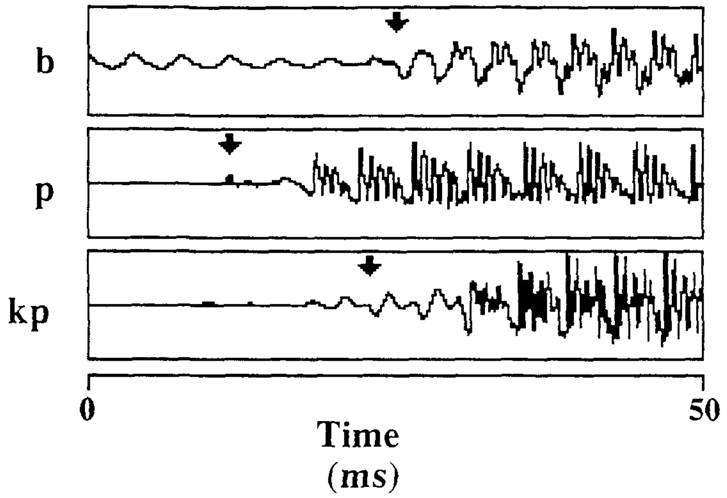


Figure 6. Waveform displays of the release of **b**, **p**, and **kp**.

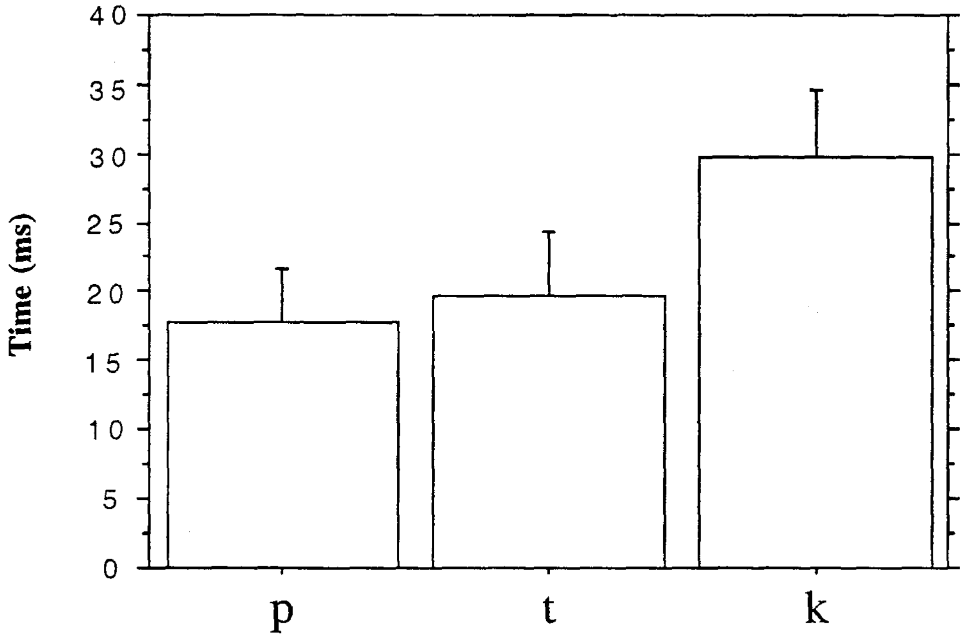


Figure 7. Means and standard deviations for the three voiceless unaspirated stops of Defaka.

The closure intervals of the voiced plosives and implosive are primarily voiced through the entire closure. There is, however, a difference in the way that the voicing is manifested. In the voiced plosives **b**, **d**, **g**, the voicing occurs throughout the entire closure, sometimes diminishing slightly during the latter part before the release. In the case of the implosive **ɓ**, there may be a tendency for the amplitude of voicing to gradually increase, but this was not evident in most of our samples. The implosive in Figure 7, in fact, shows a gradual decrease in the amplitude of voicing. The labial velar **gb** is characterized by greater amplitude of voicing during the closure. Moreover, the amplitude of the voicing generally increases toward the release of the closure. These patterns of closure voicing are illustrated by the waveform displays in Figure 8.

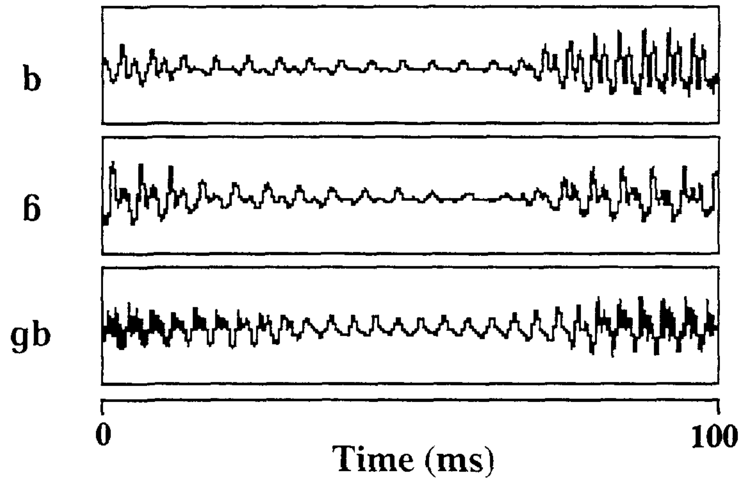


Figure 8. Waveform displays illustrating the voiced closures of **b**, **ɓ**, and **gb**.

The duration of the stop closure was measured for the voiced and voiceless unaspirated stops and the labial implosive. The closure duration of the labial velar stops was not measured because the collected data did not contain words with these consonants in the appropriate environment. The words used in the investigation of closure duration appear in Table 11. The mean closure durations are presented in Figure 9. As is generally found, the intervocalic voiceless stops are slightly longer than their voiced counterparts, although in this language the difference is statistically significant only in the case of the bilabial stops.

Table 11. Words used in closure duration measures.

p	**ápá**	'wing'
b	**èbè**	'pig'
ɓ	**èɓè**	'carve'
t	**ìtɔ̀**	'stomach'
d	**ìdò**	'farm'
k	**ɓéké**	'fall'
g	**gògó**	'namesake'

3.2. Nasals

There is a single nasal phoneme in Defaka: the bilabial **m**. The alveolar **n** occurs as an allophone of **l** in the presence of contrastive nasalization on a following vowel. The velar **ŋ** occurs in homorganic consonant closures with velar and labial velar stops. The fact that Defaka has a single nasal phoneme and that this nasal is bilabial is not unusual in West African languages, but is cross-linguistically rare. The nasals of Defaka are illustrated in Table 12.

Table 12. Words illustrating the Defaka nasals.

m			**ììmà**	'heavy'
n	**nṍm**	'person'	**ónṹmã́**	'bush'
ŋ	**ŋgbè**	'pick'	**ìŋgì**	'axe'

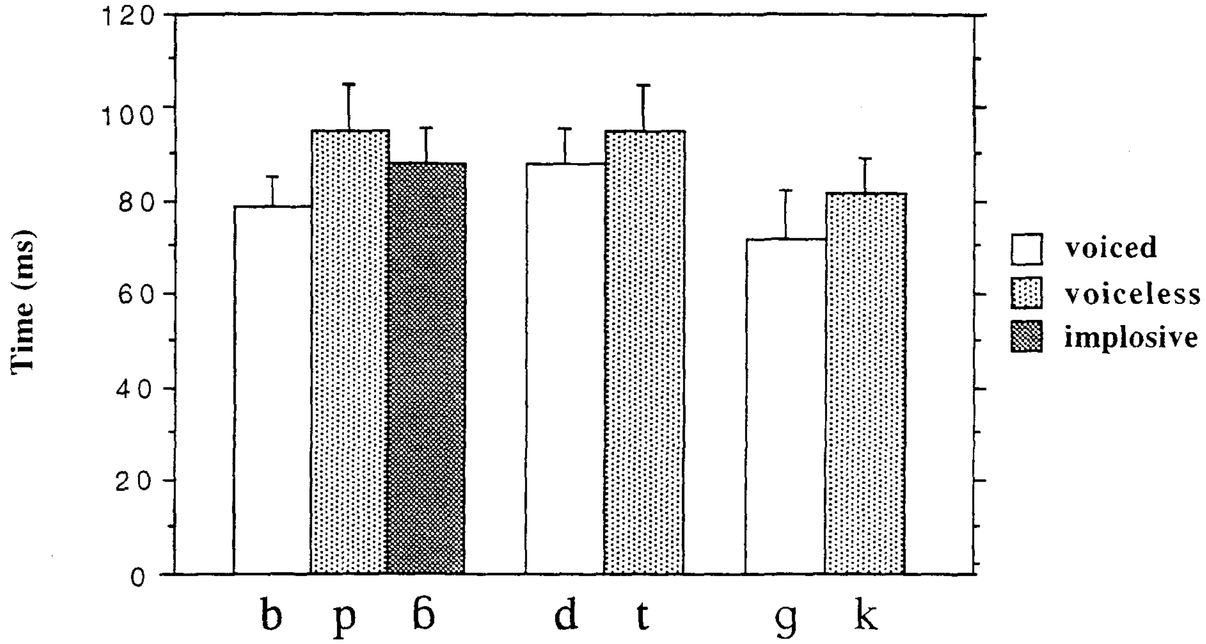


Figure 9. Mean duration (ms) of the intervocalic stops.

3.3. Affricate

There is a single voiced affricate **dʒ** in Defaka. It occurs intervocalically in our data, but it is reported to occur word-initially as well by Bob-Manuel (1990). As noted above, **z** occurs as an allophone of **dʒ** for some speakers. Two speakers consistently produced **z** in the place of **dʒ**. Six speakers produced **dʒ** for each of the four words in the corpus with **dʒ**. The remaining speakers produced instances of **dʒ** as well as **z**.

Table 13. Words illustrating the affricate **dʒ** in Defaka.

dʒ	**ídʒá**	'cooking pot'
	ìdʒè	'death'
	ádʒáré	'rust'
	ìdʒùlè	'father'

There is voicing through the entire closure interval of **dʒ**. Moreover, there is a tendency for the closure to be only partially formed resulting in frication and a greater amplitude of voicing. These differing degrees of constriction are illustrated by the spectrograms of **dʒ** in the word **ádʒáré** 'rust' in Figure 10.

3.4. Fricatives

There are three fricatives in Defaka: the voiceless and voiced labio-dental fricatives **f**, **v** and the voiceless alveolar fricative **s**. The following table illustrates these fricatives.

Table 14. Words illustrating the Defaka fricatives.

f	**fàà**	'say'	**òfòrò**	'blow with mouth'
v	**válà**	'sail'	**ìvè**	'throw'
s	**suu**	'stretch'	**ɔ́sɔ́ɾɔ́**	'navel'

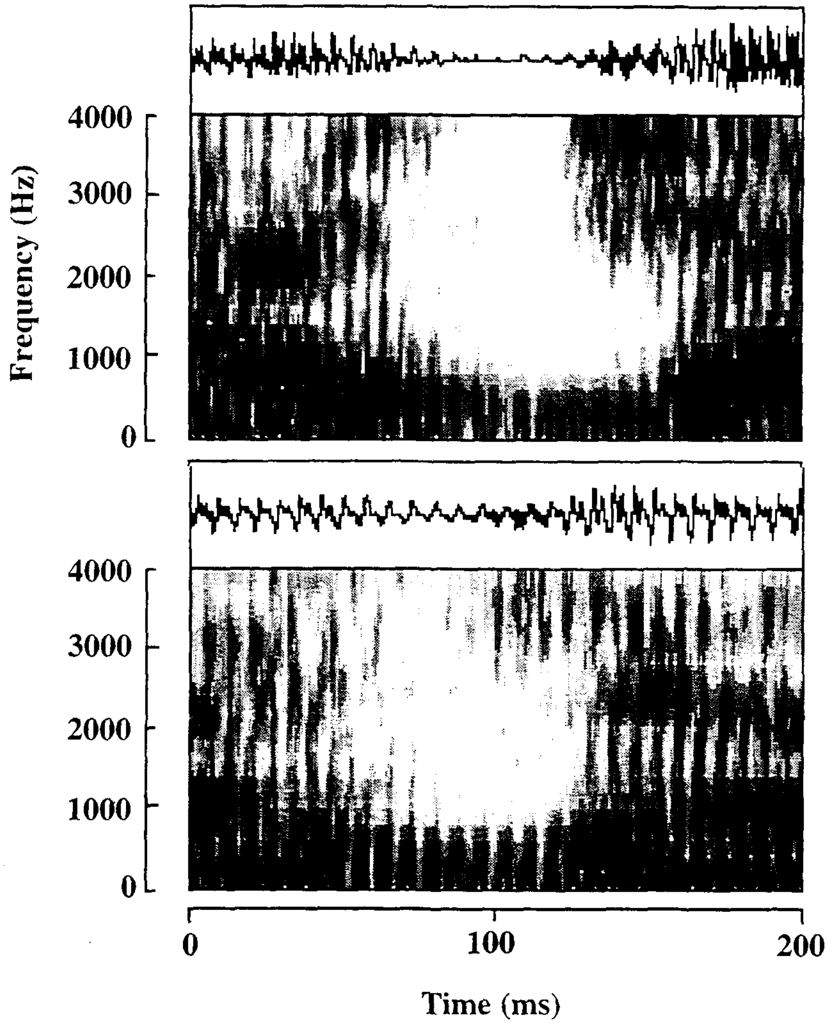


Figure 10. Spectrograms of **dʒ** with complete closure (upper figure) as opposed to incomplete closure (lower figure).

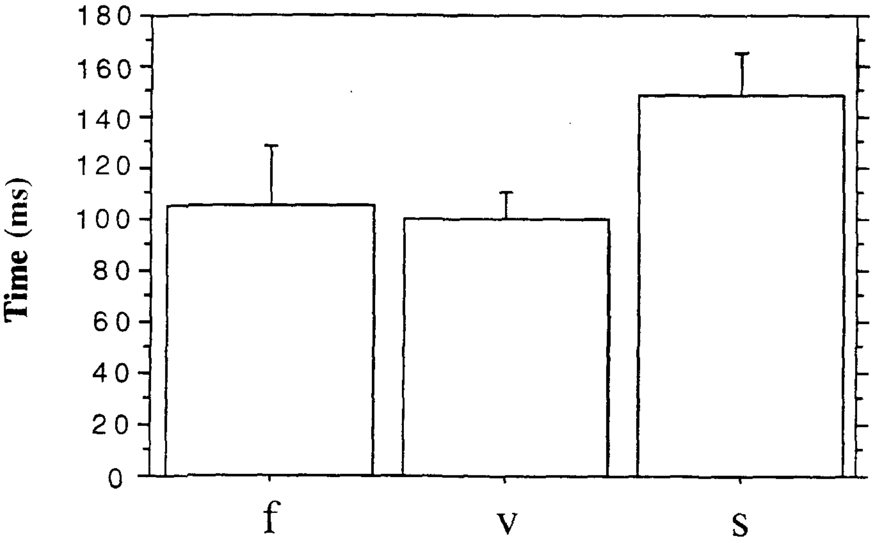


Figure 11. Means and standard deviations (ms) of fricatives in word-initial position.

The duration of the fricatives in initial position was measured. The three words in Table 14 were used for this purpose. The results are presented in Figure 11. The sibilant is longer than the labial fricatives by 46 ms. The difference is significant as determined by an ANOVA with the factor place (labial vs. alveolar), (F[1,45]=23.404, p<.0001). The greater duration of sibilant fricatives has been previously noted in the literature (Behrens and Blumstein 1988, You 1979).

3.5. Tap, lateral and approximates

Defaka has a single rhotic, an alveolar tap **ɾ**. There is an alveolar lateral **l** as well. The approximants include the bilabial **w** and the palatal **j**. These consonants appear in Table 15.

Table 15. Words illustrating the tap, lateral and approximates in Defaka.

ɾ			**pɑ́ɾɑ̀**	'leg'
l	**lɛ̀lɛ̀**	'sell'	**àálà**	'the woman'
j	**jɛ̀ɛ̀**	'bird'	**ɔ́jɔ̀**	'eye'
w	**wá**	'we'	**ɛ̀wɛ̀ɾɛ̀**	'children'

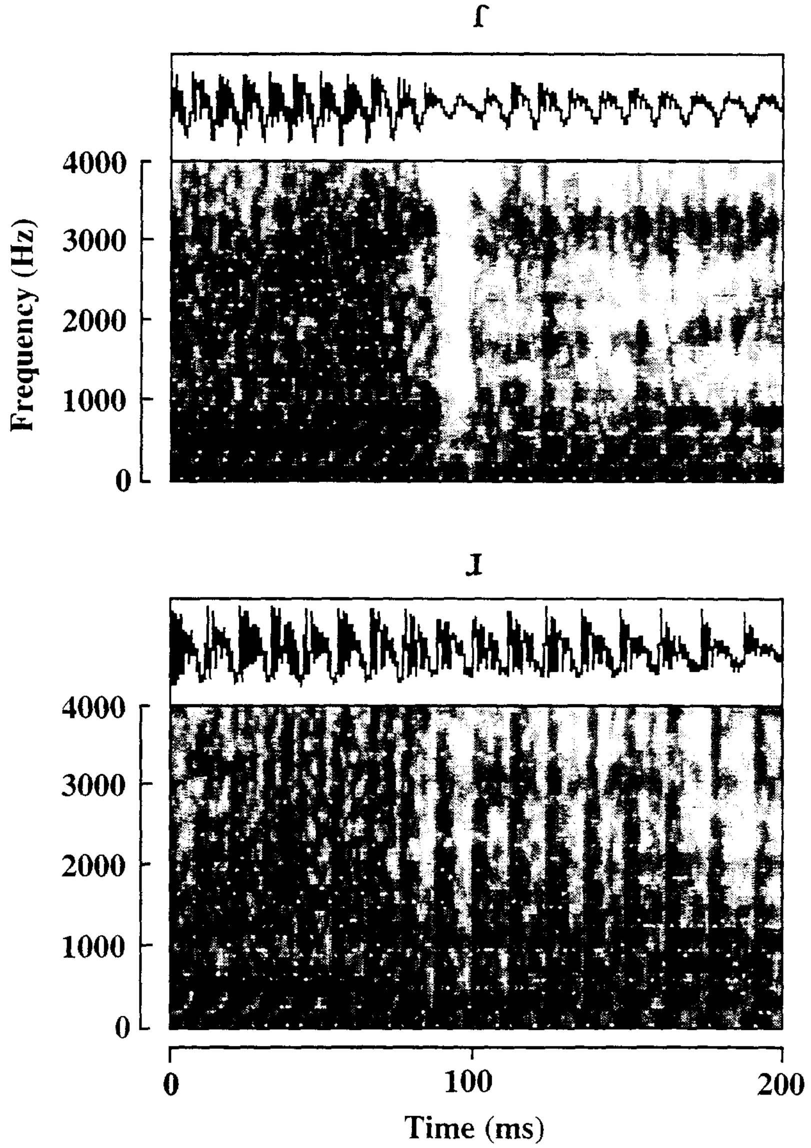


Figure 12. Spectrograms of **ɾ** (upper figure) and **ɹ** (lower figure) in **páɾà** 'leg'

The alveolar tap occurs intervocalically. Bob-Manuel (1990) reports that the tap occurs word-finally as well; however, we were not able to confirm this observation. A number of speakers produce an approximant ɹ. The spectrogram in Figure 12 shows the alveolar tap ɾ and the approximant ɹ in the word **páɾà** 'leg'.

As noted above, in the presence of contrastive nasalization of the following vowel, **l** is realized as **n**. The approximants **w** and **j** are not realized as nasals in the presence of a contrastively nasalized vowel, but they are also nasalized in this environment. The spectrograms in Figure 13 contrast oral and nasal counterparts of **j**.

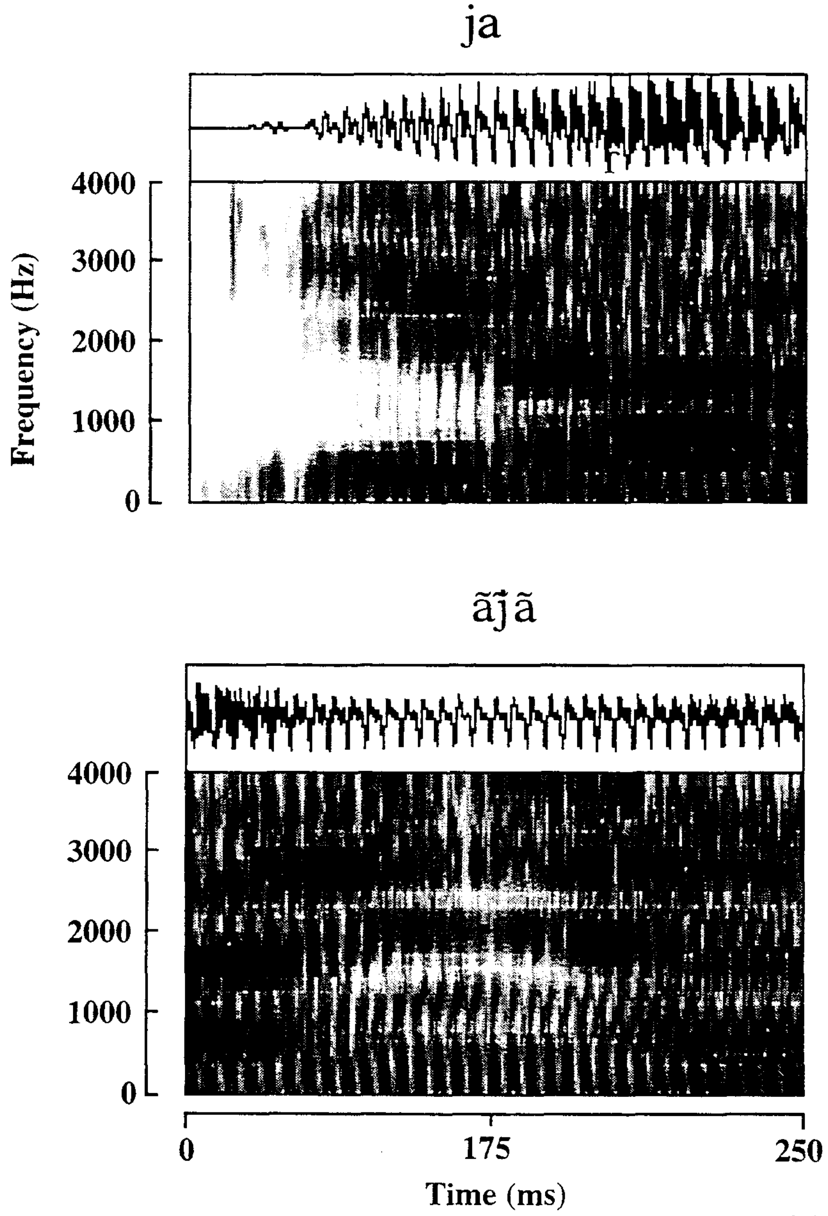


Figure 13. Spectrograms of **ja** in **jáà** 'thing' (upper figure) and **ãj̃ã** in **àtã́j̃ã̀** 'good things' (lower figure)

Compare the formants of **j** and **j̃** in Figure 13. Note the additional nasal formant in **j̃** at approximately 1200 Hz. As for the following vowels **a** and **ã**, the first and second formants have comparable values although the formants in the nasalized vowel are slightly higher.

4. Tone

There are two contrastive tones in Defaka: High and Low. There are also two contour tones, High-Low and Low-High, which arise from the combination of the two contrastive tones on a monosyllabic word containing a long vowel or diphthong. Finally, Bob-Manuel (1990) reports a downstepped High tone, **ꜜH**, resulting from a High Low High sequence of tones. The following table illustrates these tones of Defaka.

Table 16. Tones and tonal contours in Defaka.

H	**wá**	'we'	**óó**	'salt'	**ápá**	'wing'
L	**ì**	'I'	**tèè**	'play'	**èbè**	'pig'
HL			**jáà**	'thing'	**ɓátà**	'feaces'
LH			**dàá**	'father'	**gògó**	'namesake'
HꜜH			**kpáꜜá**	'wind around'	**óbꜜú**	'door'

In order to investigate the phonetic characteristics of these tones, the fundamental frequency (F0) of these distinct tones and tonal contours was measured. The words used, all of which contain the vowel **a**, are presented below in Table 17. The measurements were taken from a narrow band spectrogram display of the vowel. The F0 was calculated from the frequency of the 10th harmonic or the highest distinguishable harmonic at the onset, middle, and offset of the vowel. We limited our measurements to the data collected from the four speakers at the University of Port Harcourt because these recordings were of higher quality and three of the five words investigated were not collected from the speakers recorded in Nkọrọọ. A word with a High tone and the long vowel **aa** does not occur in the data set for the four speakers in question. Therefore, we measure the F0 of the first vowel in **ápá** 'wing'.

Table 17. Words used in investigation of F0.

H	**ápá**	'wing'
L	**bàà**	'day'
HL	**káà**	'hand'
LH	**dàá**	'father'
HꜜH	**kpáꜜá**	'wind around'

The mean F0 of these tones and contours appear in Figure 14. Note that there are essentially three F0 contours: level as in the case of **H**, rising as in the case of **LH**, and falling for **HL**, **HꜜH**, and **L**. The falling F0 patterns of **HL**, **HꜜH**, and **L** are not significantly distinct at the onset, middle, or offset of the vowel as determined by a series of single factor ANOVA's: in the case of **HL** and **HꜜH**, F[1,14]=0.65, p=.4337 at the onset, F[1,14]=0.461, p=.5082 at the middle, and F[1,14]=0.496, p=.4927 at the offset. For **HꜜH** and **L**, F[1,14]=0.062, p=.8071 at the onset, F[1,14]=0.52, p=.4826 at the middle, and F[1,14]=0.417, p=.5289 at the offset. For **HL** and **L**, F[1,14]=1.451, p=.2483 at the onset, F[1,14]=2.646, p=.1261 at the middle, and F[1,14]=2.165, p=.1633 at the offset. These findings may draw into question the phonological status of the proposed contrasts between, for instance, **HL** and **HꜜH**. More likely, though, these findings indicate that more controlled data is required to determine the phonetic differences between these tonal contrasts.

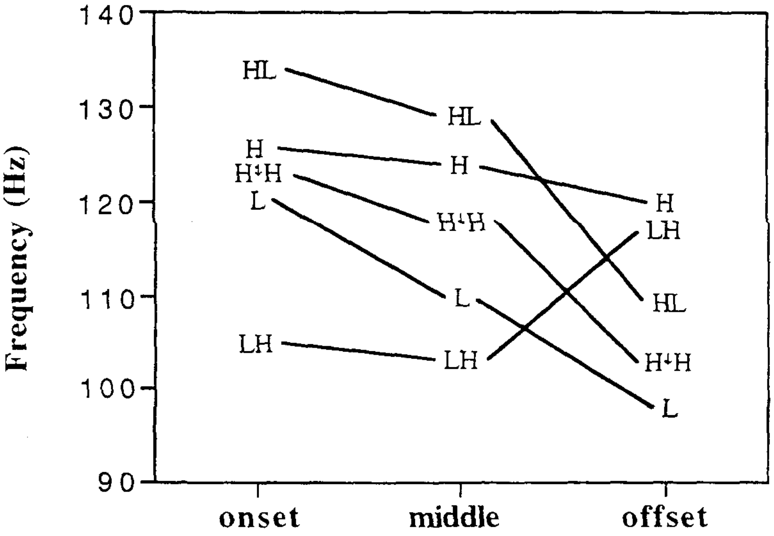


Figure 14. Mean F0 (Hz) at onset, middle, and offset of vowel in Defaka.

5. Conclusion

This paper has investigated the principal phonetic properties of the segmental and suprasegmental contrasts of Defaka. The results reported here represent the first phonetic study of the language. Moreover, this is one of the first acoustic studies of an Ijoid language. In several respects Defaka exhibits the phonological and phonetic structures common to other languages of West Africa. The vowel inventory consists of 7 vowels: **i, e, ɛ, a, ɔ, o, u**. There are also five nasal vowels. The mid vowel sets **e**, **ɛ** and **o**, **ɔ** are distinguished by the value of F1, indicating a height difference. They may also differ in the position of the tongue root. The consonant inventory of Defaka includes voiced and voiceless unaspirated stops, labial velars, and an implosive. There is a single affricate, **dʒ**, which is realized as **z** for several speakers. The nasal inventory consists of a single phoneme, **m**. **n** and **ŋ** occur as allophones of **m**. The fricatives include **f**, **v**, **s**. The language has an alveolar tap **ɾ** which has an approximant allophone, **ɹ**. The remaining sonorants include **l**, **w**, **j**. **l** has the allophone **n** in the presence of contrastive nasalization on a following vowel. The approximants **j** and **w** are nasalized in this environment. Finally, there are two tones in Defaka which give rise to five distinct tonal patterns: **H, L, HL**, **LH**, and **H$^{\downarrow}$H**. The **L, HL**, and **H$^{\downarrow}$H** tonal contrasts are similar in that they are realized phonetically as a falling F0 contour.

Acknowledgments

This paper is dedicated to the Defaka people. We appreciate their patience and generous assistance which made this study possible.

References

Behrens, S. and S. Blumstein. 1988. Acoustic characteristics of English voiceless fricatives: a descriptive analysis. Journal of Phonetics 16.295-98.

Bob-Manuel. O. S. 1990. Aspects of the phonology of Defaka. Port Harcourt: Long essay, Department of Linguistics and African Languages , University of Port Harcourt.

Connell, B. 1994. The structure of labial-velar stops. Journal of Phonetics 22.441-76.

Fant, G. 1973. Speech sounds and features. Cambridge, MA: MIT Press.

Fulop, S., E. Kari, and P. Ladefoged. 1995. Articulatory modeling of so-called advance tongue root vowels. Paper presented at the 129th Meeting of the Acoustical Society of America.

Garnes, S. 1975. An acoustic analysis of double articulation in Ibibio. Ohio State Working Papers in Linguistics (Proceedings of the Conference on African Linguistics) 20.44-55.

Jenewari, C. 1983. Defaka: Ịjọ's closest linguistic relative. Delta Series No. 2. Port Harcourt: University of Port Harcourt Press.

Jenewari, C. 1989. Ịjoid. The Niger-Congo Languages, ed. by J. Bendor-Samuel, 105-18. New York: University Press of America.

Ladefoged, P. 1964. A Phonetic Study of West African Languages: an auditory-instrumental study. Cambridge: Cambridge University Press.

Ladefoged, P., I. Maddieson and M. Jackson. 1988. Investigating phonation types in different languages. Vocal physiology: voice production, mechanisms and functions, ed by O. Fujimura, 297-317. New York: Raven.

Pettorino, M. and A. Giannini. 1995. Labial-velar stops in Dagbani. Proceedings of The XIIIth International Congress of Phonetic Sciences, ed. by K. Elenius and P. Branderud, 578-81. Stockholm: Grasfiska.

Williamson, K. 1973. Some reduced vowel harmony systems. Research Notes 6.145-69. Ibadan: Department of Linguistics and Nigerian Languages, University of Ibadan.

Williamson, K. 1989. Niger-Congo overview. The Niger-Congo Languages, ed. by J. Bendor-Samuel, 3-45. New York: University Press of America.

Williamson, K. forthcoming. Comparative Ijọ. Port Harcourt.

You, H.-Y. 1979. An Acoustic and Perceptual Study of English Fricatives. Calgary: University of Alberta M.A. thesis.

Appendix: Defaka word list

p	páɾà	leg
	pèkù	bat
	ápá	wing
	èpèlè	trick
b	bã́jã́	catarrh
	bàà	day
	bǐè	ask
	bǐɔ̀	river
	bɔ̀m̀	child of someone
	bṹṹ	flowing tide
	èbèɾè	dog
	îbà	yesterday
	ǐbà	it is done (of food)
	ǐbǐ	sleep
	ǐbò	big
	óbꜜú	door
	òbò	back
t	táátó	three
	tàà	wife
	tèè	play
	túó tùò	sing a song
	túúnɔ̃̀	five
	tùò	song
	àtã́jã̀	good thing
	ètẽ́jã̀	possessions
	ǐtà	take
	ǐtɔ̀	stomach
	ɓátà	faeces
	ɔ́lɔ̀tî	old person
d	dàá	father
	dẽ́ẽ́	chief
	àkàdà	pepper
	ɛdɛ̀lɛ̀	vulture
	ǐdò	farm
k	káà	hand
	kàɾǐ	curse
	kéé	cut
	àkàdà	pepper
	ókúnã́	compound
	òkùnã̀	fowl
	pèkù	bat
	ɓéké	fall
	kpǐkpǐkǐ	owl
g	gògó	namesake
	àgáɾà	lizard
	áŋgî	climb
	ǐŋgǐ	axe
	álàgà	madness
kp	kpáꜜá	wind around
	kpǐkpǐkǐ	owl
	ákpã́nã́kpá	hawk
gb	ŋgbé	pick
	àgbà	bite
	ègbèlègbè	(a type of) masquerade
	ɔ́ŋgbĩ́ã̀	robe
dʒ	ádʒáɾé	rust
	îdʒá	cooking pot
	ǐdʒè	death
	ǐdʒùlè	father
ɓ	ɓáá	kill
	ɓátà	faeces
	ɓéké	fall
	èɓè	carve (vb)
	óɓɔ̃́ɔ̃́	400

m	ĩ̀ima	heavy
	bɔ̀m	child of someone
	lòm	palm fruit
	nɔ̃́m	person
	áfàmà	she said
	ónṹmã́	bush
	ònṹmã́	get out
	ònũ̀mã̀	beat (drum)
n	nɔ̃́m	person
	túúnɔ̃̀	five
	ónṹmã́	bush
	ònṹmã́	get out
	ònũ̀mã̀	beat (drum)
	òkùnã̀	fowl
	ókúnã́	compound
	ákpã́nã́kpá	hawk
ŋ	ŋgbé	pick
	áŋgî	climb
	ìŋgì	axe
	ɔ́ŋgbĩ́ã̀	robe
ɾ	àɾì	catch
	fɛ́ɾɛ́fɛ́ɾɛ́	light
	ìɾì	thread
	kàɾì	curse
	òɾò	hot
	páɾà	leg
	àgáɾà	lizard
	èbèɾè	dog
	ɛ̀wɛ̀ɾɛ̀	children
	ɛ́wɛ̀ɾɛ̀	spirit
	òfòɾò	blow with mouth
f	fàà	say
	fɛ́ɾɛ́fɛ́ɾɛ́	light
	áfàmà	she said
	òfòɾò	blow with mouth
v	válà	sail
	èvè	throw
s	sã̀ã̀	urine
	sĩ̀ĩ̀	sink
	sũ̀ũ̀	stretch
	ɔ́sɔ́ɾɔ́	navel
l	lɛ̀lɛ̀	sell
	lòm	palm fruit
	àálà	the woman
	álàgà	madness
	àlĩlá	plate
	ɔ́lɔ̀tí	old person
	válà	sail
	èpèlè	trick
	ìdʒùlè	father
	ègbèlègbè	(a type of) masquerade
w	wá	we
	àwò	young person, child
	ɛ̀wɛ̀ɾɛ̀	children
	ɛ́wɛ̀ɾɛ̀	spirit
	ɛ́w̃ɛ̃́	steal
j	jáà	thing
	jɛ̀ɛ̀	bird
	bã́jã́	catarrh
	ĩ̀jĩ̀jã̀	horse
	ɔ́jɔ̀	eye
	àtã́jã̀	good thing
	ètẽ́jã̀	possessions
i	íbà	yesterday
	ídʒá	cooking pot

	ìbà	it is done (of food)
	ìbì	sleep
	ìbò	big
	ìdò	farm
	ìdʒè	death
	ìdʒùlè	father
	ĩima	heavy
	ìŋgì	axe
	ìtà	take
	ìɾì	thread
	áŋgî	climb
	àɾì	catch
	bìè	ask
	bìɔ̀	river
	kàɾì	curse
	kpìkpìkì	owl
î	ì̃jì̃jã̀	horse
	sì̃ì	sink
	ɔ́ŋgbĩ́ã̀	robe
e	èbè	pig
	èbèɾè	dog
	èè	eat
	ègbèlègbè	(a type of) masquerade
	èpèlè	trick
	ètẽ́jã̀	possessions
	èvè	throw
	èɓè	carve (vb)
	ádʒáɾé	rust
	bìè	ask
	ɓéké	fall
	ìdʒè	death
	ìdʒùlè	father
	kéé	cut
	ŋgbé	pick
	pèkù	bat
	tèè	play
ẽ	dẽ́ẽ́	chief
	ètẽ́jã̀	possessions
	ɛ́w̃ɛ̃́	steal
ɛ	ɛ̀dɛ̀lɛ̀	vulture
	ɛ́w̃ɛ̃́	steal
	ɛ́wɛ̀ɾɛ̀	spirit
	ɛ̀wɛ̀ɾɛ̀	children
	fɛ́ɾɛ́fɛ́ɾɛ́	light
	jɛ̀ɛ̀	bird
	lɛ̀lɛ̀	sell
a	ádʒáɾé	rust
	áfàmà	she said
	ákpã́nã́kpá	hawk
	álàgà	madness
	ápá	wing
	àálà	the woman
	àgáɾà	lizard
	àgbà	bite
	àkàdà	pepper
	àl̃ilá	plate
	àwò	young person, child
	àɾì	catch
	bàà	day
	ɓáá	kill
	ɓátà	faeces
	dàá	father
	fàà	say
	îbà	yesterday
	îdʒá	cooking pot
	ìbà	it is done (of food)
	ĩima	heavy
	ìtà	take
	jáà	thing
	káà	hand
	kàɾì	curse

	kpá↓á	wind around
	páɾà	leg
	táátó	three
	tàà	wife
	válà	sail
	wá	we
	ókúnã́	compound
	ónṹmã́	bush
	òkùnã̀	fowl
	ònṹmã́	get out
	ònũ̀mã̀	beat (drum)
ã	ákpã́nã́kpá	hawk
	àtã́jã̀	good thing
	bã́jã́	catarrh
	ètẽ́jã̀	possessions
	ĩ̀jĩ̀jã̀	horse
	ókúnã́	compound
	ɔ́ŋgbĩ́ã̀	robe
	ónṹmã́	bush
	òkùnã̀	fowl
	ònũ̀mã̀	beat (drum)
	ònṹmã́	get out
	sã̀ã̀	urine
ɔ	ɔ̀ɔ̀	roast
	ɔ́jɔ̀	eye
	ɔ́lɔ̀tĩ́	old person
	ɔ́sɔ́ɾɔ́	navel
	ɔ́ŋgbĩ́ã̀	robe
	bìɔ̀	river
	bɔ̀m	child of someone
	ìtɔ̀	stomach
o	ób↓ú	door
	ókúnã́	compound
	ónṹmã́	bush
	óɓṍṍ	400
	òbò	back
	òfòɾò	blow with mouth
	òkùnã̀	fowl
	ònũ̀mã̀	beat (drum)
	ònṹmã́	get out
	oo	salt
	òɾò	hot
	gògó	namesake
	ìbò	big
	ìdò	farm
	lòm	palm fruit
	táátó	three
õ	óɓṍṍ	400
	nṍm	person
u	ìdʒùlè	father
	ób↓ú	door
	ókúnã́	compound
	ónṹmã́	bush
	òkùnã̀	fowl
	ònṹmã́	get out
	ònũ̀mã̀	beat (drum)
	pèkù	bat
	túó tùò	sing a song
	tùò	song
	túúnɔ̃̀	five
ũ	bṹṹ	flowing tide
	sũ̀ũ̀	stretch