Fieldwork Studies of Targeted Languages

UCLA WPP Volume 84
June 1993

UCLA WORKING PAPERS IN PHONETICS, VOL. 84

June, 1993

FIELDWORK STUDIES OF TARGETED LANGUAGES

Table of Contents

Peter Ladefoged	Linguistic Phonetic Fieldwork: a Practical Guide	1
Ian Maddieson Siniša Spajić, Bonny Sands, Peter Ladefoged	Phonetic Structures of Dahalo	25
Bonny Sands, Ian Maddieson, Peter Ladefoged	The Phonetic Structures of Hadza	67
Michael Shalev, Peter Ladefoged, Peri Bhaskararao	Phonetics of Toda	89
Barbara Blankenship, Peter Ladefoged, Peri Bhaskararao, Nichumeno Chase	Phonetic Structures of Khonoma Angami	127
Joyce McDonough, Peter Ladefoged, Helen George	Navajo Vowels and Phonetic Universal Tendencies	143
Joyce McDonough, Peter Ladefoged	Navajo Stops	151
Kenneth deJong, Joyce McDonough	Tone and Tonogenesis in Navajo	165

The major portion of support for the present work is from
NSF grant #SBR-9107004

Introduction

The papers in this issue of *Working Papers in Phonetics* form a preliminary report on the results of our first year and a half's work on an NSF supported project for recording the phonetic structures of languages of special phonetic interest that are spoken by limited numbers of speakers. We are very conscious of the fact that every language that dies represents a significant loss to human culture, and a small part of this loss consists of the disappearance of this language as a resource for the scientific study of language. Our notions of what constitutes a possible human language are greatly influenced by what we know to occur in actual human languages, and many linguistic concepts are shaped by ideas of what is normal or unmarked in a language. This is as true of phonetics as it is of other subfields of linguistics. Yet we know so very little about the phonetic patterns of the large majority of the world's languages. In a few decades, or sooner, the opportunity to study many of them will no longer be available. We are pleased that the National Science Foundation has made it possible for us to document the major phonetic features of a number of the languages that are endangered (or, at least, will not be spoken in the same way) in the comparatively near future. It is important from a general linguistic point of view to salvage as much knowledge as we can concerning a wide variety of languages.

Within this project, our aims include making high quality recordings of all the principal sound types occurring in the languages, and of controlled sets of examples to illustrate some of the prosodic features and syntagmatic phonetic patterns. The recordings are available to anyone who wishes, so that they also may use them in quantified acoustic analyses of segmental contrasts, and in studies of pitch and duration patterns and coarticulatory adjustments between adjacent segments. We have supplemented the recorded data by physiological investigations of those sounds that merit description in greater detail. All the recordings are fully documented not only by provision of the text and transcription of the material recorded, but also by interpretative comments on the instrumental records. Linguists can manage with written records of dead languages for many aspects of their work. But in phonology and phonetics we must build our theories on explicit accounts of what people do, and what they sound like. Since many of these theories are built on notions of universality, it is incumbent on us to provide material for as many languages as possible. If this is not done soon, the number of languages on which our ideas can be based will be severely limited.

We are, of course, aware that not all the languages discussed are dying. Navajo is a vital language, used by thousands of first language speakers in their daily lives. Even very much smaller communities, such as the Hadza, are successfully maintaining their linguistic identity at this time. In deference to this, we have called this collection 'Fieldwork studies of targeted languages.' We have included the papers on Navajo as the challenges we faced in collecting and describing the data were very similar to those facing us in our work elsewhere. Most of the work on Navajo was supported by other funds, but research on this language, which has never been systematically investigated by phoneticians, is very much in the spirit of the NSF project.

The studies in this volume are presented in the order in which the fieldwork was undertaken. The first, on Dahalo, a Cushitic language spoken in Kenya, is largely a qualitative account of the phonetics and phonology of this language. One of the most valuable parts of this study is the large lexicon, accurately transcribed on the basis of observations of six speakers. In addition this study establishes the importance of this little known (and clearly vanishing) language as showing the way in which clicks can occur not only as borrowed sounds in languages within the Bantu family such as Zulu and Xhosa, but also within a Cushitic language. We hope that the material we have collected will assist in determining whether the clicks in Dahalo are also borrowed, and if so from where. Considering the rarity and distinctive nature of clicks, the oddity of their appearing in diverse languages is an important consideration in the study of the nature of language.

The second paper outlines the phonetics and phonology of Hadza, a language spoken in Tanzania whose family affiliation is not yet clear. The fieldwork on this language provided a basis for subsequent fieldwork by the first author, Bonny Sands, who returned to Tanzania for a more extended visit to gain material for her Ph.D. dissertation. Hadza is also a click language, and Sands' dissertation will provide more solid evidence of the relation (or lack of it) between this language and the Khoisan languages. Again part of the value of the paper presented here is that it provides more data on a range of click sounds that had not been previously studied by phoneticians.

The third paper, on Toda, a Dravidian language, is very different. The structure of Toda has been well described by Emeneau (see paper for references), which made our data gathering task much easier. But this also resulted in our being able to collect a great deal of phonetically interesting data, so that the paper on Toda examines only part of the data we have available. A second paper which is in preparation will examine the remainder of our corpus, which includes additional palatographic data and as yet unanalyzed recordings of six male and six female speakers.

The next paper is also on a language spoken in India, but from a different language family. Angami is a Tibeto-Burman language, with a smaller range of segments than Toda, but with some sounds — word initial voiceless aspirated nasals — that have not been reported in any other language. We were considerably aided in our work by being able to obtain the collaboration of Nichumeno Chase, who was writing a grammar of this language. In addition, we were pleased to have the assistance of P. Bhaskararao, one of India's leading phoneticians, who also played a major role in the research on Toda.

The final papers in this collection are on Navajo, which, as we have noted, is not a dying language. These papers use similar techniques and provide comparable results to the studies on other languages in this volume. We were very surprised to find that there were vitually no phonetic studies of this vital and easily available language. We now have good data on 15 native speakers, and hope to produce a number of substantial analyses.

All the papers in this volume (except the first) are being submitted for publication to a journal in the appropriate language area. (The first paper, a general introduction to phonetic fieldwork, has been accepted for publication in a slightly modified form in *A Handbook of Phonetics* (edited by W. Hardcastle and J. Laver), Oxford: Blackwell, 1994.) We hope that the general phonetic implications of these studies will become known to phoneticians by their dissemination in the present form. We consider it important for phonetic descriptions to be readily available to scholars concerned with individual languages, so we will not be submitting any of these papers to the more specialized phonetic journals.

We should conclude by emphasizing that we are well aware that much more remains to be done on all the languages described here. But in each case the work reported constitutes substantial phonetic data on six or more speakers, providing an account of a small but reasonable sample of the population speaking the language. This is something that has been lacking in much previous phonetic work (including some of our own). We would be grateful for any comments or criticisms of this work. We regard it as one of the most important (and challenging) tasks that we have undertaken, and would like to be able to produce work that the whole linguistic community can feel proud of.

Peter Ladefoged and Ian Maddieson

Linguistic phonetic fieldwork: a practical guide

Peter Ladefoged
Phonetics Laboratory, Linguistics Department, UCLA, Los Angeles, CA 90024-1543

Abstract

Gathering data for a description of the phonetic structures of a language involves first determining what these structures are by closely examining the phonology of the language. A list must be prepared illustrating all the phonological contrasts in appropriate ways. Phonetic data must then be collected from at least six people with good speaking voices. Data collection should include not only high quality audio recordings, but also aerodynamic, palatographic and electrolaryngographic records, when relevant. Computerized procedures for recording and calibrating such data are discussed.

Introduction

There are remarkably few extensive phonetic descriptions of languages. A few well known languages, such as English, French, and Japanese have recently become well documented from the phonetic point of view, because of the needs of speech synthesis systems. But in places where there is no commercial need for detailed phonetic knowledge, often only the most prominent phonetic features of the language have been described. There have been no comprehensive phonetic investigations of many major languages, such as Navajo (and, indeed, virtually all American Indian languages), Swahili, Malayalam, and hundreds of others. The lack of phonetic work is especially unfortunate in the case of less prominent languages that are no longer being spoken as a first language, and will not be available to be recorded in the near future. Hale, Krauss, Watahomigie, Yamamoto, Craig, Jeanne, & England (1992) estimate that over half the 6,000 languages currently spoken in the world will not be spoken by the end of the next century. This paper will consider what kind of phonetic fieldwork must be undertaken in order to record a substantial part of the phonetic structures of such languages.

Most phonetic fieldwork differs from typical linguistic fieldwork not only by being more limited in scope, but also by being more limited in time. Linguists going out to examine a single language are usually prepared to spend several months on the process. But with adequate preparation phoneticians can record the phonetic structures with only a comparatively short time in the field. There are four basic tasks in making a description of the phonetic structures of a language. First one must decide what to describe; second, suitable speakers must be found; third (the major concern of this paper), the necessary phonetic data must be recorded and analyzed. Lastly the results must be written up in a coherent whole, a topic that will not be discussed in this paper, but which should always be held in mind. For ease of exposition we will consider each of the first three tasks separately. In practice, all these components overlap (especially the writing, which should be carried out throughout the fieldwork; it should never be left till the investigation is completed). We will focus principally on aspects of phonetic fieldwork that have not been described in readily available literature, and have little to say about well known techniques such as spectrographic analysis and other laboratory techniques that are familiar to all phoneticians.

1. Linguistic phonetic structures

The first requirement for a phonetic description of a language is a good account of the phonology. Of course, it is also true that in order to describe the phonology one needs to know the phonetics. The two things go hand in hand, providing us with a chicken and egg problem as to which must come first. Obviously they actually evolve together. But it is usually phonological knowledge that precedes detailed phonetic observation. Some people still have a lingering feeling, based on older works such as Pike's (1947) *Phonemics*, that a linguist should go into the field, make detailed phonetic transcriptions of some speakers, and then sort out the sounds into phonemes. But it seldom works out like that. Only rarely does one notice a phonetic difference between two sounds in a corpus, and then examine the corpus to see if they are in contrast.

Most of a linguistic phonetician's work involves describing the sounds of languages that have been the subject of previous linguistic work. On the few occasions when I have myself described a language that has not been the subject of previous linguistic analysis (e.g. Igbirra, as described in Ladefoged 1964), it has clearly not been a case of my observing phonetic differences, and then seeing if they are phonemically contrastive. Instead, in my attempts to learn the language, it has been the native speaker who has pointed out that certain utterances mean different things, usually by commenting on my mispronunciations when I have managed to say one word while attempting to say another. After the phonological contrasts have been observed, then the phonetic differences involved can be described.

When trying to record the phonetic structures of a language, the ideal situation is if the fieldwork can be undertaken together with a linguist who has already analyzed the language. It is often possible to cooperate with local working linguists and their consultants to get phonetic data that they might otherwise disregard. But even if there are no local working linguists, phoneticians can usually prepare themselves reasonably well by reading everything that is available; fieldwork should never be undertaken without previous knowledge of all the literature. There will no doubt still be phonological points to be cleared up when the particular speakers being recorded do not behave as anticipated. But the phonological description of a language must have been fully worked out before an adequate phonetic description can be made. Linguistic phonetics follows phonology rather than preceding it.

This is a different attitude to fieldwork from that taken by some linguists. Kelly and Local (1989), for example, advise fieldworkers to avoid reading anything about a language they are going to investigate. I find this attitude presumptuous. Other people usually know more than I do. Previous linguists may have missed some phonetic points that I might be fortunate enough to note. They may also have described some sounds incorrectly, but that does not hinder me from making my own observations. My own practice is to try to find out as much as I can about a language, reading all the available literature on it and its neighbors.

A complete phonetic description of a language would be one that was sufficient for a computer to turn the output of a phonological description of an utterance into natural sounding speech. Making a description of this kind is an enormous undertaking, which has only recently been attempted for languages such as English. The kind of detailed phonetic description required for synthesizing natural sounding speech is well beyond the scope of the fieldwork we envisage. We will assume that when describing the phonetic structures of a language it is sufficient (at least to begin with) to consider each of the phonologically contrasting items in comparable contexts. This would involve recording the complete set of vowel and consonant contrasts of each language as well as the principal dynamic phonetic patterns.

It would be nice if it were possible to record all the desired items in naturally occurring speech, but this is impossible. There is no way in which one can sit around waiting for half a dozen speakers to use each of a set of words such as "heed," "hid," "head," "had," etc. in spontaneous speech. Accordingly it is necessary to prepare word lists. The form of the word list developed will vary from language to language, depending on particular features of the language. The general structure of the basic word list should be to place all the vowels and all the consonants in matched environments so that contextual influences are equalized. It is often appropriate to use words illustrating consonants initially before a low central vowel such as [a], and at least two sets of words illustrating vowels, one following a dental or alveolar stop, and one following a labial stop. Supplementary word lists are also needed to look at additional contextual variation in vowels, the relative durations of segments, and features such as the spreading of nasality from [+nasal] segments. Lexical suprasegmental features must be recorded. Tone and stress contrasts are usually fairly easy to illustrate in citation forms, but one has to ensure that the variation that occurs in different prosodic domains is also captured. Therefore phrases or sentences illustrating features such as tonal sandhi and stress shifts must be included.

It cannot be emphasized too strongly that the preparation of good word lists is the key to linguistic phonetic fieldwork. However much one prepares one inevitably finds that on returning from the field and doing the analysis of the data some things have been omitted or not adequately covered. I am often only too aware that I should have recorded additional material further illustrating contrasts or allophonic distributions that I had not considered important when out in the field. But adequate preparation can avoid many pitfalls. One cannot spend too much time checking word lists and making sure that all the contrasts and all the interesting allophonic distributions are included. I find it useful to make consonant, vowel and suprasegmental charts, and words illustrating each item (much as the "Illustrations" of languages in the *Journal of the International Phonetic Association*) even before going out into the field. When the time comes to work with speakers of the language, it often turns out that they do not know some of the words, or pronounce them differently. But it pays to have thought the whole process through before beginning work with native speakers.

Citation forms and carefully controlled sets of phrases and sentences provide the best overall view of the phonetic structures in a language, but they do not provide sufficient information on the rhythm of the language, and on larger units such as contrasting intonation patterns. In an effort to record at least some material that would illustrate these properties, a short narrative or other material containing longer stretches of speech should be included. This might be a local folk tale, a short autobiographical narrative, or perhaps something more standardized. We have found that a suitable way of recording a short text is to ask speakers to repeat a story that has been told to them. If the story is illustrated by a series of pictures, then the pictures can be used as prompts in the subsequent re-telling.

2. Linguistic consultants

Probably the questions asked most often by people considering doing phonetic fieldwork is "How many speakers do I need, and how do I find them?" There are no set answers to these questions. It is clear that it is no longer appropriate to present data based on instrumental records of only a single speaker, as some misguided phoneticians have done in the past (Ladefoged 1968). But beyond that it depends on the language and the local circumstances. One wants a group of people who all speak in the same way, so that it is a single dialect of a language that is being described, but who do not all belong to the same family, so that there is sufficient diversity for any familial idiosyncrasies to be noticeable. For quantitative phonetic work an absolute minimum of three people of each sex is essential; and it is definitely preferable to have at least half a dozen women and men. Of course, in the case of languages with few native speakers, or when the field worker has limited time in the field, it may not be possible to get sufficient numbers; but it is difficult to provide meaningful measurements without a sample of at least six people of each sex.

The best speakers for phonetic research are usually not the old people who know the language well. These people may be the most helpful for finding appropriate illustrative words; but even when eliciting phonological contrasts, they can be troublesome because they have their own agenda. They are often eager to tell stories or sing songs, and they nearly always want to make sure that all the details of the meaning are understood, and that notice is taken of all the homophones of each word. These points may be useful and fascinating for other linguists, but usually all the phonetician wants is to be sure that the gloss has been noted faithfully enough to allow another linguist to be able to repeat the data gathering by eliciting the same form. Once the material to be recorded has been established, it is preferable to use younger speakers, with stronger voices (and all their teeth). School children in their teens often make excellent consultants; they enjoy playing teacher and correcting the fieldworker's mistakes.

3. Instrumental techniques

The major concern of this paper is the instrumental phonetic techniques that are needed for a good description of the phonetic structures. The use of phonetic instrumentation in the field is comparatively recent. There is a story about Daniel Jones, the great British phonetician who dominated the field in the first half of this century. When he was about to go off on a field trip someone asked him what instruments he was going to take with him. He pointed to his ears and said: "Only these." It is surely true that by far the most valuable assets a phonetician can have are a trained set of ears. It is also true (and Daniel Jones would certainly have agreed) that the ears should be coupled to highly trained vocal organs that are capable of producing a wide range of sounds. There is no substitute for the ability to hear small distinctions in sounds. There is also no substitute for the ability to pronounce alternative possibilities, so that one can ask a speaker which of two pronunciations sounds better. One of the most efficient procedures for getting results in the field is to test different hypotheses by trying out various vocal gestures of one's own.

Nevertheless, however well trained they might be, phoneticians who now go out with only their ears and their own vocal apparatus are doing themselves a dis-service. There are three ways in which instrumental aids can be valuable supplements to the field phonetician. Firstly they can sometimes suggest new descriptive possibilities; we have, for example, learned a number of facts about the unusual voiceless nasals that occur in the Tibeto-Burman language Angami by observing records of the airflow from the mouth and the nose (Bhaskararao and Ladefoged, 1991). Secondly instruments allow permanent records to be made so that one can demonstrate the facts to those who do not have access to speakers of the language being described. Thus readers might believe descriptions of the curious consonant clusters that occur in Khoisan languages without any instrumental evidence; but it is nice to have records so that they can see for themselves (Ladefoged and Traill, 1993. Thirdly, instruments enable one to make quantitative descriptions. However good one's ears one cannot, for example, measure and report the duration in milliseconds of the aspiration of stops in different contexts; and without measurements one cannot prove that there is a statistically significant difference between one group of stops and another, or between the sounds of one language and another.

3.1 Acoustic data

For almost half a century the most important instrument available to the phonetician has been the tape recorder. Before World War II phoneticians had to rely on their notes and their skill in making phonetic transcriptions. But since that time few phoneticians would start a fieldwork investigation of a language without having some means of making recordings. We will begin this section with a discussion of techniques for making recordings.

Recordings can be like audible notebook jottings, or like more formal finished papers that demonstrate the sounds of a language. In either case a recording should *always* begin with a statement of the date, the place, the names of the speakers, and the material. These things will no doubt be written on a label; but labels can come off and tapes can easily get put into the wrong box, and become virtually useless unless they have been properly identified. When recording on a portable tape recorder it is also advisable to record a known signal such as a tuning fork at the beginning and the end of a recording session. This provides a useful check on the speed of the recording.

If it is possible, all the preliminary labeling procedures should be carried out before setting up to record the language consultant. The whole process of making a fieldwork recording should be done with as little fuss and as inconspicuously as possible. (Of course, recordings should not be made surreptitiously. The speakers must be asked if they mind being recorded.) One can often make good recordings unobtrusively by walking in with a tape recorder in one pocket or slung over the shoulder, and a microphone in another pocket, already connected and ready to be produced at the appropriate moment.

Tape recordings should be of the highest quality possible. Interviews sometimes have to take place amidst noises such as background talking and household sounds that the fieldworker can ignore. But on a recording for linguistic phonetic purposes anything other than the voice of the language consultant can be a distraction. We have found that when trying to make a high quality recording it is often a good idea to go off outside. But this may prove less convenient on windy days. It is possible to protect the microphone itself from the effect of the wind by using a wind-shield, a protective covering of foam rubber or a similar substance. But it is more difficult to deal with the noise of the wind in the trees, although much can be achieved by finding a large enough open space. Both in everyday parlance and in the technical acoustic terminology, the best recordings are often those made in a free field, rather than in a noisy, reverberant room.

The intensity (loudness) of the speech being recorded should be as high as possible without overloading the tape recorder. This requires proper use of the microphone. The intensity varies with the square of the distance between the source of the sound and the microphone. A baby crying ten feet away may be one quarter of the loudness of a speaker who is five feet from the microphone; but the same baby will be only one hundredth of the loudness of a speaker who is one foot from the microphone. If recordings must be made in circumstances where there is a lot of background noise, make sure that the speaker is a close but constant distance away. We have found so-called lip-microphones fastened to a headset to be excellent for recording one speaker at a time. They are usually designed so that they cancel sounds coming from the other side, which makes them suitable for recording even in a slightly noisy room. When recording a group of people, arrange them in a small circle. If you have somebody else to operate the tape recorder and watch the signal level, then you can move around and control the recording by holding the microphone just in front of the speaker who should be speaking. If you are alone, place the microphone on something soft at least a few feet from the tape recorder.

Check the signal level frequently while making the recordings, as speakers often vary their loudness. The signal should always be at the highest permissible level, so as to make sure that it is as far above the background noise as possible. Finally, after making a recording, play it back and verify that you have proper written notes to explain what is on it. As soon as you can, play all your recordings right through, checking them against a complete text. (We try to do this every night, so that we never accumulate data that is not useful).

What sort of machine should be used for making field recordings? DAT (digital) recorders are superior to analog (cassette) recorders; virtually any DAT recorder will be more than adequate for linguistic phonetic purposes. Cassette recorders can vary widely in quality, and it is important that the one chosen should meet certain specifications. Firstly it should have a good frequency response – that is to say it should be able to reproduce the complete range of frequencies recorded with the same relative intensities as when they were originally produced. People in their prime can hear sounds with frequencies around 20,000 Hz, but this ability declines with age. No languages relies on distinctions that can be heard only by young adults, and a frequency range up to 12,000 Hz should be sufficient for linguistic purposes. Any DAT recorder will be able to record this range, but only so-called professional cassette recorders will. Secondly the tape recorder should have a good S/N (signal to noise) ratio. The difference between the maximum signal that can be recorded without overloading and the noise that is present in the absence of any signal should be at least 45 db. DAT recorders have about twice this range, but even professional quality cassette recorders often only just exceed it. Thirdly speed variations, whether of the short term kind known as flutter, or the day to day kind due to variations in the batteries and motors, should be less than 0.1%. This is usually no longer a problem for any recorder, as long as the batteries are not run down. In addition, the tape recorder, like any other piece of apparatus taken into the field, should be tough and reliable. It should be possible to drop it on the floor and kick it around for a few minutes without too much damage. We have heard stories of professional tape recorders that have been fished out of rivers, dried out, and found to be still working. The stories are probably apocryphal, but they provide a good standard to aim for.

Before going into the field, the tape recorder should be thoroughly checked out. If it has rechargeable batteries, make sure that they are not only fully charged, but also are capable of maintaining their charge for as long as is required. In any case, spare batteries are a necessity that should not be forgotten. We have found rechargeable batteries useful when we have a vehicle available, as they can be recharged with the aid of a unit that plugs into the dashboard. On analog recorders (as opposed to DAT machines) the heads should be cleaned and demagnetized (simple procedures that may increase the frequency response by as much as 5,000 Hz), and the frequency response checked. Record frequencies from 50 Hz to 15,000 Hz at about half octave intervals, checking with a meter that they are all being produced by the signal generator at the same level. Then observe on a meter the relative intensities of these signals when they are played back, so as to make sure that the tape recorder really has a full frequency range. A plot of the frequency response of a Sony Professional cassette recorder used on a recent field trip is shown in Figure 1. In addition, check the S/N ratio, and, if the necessary equipment is available, the speed constancy.

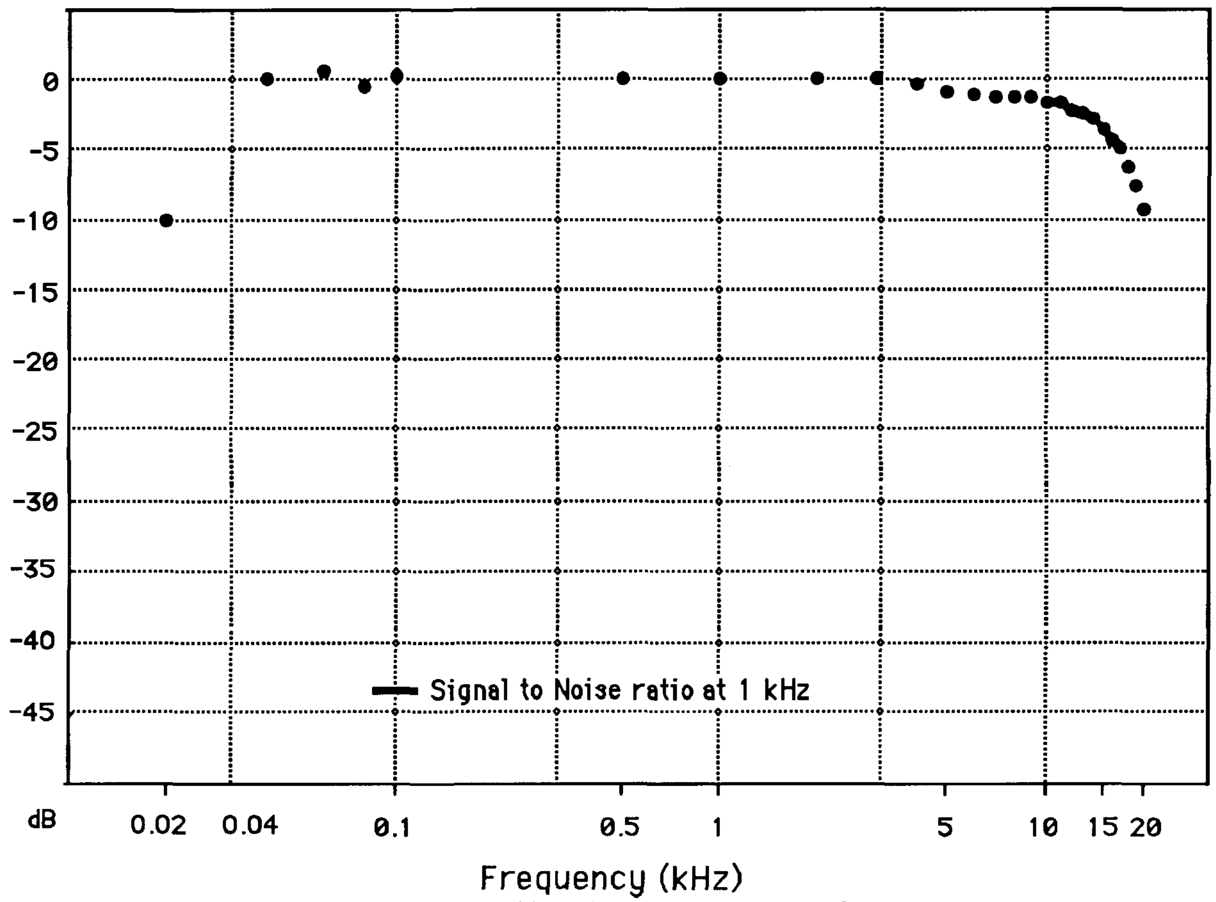


Figure 1. Frequency response and Signal to Noise ratio of a UCLA Sony used in recent fieldwork.

As portable computers become more available, the days of dependence on tape recorders may be passing. Direct recording onto portable computers may be used, with the tape recorder being regarded simply as a backup. As with the tape recorders discussed above, the computer battery can be kept charged with the aid of a unit plugged into a vehicle's dashboard. Even when considered just as devices for reproducing sounds, computers are much more versatile than tape recorders. Fieldworkers want to be able to record word lists or short paragraphs and then to play back selected pieces over and over again, so that they can hear subtle nuances of sounds that are new to them. They also want to be able to hear one sound, and then, immediately afterwards, hear another that may contrast with it. Both tasks can be done somewhat cumbersomely and tediously using tape recorders. But they are trivial, normal operations on any computer equipped with a means for digitizing and editing recorded sounds. Figure 2 is an illustration of a computer screen while using the commercially available program SoundEdit™ with a MacRecorder. Clicking on the icon representing a speaker in either of the windows on the screen will play the utterance in that window. This makes it easy, or, at least, easier to hear the difference between the two words shown in this illustration, exemplifying the contrast between so-called stiff and slack syllables in Bruu, a Mon Khmer language spoken in Thailand. Comparison of the two words is also aided by being able to see from the waveforms that the stop at the beginning of the word in the lower window is slightly more aspirated than the stop in the upper window.

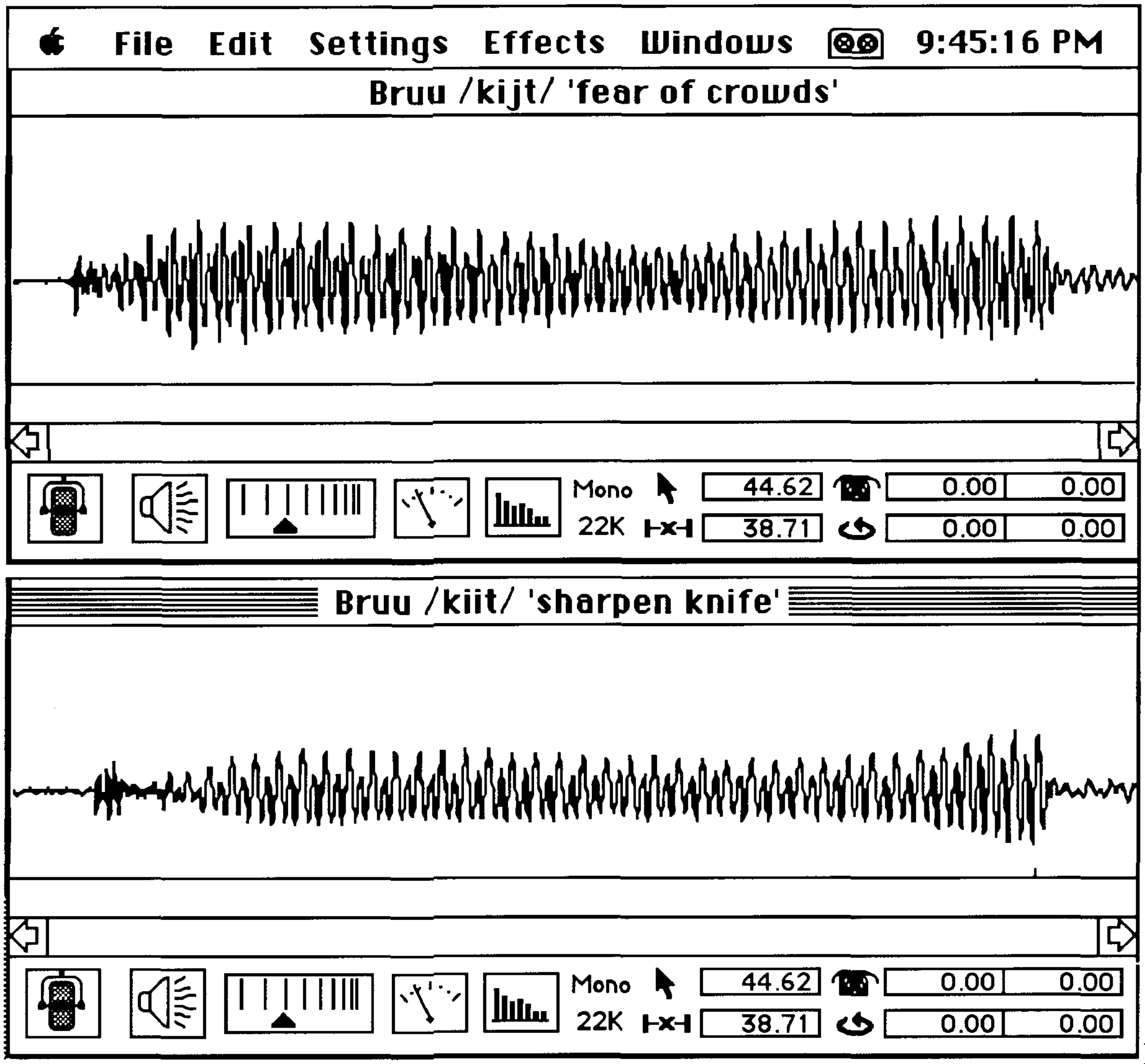


Figure 2. Computer screen showing two words in Bruu, a Mon Khmer language. Either word can be played by clicking on the speaker icon in its window.

It is worth considering at this point some of the specifications of a computer system suitable for use in linguistic fieldwork. Sound waves have to be stored on a computer as a series of numbers (samples) representing the amplitude of the wave at regular intervals. There are two factors affecting the fidelity with which a given wave is stored. The first is the rate at which the amplitude of the wave is sampled. The sample rate must be at least twice the frequency of the highest frequency component in the wave. The second is the accuracy of the representation of the amplitude of each sample. If each amplitude can be represented as any one out of 1024 possible numbers, then the stored wave can be more like the original sound wave than if each amplitude has to be represented as one out of only 256 possible values.

Bearing this in mind, we can now determine the sample rate and sample size required for a computer system that is to be used for phonetic fieldwork. As we noted earlier in discussing tape recordings, a frequency range from 60 Hz to 12,000 Hz is sufficient for most linguistic investigations. In fact, although some speech sounds (particularly voiceless stops and fricatives) may differ significantly in the amplitude of the frequency components in the 10,000 to 20,000 Hz range, the major differences (even in voiceless stops and fricatives) are below 8,000 Hz; and for the analysis of vowels and similar sounds it is preferable to record as digital signals *only* frequencies below 4,000 Hz, so that (as will be discussed below) the recorded frequency range can be analyzed using the smallest possible steps. The implications of these requirements are that it is advisable to be able to record speech onto a computer in two different ways, one with a comparatively high fidelity from a speech point of view, capturing frequencies up to 12,000 Hz (although speech high fidelity could perhaps be regarded as up to only 8,000 Hz), and the other ensuring that only frequencies up to 4,000 Hz were recorded. In order to allow some tolerance in the system, the sample rate should be 2.5 times the highest frequency present. Accordingly, the computer system should be capable of sampling speech at 20-24,000 Hz for high quality listening and analysis, and at 10,000 Hz for the analysis of vowels and similar sounds.

The minimum sample size is more difficult to determine. DAT recorders and compact discs used in high fidelity audio systems specify amplitudes in terms of 32,000 possible levels (16 bits), which allows for the maximum signal level recorded to be 96 dB above the system noise. There is no doubt that a Signal/Noise ratio of this magnitude is highly desirable, but it is fully useful only in studio recordings. In most fieldwork situations, it is difficult to record in an environment in which the background noise is as much as 48 dB below the level of a speaker who is as close as possible to the microphone. There is always a baby crying in the next room, or the wind in the trees, or some other noise that cannot be avoided. Accordingly, when recording onto a computer from a well-made analog recording, 256 possible levels (8 bits) providing a 48 dB Signal/Noise ratio may be sufficient, provided that a great deal of care is taken to ensure that the full range is used (i.e. that the original signal is always re-recorded at the maximum level possible without overloading). In linguistic fieldwork, in which citation forms or specific sentences are produced in controlled circumstances, it is often possible to ensure that a maximum signal is recorded. In these situations, 8 bit systems providing a 48 dB Signal/Noise ratio (such as the MacRecorder used by the SoundEdit program illustrated above) are satisfactory for preliminary work. As we noted above, 48 dB is about the Signal/Noise ratio of a good laboratory tape recorder. Few portable cassette recorders (even so-called professional models) have a better Signal/Noise ratio. But it is often not possible to get the speaker to maintain a constant level, or multiple speakers may be being recorded, some having strong voices and others not. In these situations, it is essential to have a system with as great a Signal/Noise ratio as possible, making a 16 bit system highly preferable. In any case, the added margin of safety provided by 16 bit systems is always desirable.

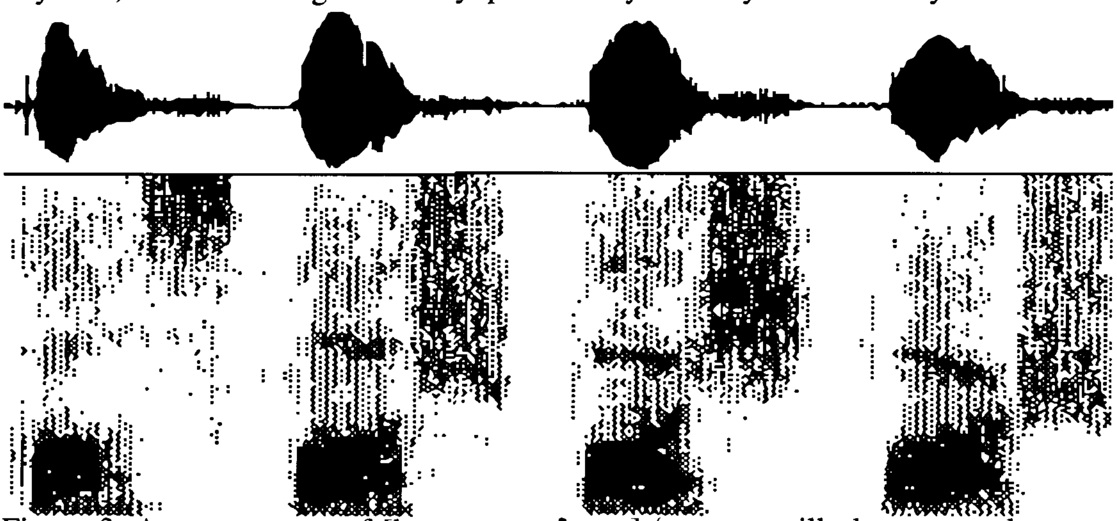


Figure 3. A spectrogram of [kos̪, pos̺, poʃ, poʂ] 'money, milk, language, clan name' in Toda, made under field conditions.

In addition to being useful as a sophisticated playback device, a computer can provide several types of analysis that a fieldworker might find convenient. The most useful display of the general acoustic characteristics of a sound is a spectrogram. Figure 3 shows the kind of spectrogram that can be produced on a portable computer without a color (gray scale) screen. The figure is an unretouched copy of the printout from a battery operated printer used in the field. As portable computer screens and printers improve so that they produce better gray scale or color output, this kind of display will become even more useful.

The display in Figure 3 was created by another commercially available program, Signalyze. Spectrograms generated by this program on a color screen on a laboratory computer are much more impressive. But even the quality of the display in Figure 3 can be very useful to the fieldworker. The words shown illustrate the four contrastive sibilants that occur in Toda, a Dravidian language spoken in the Nilgiri Hills in India. Each of these words ends in a different sibilant. The overall spectral characteristics of these sibilants are evident . The laminal dental sibilant at the end of the first word has the highest frequency, and the retroflex sibilant at the end of the last word has the lowest. The apical alveolar and (laminal) palatoalveolar sibilants at the ends of the second and third word have very similar spectral characteristics. (The lowering of the spectral energy peak at the end of the second word is a non-distinctive feature, being simply due to the closure of the lip for the consonant at the beginning of the next word.) These two sibilants are distinguished primarily by their on-glides. The increasing second formant at the end of the third word is due to *the raising of the blade and front of the tongue for this laminal sound. In the last word, the* lowering of the third formant is probably due to the sublingual cavity that is formed by raising the tip of the tongue for this retroflex sibilant. A great deal of information can be obtained even from these low quality spectrograms, actually produced under field conditions and computer cut and pasted into this paper. Of course, still more information can be obtained from high quality spectrograms produced by this or another program on a laboratory computer at a later date.

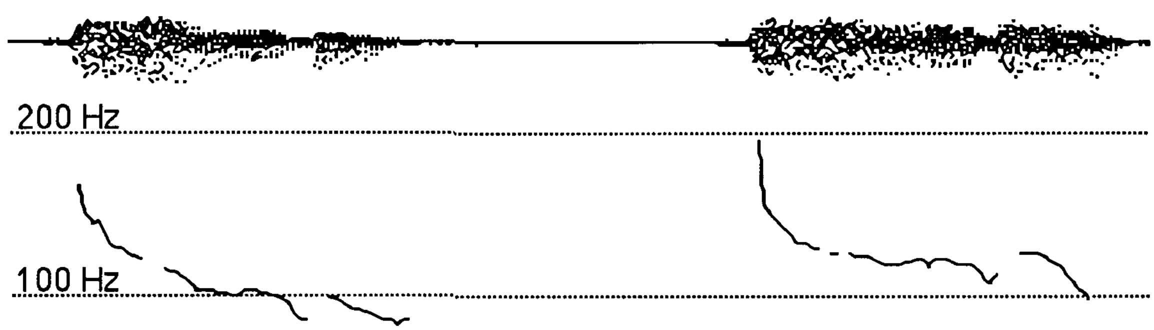


Figure 4. The tonal contrast between /ku!laamba/ 'to lick', and /kulaamba/ 'to be dear, in Sukuma, a Bantu language spoken in Tanzania.

Another kind of analysis that is very useful to the fieldworker is one that indicates the pitch. The Signalyze program discussed above will generate good displays of the fundamental frequency (and it will produce narrow band spectrograms, which are sometimes even more useful for pitch analysis when a creaky voice quality or other unusual spectral characteristics are involved). But a number of other programs will also provide similar information. Figure 4 shows the fundamental frequency in a pair of words with contrasting tones in Sukuma, a Bantu language, as analyzed by a public domain modification of SoundWave, written at the University of Uppsala, Sweden.

The final kind of computer analysis of speech sounds that will be illustrated here is one for determining the formant frequencies, the principal aspects of vowel quality. A common way of obtaining formant frequencies is by inspection and peak picking using superimposed LPC and FFT displays of the kind shown in Figure 5. The original display was also produced by the Uppsala version of SoundWave mentioned above (slightly further modified at UCLA). The figure shows analyses of two Assamese (Indo-Aryan) vowels that are very similar in quality; one is between the IPA reference vowels [ɔ] and [o], and the other is a vowel with a tongue position like that in [ɑ], but a lip position more like that in [u]. In each of these two vowels, there are two formants close together; but the more rounded vowel in the lower part of the figure has a very sharp decrease in spectral energy immediately above the second formant. The third formant is not clear in either of these vowels.

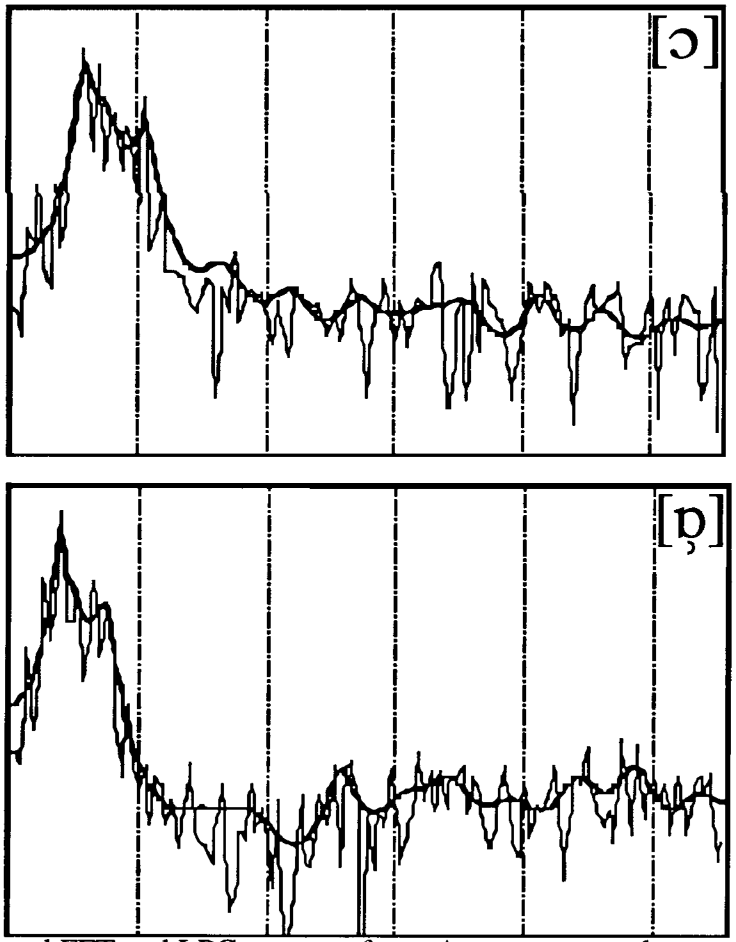


Figure 5. Superimposed FFT and LPC spectra of two Assamese vowels.

When making an FFT it is important to remember the system limitations. In effect, an FFT provides the amplitudes of the spectral components that are present on the assumption that these components are all multiples of a wave with a frequency depending on the number of points in the FFT. The greater the number of points in the FFT, the longer the wave length, thus the lower the frequency of this wave, and the smaller the interval between calculated components. But any program calculating an FFT will have a certain maximum number of points permissible (usually something like 512 or 1024). Accordingly, the only way to further increase the accuracy in the frequency domain (i.e. to decrease the interval between measured components) is to *decrease* the sample rate. This will have the effect of decreasing the range of frequencies that can be observed. But, as the same number of components will be calculated within that range, they will be closer together. Given a 512 point FFT and a sample rate of 20,000 Hz, there will be 256 components spaced about 40 Hz apart in the range up to 10,000 Hz. But if the sample rate is reduced to 10,000 Hz, the 256 components will be spaced about 20 Hz apart in the range up to 5,000 Hz. It was for this reason that it was suggested earlier that if vowel formants were being studied it is advisable to use a lower sampling rate. The alternative would be to use an FFT with a larger number of points, but no analysis system will permit the maximum number of points to be increased beyond some fixed limit.

3.3 Aerodynamic data

Acoustic analyses made from good quality tape recordings or direct recording onto a computer can provide large amounts of data. But they often do not indicate in an unambiguous way important articulatory facts such as the direction of the airstream or the timing of movements of the vocal organs, particularly those during voiceless closures. The best way of gaining information on these phonetic parameters is by recording a number of aerodynamic parameters. Records of the pressure of the air in the mouth and of the airflow from the nose and the lips have been used for many years in phonetic research, dating back, in an unquantified way, to the kymograph tracings of Rousselot (1924-5), Scripture (1923) and other early experimental phoneticians. For many years we have been able to make good, calibrated, records of these variables (Ladefoged 1967). Now this ability is available to the field phonetician.

The principal aerodynamic parameters with which we will be concerned are: (1) the pressure of the air in the mouth behind any bilabial closure; (2) the pressure of the air in the pharynx, behind any alveolar, velar or uvular closure; (3) the subglottal pressure; (4) the flow of air in and out of the nose; (5) the flow of air in and out of the mouth.

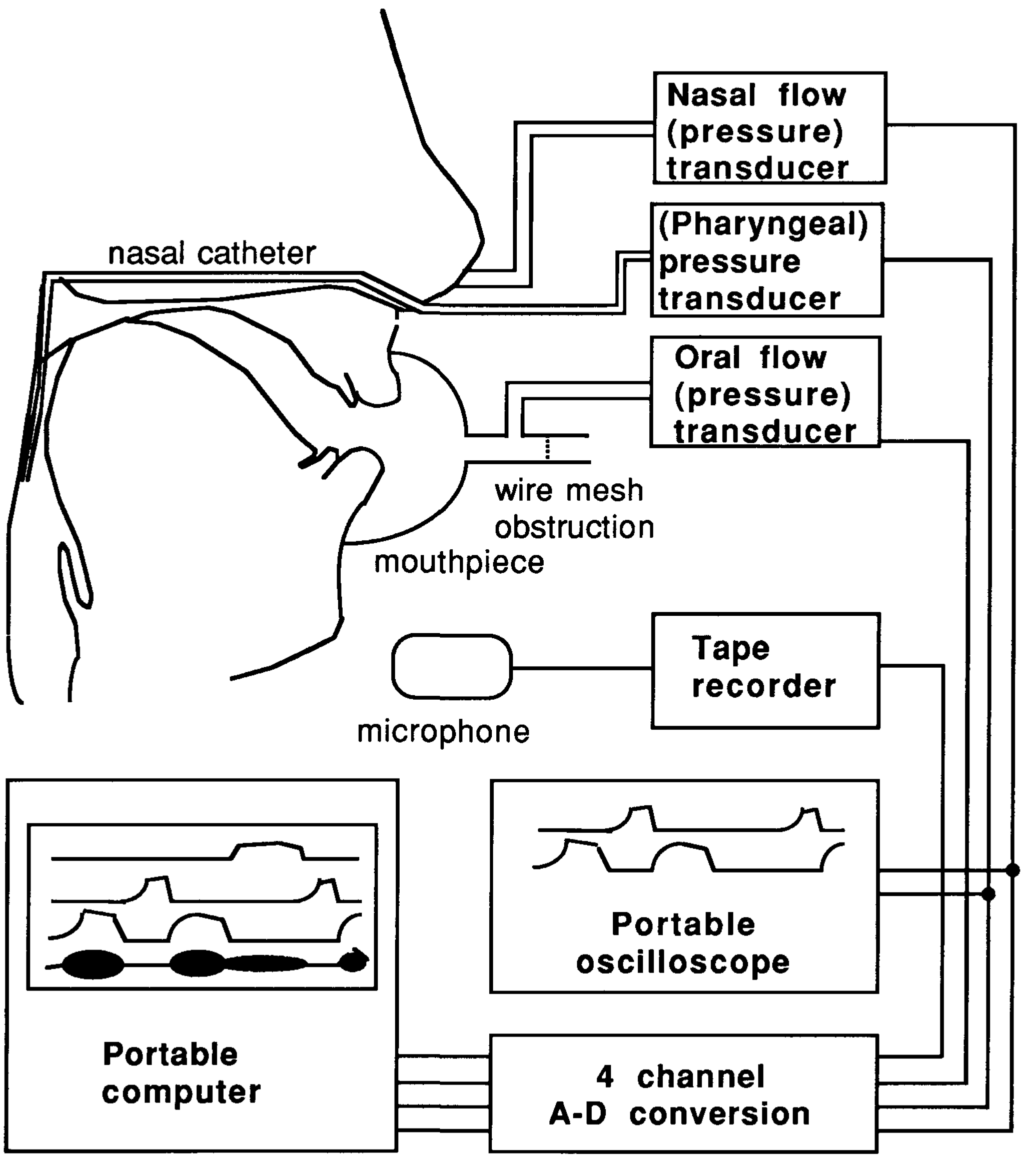


Figure 6. Apparatus for obtaining aerodynamic records in the field.

Examples of the use of some of these parameters for elucidating articulatory descriptions of West African languages have been given by Ladefoged (1968). In these examples the data were recorded live by subjects in the laboratory using a four channel inkwriter. Nowadays it is possible to record such data in the field using a portable computer. One type of system for use in the field is shown in Figure 6. It will record the audio signal and up to three physiological signals. Typically these include either the pressure of the air in the pharynx obtained by passing a tube through the nose, or the pressure of the air in the mouth using a more convenient tube between the lips, and the oral and nasal air flow. In investigations of prosodic features it is sometimes appropriate to record an approximation to the subglottal pressure by means of a tube with a small balloon on the end of it in the esophagus.

We will begin our discussion of a system for recording aerodynamic parameters in the field by noting some general requirements of such a system. It should obviously be light weight, compact, and battery operated. As with the tape recorder, it should be rugged enough to withstand the pounding it is likely to get while being transported into places where there are no roads. The air pressure devices should be capable of measuring $\pm$ 25 cm H_2O. The air flow devices should be constructed so that they measure ingressive and egressive air flow rates up to 2 litres/second. They should have a flat frequency response, so that they do not give different readings for sounds with identical flow rates said on different pitches.

There are also a number of practical points to take into consideration when making pressure and flow recordings. The tubes used should be short and as thick as the speaker can conveniently tolerate, so that they have a high enough frequency response to show some voicing vibrations. Long thin tubes act as acoustic filters and cut out the small pressure variations associated with voicing. Generally, tubes should be sealed at the end, with small holes at the sides near the tip to let the pressure in. Tubes that are open at the end often become full of mucous. We have found infant feeding tubes, size 12 French, to be suitable.

When recording the pressure of the air in the mouth one end of the tube should be connected to the pressure transducer, and the other end should be held (by the speaker) so that it is just behind the lips. The speaker should be discouraged from sucking on the end of the oral tube, as this leads to the tube becoming full of saliva, which will lower its frequency response.

The pharyngeal pressure tube should be inserted through the nose so that its open end rests on the back wall of the pharynx about 1 cm below the uvula. The speaker should first be given a practical demonstration by the field-worker of how easy it is to pass a tube through the nose into the pharynx. Take a clean, sterile, tube. (Tubes should be cleaned with disinfectant and boiled for 30 minutes if they have to be reused.) Hold it about 12 cm from the tip, and moisten it with saliva. If it has a slight natural curve, make sure that this curve is pointing downwards. Then, still holding the tube about 12 cm from the end, push it straight (i.e. horizontally, not upward) back into the nose until some obstruction is reached. It is advisable to be fairly rapid about this part of the proceedings. Many people find the most difficult part of the process to be the irritation that may occur while they are tentatively pushing just inside the nostril. Once an obstruction has been reached, take a mouthful of water, and while pushing the tube gently further in, swallow a small quantity. Keep pushing gently, swallowing, and breathing in through the nose until the tube has passed over the top of the velum into the pharynx. A large ingressive breath will cause the velum to lower more, and may help the tube go round and into the mouth. If there is any difficulty in getting the tube to pass round the velum, twist it slightly, first one way then another, all while swallowing or taking a deep breath through the nose, and gently pushing. Remember to hold your hand high above the level of the nose, so that the tube is never pushed upward, but always goes straight back (and, hopefully, down). Nearly everybody finds it easier to pass the tube through one nostril rather than the other, so if at first you don't succeed, try again with the other one.

For an average male adult about 15 cm of the tube should be inserted into the nose in order to locate the open end properly. It is a good idea to make a mark by putting a thread around the tube at slightly more than this distance before beginning the experiment. As soon as the tube is located properly in the pharynx the speaker should relax quietly for a moment in order to avoid a gag reflex. Check the location of the tube by using a flashlight (a small, focused pen light) while looking into the speaker's mouth. It may be necessary to use some object such as a spoon or a spatula to depress the speaker's tongue in order to see the end of the tube. If it is correctly located about 1 cm below the uvula, fasten the tube to the cheek just outside the nose with a small sticking plaster. When recording pharyngeal pressure through a nasal catheter, it is necessary to keep the tube free from mucus. Before attaching the end to the pressure transducer, connect it to a rubber bulb and blow some air through it. Repeat this procedure at frequent intervals throughout the recording.

A similar procedure can be used for passing a tube into the esophagus so as to estimate the subglottal pressure. In short phrases the esophageal pressure is a very good indication of the pressure below the vocal cords; but in longer sentences there will be a shift of the baseline upwards due to the changes in the relaxation pressure of the air in the lungs. In any case it is usually possible to get a fairly precise measure of the relation between the pressure recorded in the esophagus and that in the trachea by recording the pressure in the mouth with a tube between the lips during the pronunciation of voiceless bilabial stops. During a voiceless stop the pressure in the mouth will be the same as that in the trachea, so the esophageal pressure record can be calibrated from the oral pressure record. For recording subglottal pressure it is necessary to pass 30 - 35 cm of tube through the nose. Put markers on the tube at these distances before beginning, so that the position of the end can be estimated. A tube that is to be passed into the esophagus should be sealed by having a small balloon about 1 cm on the end of it. The cut off end of a condom can be used for this purpose, tying it on to the end of the tube with thread. When the tube is in the esophagus the balloon end should be inflated slightly.

In the system we have been describing, the pressure in the esophagus, or the pharynx, or behind the lips, has been recorded through tubes connected to pressure sensors mounted on the apparatus outside the mouth. An alternative technique is to use miniature pressure transducers which can be placed at the ends of the tubes passed through the nose into the esophagus or the pharynx or between the lips into the oral cavity. Miniature transducers are more convenient in many ways, particularly for recording esophageal or pharyngeal pressures; but they require more complex electronics, are more expensive, and have to be treated with more care than fieldwork conditions often allow. In general, for recording gross changes of air pressure of the kind that are associated with articulatory movements, externally mounted transducers are sufficient.

Pressure recordings may be calibrated using a water manometer, or more simply (and only slightly less accurately) by connecting one arm of a tube in the form of a T piece to the pressure transducer, and the center of the T to a tube attached to a ruler in a container of water as shown in Figure 7. Blow into the other arm of the T until bubbles come out of the tube. At this time the pressure exerted on the transducer will be equivalent to the depth of the open end of the tube in the water. Repeat this procedure with varying depths, so that signals are produced over the entire range of interest (probably up to 15 cm H_2O). When these signals are reproduced a scale can be created showing the relation between the variations in the signal and the set of known pressures.

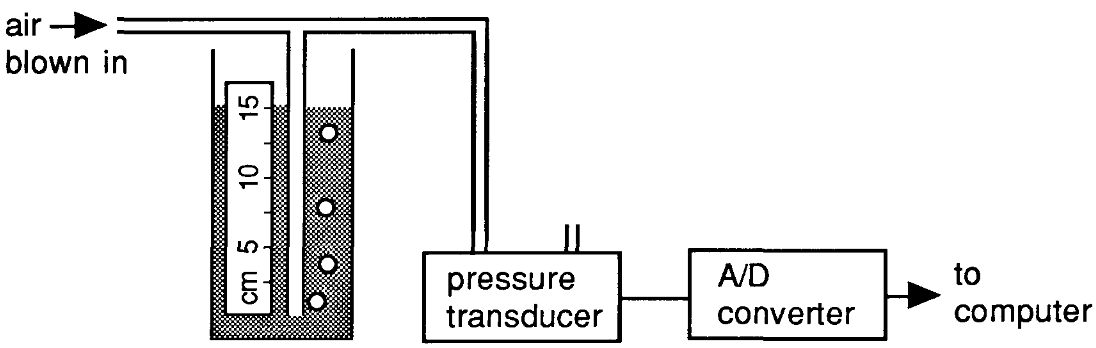


Figure 7. Calibrating a pressure transducer.

Variations in airflow are slightly more difficult to record than variations in air pressure. The most suitable way of measuring the rate of air flow in speech is by measuring the slight increase in pressure that occurs when the air flows through a fine wire mesh. Rothenberg (1973, 1977) has described a mask with built in stainless steel gauze that provides the appropriate kind of resistance. This system measures egressive or ingressive flows by converting them to small positive or negative pressures which can be transduced. It can be calibrated by connecting it to a commercially available flow meter, a device containing a ball in a tube; the greater the airflow, the higher the ball rises in the tube. With a little practice one can learn to blow steadily at a number of different rates, so that one can observe the flow meter readings that correspond to different signals, as shown in Figure 8. More accurate measure can be made in the laboratory by using a vacuum cleaner controlled by a variable voltage, reversing the drive to the fan, so that the device blows instead of sucking.

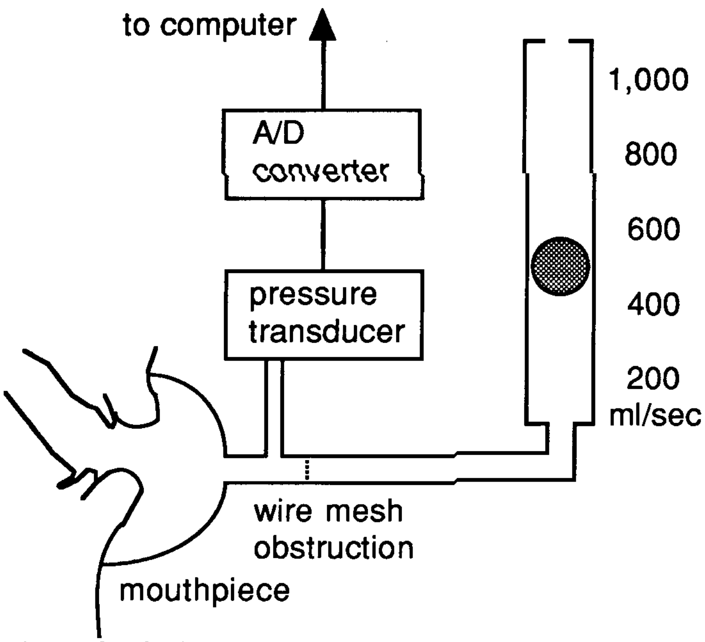


Figure 8. Calibration of oral airflow.

In fieldwork situations it may be necessary to have a face mask that can be used by speakers with very different features from those of the standard speaker of American English for whom the commercially available form of the Rothenberg mask was designed. Suitable masks can be made by taking a small rubber mixing bowl (available from a dental supply house) and cutting an opening in the bottom so as to connect it to a tube containing a stainless steel mesh. The frequency response of the system is not as good as that of the Rothenberg mask; but it is often more satisfactory for recording the overall airflow from the mouth.

For recording airflow from the nose a divided Rothenberg mask can be used. Alternatively, because it is often especially difficult to provide an airtight division for speakers with different features, nasal airflow can be recorded separately. The flow of air from the nose is usually very small indeed, and it is simplest to gather it by two tubes ending in bulbs (nasal olives) lightly inserted into the nostril. A good fit into the nostril can be achieved using malleable earplugs with tubes going through them. These tubes are then joined and led out from the face mask into a tube containing a wire gauze. In practice it is difficult to collect air from both nostrils if one nostril is also being used for a catheter in the esophagus or the pharynx. If a calibrated record of the nasal airflow is required, this nostril must be sealed off completely, and the nasal flow recorded through the other one. Alternatively, if it is sufficient to record just the relative amount of nasal airflow at different moments (as, for instance, when recording different degrees of nasalization), then the build up of the air pressure in one nostril can be recorded directly, the other nostril through which the air may be flowing providing the (uncalibratable) resistance. In any case, whether one nostril or both are used, the small pressure build up is measured and used as an indication of the flow.

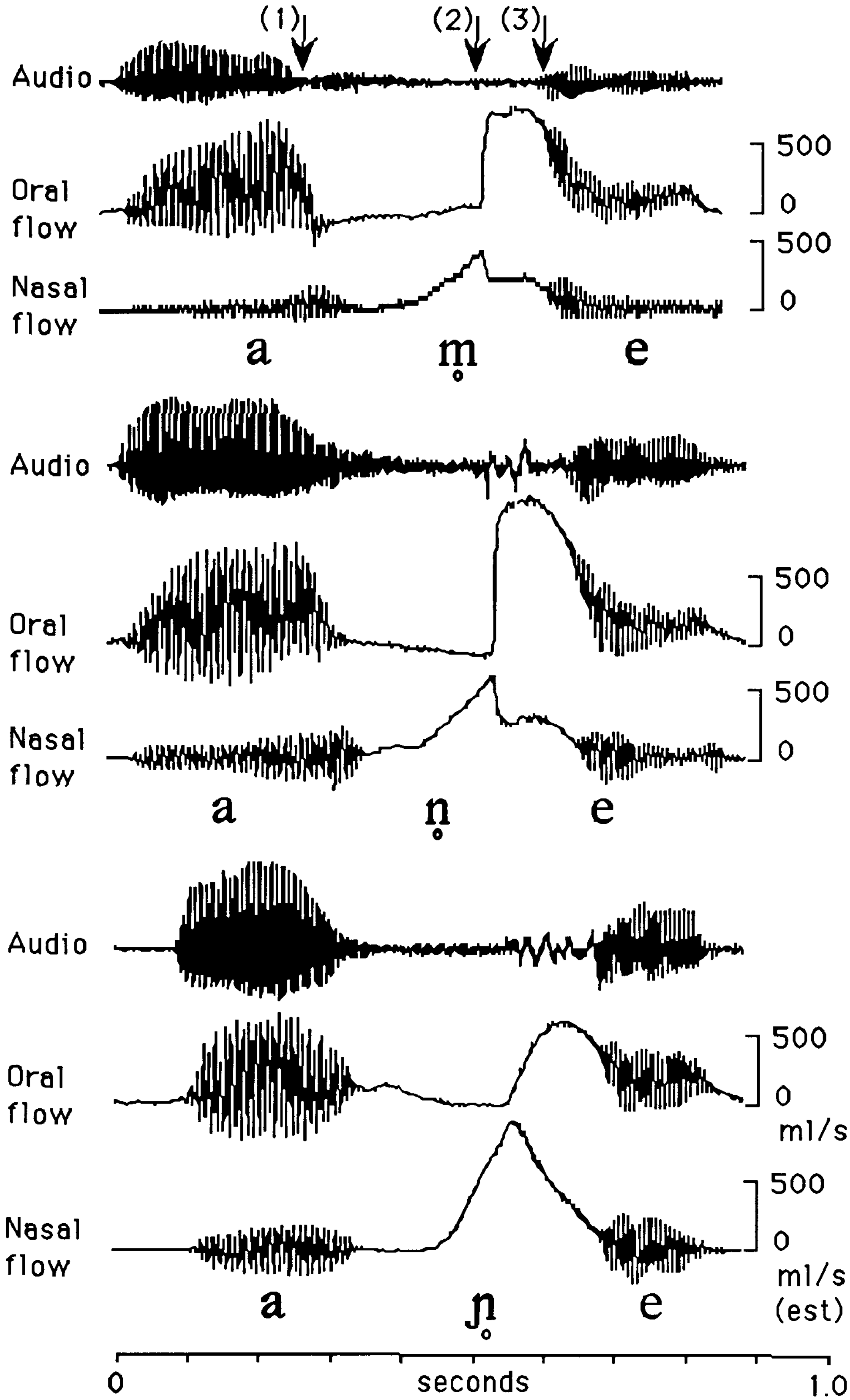


Figure 9. The voiceless aspirated nasals in Angami

Figure 9 illustrates nasal and oral airflow recorded in this way, so as to show the aerodynamic features of Angami voiceless nasals (Bhaskararao and Ladefoged, 1991). In most South East Asian languages with voiceless nasals (e.g. Burmese), the last part of phonologically voiceless nasals is actually voiced, the vocal cords vibrate before the oral closure is released, while there is still nasal airflow. But as can be seen in Figure 9, in Angami these nasals are aspirated, with the oral closure being released while there is voiceless nasal airflow. These sounds are discussed in Blankenship, Ladefoged, Bhaskararao and Chase (1993).

We can see that the value of a portable computer in phonetic fieldwork extends beyond its use in recording and analyzing sounds. It can also be an important part of a system for recording aerodynamic data. The parameters we have been discussing can be digitized along with the audio signal from a microphone. The sample rate required for digitizing the physiological parameters depends on the kind of information that is being sought. Records of the oral airflow can be filtered so as to remove the formant resonances and reveal the shape of the waveform produced at the glottis (Rothenberg 1977, Javkin, Antoñanzas-Barroso and Maddieson 1987). If this is the intent, then an appropriate face mask such as that described by Rothenberg (1977) will have to be used, and the oral airflow should be digitized at 5,000 Hz. But if the intent is to use the physiological parameters to show the gross movements of the vocal organs, then a sample rate of 500 Hz is sufficient.

The system in Figure 6 also shows two of the channels being monitored on a battery operated portable oscilloscope. This is a luxury that it is very pleasant to have when recording aerodynamic data in the field. This device makes it possible to check the calibration voltages and signals as they are being recorded, and to do simple maintenance tasks when required. (We used one to find a fault in the ignition system of our vehicle when a short circuit occurred while out in the Kalahari Desert.) Equally importantly a portable storage oscilloscope combined with a camera for photographing the screen, provides a back up system for recording in the event of a computer failure. The value of back up systems cannot be overemphasized. As all fieldworkers know, things often go wrong. We routinely go on fieldwork trips with a second small (cheaper) tape recorder, two cameras, spare pressure transducers, and spare parts for anything that might break.

3.4 Laryngeal actions

Many aspects of glottal activity can be deduced from aerodynamic records, or observed in acoustic analyses. We have seen above how airflow records can be used to illustrate characteristics of aspiration in Angami; and Ladefoged, Maddieson and Jackson (1988) have exemplified ways of quantifying phonation types such as breathy voice and creaky voice using acoustic analyses.

More direct observations of laryngeal activity can be made using an electroglottograph, which (substantially) provides a signal proportional to the degree of contact between the vocal cords. The signal actually reflects the electrical impedance between two electrodes placed on the surface of the neck, one on either side of the larynx. Consequently, it is affected by movements of the larynx beneath the skin, as well as by the opening and closing of the vocal cords, and by the degree of force with which they are thrust together.

The electroglottograph signal can be recorded on a computer in the same way as the aerodynamic signals described above. Figure 10 shows the aerodynamic and laryngeal activity that occurred during the pronunciation of the Montana Salish phrase /tʃ'tʃen'/ 'Where to'. The top line shows, on an expanded time scale, the laryngeal activity that occurred in the final laryngealized nasal /n'/. In this case this sound is realized with final creaky voice and a glottal stop. (Other speakers often produce this sound with a few periods of creaky voice after the glottal stop as well.)

The larynx record also reflects the laryngeal movements associated with the ejective [tʃ']. It cannot be taken as a direct indication of larynx raising and lowering, both because the gross movements of the larynx do not affect glottal impedance in a way that is directly proportional to larynx movement, and because this record has been band-pass filtered (30-5,000 Hz). If it had not been filtered, the small changes due to the opening and closing of the glottis (which were the major focus of the investigation) would have appeared insignificant in comparison with the large changes associated with the movements of the larynx. Nevertheless, the record clearly shows that there is greater laryngeal activity during the closure of the ejective than there is for other sounds.

Both the larynx record and the nasal flow record show that there is an epenthetic (non-contrastive) nasalized vowel after the ejective.

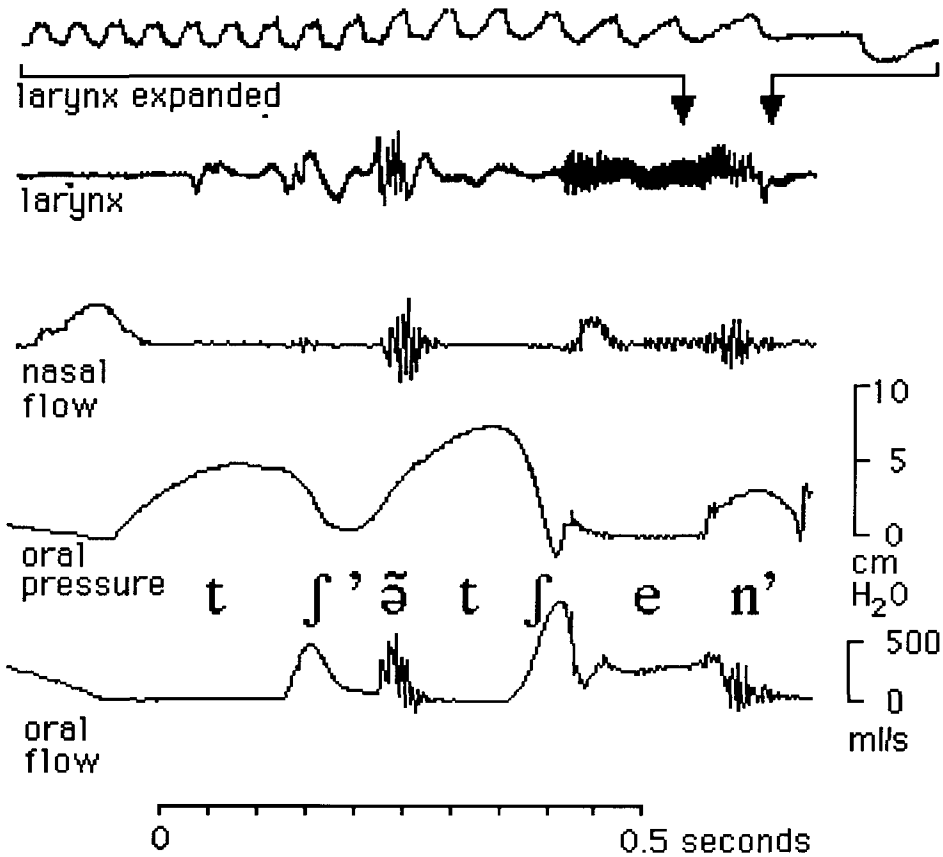


Figure 10. Laryngeal and aerodynamic records of Montana Salish.

The aerodynamic records show that in this case the ejective [tʃ'] did not have a higher oral pressure than the pulmonic affricate [tʃ] that follows it. The comparatively slow decreases in the pressure and the corresponding increases in the oral flow are typical of affricates.

3.5 Recording the place of articulation

Fieldworkers want to know not only the manner but also the place of articulation. Photographs of the lips can be very informative particularly if a mirror is used so that a full face and side views are recorded simultaneously. Figure 11 shows the contrast between the bilabial and labiodental fricatives /ɸ, f/ in Kwangali (Bantu) recorded in this way, by means of sketches of the articulators, traced from the original photographs so as to show the lip gestures more clearly. As discussed more fully in Ladefoged (1990), a salient aspect of the difference is that, for the labiodental, the lower lip is drawn back behind the upper teeth. This may be observed most clearly by reference to the stretching of the skin in the area of the arrow in the lower sketch, which does not occur at the corresponding point indicated by the arrow in the case of the bilabial.

The best way of recording lip positions is with a video camera. But with a little practice it is quite possible to take photographs at appropriate moments in the words being investigated using an ordinary camera. Ask the speaker to repeat the sound in a phrase, over and over again, while you get ready to take the photograph. Have a tape recorder running, and make sure that the camera is close to the microphone, so that there is a clear recording of the click of the shutter when you eventually take the photograph. You can find out roughly when the photograph was taken by playing the recording at half speed. A more precise determination of the time can be made from examination of the waveform or spectrogram of the utterance. If a computer system is being used, this can easily be done in the field.

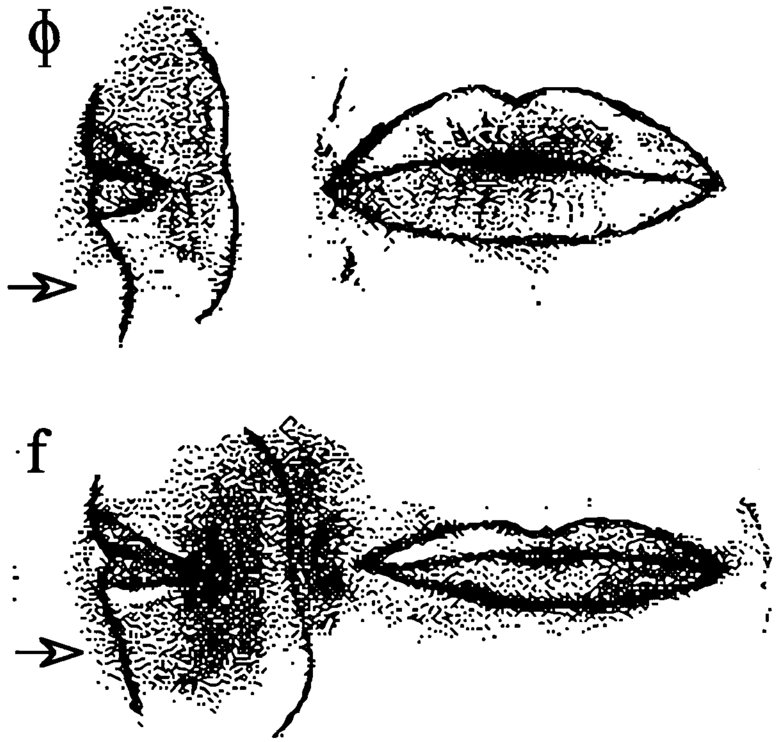
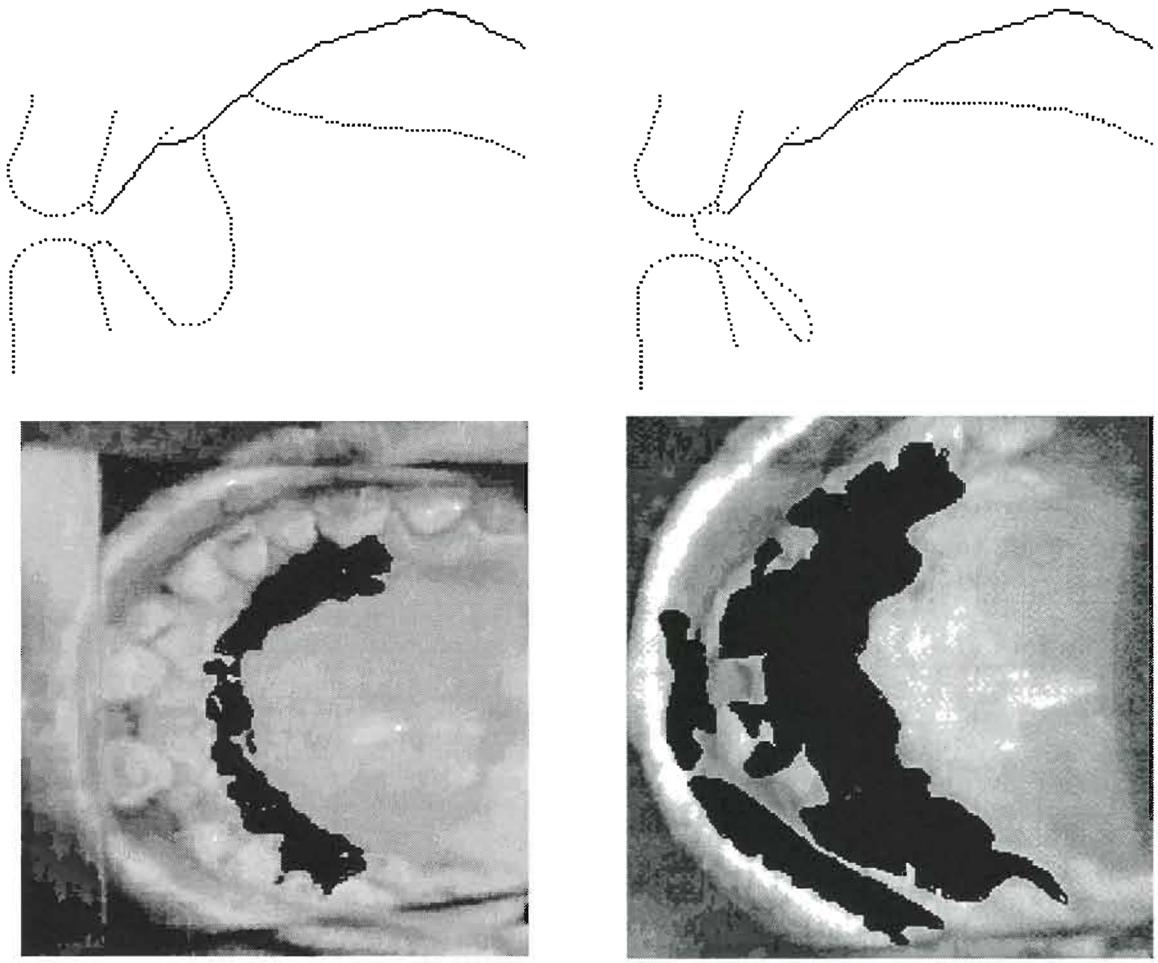


Figure 11. Sketches based on photographs of the lips during the pronunciation of bilabial and labiodental fricatives /ɸ, f/ in Kwangali.

Palatography is a well known traditional method of obtaining articulatory data (Abercrombie 1957, Ladefoged 1957). The best method of determining the region of the upper surface of the vocal tract contacted by the tongue in a given word is to coat the tongue with a mixture of equal parts of olive oil and powdered charcoal (Dart 1991). Then ask the speaker to say a word containing the sound to be investigated and no other consonants made in the same articulatory region. When this has been done, the marking medium on the tongue will have been transferred to the upper articulator. Insert a mirror into the mouth, and use it to view and photograph the place of articulation. When all the required pictures of the roof of the mouth have been obtained, the procedure can be reversed. Paint the upper surface of the mouth with the olive oil and charcoal mixture, and observe (and photograph) the part of the tongue that is making the contact.

Figure 12 (from Spajić, Ladefoged, Maddieson and Sands, 1993) illustrates a pair of contrasting words recorded during fieldwork on Dahalo. The tongue has contacted the roof of the mouth and has transferred the black marking medium to the post-alveolar area in the case of the word on the right, and to the denti-alveolar area for the word on the left. Note that in these and in all such photographs, the contact areas reflect the sum of the articulatory contacts that occurred in the pronunciation of the words investigated; they do not show the position at any one particular moment.



[tʰokke]　　　　[t̪a:mi]

Figure 12. Palatograms and sagittal sections showing the contact between a laminal dental and apical alveolar stop in Dahalo. The area of contact has been made darker to aid in making the contrast visible, but it is otherwise identical to the unretouched photograph.

Polaroid makes a suitable dental camera, model CU5, which comes with mirrors and attachments making it possible to get accurate life size pictures, as in Figure 12. Alternatively, a video camera can be used. There are a number of advantages to using a video camera. Firstly it is cheaper to buy and to use than the special purpose Polaroid camera. Secondly it can be held further away from the subject's face (about 120 cm, using an 8x zoom lens) and still fill the frame; this makes it less intimidating for the speaker. Thirdly, the image can be fed directly into a computer for further processing.

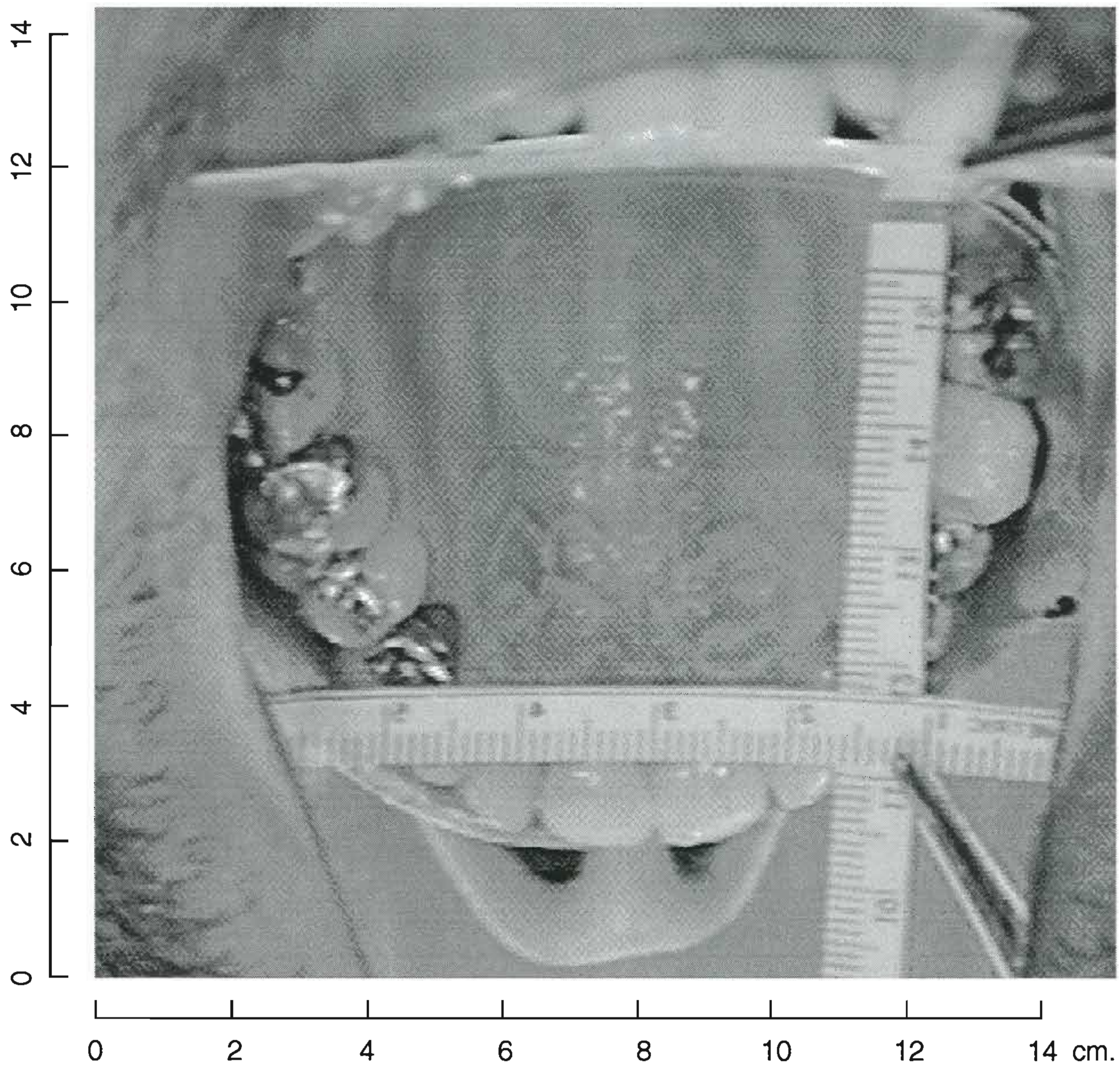
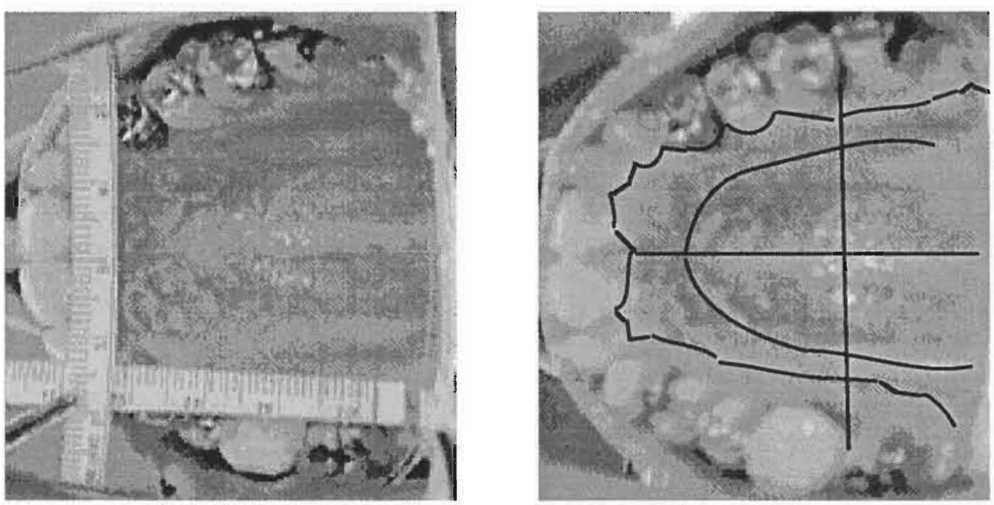


Figure 13. A computer image of a mirror placed so as to show the roof of a speaker's mouth, and rulers that have been placed in the plane of the teeth. The scales around the outside of the photograph were added in the computer graphics program.

Figure 13 shows a picture of the upper surface of a speaker's mouth, exactly as it was captured on the computer screen, except for the scales that have been added around the outside of the photograph. Two rulers were held against the speaker's teeth, one in the sagittal and one in the coronal plane. As the mirror was not exactly at a 45° angle to the plane of the teeth, it may be seen that the posterior anterior dimension is slightly foreshortened. In the horizontal dimension, 10 mm on the ruler in the photograph is equivalent to 20 mm on the scale; but in the vertical dimension, 11 mm on the ruler is equivalent to 20 mm on the scale. This can be corrected quite easily by using a graphics program that permits separate horizontal and vertical scaling of the image. In the particular case shown in Figure 13, reducing the horizontal dimension by 50% and the vertical dimension by 55% will produce a life-size image, as shown in the lefthand half of Figure 14.

Palatograms should be accompanied by diagrams showing the shape of that particular speaker's mouth in the form of a traditional sagittal section. Diagrams of this kind can be based on dental impressions of the oral cavity made in the field. It is best to use an alginate impression material, as other substances which set harder cannot be manipulated suitably. There is no reason for the impression to be made using a tray of the kind that dentists use. All that is needed is an impression of the inner surfaces of the roof of the mouth; as the outer surfaces of the teeth play no role in the production of speech, we can simply neglect them. The easiest way to make an impression is to mix a sufficient quantity of the material and place it on the back of the palatography mirror. Get the speaker to lean slightly forward, and then insert the mirror with the material on it into the mouth. Press the mirror firmly against the upper teeth, allowing some of the material to flow out of the mouth around the upper lip. A good impression for phonetic purposes should be made with sufficient material to indicate (at least roughly) the shape of the upper lip and the curvature of the soft palate. The palate will, of course, be in a lowered position, as the speaker will have been breathing through the nose while the impression material is setting. The impression material around the lips sets slightly more slowly than that inside the mouth, where it is slightly warmer. When the material around the lips is firm, it is quite safe to remove the mirror from the mouth, first rocking it back and forth, raising and lowering it slightly, so as to break the seal.

If an alginate impression is to be kept for any length of time it must be immersed in water so as to prevent it from drying and shrinking. Otherwise, take it off the mirror and trim the base flat so that it is parallel to the plane of the teeth. If the mirror really was pressed firmly against the upper teeth while the impression was being made, this should involve no more than the removal of excess material from around the sides. The impression may then be cut in half in the mid-sagittal plane, so as to form an outline for drawing the mid-line of the upper surface of the vocal tract. The exact positions of movable structures such as the lips and the soft palate have to be estimated, but if care has been taken to have sufficient impression material around the lips and as far back in the mouth as possible, the sagittal diagram will be reasonably accurate. Palatograms should always be accompanied by diagrams of this kind, as in Figure 12. It has long been established that sagittal sections provide the most useful representations of speech sounds.



(a) (b)

Figure 14. Photographs of the roof of a speaker's mouth, (a) as in Figure 13, but reduced and transformed into an undistorted life-size view, and (b) with contours drawn as explained in the text..

It is possible to provide even more information from dental impressions of the roof of the mouth. The right hand portion of Figure 14 shows records in which the impression material has been used to construct contour lines at fixed distances from the plane of the teeth. Lines showing points 6 mm and 12 mm above the plane of the teeth have been superimposed on the undistorted life-size photograph. In order to draw contour lines of this sort, put the two halves of the impression material together again, and place them between two spacers with a known height. Using these spacers as guides, cut the impression material horizontally, with the blade of the knife parallel to the surface corresponding to the plane of the teeth. Draw a line around the cut edges of the impression material, which will represent a contour at the height of the spacers up from the plane of the teeth. Repeat the process with the remainder of the impression material so as to get additional contour lines. In order to place the contour lines accurately on the photograph it will be necessary to draw two guide lines on both them and the photograph, one corresponding to the midline between the frontal incisors, and the other to a line between two teeth that can be seen on both the impression material and the photograph. These lines are also shown in the righthand photograph in Figure 14.

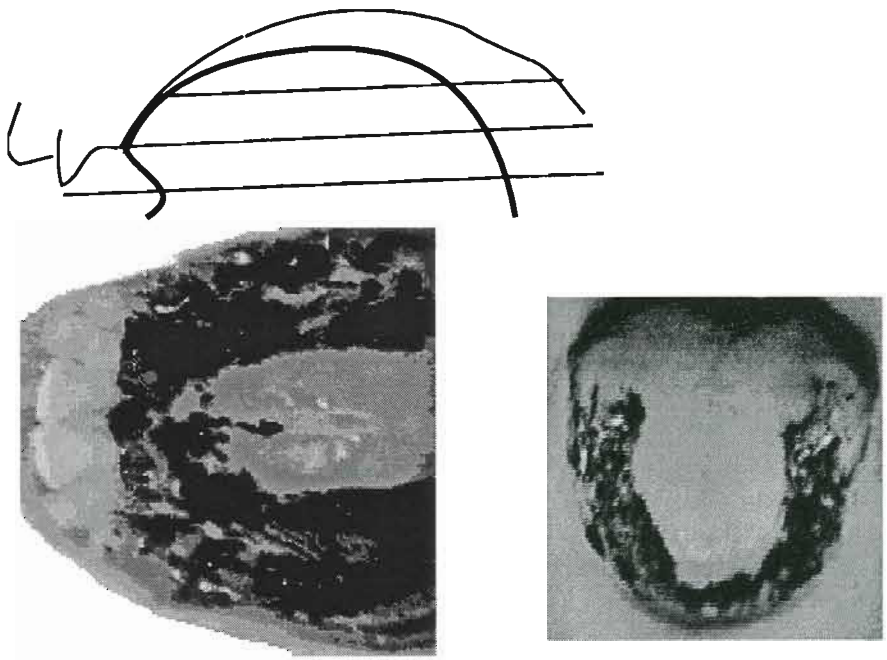


Figure 15. Palatogram (bottom left) and linguagram (bottom right) produced during an utterance of the Serbian word [dʒaːk] 'student'. The upper part of the figure shows a reconstructed sagittal section (see text).

All this information can be put together as shown in Figure 15, which is part of an investigation of palatalization in Serbian (Spajic, forthcoming). In interpreting such records it must be remembered that although the upper surface of the palate may have been accurately mapped, no such accuracy is possible when considering the photographs of the tongue. The photograph of the tongue in this figure shows it when it has been slightly stuck out of the mouth, and is therefore not in the same shape as it was when producing any of the sounds. However, consideration of the contact areas on both the tongue and the palate allows us to reconstruct where the tongue must have moved at some point during the word. It may be seen that the tip an forward part of the blade of the tongue are the articulators. Close to the mid-sagittal plane the contact on the roof of the mouth was on the posterior part of the alveolar ridge, at a level corresponding to the area between the two contour lines. There has also been some contact on the roof of the mouth slightly further back, particularly on the midline itself. The speaker had a number of prominent rugae (small ridges) in this area, which have not been indicated on the diagram of the midsagittal section. the sides of the tongue contacted the roof of the mouth at a level above that of the 12 mm contour. Accordingly, this must be a palatalized sound, with the midline of the tongue being in the vicinity of the area shown by the heavy black line.

We may conclude this section by noting a few practical points in connection with palatography. Firstly, care should be taken in selecting appropriate words. We are often interested in comparing the places of articulation of different sounds. Accordingly words must be chosen that contain these articulations, and do not contain any other similar articulations that might overlap with them. Thus when investigating the difference between [s] and [ʃ] in English one should use words such as "sop-shop" rather than "sot-shot." Similarly one should use either a range of vowels ("seep-sheep, sip-ship, same-shame, Sam-sham, sop-shop, etc.") or, if this is not possible, just open vowels which will not obscure the consonant contacts. As with all instrumental phonetic investigations, time spent selecting suitable words is a good investment.

When doing palatography, one should allow the speaker to practice the task extensively. It is important to get the speakers to relax after the upper surface of the mouth has been painted, so that when they say the word being investigated they do so naturally. It also requires practice to stick the tongue out of the mouth the same way every time. It is obviously important to date and label the photographs as soon as they are taken. When using a video camera, photography can be *begun just before the speaker says the word, thus ensuring an audio record of the particular token* photographed. In addition, it is preferable to make records of several different speakers saying a few utterances rather than one or two speakers repeating a large number of different utterances. Ideally one would like to get a dozen speakers of the same dialect each repeating a dozen times all the contrasts to be investigated. But making palatographic records is fairly time consuming, and in a world in which resources are limited one may have to be satisfied with half a dozen speakers saying each word once. We hope, however that gone are the days when phoneticians such as Ladefoged (1964) made general statements about a language based on the palatographic records of a single speaker. We need to find out the properties of the language that a group of speakers have in common, rather than the details of an individual's pronunciation.

4. Concluding comment

The best kinds of instrumental investigations are those that are quantifiable, and to which one can apply statistical techniques such as analysis of variance. Differences *between* individuals have been shown to be much greater than differences *within* an individual's repetitions (Johnson, Ladefoged and Lindau, 1993), which is why, if a choice is necessary, it is preferable to record six different individuals saying a single utterance rather than one individual repeating an utterance six times. What we are interested in as linguistic phoneticians are the differences between languages, which means that we must have records of groups of speakers who can be considered to represent the different language. When we are trying to show that there are differences between particular sounds in different languages, we need to show that these differences are statistically large in comparison with differences among the speakers in each group, and also that they are large in comparison with differences among repetitions by a single speaker. When we have done this, we can say that the sounds in the languages represented by the groups really are phonetically different. And that's what linguistic phonetics is all about.

Acknowledgments

This paper includes some material from a short presentation at the XIIth International Congress of Phonetic Sciences (Ladefoged, 1991), as well as some material from an earlier paper by Ladefoged and Traill (1984). My thanks are due to Tony Traill for his wonderful collaboration in the earlier paper. Many members of the UCLA Phonetics Lab have also contributed helpful comments. My colleague Ian Maddieson has offered much useful advice, and Sinisa Spajic has been particularly helpful in developing palatographic techniques.

References

Abercrombie, D. (1957). "Palatography." *Zeitschrift für Phonetik* 10: 21-25.

Bhaskararao, P. and Ladefoged, P. (1991). "Two types of voiceless nasals." *Journal of the International Phonetic Association* 21: 80-88.

Blankenship, B., Ladefoged, P., Bhaskararao, P and Chase, N. "Phonetic structures of Khonoma Angami." *UCLA Working Papers in Phonetics* 84: 127-142.

Dart, S. (1991). Articulatory and Acoustic Properties of Apical and Laminal Articulations. *UCLA Working Papers in Phonetics* 79.

Hale, K., Krauss, M., Watahomigie, L. J., Yamamoto, A., Craig, C., Jeanne, L. M., and England, N. C. (1992). "Endangered languages." *Language*. 68: 1-42.

Javkin, H., Antoñanzas-Barroso, N. and Maddieson, I. (1987). "Digital inverse filtering for linguistic research." *Journal of Speech and Hearing Research* 30: 122-129.

Johnson, K., Ladefoged, P. and Lindau, M. (1993). "Degrees of freedom in the production of vowels". *Journal of the Acoustical Society of America* (in press).

Kelly, J. and Local, J. (1989). *Doing Phonology: Observing, Recording, Interpreting.* Manchester and New York: Manchester University Press.

Ladefoged, P. (1957). "Use of palatography" *Journal of Speech and Hearing Disorders.* 22:764-774

Ladefoged, P. (1964). "Igbirra notes and word-list." *Journal of West African Languages* 1: 27-37.

Ladefoged, P. (1967). *Three Areas of Experimental Phonetics.* London: Oxford University Press.

Ladefoged, P. (1968). *A Phonetic Study of West African Languages.* Cambridge: Cambridge University Press.

Ladefoged, P. (1991). "Instrumental phonetic fieldwork: techniques and results." In *XII International Congress of Phonetic Sciences,* 4: 126-129. Aix-en-Provence: Université de Provence.

Ladefoged, P., Maddieson, I., and Jackson, M. (1987). "Investigating phonation types in different languages". In O. Fujimura (Ed.), *Vocal Physiology: Voice Production, Mechanisms and Functions* (pp. 297-317). New York: Raven.

Ladefoged, P., and Traill, A. (1984). "Instrumental phonetic fieldwork." In R. Thelwall & J. Higgs (Eds.), *Studies in Experimental Phonetics in Honour of E. T. Uldall* Coleraine, New University of Ulster.

Ladefoged, P. and Traill, A. (1993). "Clicks and their accompaniments." *J.ournal of Phonetics.* (in press)

Maddieson, I., Spajić, S., Sands, B. and Ladefoged, P. (1993). "Phonetic Structures of Dahalo." *UCLA Working Papers in Phonetics* 84: 25-65.

Pike, K. L. (1947). *Phonemics: A Technique for Reducing Languages to Writing.* Ann Arbor, Michigan: University of Michigan Press.

Rothenberg, M. (1973). "A new inverse-filtering technique for deriving the glottal airflow waveform during voicing." *Journal of the Acoustical Society of America* 53: 1632-1645.

Rothenberg, M. (1977). "Measurement of airflow in speech." *Journal of Speech and Hearing Research* 20: 155-176.

Rousselot, P. J. (1924-5). *Principes de phonétique expérimentale.* Paris: Didier

Scripture, E. W. (1923). *The Study of English Speech by New Methods of Phonetic Investigation.* London: Oxford University Press.

Spajić, S. (forthcoming). "Serbian affricates."

Phonetic Structures of Dahalo

Ian Maddieson, Siniša Spajić, Bonny Sands and Peter Ladefoged

1. Introduction

Dahalo [ɗaha:lo] is a Cushitic language spoken by a small population living on the northern coast of Kenya near the mouth of the Tana river. In addition to an extensive consonant inventory inherited from Cushitic it has a number of phonological characteristics borrowed from languages with which it has been in contact, including members of the Bantu group, other Cushitic languages such as Aweera (Boni), and a putatively Khoisan language formerly spoken in East Africa from which clicks were borrowed. It thus has the distinction of being the only language with clicks that is known not to be Khoisan or Bantu. Its diverse repertoire of consonants make Dahalo an object of especial phonetic interest.

Cushitic is one of the branches of the large Afro-Asiatic family of languages. Cushitic is usually subdivided into Central, Eastern and Southern sub-branches. Greenberg (1963) classified Dahalo as a Southern Cushitic language, relating it to Ma'a (Mbugu) and the Rift group, consisting of Kw'adza and Asa (both probably extinct) and Iraqw, Burunge, and Alagwa. This affiliation is widely accepted (e.g. by Elderkin 1978, Ehret 1980, Nurse 1986), but is questioned by Tosco (1989, 1991) who prefers to see Dahalo as a divergent Eastern Cushitic language, perhaps closest to Yaaku (Heine 1974).

Dahalo is spoken in an area roughly bounded on the west side by the Tana river delta up to about twenty miles inland near Garsen, and bounded on the east side by the coast of Lamu Bay (Elderkin 1976, Nurse 1986, Tosco 1991, and our own data). The principal settlements in this area include Witu on the main highway to Lamu, Kipini at the mouth of the Tana river, Mokowe and Mkunumbi on Lamu Bay, and Lake Kenyatta. The majority of the population in this area are speakers of the Bantu languages Swahili and Lower Pokomo, or of Aweera, with the Dahalo speakers living widely dispersed among these larger linguistic groups. Nurse (1986) estimated the number of remaining Dahalo speakers to be under 500, while Tosco (1991) thinks that "the figure of 400 cannot greatly exceed the truth" (p. xi). The estimate of 3000 speakers cited in Grimes (1992) is wildly excessive. It is likely that the number of competent Dahalo speakers is still declining as it was a very difficult task to find even six speakers to record the data for this paper and we did not observe any children acquiring Dahalo; instead Swahili seemed to be the most commonly used language among the younger Dahalo.

Nurse (1986) reports that the name Dahalo is used freely by the Dahalo and their neighbors, but we found that this word was not used by any other Kenyans we encountered, only by the Dahalo themselves. The term most commonly used by others to refer to the Dahalo is (Wa-)Sanye. This word is used to describe a number of minority groups in the coastal area, particularly speakers of Waata, an Eastern Cushitic language of the Oromo ("Galla") group. Hence it is very ambiguous. Our speakers used both the names 'Dahalo' and 'Sanye' (in Swahili 'KiDahaalo' and 'KiSanye') to refer to their language. Elderkin (1976) notes that he was unable to elicit a proper name for either the people or the language from his Dahalo informants. They preferred to call themselves 'guɦo garima:ni' "people of the outside", or 'guɦo gwits:o' "the little people" and to refer to their language as 'numma guɦo:ni' "the lip of men". Nurse (1986) notes that his Dahalo informants spoke of two groups of Dahalo: 'dahalo' and 'sanye' and that Aweera informants spoke of two groups: 'dako' and 'denk', but we do not know if this reflects the existence of separate dialects or bears in some other way on the language that is presented here. We found no evidence that the name Dahalo was perceived to be pejorative, as is reported by Grimes (1992).

The most complete description of Dahalo is the 'Grammatical Sketch' by Tosco (1991). An impressionistic description of the segmental characteristics appeared in Tucker and Bryan (1977). A more analytical (and accurate) report on the phonology and morphology of the language is given in Elderkin (1974). Analysis and comparison of the lexicon is a major focus of Ehret (1980) and Nurse (1986). Ehret, Elderkin and Nurse (1989) published all the known lexical items of Dahalo. These last three works are all concerned in part or in their entirety with discerning the influences of other languages on Dahalo, and on the inferences for the history of East Africa that can be drawn from understanding this material.

No detailed phonetic study of Dahalo has yet been published. The present paper constitutes a first step in this direction. It reports on the major phonetic characteristics of the language, based pricipally on auditory and acoustic analyses of data recorded in the field in July 1991. Our prinicipal informant was Ali Mena, an elderly resident of Witu. He helped with the preparation of a list of over 400 words illustrating all the phonological contrasts in the language. We recorded the whole of this material as spoken by Ali Mena himself, and by Fatuma and Karata Papureki, two of his female relatives in Witu. This wordlist, as transcribed from the tape, is included as Appendix 1 to this paper. The transcribed forms represent the concensus of all four authors of the present paper. We also recorded a subset of the list — the words being chosen so they still illustrated all the major phonological contrasts in the language — as spoken by three other speakers living in the vicinity of Kipini. These were Swalehi Bakari, a male probably in his thirties, and two slightly older speakers, one male and one female. In addition, palatographic and linguagraphic data was obtained from Swalehi Bakari. Our report is organized around an analysis of the segments that would be considered distinctive in a rather straightforward phonological analysis.

2. Consonants

The 64 consonants of Dahalo that we recognize are listed in Table 1 below. Generally similar inventories of consonants are reported by Elderkin (1974, 1976), Nurse (1986), Ehret, Elderkin, and Nurse (1989), and Tosco (1991), although none of these reports agree exactly with the analysis we have made. The discrepancies arise from three types of reasons; some are merely matters of interpretation, some may reflect lacunae in our data-gathering or other 'accidental' factors, but in a number of cases we feel that we have been able to add to or correct the previous accounts.

2.1 Places of articulation

Dahalo distinguishes consonants at eight places of articulation. As the chart makes clear, Dahalo is among the minority of languages that makes use of a distinction between dental and alveolar places. Plain and prenasalized plosives and ejective stops all reflect this place distinction. Previous publications on Dahalo have not noted its occurrence with ejective stops, reporting instead a contrast between dental ejective stops and affricates. The dental vs alveolar contrast is not found among coronal consonants of other types, such as nasals, implosives or clicks.

In the chart the clicks have been placed below the heading "dental", while nasal, lateral, implosive and other coronal consonants lacking a distinction between dental and alveolar places have been placed below the heading "alveolar". In order to highlight the absence of contrast, no line is drawn to separate dental and alveolar in these cases. The positioning of the symbols is, however, based on our impression of the most typical productions of these segments. This does not rule out the possibility that the place may be more variable where it is not contrastive than in contrastive instances.

Table 1. Chart of Dahalo consonants

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Epi-glottal	Glottal
Plosives	p ɓ		t̪ d̪	t̠ d̠			k g kw gw	ʡ	ʔ
Affricates				ts dz dzw	tʃ dʒ				
Prenasalized plosives	mp mb		n̪t̪ n̪d̪	nt̠ nd̠ nd̠w			ŋk ŋg ŋkw ŋgw		
Prenasalized affricates				nts ndz	ntʃ ndʒ				
Ejective stops	p'		t̪'	t̠'			k' kw'		
Ejective affricates					tʃ'				
Lateral ejective affricates				tɬ'		cʎ̥'			
Implosives	ɓ			ɗ					
Nasalized clicks			ŋ̊ǀ ŋǀ ŋ̊ǀw ŋǀw						
Nasals	m			n		ɲ			
Fricatives		f		s (z)	ʃ			ʜ	h
Lateral fricatives				ɬ ɬw		ʎ̥			
Approximants	w					(j)			
Lateral approximants				l					
Trill				r					

The articulatory nature of the dental/alveolar place distinction is illustrated in Figure 1, showing palatograms of /t̪áːmi/ "grass, thatch" (271) and /t̠oke/ (a species of black bird) (288) together with diagrams showing the inferred articulatory position in the mid-sagittal plane. (The parenthesized numbers refer to the position of the word in the word list included as Appendix 1. Note that the second of these words is pronounced as /nt̠oke/ with prenasalization by our Witu speakers). The articulatory diagrams show a solid line indicating the shape of the roof of the mouth of this speaker, traced from a mid-saggital section of a dental impression made in the field. The positions of moveable structures are shown as dotted lines with the positions of the tongue estimated from the contact areas shown in the palatograms, as well as from observation of the contact areas on the tongue seen in linguagrams made of the same words. The area of contact has been made darker in this figure to aid in making the contrast visible, but it is otherwise identical to the unretouched photograph.

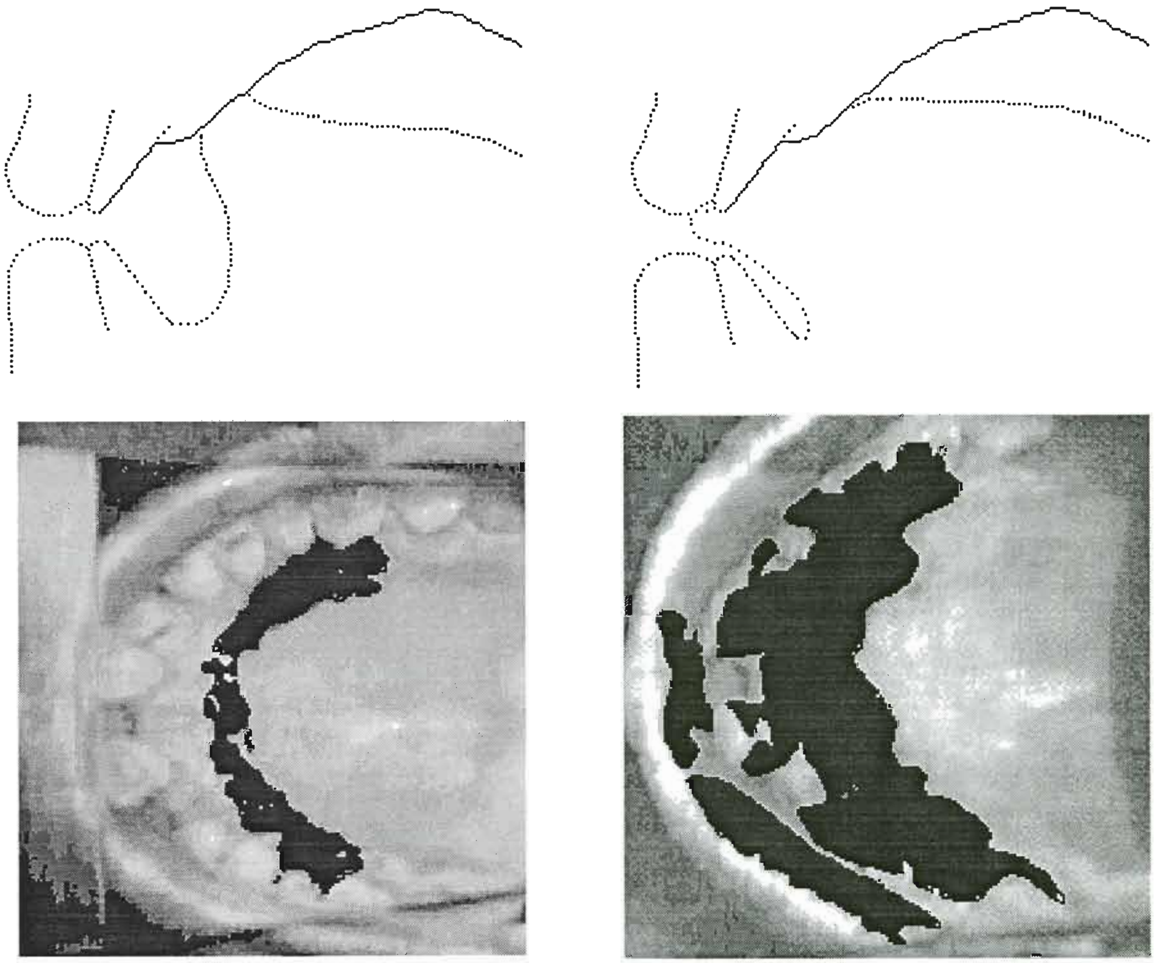


Figure 1. Palatograms and sagittal sections of an apical alveolar (left, /t̠oke/ (a species of black bird) and a laminal dental stop (right, /t̪ami/ "grass, thatch") in Dahalo.

These palatograms demonstrate that there are differences for both the active and passive articulators in these stops. In /t̠oke/ the stop closure is made toward the rear of the alveolar ridge (which is quite prominent for this speaker) in a location that might almost be called retroflex. The stop is appropriately called apical in that there is a very small area of contact produced by the tip of the tongue. On the other hand, the stop in /t̪ami/ is laminal and the contact spans from the alveolar ridge to the tip of the front teeth, extending even onto the inner surface of the upper lip. It might more accurately be called denti-alveolar since both the teeth and the alveolar ridge are involved in the articulation. A tendency to pair apicality with alveolar place and laminality with dental (or denti-alveolar) place is the norm across languages (Ladefoged and Maddieson 1986, Dart 1991).

What is somewhat surprising is the distribution of friction noise associated with the releases of these consonants. One would expect that laminal stops might have a longer and noisier release than apical stops due to the greater contact area involved. However, in Dahalo we find the opposite to be true. Figure 2 provides sample spectrograms of /t̪aːmi/ and /t̠aʔadi/ "fruit of *ʃitinke* palm" (270) illustrating the difference. The noisy part of the release of the alveolars is roughly 3 times as long as that of the dentals. In this respect Dahalo is like Temne (Ladefoged and Maddieson 1986), and unlike most of those languages of India, Australia and the Americas in which dental/alveolar contrasts are found.

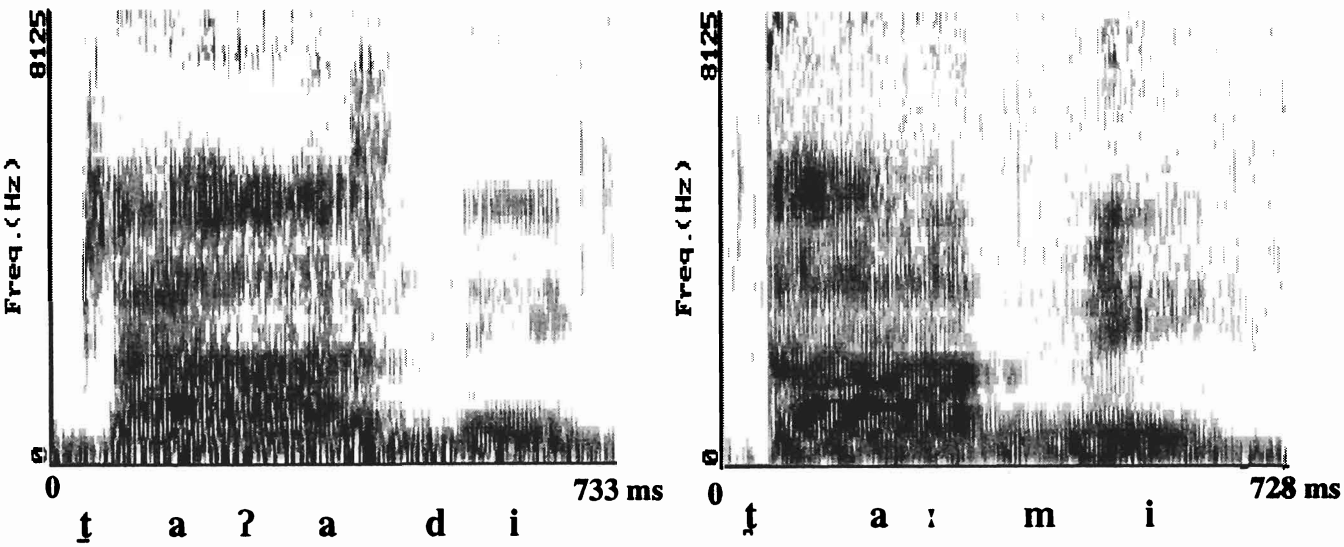


Figure 2. Spectrograms of /t̠aʔadi/ "fruit of *fitinke*" and /t̪aːmi/ "grass, thatch", illustrating longer and noisier release for alveolar than for dental plosives.

To quantify this difference we measured the release noise and any associated aspiration in two repetitions by each of the three Witu speakers of two representative words beginning with voiceless bilabial, laminal dental, apical alveolar and velar plosives. The mean durations of the interval from the release transient to the onset of voicing for the following vowel for these four places are given in Table 2. The alveolars have a substantially longer voice onset time (VOT) than plosives at other places.

Table 2. Mean VOT (in ms) of Dahalo voiceless stops at four places of articulation.

p	t̪	t̠	k
19.7	15.3	42.1	26.8

In a one-factor analysis of variance a significant main effect of place was found ($F(3, 47) = 34.358$, $p < .0001$). In post-hoc analyses the alveolar plosives were distinct from all the others at at least the .01 level of significance, the velars were significantly different from all others at at least a .05 level, but the dental and labial stops were not significantly distinct from each other. The noise pattern associated with the alveolars is even more striking when compared to the typical pattern of voice onset time (VOT) for voiceless stops among the world's languages. The typical rule for VOT is, in general, the further back the place of articulation in the mouth, the longer the VOT (Fischer-Jørgensen 1954). But this is not the Dahalo pattern. It is likely that their relatively long friction and aspiration duration could serve as a significant cue distinguishing alveolar plosives from the others.

The place of articulation for the affricate /ts/ in the word /tsaːka/ "hot season" (251) is illustrated in the palatogram in Figure 3. (Recall that Dahalo affricates show a distinction between alveolar and palato-alveolar places, but do not include distinct dental and alveolar types). Here the contact is more extended in the sagittal plane than for the alveolar stop in Figure 1, but is less extended than is seen for the dental stop. The contact does not include the incisors nor reach the lips. In other words, this articulation appears somewhat intermediate between the distinctive dental and alveolar categories illustrated above. We cannot be sure whether this is a consequence of the affricate manner of production of this segment, or indicates a general tendency to default to less distinctive articulations when contrast does not need to be maintained.

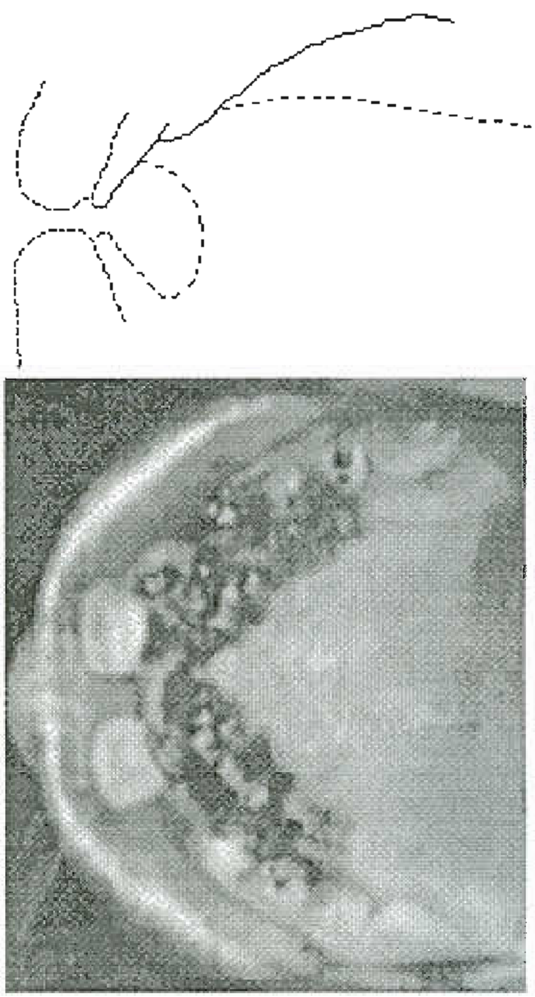


Figure 3. Palatogram of /tsa:ka/ "hot season" (unretouched).

The other unusual feature among places of articulation in Dahalo is the use of the epiglottal place for a pair of contrastive segments that we represent as /ʡ/ and /ʜ/. The segments in question have previously been described as voiced and voiceless pharyngeal fricatives and represented as /ʕ/ and /ħ/. In current IPA usage these symbols represent articulations made in the upper pharyngeal region. The symbols /ʡ/ and /ʜ/ represent articulations made in the lower pharyngeal region at the level of the epiglottis, /ʡ/ being used for an epiglottal stop and /ʜ/ for an epiglottal fricative. The most typical pronunciation of the segment we write /ʡ/ is as a stop. On anatomical grounds, we may be reasonably certain that the articulation must be epiglottal in this case, since the pharyngeal constrictor muscles are unlikely to move the back wall of the pharynx forward to contact the upper tongue root, and extensive retraction of the tongue root will carry the epiglottis itself backward. The only plausible location for a closure in the pharyngeal region is therefore at the epiglottis; in fact, active involvement of the epiglottis to make closures in the lower pharyngeal region has been observed in Arabic and Hebrew (Laufer and Condax 1979, Laufer and Baer 1988). These epiglottal sounds have quite marked coarticulatory effects on neighbouring vowels, tending to produce a perceptual fronting of /a, o, u/ and a perceptual lowering of /i/ and /e/. Since the effect of /ʜ/ is quite similar to that of /ʡ/, we conclude that both are more precisely described as epiglottal (although, of course, they are pharyngeal in a broader sense).

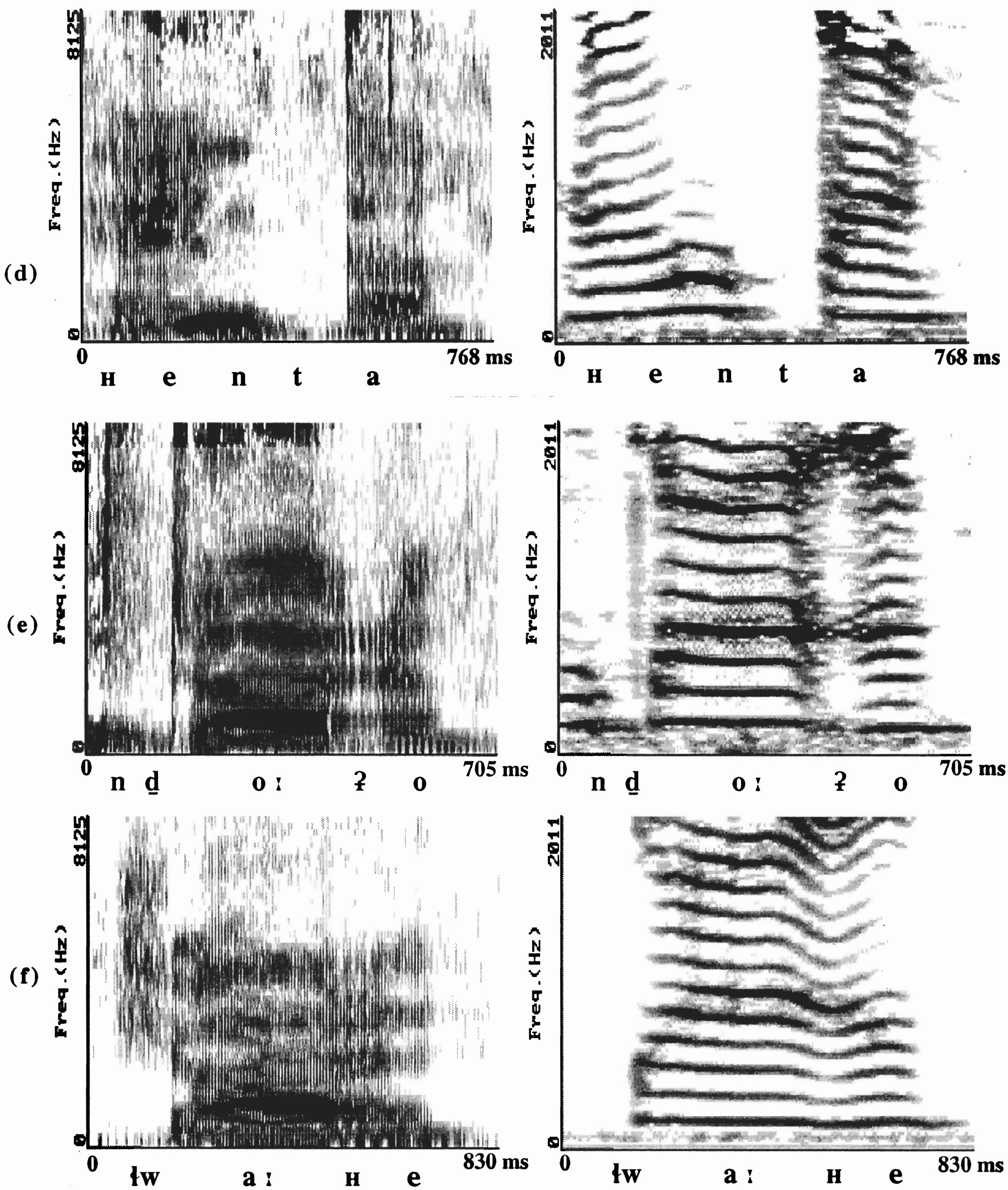
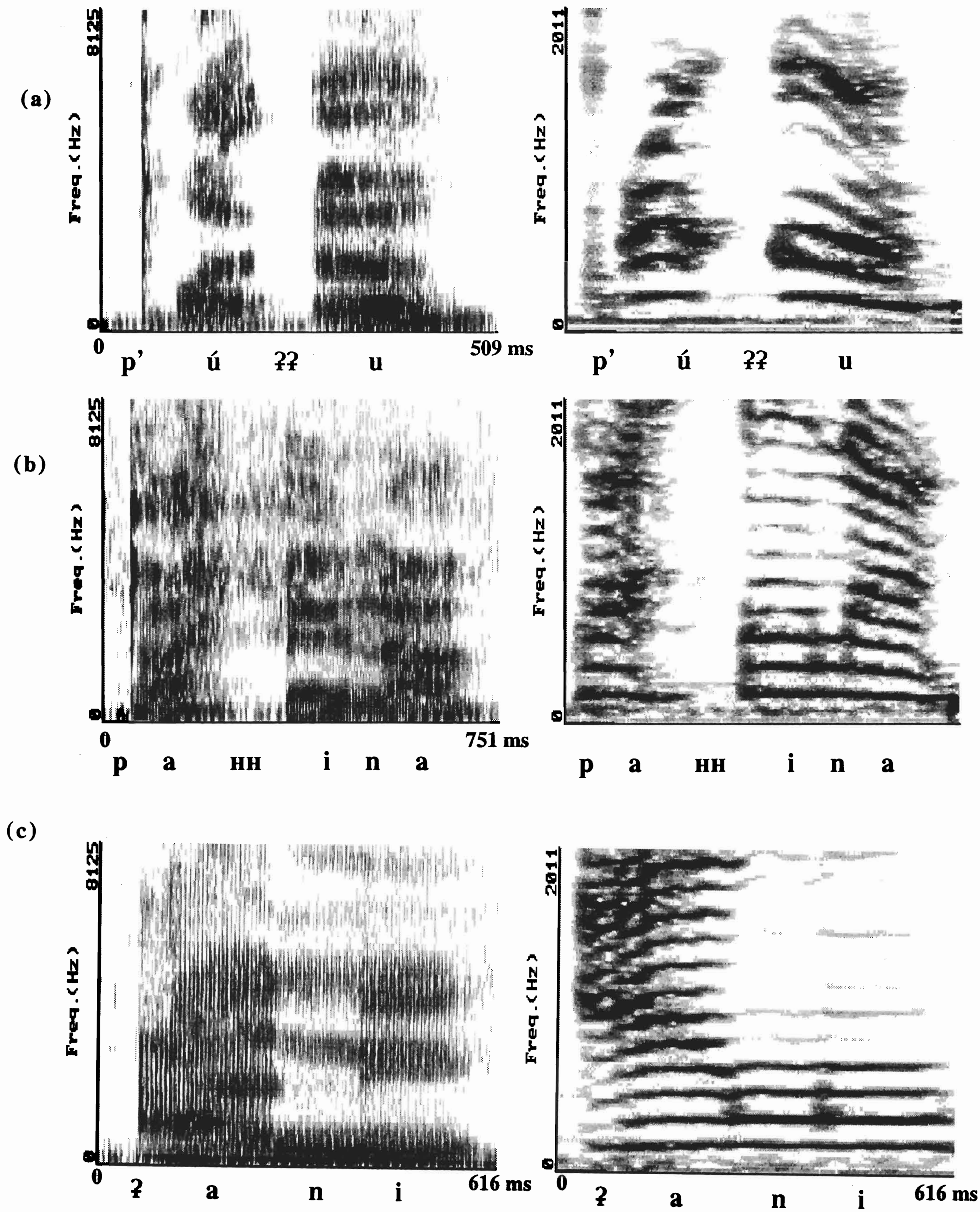


Figure 4: Wide and narrow band spectrograms illustrating the variants of /ʡ/ and /ʜ/ in **(a)** /p'uʡʡu/ "pierce" (373), **(b)** /paʜʜina/ "hit" (359), **(c)** /ʡani/ "head" (11), **(d)** /ʜenta/ "flies" (113), **(e)** /nd̠oːʡo/ "floor, mud" (167) and **(f)** /ɬwaːʜe/ "pinch" (374).


(a)
8125
Freq.(Hz)
0
0
509 ms
p' ú ʔʔ u
2011
Freq.(Hz)
0
p' ú ʔʔ u
(b)
8125
Freq.(Hz)
0
0
751 ms
p a нн i n a
2011
Freq.(Hz)
0
p a нн i n a
(c)
8125
Freq.(Hz)
0
0
616 ms
ʔ a n i
2011
Freq.(Hz)
0
0
ʔ a n i
616 ms

In Table 1 both /ʡ/ and /ʜ/ are classified as voiceless. They are indeed usually phonetically voiceless in utterance-initial position and when geminated in intervocalic positions. In these cases, /ʡ/ is an epiglottal stop with no vocal cord vibration during the closure, and /ʜ/ is a voiceless epiglottal fricative. However, both segments often occur with partial voicing in initial position and are frequently voiced throughout as singletons in intervocalic position. Under these circumstances, /ʜ/ appears as a voiced epiglottal approximant accompanied by a quite marked lowering of the fundamental frequency (F_0) of the voice during its production, and /ʡ/ appears as an epiglottal tap or as an approximant with an even greater effect on F_0. Spectrograms illustrating these variants are shown in Figure 4. In each case, wide and narrow band analyses are shown side by side, so that time and frequency effects are both clearly visible. The wide band spectrograms show a frequency range from 0-8125 Hz, and the narrow band show 0-2011 Hz. The time scales vary depending on the word length, but each spectrogram shows between about 500 and 800 ms. Spectrograms 4a and 4b illustrate geminate intervocalic examples of /ʡ/ and /ʜ/ in /p'uʡʡu/ "pierce" (373) and /paʜʜina/ "hit" (359) (inflected verb forms are not glossed precisely). The respective stop and fricative characteristics of these segments are clearly observable in such contexts. Spectrograms 4c and 4d illustrate representative utterance initial tokens, in the words /ʡani/ "head" (11) and /ʜenta/ "flies" (113). In both words a period of low F_0 with a rather abrupt termination is observable in the narrow band spectrogram. At the beginning of /ʜenta/ partial voicing of /ʜ/ is apparent. Spectrograms 4e and 4f illustrate intervocalic tokens in the words /nd̠oːʡo/ "floor, mud" (167) and /ɬwaːʜe/ "pinch" (374). Vocal cord vibration comes close to ceasing altogether during this token of /ʡ/ but does not quite do so, whereas voicing is quite uninterrupted during the token of /ʜ/ shown. The considerable lowering of F_0 accompanying these segments is clearly shown in the spectrograms. A constant relationship seems to hold between these two segments, such that in any given context /ʡ/ has a more tightly constricted articulation than /ʜ/. This is a further reason for using the symbols /ʡ/ and /ʜ/ as they represent a difference in stricture rather than a difference in voicing. The tendency for these segments to occur as approximants intervocalically and for intervocalic voicing can be seen as part of a wider process of laxing that affects many of the obstruents in the language.

2.2 Consonant manners

2.2.1 Stops

Dahalo is rich in stop consonants and uses the full range of the usual linguistically-employed airstream processes in their production, namely pulmonic egressive (plosives), glottalic egressive (ejectives), glottalic ingressive (implosives) and velaric ingressive (clicks). It has 9 types of stops: plain voiceless, plain voiced, prenasalized voiced, prenasalized voiceless, voiced implosive, voiceless ejective, prenasalized voiceless ejective and voiced and voiceless nasalized clicks. These facts place Dahalo in a very rare class of languages. It is the only language we know of that uses such a wide range of types of stops distinctively.

2.2.2 Pulmonic stops and affricates

There are 17 non-prenasalized pulmonic stops and affricates in Dahalo. These are exemplified with word-initial examples in Table 3, except for /b/ and /d̠/. We found no examples of /d̠/ in initial position, and most of the medial examples are in Bantu loanwords where they correspond to voiced coronal obstruents. The prenasalized counterpart of /d2/ does appear initially, as in the word /nd̠oːʡo/ in Figure 4e. We have only one marginal example of /b/ in initial position, in the word for "scatter". This item is discussed further below. Tosco (1991) also reports no examples of the bilabial plosive /b/ in initial position; /ɓ/ is found in this position. Medially there is a contrast between /b/ and /ɓ/.

Table 3. Words illustrating central oral plosive and affricate segments in Dahalo. (The number after each word indicates its position in the recorded word list in Appendix 1).

PLOSIVES

p	pe:o	'broom'	(153)
b	d̪abi	'meat'	(144)
t̪	t̪á:mi	'grass, thatch'	(271)
t̠	t̠ó:i	'boat'	(199)
d̪	d̪aba	'hand'	(41)
d̠	kád̠i	'work'	(417)
k	kálat̪i	'teeth'	(21)
g	gát̪'a	'beard'	(1)
gw	gwat̪'ana	'chew'	(332)
kw	kwáʔana	'run'	(383)
ʔ	ʔágaddzo	'ear'	(4)
ɂ̵	ɂ̵ani	'head'	(11)

AFFRICATES

ts	tsa:ka	'hot season'	(251)
dz	dza:ɂ̵áta	'death'	(314)
tʃ	tʃoŋgo	(a small yellow bird)	(307)
dʒ	dʒúkku	'navel'	(51)
dzw	dzwa:ʜana	'wash wound'	(419)

The voiced stops in initial position (/d̪, g/) are quite often pronounced without any phonetic voicing during the closure, but the distinction between these stops and their voiceless counterparts is not neutralized. An audible distinction between them remains, probably due to the more intense release bursts associated with the voiceless series. In intervocalic position the voiced bilabial, dental and alveolar stops /b, d̪, d̠/ frequently occur as the corresponding approximants or weak fricatives, i.e. [β], [ð̪] and [ð̠]. Thus /d̪ába/ 'hand' (41) is usually [d̪áβa], although [d̪ába] is also heard. Most of the literature on Dahalo prior to Tosco (1991) records these intervocalic variants as distinct segments. We noted one word in which we observed an initial [β]; this was [β́ilaβ́ilina] "scatter" (387). Since /w/ in Dahalo is often pronounced with friction it is possible that this word should be analyzed as beginning with /w/ rather than with /b/. The auditory distinction between [ð̪] and [ð̠] is slight but listeners familar with Danish may note the similarity between the intervocalic allophone of Danish /d/ and the Dahalo [ð̠] phone. The voiced velar stop does not undergo a similar laxing.

We prefer to recognize two series of prenasalized stops, which we interpret as being distinguished by the voicing of the stop closure. Examples are given in Table 4. In contrast, Nurse, Tosco and other sources recognize only one prenasalized stop series, a voiced one, and prefer to analyze the voiceless counterparts as a sequence. The 'voiceless' prenasalized stops (and affricates) always have a voiceless stop closure portion, and this tends to be longer than the stop closure in the 'voiced' cases. The release of the 'voiceless' cases is noisier than that of the voiced ones. However, the 'voiced' set are quite commonly pronounced with a devoiced stop closure and release, reducing the phonetic distinction between the series. That there is an underlying distinction is clear from the fact that the voiceless ones are never voiced; only the 'voiced' ones show variation. However, it is not always possible to determine which occurs in a particular word if it is heard only once.

Table 4. Prenasalized stops and affricates.

mp	mpaɲe	'rat'	(100)
mb	mberewere	'flower of maize'	(268)
n̪t̪	ʜén̪t̪a	'flies' (noun)	(113)
n̪d̪	n̪d̪a:re	'cattle egret'	(290)
n̠t̠	n̠t̠ét̠e	'maize'	(275)
n̠d̠	n̠d̠ó:ʔo	'floor'	(167)
n̠d̠w	n̠d̠wa:la	'mucus'	(16)
nts	tsintso	'vagina'	(65)
ndz	ndzo:me	'udder'	(106)
ŋk	ŋkuko	'chicken'	(291)
ŋg	ŋgíkíne	'eyebrow'	(6)
ŋgw	ŋgwale	(sp.bird with red legs)	(310)

Elderkin (1976) interprets voiceless prenasalized stops as the geminate equivalents of the voiced ones. He gives as two of his reasons the absence of voiceless prenasalized stops in word-initial position, and the absence of long vowels before them in word-medial position. We do not find these distributional reasons persuasive. Our data set illustrates the fact that Dahalo speakers have a variety of prenasalized voiceless stops word-initially. While it is true that the majority of the words involved can be fairly easily traced as Bantu loans, these items seem well-entrenched. Equally the medial cases of voiceless prenasalized stops that we have observed do not seem to be functioning as geminates - they are not noticeably long, for example, in words such as /ɟun̪t̪á:ʔa/ "small grasshopper" (126), /kóŋko:lo/ "shin" (37), and /ɬímpínna/ "blow (nose)" (328). A prohibition on long vowels before geminates is quite commonly found in languages, including many Afro-Asiatic ones, and it seems to apply in Dahalo. However, we noted at least one case where a long vowel precedes a voiceless prenasalized stop, /wu:mpi/ "dust" (226), and at least one case in which a distinctive geminated voiceless prenasalized stop occurs — /ʔúk'en̪t̪t̪ina/ "lift" (367) — in which it is the oral part of the stop that seems to be lengthened, perhaps as a result of suffixation with /-tina/.

An alternative to analyzing Dahalo as having a voiceless prenasalized stop series is to view the words in question as containing a syllabic nasal followed by a voiceless stop consonant. This suggestion is made by Elderkin, Ehret & Nurse (1989) for words like /mpáɲe/ (100) 'rat'. If these nasal elements were syllabic, it might be expected that they would have the potential for independent tonal contrast. We found no evidence for this. Only one nasal in our Witu recording was transcribed as tone-bearing, and that was in /ḿso/ "chick" (293) where the nasal is not part of a prenasalized stop. Tosco (1991: 5) also remarks that "a few Bantu words have retained a syllabic nasal", giving examples such as /m̩fállume/ "king". The cases he cites are also ones that do not involve prenasalized stops. We therefore reject this alternative. There is one further interesting fact connected with gemination processes that bears on the interpretation of prenasalized stops; we will discuss this is section 4 below.

In the plosive series, only the velar place shows a systematic contrast between plain and labialized stops. This contrast also extends to the velar ejectives.

The glottal stop can be viewed as an automatic feature at the onset of words that have no other initial consonant, and glottal stops are sometimes pronounced in word final position. In word-internal position, glottal stops are distinctive. Examples include /maʔa/ "water, rain" (241), /ɓaʔanna/ "set out to dry" (323), /t̠'iraraʔe/ "red" (427). The pronunciation of such words ranges from containing a complete glottal closure to containing only a weak increase in the degree of vocal cord constriction which modifies but does not interrupt the voicing of the vowels, and lowers the fundamental frequency over the affected interval. Contiguous vowels in separate syllables are not separated by a glottal stop, e.g. /kw'ai/ "shaft of arrow" (197), /peːo/ "broom" (153), /ɗau/ "cooking pot" (158). Tosco (1991:12) writes a glottal stop deletion rule that would optionally remove any intervocalic glottal stop. Our impression is that the stop may be greatly weakened, but that it is not actually deleted.

2.2.3 Ejective stops and affricates

Both ejective stops and ejective affricates occur in Dahalo. Examples are given in Table 5. A bilabial ejective stop is illustrated in Figure 4a above in the word /p'uʔʔu/. We also noted a small number of ejective prenasalized stops, e.g. /hant̪'inna/ "scratch" (389), /gimp'o/ (46) "knee". Only a few of the plain ejectives are matched by a prenasalized counterpart in our data, but we suspect that this is a result of accidental gaps in our data collection rather than a systematic restriction.

Table 5. Word illustrating ejective segments in Dahalo.

p'	p'áːra	'termites'	(129)
t̪'	t̪'at̪t̪a	'hair'	(10)
t̠'	t̠'irimalle	'spider'	(128)
tɬ'	tɬ'aːʔa	'lake'	(232)
cç̥'	ʔacç̥'áno	'semen'	(62)
k'	k'aba	'three'	(261)
kw'	kw'aːi	'shaft of arrow'	(197)

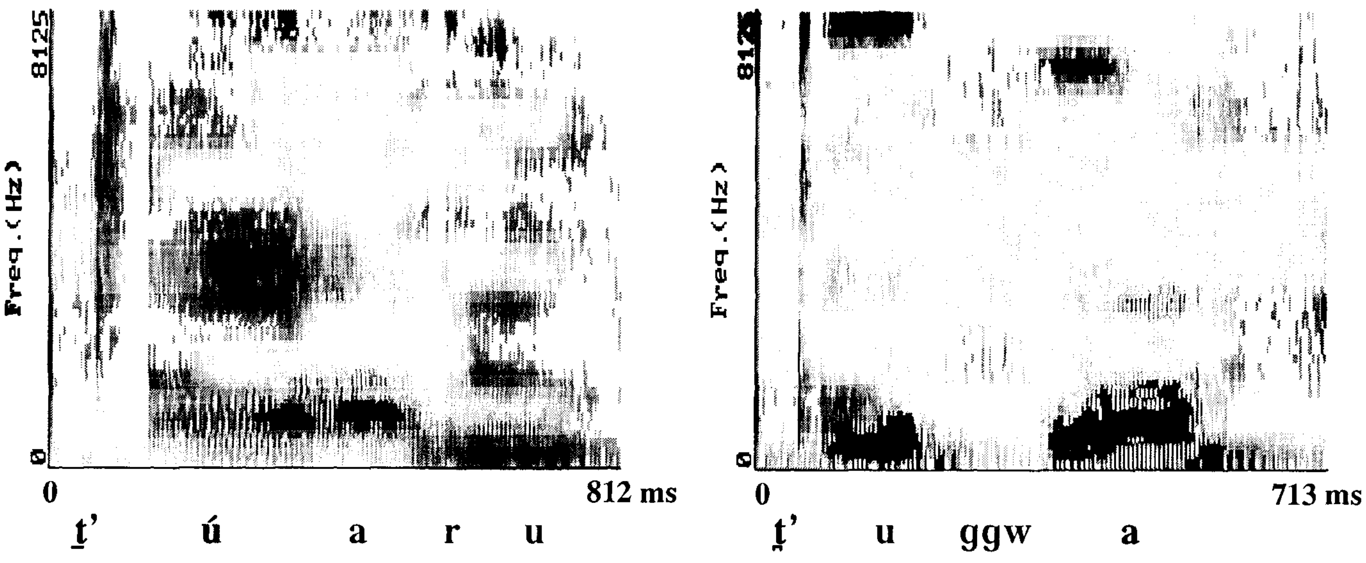


Figure 5. Spectrograms to illustrate alveolar and dental ejective stops in /t̠'úaru/ (species of ruddy bird) (312) and /t̪'uggwa/ "smoke" (247).

Contrary to previous accounts, we noted both dental and alveolar ejectives. These are auditorily distinguished in a similar fashion to the dental and alveolar plosives, namely, the alveolar has a noisier and more fricated release. Spectrograms of the words /t̪'uggwa/ "smoke" (247) and /t̠'úaru/ (species of ruddy bird) (312) are shown in Figure 5 to illustrate this contrast.

A palatogram of /t̪'úːt̪'o/ (species of small antelope) (104) is shown in Figure 6. As can be seen, the dental ejective /t̪'/ spans the dental and alveolar region and the tongue makes some contact with the underside of the upper lip. Thus, the articulation is very similar to that observed for the voiceless dental plosive in Figure 1. We do not have any palatographic data on the alveolar ejective, but would expect its oral articulation to be similar to that of the alveolar plosive. The word that we have used to illustrate this sound in Figure 5 was transcribed by Elderkin as having a (palato-alveolar) ejective affricate, but we feel that the frication is too short to merit being described as an affricate. The only ejective affricates we observed were those with lateral release.

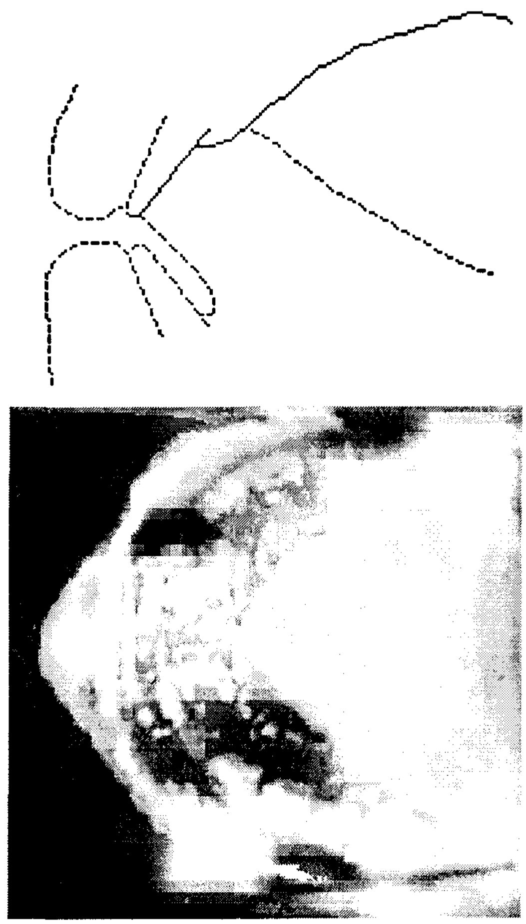


Figure 6. Palatogram of the Dahalo word /t'úːt'o/ (species of small antelope) (104).

Our impression, supported by acoustic analysis, is that there are two different lateral ejective affricates in Dahalo, one of which appears to have a more retracted articulation than the coronal consonants so far described. We find this is best characterized as a palatal lateral ejective, with articulatory contact centered in the palatal area, but extending over much of the roof of the mouth. We therefore write /cʎ̥'/ for this sound. The other we transcribe as the more customary /tɬ'/. Spectrograms of the words /ʔacʎ̥'áno/ "semen" (62) and /tɬ'aʔa/ "lake" (232) are shown in Figure 7. The extensive raising of the second formant of the vowel /a/ adjacent to /cʎ̥'/ is absent next to /tɬ'/, indicating this difference in the place of articulation. There is a parallel difference in place between two lateral fricatives.

These lateral ejectives have loud frication giving them a very marked auditory quality that makes an impression somewhat reminiscent of a click. Similar sounds are found in Hadza and Sandawe. In order to produce tokens of these sounds that were satisfactory to our informants, we found that it was necessary to make the location of the lateral escape quite far back in the mouth, maintaining the lateral seal to about the region of the third molar.

Tosco also reports the presence of a voiced laterally-released affricate (or plosive) in Dahalo, but we collected no examples.

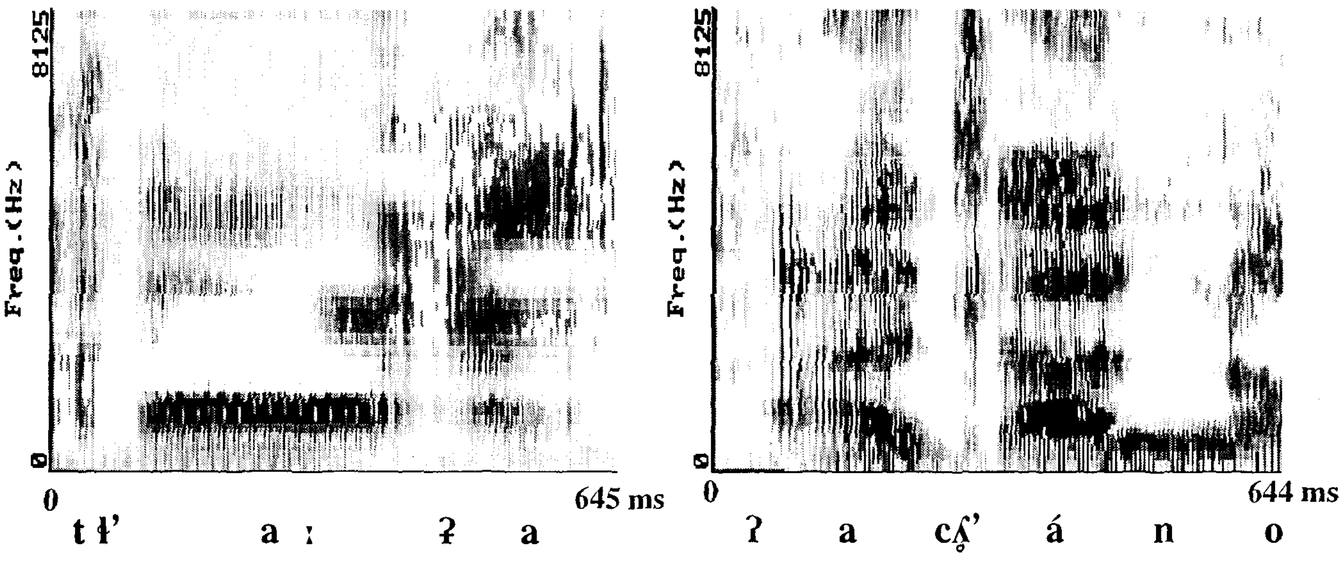


Figure 7. Spectrograms of and /tɬ'aʔa/ "lake" (232) and /ʔacɬ̥'áno/ "semen" (62), illustrating the two types of lateral ejective affricates.

2.2.4 *Implosives*

Implosives occur only at the bilabial and alveolar places of articulation, and not at the dental or velar places. Examples are given in Table 6. Our observations agree with Tosco (1991), although Ehret, Elderkin and Nurse (1989) suggest that the plosive /g/ might in fact be implosive [ɠ] in some of the Bantu borrowings. This was not found to be the case with the Dahalo speakers we worked with. The implosives are voiced throughout the closure in intervocalic position, and even in initial position voicing begins 10-25 ms before the oral release.

Table 6. Words illustrating implosive segments in Dahalo.

ɓ	ɓaːɓa	'father'	(207)	**ɗ**	ɗíme	'bee'	(110)
	ɓoːt̪i	'thigh'	(56)		ɗóːɗaːma	'topi antelope'	(105)

2.2.5 *Clicks*

A striking feature of Dahalo is the occurrence of clicks. It is the most northerly language in Africa to utilize this kind of sound in its phonology. Clicks only occur in a relatively small number of words — many fewer than in Hadza and Sandawe (the two other East African languages with clicks) — but a good proportion of these words would be considered to belong to the basic vocabulary with meanings like "saliva", "shit", "forest", "breast", "star". We elicited as many words as possible containing clicks, but could only obtain 40. These include one, /ŋǀát'u/ "constipation" (68) that had not been reported previously in the literature. In some other cases we were able to clarify the meanings and transcription. The glossary in Tosco (1991) lists only 26 words with clicks; Ehret, Elderkin and Nurse (1989) list 58 (19 of which we were unable to elicit).

Following usual practice we describe clicks in terms of a primary class or type and an accompaniment (Ladefoged and Traill 1993). Click type refers to the primary oral articulation in the front part of the oral cavity, while accompaniments are features of the back closure and the laryngeal and velo-pharyngeal settings that co-occur with the click. Dahalo has only one click type, the dental [ǀ]. The accompaniment is always nasalized, but may be voiced or voiceless with the voiceless option occurring much more frequently. Waveforms illustrating the voicing difference in the nasal accompaniment to the clicks are shown in Figure 7. In /ŋ|aba/ "forest" the voiced nasal starts substantially before the click burst, while in /ŋ̊|abate/ "good smell(ing)" a delay of about 30 ms occurs between the click burst and the onset of voicing. Voiceless nasalization can be auditorily detected principally through its coarticulatory effect on adjoining vowels, which display a brief nasal on- or off-glide or receive light nasalization.

It should be noted that considerable fluctuation was observed in the voicing of the click accompaniment. Although certain words, such as /ŋ|aba/ "forest" (257), were almost invariably pronounced with voicing, others, such as /ŋ|uʔite/ "bitter" (437) varied even for the same speaker. Furthermore, many tokens were recorded in which voice onset and click release are almost simultaneous; it was difficult to classify these as either clearly voiced or voiceless. The speakers recorded at Kipini showed greater variation than those recorded at Witu. Since these speakers were younger it is possible that the voicing distinction in words with clicks is in the process of being lost.

Clicks can also occur with labialization on the release, yielding a potential four-way contrast between voiceless nasalized [ŋ̊|], voiced nasalized [ŋ|], voiceless nasalized labialized [ŋ̊|w], and voiced nasalized labialized [ŋ|w], as shown in Table 7. The labialized cases are extremely few in number; variability in maintaining labialization was also noted. The majority of clicks occur in word initial position, but medial clicks occur in a few words such as /ʜaŋ̊|ana/ "lick" (366), /meŋ̊|ete/ "carefully" (435) and /fuŋ̊|inna/ "root up" (382).

Table 7. Words illustrating clicks in Dahalo. The number after each word indicates its position in the recorded word list

ŋ̊|	ŋ̊|ó:ne	'breast'	(26)
ŋ|	ŋ|í:t'i	'gums'	(9)
ŋ̊|w	ŋ̊|waɲi	'saliva'	(19) (in Kipini)
	ŋ̊|wá:ʔana	'knead'	(321) (in Witu)
ŋ|w			no examples in our data

A palatogram of the word /ŋ̊|aba/ "good smell" (184) is shown in Figure 8. The contact area is broader in this articulation than in the alveolar plosive illustrated in figure 1, but much less extensive than in the dental plosive in the same figure. Rather, the contact area is similar to that seen in the affricate /ts/ in figure 5. Again, the pattern seems to be that when no contrast obtains between dental and alveolar, the articulation is somewhat intermediate. A noteworthy aspect of the palatogram is the absence of any indication of the contact for the back closure that is required in the production of clicks. Clearly this contact must be quite far back on the roof of the mouth and/or quite short in the mid-saggital plane. A relatively large pocket of air remains between the two closures during the most constricted phase in the production of this click. In other languages (Ladefoged and Traill 1993; Sands, Maddieson and Ladefoged, this volume) the air pocket between the closures appears to be typically much smaller either because the back closure is centered further forward or is broad and extends further forward.

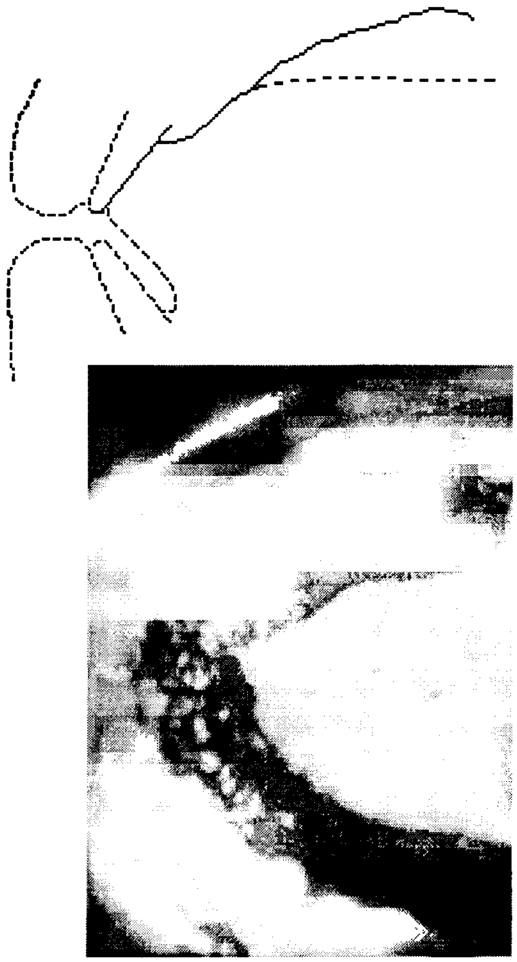


Figure 8. Palatogram of /ŋ̊|aba/ "good smell".

Spectrograms and waveforms of the words /ŋ̊|abate/ "good smell" (184) and /ŋ|aba/ "forest" (257), contrasting voiced and voiceless nasal accompaniments are shown in Figure 9. In this token of /ŋ̊|abate/ the click release can be seen to occur well before the onset of voicing. It is also possible to detect that there is a short voiced nasal portion before the velar contact is broken at the onset to the vowel. This is one common pattern in the production of the voiceless nasalized clicks. In /ŋ|aba/ on the other hand there is an early onset of the voiced nasal, well before the click is released. The nasal continues to be held for a short interval after the click release occurs.

A spectrogram of the labialized click in /ŋ̊|waːʔana/ "knead, massage" (321) is shown in Figure 10. Here again the release of the click precedes the onset of voicing and there is a brief voiced nasal portion before the vowel. At the vowel onset the second formant is lower than in the non-labialized cases in Figure 9.

2.3 Fricatives

Dahalo can be considered to contain 10 fricatives; examples are given in Table 8. /f/ and /s/ are relatively common, but /ʃ/ occurs only in recent loanwords. There are also three different lateral fricatives, plus the two rather special cases of /h/ and /ʜ/. With one marginal exception, there is no voicing contrast in fricatives. The /z/ in /ziːwa/ 'milk' (27) is the only instance we found of this segment, but it serves to establish a voicing contrast with /s/. However, this word is a Swahili loan, almost certainly of recent date. In other Bantu loanwords where Swahili has /z/, the Dahalo borrowing has /d̪/. /f/, /s/ and /ʃ/ are invariably voiceless, but /h/ and /ʜ/ show intervocalic voicing (see the discussion of epiglottal place above). The /h/ appears to occur optionally as a fully voiced segment [ɦ] even in word-initial position, whereas /ʜ/ is usually no more than partially voiced in this position..

Our impression is that the lateral fricatives display a place contrast between alveolar and palatal (or palatalized alveolar), which we transcribe as /ɬ/ vs /ʎ̥/. This has not been remarked on in the previous literature on the language, but is parallel to our observations on the lateral ejectives. In some words, notably /ʎ̥aːbu/ "leaf", there is an audible [i]-like quality at the release of the lateral. At first we were tempted to transcribe this word as containing a sequence of vowels, i.e. as /ɬiabu/, but the movement seems too rapid for this interpretation. Instead, it seems that some inherent property of the initial consonant is affecting the onset of the long /aː/. The spectrograms in Figure 11 compare /ʎ̥aːbu/ and /ɬáʔana/ "roast" (380) spoken by the same speaker. The second formant originates much higher in /ʎ̥aːbu/ than in /ɬaʔana/, and falls abruptly. We noted only one other word with /ʎ̥/, namely /ʎ̥ákane/ "sharp" (426). The labialized lateral fricative is illustrated in Figure 4f above.

Table 8. Words illustrating fricatives in Dahalo.

f	fiːt̪'e	'whistle'	(416)
s	saʡála	'four'	(262)
z	ziːwa	'milk'	(27)
ʃ	ʃóːt̪oni	'left'	(72)
ɬ	ɬunno	'stew'	(148)
ɬw	ɬwaːʜe	'to pinch'	(374)
ʎ̥	ʎ̥aːbu	'leaf'	(273)
ʜ	ʜelléʡa	'zebra'	(108)
h	hîbe	'baboon'	(79)

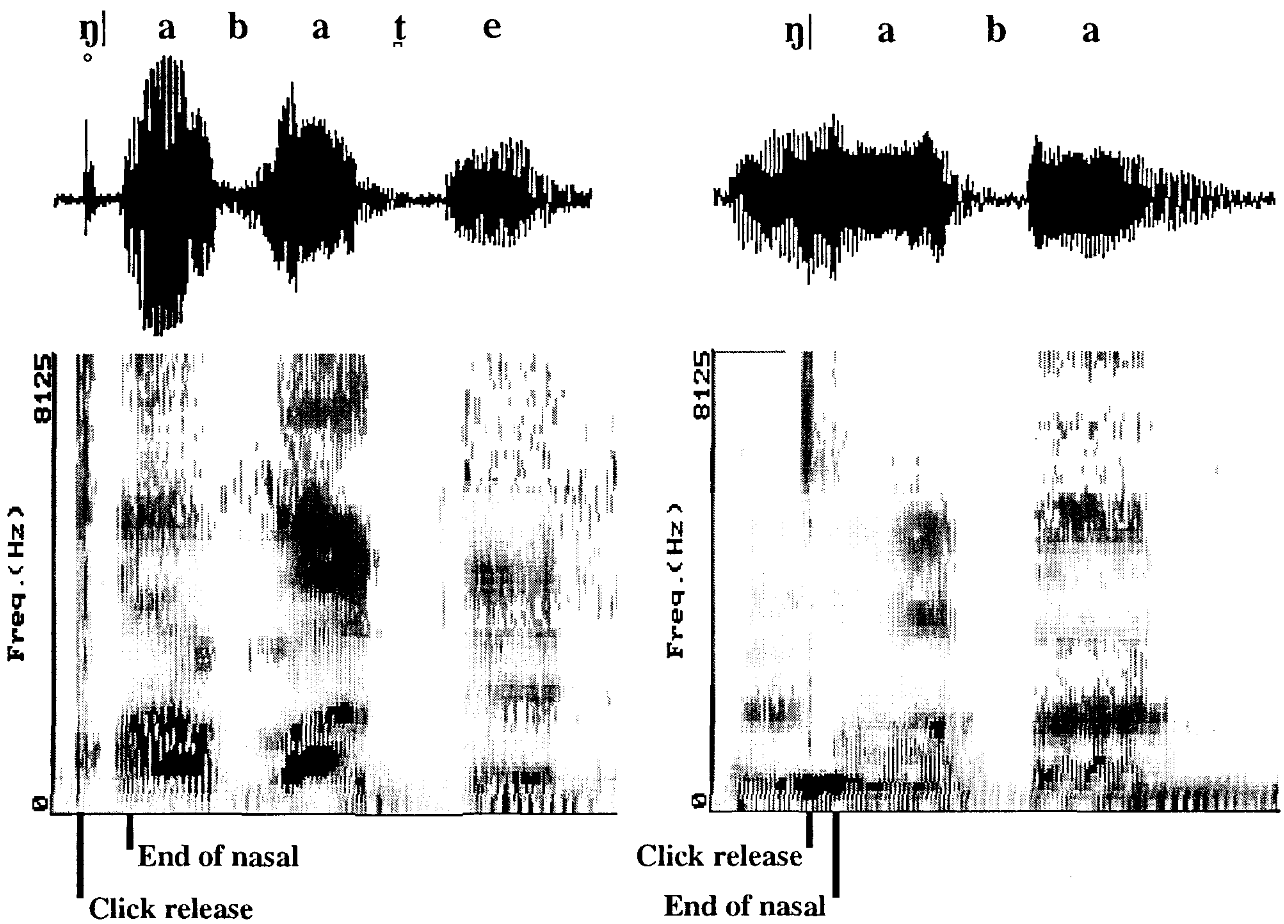


Figure 9. Spectrograms and waveforms of the words /ŋ|aba/ "good smell" (184) and /ŋ̊|aba/ "forest"(257)

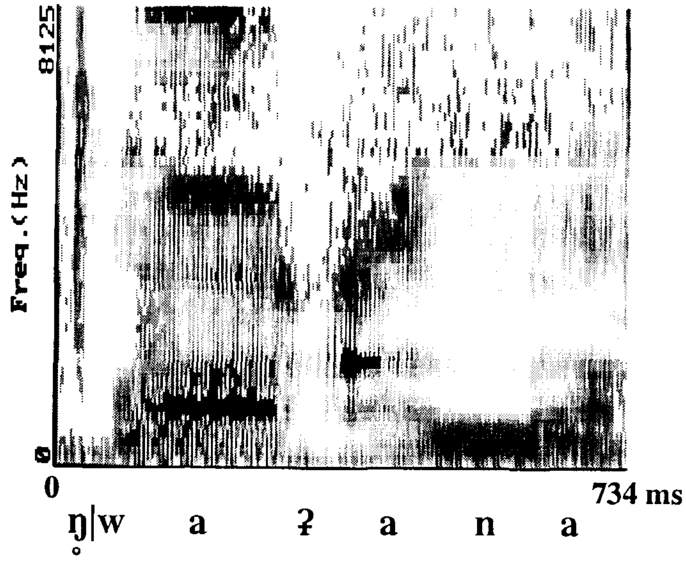


Figure 10. Spectrogram of /ŋ̊wa:ʔana/ “knead, massage”

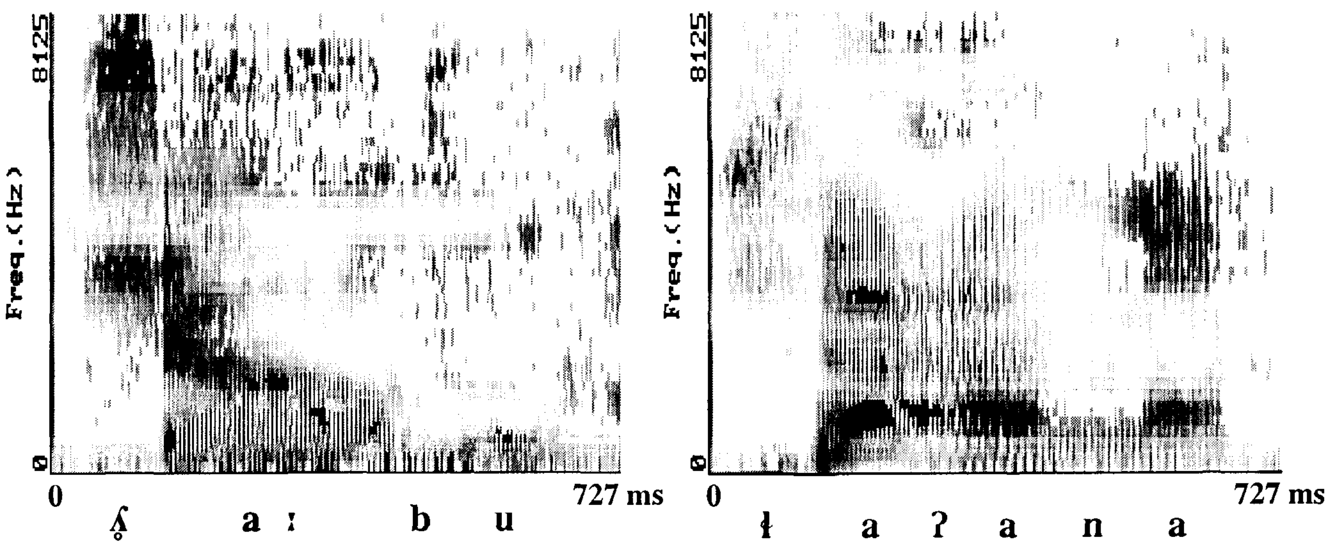


Figure 11. Spectrograms of /ʎ̥a:bu/ “leaf” and /ɬáʔana/ “roast” (380)

2.5 Nasals

Dahalo has three nasals, /m, n, ɲ/. Examples of these are given in Table 9. There is no dental/alveolar contrast among nasals, nor does a velar nasal occur, except as part of a complex consonant such as a prenasalized stop or a click. (We have already commented on the nasalization that accompanies all the clicks of Dahalo, and of course, a wider range of phonetic nasal elements occurs as the nasal part of the prenasalized stops.) Although /n/ is more palatal preceding /i/, there is nonetheless a distinction between /n/ and /ɲ/ in this environment, illustrated by the difference between /kwé:ni/ “water pot” (182) and /ɬwé:ɲi/ “arrow” (193).

Table 9. Words illustrating nasals in Dahalo.

m	muna	‘heart’	(30)
n	nóʔona	‘suck’	(408)
ɲ	ɲáɦe	‘hippo’	(89)

2.6 Liquids and Approximants

Dahalo has two liquids and two central approximants. Examples are given in Table 10. There is only one voiced lateral approximant, i.e. there is no voiced counterpart to the palatal affricate and fricative laterals. /l/ seems to be articulated less far forward than the dental stops. /r/ is a trill or single tap. /w/ has little noticeable rounding, and is often lightly fricated. Since /w/ has a specified back tongue position (though not an extreme one) it remains distinct from the intervocalic allophone [β] of /b/. Because of the frication, Tosco (1991) prefers to write /v/, but the sound is not a labiodental. The palatal approximant /j/ has a very limited distribution, occurring only in the word /jáːjo/ “mother” (212) and the derivative /jaːjoni/, glossed as “(my) parents” (215).

Table 10. Words illustrating liquids and approximants in Dahalo.

w	wat̪t̪úkwe	‘one’	(259)
r	rik’a	‘tail’	(54)
l	ĺiːma	‘two’	(260)
j	jáːjo	‘mother’	(212)

3 Vowels

Dahalo has 5 different distinctive vowel qualities, each of which occurs short and long. Examples are given in Table 13.

Table 11: Words illustrating the long and short vowels of Dahalo.

vowel	short			long		
i	ḱit̠i	‘chair’	(155)	ḱiːbu	‘calabash’	(154)
e	nt̠ét̠e	‘maize’	(276)	ʔeːga	‘fire’	(165)
a	t̪’at̪t̪a	‘hair’	(11)	tsaːka	‘hot season’	(251)
o	k’ok’o	‘throat’	(49)	ɓóːku	‘hole’	(230)
u	nd̠upa	‘bottle’	(152)	t’úːt’o	‘sp.small antelope’	(104)

We examined the formant structure in a subset of our data, illustrating these possibilities. For each of our three Witu speakers there were two tokens of each of two words of each long and each short vowel, including the items in Table 11 above. Our data set did not lend itself to producing nicely matched minimal sets, but we chose words in which there were no surrounding consonants likely to cause major allophonic differences. In particular, words with nasals, liquids or epiglottal consonants next to the vowel to be measured were avoided. Measurements of the first three formants were made, using spectrographic displays and LPC spectra on a Kay CSL system. An analysis of variance indicated that there were no significant differences between the vowel qualities of the long and short vowels, as determined by the frequencies of the first and second formants. Accordingly long and short vowels are pooled for the formant plot in Figure 12, which shows F1 (roughly equivalent to vowel height) vs. (F2 – F1) (roughly equivalent to vowel backness) for the 24 tokens of each vowel. The axes are scaled in proportion to the Bark scale, but are labeled in Hz. The means are marked by larger points, and an ellipse with a radius of two standard deviations along the principal components of the scatter has been drawn around the mean of each vowel. Considering the fact that there are three speakers, one male and two female, the scatter for each vowel is not very large.

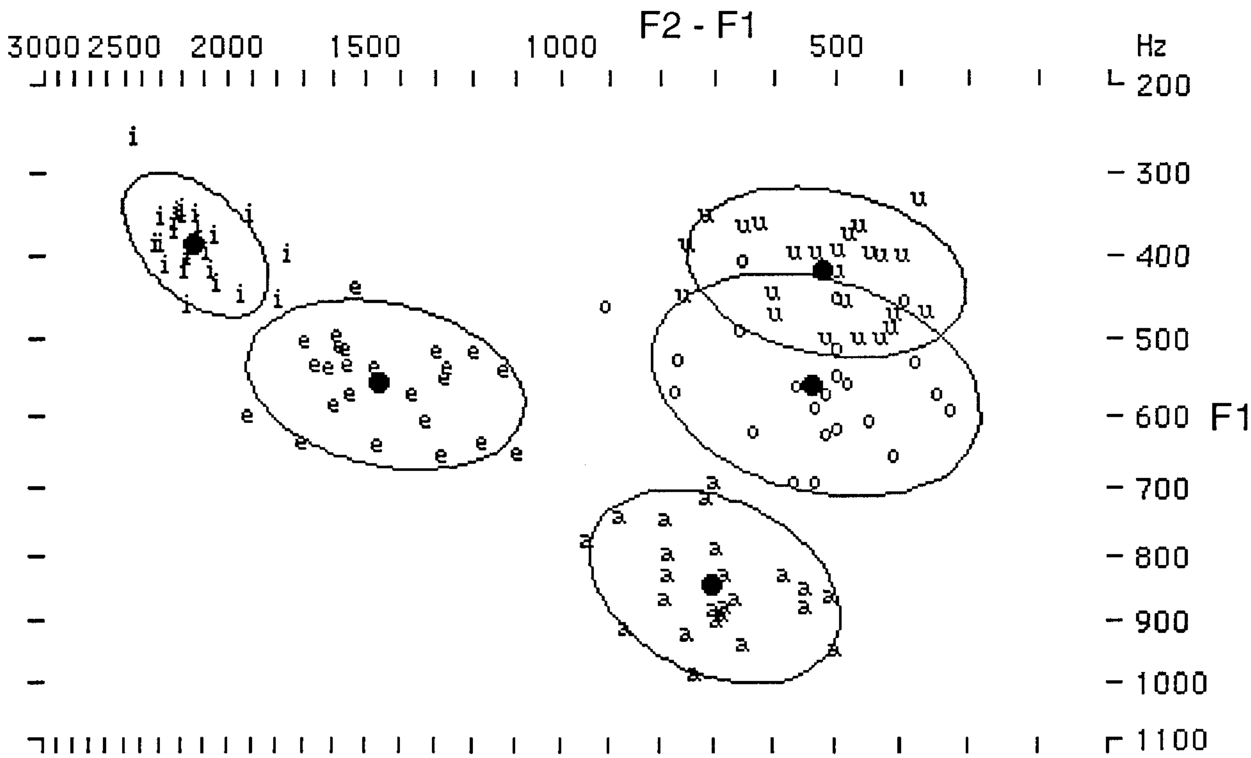


Figure 12. Formants of the Dahalo vowels from three speakers (see text for details).

The most notable point in the formant chart is the lack of separation of the two high back vowels. They are the only pair of vowels with clearly overlapping distributions. Some of this may be attributed to differences among speakers, and some to allophonic variation associated with the particular context in which each vowel appears but the contextual variation (and, of course, the differences among speakers) were equally great among the other vowels. The high back vowels are clearly very similar. We ourselves found that when listening to the language we sometimes had difficulty in deciding which of the high back vowels occurred in a given word.

The only major source of allophonic variation in vowel quality is the epiglottal consonants, which tend to color following vowels quite dramatically. For example, /a/ following an epiglottal consonant will tend to sound more fronted, like [æ]; /i/ sounds lowered and similar to /e/ in other contexts; /u/ sounds fronted and lowered; and /o/ has a fronted quality. The acoustic basis for some of these auditory impressions can be seen in the formant patterns in the words shown in Figure 4. A similar phenomenon is also seen in Arabic and in some of the Caucasian languages with pharyngealized sounds. Elderkin (1974) claimed that the vowels following the epiglottal consonants tend to be nasalized, but we did not notice this.

A specific alternation was noticed in the word /kweːni/ "water pot" (182); the labialization of the initial consonant can envelope the vowel which then occurs as the rounded vowel [øː],

masking the labialization on the consonant. We only noticed this effect with this labialized velar, and not in the structurally similar word /ɬweɲi/ "arrow" (193).

4. Gemination

There are lexical contrasts between single and geminate consonant structures in intervocalic position for most of the consonants of Dahalo, thus /dʒúkku/ "navel", /luka/ "thigh". As in most Afro-Asiatic languages, gemination of consonants also plays an important role in the morphology of Dahalo (Elderkin 1974, Zaborski 1986, Tosco 1991). For example, several patterns of noun plural formation involve a reduplicative suffixation which adds -VCCV to the noun stem. In this formula -CC- is a reduplicated copy of the last stem consonant. Some examples, drawn from Tosco (1991: 23-25) but retranscribed according to our conventions, are given in Table 12.

Table 12. Reduplicative plurals involving gemination (after Tosco 1991).

Singular	Plural	Gloss	
d̪ába	dábabbe	"hand"	(41)
nd̪é:gi	nd̪é:gaggi	"canine tooth"	not in our data
kálat̪i	kálat̪ét̪t̪o	"tooth"	(21, 86)
hé:ri	hé:rarre	"sheep"	(88)
filime	filimámmi	"comb"	(157)
gát'a	gát'att'i	"beard"	(1)
kit̠ánd̠a	kit̠and̠ád̠d̠i	"bed"	not in our data
tumpi	tumpáppi, tumpabbi	"horn"	(90)

The last two items in Table 12 illustrate the behavior of prenasalized stops in this reduplicative pattern. Only the oral stop component is copied and geminated. This provides a basis for an argument that the prenasalized stop series are, in fact, sequences of segments. However, as both voiced and voiceless prenasalized stops behave in the same fashion, it suggests that the difference between them is not between a single segment on the one hand and either a geminate or a sequence on the other.

A characteristic of verbs in a number of the tense/aspect paradigms is to show third singular masculine subject by, *inter alia*, gemination of the final stem consonant. Thus, in the 'general non-past' paradigm of the verb /lub-/ "to hit" first person singular is /lúbo/, but third person masculine is /lúbbi/ (Tosco 1991: 51-59). Geminates are also derived by assimilatory processes in a number of verbal inflectional patterns; particularly when a consonant-initial affix is added to the stem of a verb that has a final dental obstruent. Many of the geminates in the verbal forms cited in Appendix 1 arise from this process.

The durations of single and geminate stops were compared by measuring the stop closures in a small subset of words containing single and geminate dentals and velars. Two repetitions from three speakers of a pair of words in each category were measured, yielding 18 measurements per cell for comparison. The results are given in Table 12. Not surprisingly, geminates are significantly longer than single stops ($F(1, 68) = 46.908$, $p = .0001$), but by differing amounts depending on place. The mean dental geminate duration is only about 1.2 times longer than the singleton but the mean velar geminate duration is about 1.7 times the singleton duration. In this data set the interaction between place and consonant quantity is significant ($F(1,68) = 8.465$, $p = .0049$), but since the words used were not fully matched for vowel environments and syllable pattern, this result should be interpreted with caution.

Table 12. Mean durations (in milliseconds) of Dahalo single and geminate voiceless dental and velar plosives.

n = 12	single	geminate
dental	123.2	149.7
velar	89.6	155.4

The VOT of these same voiceless stops was also measured to see if Dahalo displayed any sign of the differences in the relative timing of oral and laryngeal articulations that are a feature of single vs geminate stop distinctions in a number of other languages. The results showed almost identical means for single and geminate cases, around 13.5 ms for dentals and 21 ms for velars, showing that there is no such difference in Dahalo. We may also note here that voiced geminate stops tend to devoice approximately half way through their production. This is a presumably for aerodynamic reasons, rather than due to actively changing the laryngeal setting. The vocal cords stop vibrating when the air pressure of the cavity above the glottis is close to the air pressure in the cavity below the glottis.

5. Suprasegmentals

Dahalo is a language with pitch accent. Tosco (1991) while admitting that "a lot of work remains to be done in this field" has advanced the understanding of Dahalo tonal processes considerably. In our transcriptions we have assigned each syllable a high (H) or a low (L) tone, marking only the high with an acute accent. The most frequent pattern is for a high tone to fall on the first syllable. Cases where there is more than one high tone on a non-compound word are rare, but all-low words are not unusual. The narrow-band spectrograms provided in Figure 4 enable a few of the tonal properties of the language to be noted. Figure 4c shows a LL word /ʔani/ "head" (11), while 4a shows a HL word /p'úʡʡu/ "pierce" (373). These are the only permitted disyllabic patterns. Generally the low tone immediately following a high shows a rather slow decline of pitch over the syllable.

6 Concluding Remarks.

The richness of Dahalo phonetics is far from being fully explored in this paper, but we do feel that we have been able to add to the understanding of this aspect of the language. As noted, Dahalo has an unusually rich inventory of stops, appears unique in having only nasalized clicks, and has clear epiglottal stops as well as a type of lateral ejective not previously described. We hope that we, or others, will be able to pursue a fuller study of some of these matters.

Acknowledgements

The research reported in this paper would not have been possible without the help of many people. We are very grateful to all who assisted in any way, but most especially to the Dahalo speakers who shared their knowledge of their language with us in very generous fashion. We would also like to acknowledge the generosity of Edward D. Elderkin and Derek Nurse in sharing their own work with us, Lucia Omondi for assisting with research permission for our work in Kenya, Nico Kalama for providing essential contacts in Witu, Haji Omar Abdallah and Ebrahim Khamisa for hospitality and research assistance in Kipini, and Ali Mirza for assistance with transportation. The research reported here was supported by a grant from the National Science Foundation, as described in the Preface.

References

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

Ehret, C., Elderkin, E.D., and Nurse D. (1989). Dahalo Lexis and its Sources. *Afrikanische Arbeitspapiere*. 18: 5-49.

Ehret, C. (1980). *The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary.* Dietrich Reimer Verlag, Berlin.

Elderkin, E.D. (1974). *The Phonology of the Syllable and the Morphology of the Inflected Word in Dahalo*. Unpublished M.A. thesis, University of Nairobi, Nairobi.

Elderkin, E.D. (1976). Southern Cushitic. In *The Non-Semitic Languages of Ethiopia* (ed M. L. Bender). African Studies Center, Michigan State University, East Lansing, MI: 278-297.

Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica* 2: 42-59.

Greenberg, Joseph. H. (1963). *The Languages of Africa.* Indiana University, Bloomington IN.

Grimes, Barbara F (ed). (1992). *Ethnologue: Languages of the World.* Summer Institute of Linguistics, Dallas.

Heine, Bernd. (1974). Notes on the Yaaku language. *Afrika und Übersee* 58: 27-61.

Ladefoged, Peter and Maddieson, Ian. (1986). (Some of) the Sounds of World's Languages. *UCLA Working Papers in Phonetics* 64.

Ladefoged, Peter and Traill, Anthony. 1993. Clicks and their accompaniments. *Journal of Phonetics.* 21.

Laufer, Asher and T. Baer. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech* 31: 181-205.

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

Maddieson, Ian. (1984). *Patterns of Sounds.* Cambridge University Press, Cambridge.

Nurse, Derek. (1986). Reconstruction of Dahalo history through evidence from loanwords.*Sprache und Geschichte in Afrika (SUGIA)* 7.2: 267-305.

Tosco, Mauro. (1989). The classification of Dahalo: another perspective. Paper presented at the Second International Symposium on Cushitic and Omotic Languages, Turin, November 16-18, 1989.

Tosco, Mauro. (1991). *A Grammatical Sketch of Dahalo.* Helmut Buske Verlag, Hamburg.

Tucker, Archibald and Bryan, Margaret A. (1977). The East African click languages: a phonetic comparison. In *Zur Sprachgeschichte und Ethnohistorie in Afrika* (ed. W.J.G. Möhlig, F. Rottland, B. Heine). Dietrich Reimer, Berlin: 300-323.

Zaborski, Andrzej. (1986). *The Morphology of Nominal Plural in Cushitic Languages* (Beiträge zur Afrikanistik 28). Institut für Afrikanistik und Ägyptologie, Vienna.

Appendix 1.

Dahalo word list

The following appendix provides a broad transcription of a recording made in Witu, Kenya in July 1991, together with the transcriptions provided by other authors for the same words. The first column is the number of the word in the recorded sequence. The second column contains the transcriptions used in this paper (MSSL) Maddieson, Spajić, Sands, and Ladefoged (1993), with the accompanying glosses in the third column. In a few case variants heard in the shorter list recorded in Kipini are noted. Additional columns show the forms of these words reported by other linguists. The fourth and fifth columns contain the transcriptions and glosses found in (EEN) Ehret, Elderkin, Nurse (1989), which is a compilations of the work of several authors. The sixth column notes the source of their transcription; (DE)=Derek Elderkin, (CE)=Christopher Ehret, (DN)=Derek Nurse, (AT)=Archibald Tucker. All [j]'s in their transcription have been replaced with [dʒ], and length on consonants has been marked by doubling the consonant rather than by using length marks. The seventh and eighth columns contain transcriptions and glosses from Tosco (1991). Note that he marks alveolars, whereas dental coronals are left unmarked. Nouns are given in both singular and plural forms in the EEN and Tosco columns when both are provided in these sources.

The verbs in the second column, glossed as bare verb stems in the third column, were elicited using infinitive Swahili forms, but in most cases our consultants gave inflected forms. The majority of these seem to be first person past perfect forms, but as there are other possibilities we were not able to fully gloss the forms without having elicited a more complete paradigm.

Copies of this tape recording will be made available to other interested researchers.

#	MSSL	MSSL gloss	EEN	EEN gloss	EEN	Tosco (sg, plu.)	Tosco gloss (sg, plu.)
1	gát̪ʼa	beard	gát̪ʼa	beard	DE	gátʼa,gátʼattʼi	beard
2	kʼúba	cheek	kʼúːβa	cheek	DE	kʼúúba, kʼúúbaddi	cheek
3	gákʼane	chin	gákʼane	chin	DE	gákʼane, gákʼanúúta	chin
4	ʔágaddzo	ear	ʔágaddzo	ear	DE	ʔágaddzo, ʔágaddzi	ear
5	ʔílla	eye	ʔíla, ʔílla	eyes, eye	DE	ʔila, ʔílla	eye
6	ŋgíkine	eyebrow	ŋgíkine	eyebrow	CE	ⁿgíkine, ⁿgíkinuuta	eyebrow

7	he:ŕífa	eyelash	ɦírifane, ɦiŕífa	eyebrow	CE	ʔirifa, ʔirifamunte or ʔirifáfi	eyelash
8	summa	forehead	suma	face, forehead	DE	summa	in front of, before
9	ŋǁi:t̪'i	gums	ǁi:t̪'i, (ǀé:t̪'i)	gums	(CE)		
10	t̪'at̪ta	hair	t̪'át̪a	hair	DE	t'áttane,t'átta	hair
11	ʔani	head	ʕani, ʕánu:t̪a	head	CE	ʕani, ʕánuuta	head
12	ʔáŋkoma	jaw					
13	numme	lips	númme	upper lip	DE		
14	ʔákake	molar tooth	ʔákake, ʔákaku	molar tooth	CE		
15	ʔá:fo	mouth	ʔáfo, ʔáfud̪d̪a	mouth	DE	ʔáfo, ʔáfudda	mouth
16	nd̪wa:la	mucus	nd̪ʷa:la	mucus	DE		
17	sína	nose	sína, sinanne	nose	DE	sína, sínadde	nose
18	ʔaŋkolle	palate					
19	ŋ̊ǁaɲi, ŋ̊ǁwaɲi (Kipini var.)	saliva	ǀwá:nyi	saliva		ǀaaɲi	saliva
20	ʔilíma	tears	ʔilíma	tear	DE	ʔilímá, ʔilimámi	tear
21, cf. 86	kálat̪i	teeth	kálat̪i, kalát̪et̪to	tooth	DE	kálati,kálatétto	tooth
22	ʔe:na	tongue	ʕé:na	tongue	DE	ʕééna, ʕéénaddi	tongue
23	ɗi:ga	blood	ɗi:ga	blood	DE		
24	miccʎ'o	bone	mítl'tl'o	bone	DE		
25	ɓóŋko	brain	ɓóŋko	brain	DE	ɓonko, ɓónkagááme	brain
26	ŋ̊ǁó:ne	breast	ǀó:ne, ǀo:nakke	breast		ǀóóne, ǀoonu or ǀóónudda	breast
27	z̪iwa	milk	ðí:wa	milk	DE	zíva	milk
28	la:be	coagulated blood					
29	d̪ú:ra númo	gut, intestines	d̪u:ra	intestines	DE	dúúra, dúúrari	bowels
30	muna	heart	muna, múne:ka	heart (also back?)	DE	muna, múnadde	heart
31	sáre	spare ribs				sáre	back
32	nábaro	rib	náβaro	rib	DE		
33	ŋ̊ǁuʔu	shit	ǀuʔu, (ŋǀuʔu)	excrement	(CE)	ǀuʔu	excrement
34	lé: munáni	spine					
35	fúnt̠o:no	sweat	fúnto:no	sweat, heat	DE	fúnt'oono	warm sweat
36	ʜunna	sweat					
37	kóŋko:lo	shin	kóŋko:lo	shin	CE	kónkoolo, kónkooláli	leg (from knee to foot)

38	goŋgótɬ'ome	ankle	gó:gontl'ima	ankle	DE		
39	t̪aʜára	calf of leg	t̪ahára	calf of leg	DE	táhara, táharúrra	heel
40	míŋki	arm	míggi	arm, wing	DE	miggi,míggine	arm
41	d̪ába	hand	d̪aβa	hand	DE	dába. dábabbe	hand, paw
42	l̪i:be	buttocks	li:βe, lí:βut̪t̪a	buttock	CE	ɗíibe, ɗíibuuta	buttock
43	ɓú:ba	chest	ɓú:βa	chest	DE	ɓúúba, ɓúúbaddi	chest
44	tʃá:nd̪a	finger	tʃá:nd̪a	finger	DE	tʃaa^{n}da (ʃaa^{n}da), tʃáá^{n}dudda or tʃáá^{n}duka	finger
45	tso:lo	fingernail	tso:lo	fingernail	DE	tsoolo, tsóóle or tsóóludda	nail, claw
46	gimp'o	heel	gimp'o	heel	DE	gimpo, gímpoma	carpus, malleolus
47	k'áno	hip					
48	gilli	knee	gilli	knee	DE	gilli,gíllibe	knee
49	k'ok'o	throat	k'ok'o	throat	DE	k'ok'o, k'ók'addi	throat
50	d̪akáʡa	foot	d̪akáʕa	leg, foot	DE	dakáʕa, dakáʕadde	leg, foot
51	dʒúkku	navel	dʒúkku	navel	DE	dʒúko	navel
52	ɗáʡʡe:ro	neck	ɗá:ʕe:ro	neck	DE	ɗááʕeero, ɗááʕoorúdda	neck
53	dʒékkele	shoulder	dʒékkele, dʒékkeláli	shoulder	DE	dʒékkele, dʒékkeláli	shoulder extremity
54	rik'a	tail	rik'a, ŕik'e:ma, (ŕik'ak'k'e)	tail	DE, CE	rik'a	tail
55	luka	thigh	luka, lúke:ma	thigh	CE	lúk'a, lúk'agááme	leg (from thigh to knee)
56	ɓo:t̪i	thigh				ɓooti,ɓóótime	thigh
57	ŋ̊ǀút̪u	waist	ǀú:t̪u	waist of humans, joint of tool		ǀútu	waist
58	kíwi:ko	wrist				kííviko, kíívikáki	wrist
59	rúga	root of penis					
60	níŋka	clitoris	níŋka	clitoris	CE	nínka, nínkaggi	clitoris
61	lad̪d̪a	penis	lad̪d̪a	penis	CE	ladda	penis
62	ʔacʎ̥'áno, ʔácʎ̥'ano	semen					
63	kámpo:re	testicle	kámpo:re	testicles	CE	kámpore, kámporári	testicle
64	gid̪d̪a	semen	gid̪d̪a	semen	CE	gídda	sperm
65	tsintso	vagina	tsintso	labia, vulva	CE	tsintso	vagina

66	miɬo	body	miɬo	living body	DE	miɬo,miɬa	body
67	ɓágama	stomach, belly	ɓágama	belly	DE	ɓágama, ɓagamudda	belly
68	ŋ̊ǀát̪'u	constipation					
69	t̪̊'ókkome	fever, cold	t̪'ókkoːme	fever, cold	DE	t'ókkoome	cold
70	kít̪t̪io	leprosy					
71	(d̪ába) lóaːni	right (hand)				dába lua	right (hand)
72	(d̪ába) ʃóːt̪oːni	left (hand)				dába ʃoto	left (hand)
73	ruk'o	sickness	ruk'o	illness, pain	DE	ruk'o	sickness
74	gino	skin	gino	skin	DE		
75	had̪d̪úra	sleep	had̪d̪úra	sleep	DE	haddur-	to sleep
76	gwiʔi	thirst	gʷiʔi	thirst	DE	gʷiʔi	thirst
77	ʔút̪unu	wound	ʔút̪unu	wound	DE	ʔútunu, ʔútunáni	wound
78	ɓáɓaʔane	sp. large antelope (kungu)	ɓáβaʔaːne, (-i), ɓáβaʔaːnuːt̪a	bushbuck	DE (CE)	ɓábaʔaane, ɓábaʔaanuuta	topi
79	híbe	baboon	híβe, hiβe	baboon	DE	híbe, hibe or híbema	baboon
80	ɓeːʡʡa	buffalo	ɓeːʕa, ɓeːʕádʒu	buffalo	DE	ɓeeʕa, ɓééʕamunte	buffalo
81	dʒáʔawu	monkey	dʒáʔawu, (dʒaʔaβu, ʋaʔaβuːt̪a)	vervet monkey	CE	dʒáʔawo, dʒáʔawuuta	vervet
82	dʒáːgo	cow	dʒáːgo, dʒaːgu	cow, cattle	DE	dʒáágo, dʒáágu	cow
83	náʡʡeːt̪e	dog	náʕeːt̪e	dog	DE	náʕeete, náʕeeto	dog
84	puːnd̠a	donkey	puːnda	donkey	DE		
85	d̪okóːmi	elephant	d̪okkóːmi	elephant	DE	dokóómi, dókoomámi	elephant
86, cf. 21	kálat̪i	tooth, tusk	kálat̪i, kaláʈet̪t̪o	tooth	DE	kálati,kálatétto	tooth
87	kiri	giraffe	kiri	giraffe	DE	kiri, kírima	giraffe
88	héːri	goat	héːri	goat, sheep	DE	hééri, héérarre	goat
89	ɲáʜe	hippo	nyáɦe	hippo	DE	ɲáħe, ɲáħudda	hippo
90	t̪umpi	horn	t̪umpo	horn	DE	tumpi, tumpábbi	horn
91	fáraːsi	horse					
92	wáraːba	hyena	wáraːβa	Striped hyena	DE	wáraaba, wáraabuuta	hyena
93	ɓwéha	jackal					
94	ɲímaŋkálo	leopard	nímaŋkalo	leopard	DE	nímankalo, -uuta or -áli	leopard

95	ndu:ma	leopard, cheetah					
96	ŋ̊e:ɬe	lesser galago	|ʷeɬe, (|ʷete)	bushbaby			
97	ɓáʔi	lion	ɓaʔi	lion	DE	ɓáʔi, ɓáʔima	lion
98	dʒáme	pig	dʒáme, dʒámeni	pig	DE	dʒáme, dʒámema	warthog
99	dʒara	porcupine	dʒara, dʒérema	porcupine	DE	dʒara, dʒárema	porcupine
100	mpáɲe	rat	mbánye	rat	DE		
101	wa:la	rhino	wâ:la, wá:lu:ka	rhino	DE	waala	rhino
102	mo:dʒo	sp. large antelope	mó:dʒo, (mo:dʒu)	eland/greater kudu	CE		
103	hé:rilibe:ni	sheep (cf 88 'goat', 42. 'buttocks')					
104	t'ú:t'o	sp. small antelope	t'u:t'o, t'u:t'e:ma	waterbuck (?)	DN	t'úúto, t'úútema	waterbuck
105	ɗó:ɗa:ma	topi antelope	ɗó:ɗa:ma	topi antelope (?)	DE		
106	ndzó:me	udder					
107	guʔá:t'e	warthog	guʕá:t'e	bushpig	DE		
108	ʜelléʔa	zebra	ɦelleʔa	zebra	DE	ħélleʕa	zebra
109	t'át'i	ant	t'át'e, t'at'i	sp. ant	DE	t'át'e,t'át'ima	sp. small ant
110	ɗíme	bee	ɗíme	honey bee	DE	ɗíime, ɗíimudda	bee
111, cf. 117, 134, 135	ŋ̊á:dʒu	creature that lives in and out of water (pl.)	|á:dʒume, |á:dʒu	lungfish			
112	ɲó:ro	intestinal worm					
113	ʜénta	flies	ɦíntote	house fly	DE	ʕíntone, ʕinta	housefly
114	húmu:me	frog	húmume	frog	DE	hómome, hómomúúta	frog
115	gurúbi	gecko	gurúβi	skink	DE		
116	ɓáki ʔaŋkóma	sp. spotted lizard					
117, cf. 111, 134, 135	ŋ̊á:dʒume	creature that lives in and out of water (sg.)	|á:dʒume, |á:dʒu	lungfish			
118	ʔitta	lice	ʔitto:te, (ʔittó:ni), ʔitta, (ʔitta)	louse	DE (CE)	ʔittone, ʔitta	louse
119	ɲoŋkó:lo	millipede					

120	t̪'ó:nane	mosquito	t̪'ó:nane:te, t̪'ó:nane	mosquito	DE		
121	ŋ̊ǀububuʔe	moth	ǀuβuβuʕe	moth		ǀúbubuʕe, ǀúbubúʕa	moth
122	ŋǀé:no	python	ŋǀé:nu, ŋǀé:nu:ma	python		^{n}ǀéénu, ^{n}ǀéénanni	python
123	ha:ʃíad̪a	short, poisonous snake					
124	ŋ̊ǀúʜi	siafu ant	ǀoɦe, (ŋǀoɦe)	siafu ant	(CE)	ǀúʕe, ǀúʕite	brown ant
125	tsúŋki	small black bug	tsúŋke, tsuŋki	sisimizi ant	CE	tsúnke, tsúnkima	soldier ant
126	ŋ̊ǀunt̪á:ʔa	small grasshopper	ŋǀt̪a:ʕe, ǀunt̪a:ʕe	grasshopper, small grasshopper		ǀuntáʕe, ǀuntaʕe	cicada
127	ŋ̊ǀámpirikit̪t̪e, ŋ̊ǀímpirikit̪t̪e	sp. small lizard	ǀímbirik'ít̪t̪e	sp. lizard		ǀ́^{n}birik'itte (sg?)	gecko
128	t̠'írimalle	spider					
129, cf. 244	p'á:ra	termites	p'á:re:t̪e, p'á:ra;; p'á:rane	termite	CE;; DE		
130	k'úbe	tortoise	k'óβe, (k'úβi), k'óβet̪t̪a	tortoise	CE, (DE)	k'óbe, k'óbaddi	tortoise
131	d̪ó:ruáni	wasp	d̪ó:roá:ni	large stinging insect	DE		
132	hímma	worm					
133	dʒáribo		dʒáriβo	sp. fish (tundi)	AT	dʒáribu	sp. fish
134, cf. 111, 117, 135	ŋ̊ǀá:dʒume	creature that lives in and out of water (sg.)	ǀá:dʒume, ǀá:dʒu	lungfish			
135, cf. 111, 117, 134	ŋ̊ǀá:dʒu	creature that lives in and out of water (plu.)					
136	mpálabe	sp. fish	mbaláβe	fish	DE	^{n}balábe, ^{n}bálabéni	sp. fish
137	ŋgé:re	sp. lungfish	ŋgerre	female of ŋgi:βu	DE		
138	ʔégo:e	egg	ʕógo:e, (ʕégo:e), ʕego:i	egg	DE (CE)	ʕógohi, ʕogóhi	egg
139	ʔáɬi	fat, oil	ʔáɬi	oil	DE	ʔáɬi	fat, oil

140	ʡaga	food	ʕaga, (hágâ), ʕáge:ma	food	DE (AT)	ʕaga	food
141	nala	honey	nala	honey	DE	nala	honey
142, cf. 267	ɓuru	maize	ɓuru	maize	DE	ɓuru	maize
143	móla	honey beer	móla	(honey)m ead	CE	móla,mólalle	mead
144	d̪abi	meat	d̪aβi	animal, meat, flesh	DE	dabi, dábima	animal, game (cf. Br. daaba)
145, cf. 281	ŋ̊|ó:d̪e, ŋ̊|óʔod̪e (Kipini var.)	sp. tree or its wood					
146	má:puŋka	rice				mápunka	rice before husking
147	ʔút̠a:t̠i	ugali, stiff cornmeal porridge				ʔút̠aat̠i, ʔút̠aat̠át̠i	ugali
148	ɬunno	stew	ɬu:no, ɬú:ne:ma	relish	CE		
149	ʔi:bu	ashes	ʔi:βu	ashes, gray	DE	ʔíivu	ash
150	kapu	basket	kápu, kápape	basket	DE	kapu, kápapánne	basket
151	mútsuŋki	water pot	mútsuŋki, mútsuŋkakí	earthen water pot	DE	mútsunki, mútsunkággi	waterpot
152	nd̪upa	bottle	nd̪upa	bottle	DE	^{n}d̠upa, ^{n}d̠úpema	bottle
153	pe:o	broom	pé:o	broom	DE	péélaadiini, peeláádiide	broom
154	kí:bu	calabash, gourd	kíβo, kíβod̪d̪a	gourd	DE	kíbo, kíbudda	calabash, gourd
155	kít̠i	chair	kíti	chair	DE		
156	t̪'ilíʜʜa t̪'ilíʜʜe	charcoal (pl.) charcoal (sg.)	t̪'iliɦa	embers	CE	t'ilíɦe,t'ilíɦa	embers
157	fílime	comb	fílime	comb	DE	fílime, fílimámmi	comb
158	ɗau	cooking pot	ɗau, ɗáudʒe	pot	DE	ɗau, ɗáudʒe	pot
159	kúpini	lid, cover	(kupi)	(to cover)		kupid-	to cover
160	ŋ̊|ápat̪i, ŋ̊|ápila (Kipini var. sp. 6)	lid, cover	|áppattsi	lid			
161	ko:mbe	cup					
162	ŋ|ákwi	deserted home	ŋ|ákʷi, ŋ|ákʷi:ma	deserted homestead			

163	mún̪t̪a	cultivated field, shamba	múnta, múntate	field	DE	mún̪t̪a, mún̪t̪eka	farm, shamba
164	po:bu	deserted farm					
165	ʔe:ga	fire	ʔe:ga	fire, firewood	DE	ʔééga	fire
166	ɓat̪á:li	fist-sized calabash				(ɓááre, ɓáárema)	(big calabash)
167, cf. 238	nd̪ó:ʔo	floor					
168	ŋǁikína	forehead bead decoration	ŋǁikí:na	bead bracelet			
169	mpi:ni	handle	mpíni	handle	CE		
170	kó:fia	hat	kó:fia	hat	DE		
171	me:ko	hearth, etc.	me:ko, (mé:ko, mé:ko:ma)	fireplace, cooking stones	DE, (CE)		
172	mini	house	mini, míddzi	(lean-to) house	DE	mini,míddzi	house
173	ŋgú:d̪o	house post					
174	mígau	kind of spoon	míga:wa, míga:wáwi	ladle	CE		
175	mi:ko	kind of spoon	mí:ko, mi:ko:ma	spoon	CE		
176	ʜálit̪e	knife	ɦáli:t̪e, ɦáli:t̪úmu	knife	DE	ʕálite, ʕálitúmu	knife
177	pú:nd̪e	long calabash	pu:nt̪e	small round calabash	CE		
178	gárima	outside	gárima	outside	DE	gárima	outside
179	nd̪á:ni miníni	inside (the house)	ndá:ni	inside	CE	nd̪ani	inside
180	ŋ̊ǀíme, ŋ̊ǀíma	pot scraper	ǀime	scraper, shell for scraping			
181	páa	roof	páa	roof	DE		
182	kwé:ni	water pot					
183	nd̪úfuro	whisk	nd̪úfuro, nd̪úfurari	porridge twirling stick	CE		
184, cf. 431	ŋ̊ǀábat̪e	good smelling	ǀaβa	nice smell (of oil)			
185, cf. 189	ʜaso	bowstring	ʔa:so, (ɦaso)	bowstring, animal tendon	DN, (CE)	hááso, háásooma	bow-string
186	ʜa:ɗo	arrow	ʔa:ðo, (ɦa:ɗo)	arrow head for small animals	CE, (DE)	ħaaɗo, ħááɗa	arrow

187	ɓod̪a:ni	type of arrowhead	ɓoðani	arrowhead for big animals	DN	ɓóóɗaani, ɓóóɗaanánni	not-poisoned arrowhead
188	ʡá:la	bow	ʕá:la	bow	DE	ʕáála, ʕááluuke or ʕáálali	bow
189, cf. 185	ʜaso ʡá:lani	bowstring	ʔa:so, (ɦaso)	bowstring, animal tendon	DN, (CE)	hááso, háásooma	bow-string
190	ru:ŋku, ru:ŋgu (Kipini var.)	club	ruŋku	club	CE	ruu^{n}gu, rúú^{n}gume	club, knobbed stick
191	ʔút̪o:ma, t̪ómá (Kipini var.)	fish spear	t̪omá	fishing spear	DN		
192	wáraʜa	fish spear	wáraɦa	spear	DE	wáraħa, wáraħuuke	spear
193	ɬwé:ɲi	kind of arrow	ɬʷe:ni	poisoned arrow for med.sized animals	DE/ DN		
194	t̠'ó:k'oma	point	|ók'oma	point			
195	t̪'aʡa	arrow poison	t̪'áʕa	arrow poison	CE	t'áʕa, t'áʕema	poison
196	kíraŋkat̪e	quiver	kíraŋkat̪e, kíraŋkat̪u:ke	quiver	CE	kíra^{n}gáti, kíra^{n}gatáti	quiver
197	kw'ai	arrow shaft	kʷ'ai, kʷ'ái:ma	arrow shaft	CE		
198	puré:ki	big hole					
199	t̠ó:i	canoe, *fishing boat*					
200	dʒémpe	hoe	dʒémpe	hoe	DE	tʃ'émpe, tʃ'émpabbi	hoe
201	si:mbo	kind of stick	si:mbo	stick	DE	síí^{n}bo	stick
202	rúpaŋka	machete	rúpa:ŋga	machete	DE	rúpa^{n}ga, -ággi	panga
203	ɓá:kora	kind of stick					
204	ʔánno	aunt	ʔanno	father's sister	DE	ʔánno, ʔánnoni	father's sister
205	ɓó:re:t̪e	boy	ɓó:re:t̪e	boy	DE	ɓóóra, ɓóóreete (sg)	boy
206	gwíttsa	child	gʷíttsa, gʷíttso	child	DE	gʷíttsa, gʷíttso	child
207	ɓa:ɓa	father	ɓá:ɓa	father	DE	ɓááɓa, ɓááɓani	father
208	d̪é:la	girl	d̪é:la, d̪e:le	girl	DE	dééla, dééle	girl (young unmarried woman)
209	ʔa:bo	grandmother	ʔa:βo	mother's sister	CE	ʔáábo, ʔááboni	grandmother
210	ʔó:ra kudzi	male rel., nephew					

211	ʜá:dʒo	man	ɦá:dʒo, ɦa:ʕi or ɦádʒo:ma	man, husband, judge	DE	ħáádʒo, ħáʕi	man
212	já:jo	mother	yá:yo	mother, alive	DE	yááyo	mother (alive)
213	ʔá:dʒi	older brother	ʔá:dʒi	older brother	DE	ʔáádʒi, ʔáádʒadʒdʒi or ʔáádʒini	elder brother
214	gá:na	adults	ga:na	woman	AT	gaano, gáána	large, big; grown up person
215	já:jo:ni	(my) parents					
216	ʡáme	male rel., grandfather	ɦáme ʔá:ma	father's brother mother's brother	CE DE	 ʔááma, ʔáámani	 mother's brother
217	gwítso pa:t̪a	twins	pa:t̪a	twins	DE		
218	nat̪'étsa	female	nát̪'etsa	cow	DN	nát'ettsa	female
219	nát̪'a	woman	nat̪'a, nat̪'e	woman, wife	DE	nat'a, nat'o	woman
220	gud̪d̪e	bush	gud̪d̪e, (gut̪t̪e)	land	DE, (DE 1976)	gudde	bush
221	máwi:ŋku	clouds	máwi:ŋgu, (wíŋkune, wíŋku)	cloud	DE, (CE)	máviⁿgu	clouds
222, cf. 69	t̪'okkome	cold	t̪'ókko:me	fever, cold	DE	t'ókkoome	cold
223, cf. 250	ʔad̪d̪o	day, sun	ʔad̪o	sun	DE	ʔaddo	sun, day
224	ʔad̪d̪o ɓúat̪e	midday					
225	ŋ̊ǀimi	dew	ǀimi, (ǀémi)	dew			
226	wu:mpi	dust					
227	ɓúrune	dust	ɓúrune	dust	DE	ɓúrune	dust
228	ɓáriddzina	dawn, sunrise	ɓariddzina	sunrise	CE	ɓaridʒ-	to depart on dawn
229	het̪t̪o	evening	hed̪d̪o	evening	DE	heddo	evening
230	ɓó:ku	hole	ɓó:ku, ɓo:kakke	hole	DE	ɓóóku, ɓóókakki	hole
231	ragáma	hole	ragáma	grave	DE	ragáma, rágamuuta	grave
232	tɬ'a:ʡa	lake	tl'á:ʕa, (tl'á:ʔa, tl'á:ʔut̪t̪a)	lake	DE, (CE)	tl'ááʕa, tl'ááʕudda	river, lake
233	birik'inna	lightning	ɓírik'k'ína	lightning	DE	ɓírik'inna	lightning
234	háge	moon	háge, hágega:ma, (hageut̪t̪a)	moon, month	DE (CE)	háge	moon

235	ɓurra	morning	ɓurra	morning	DE	ɓura	morning
236	múlime:t̪e	mountain	múli:ma	mountain	DE		
237	ɗokk'e	mud	ɗok'k'e	dust, mud (?)	DE		
238, cf. 167	nd̠ódoʡo	kind of mud	ndodóʕo	mud	CE	ɗóɗoʔa	mud
239	hima	night	hi:ma	night	DE	hííma	night
240	hélleʔémo	hole with water					
241	maʔa	water, rain	maʔa	water	DE	maʔa	water
242	maʔa lúbbi	it's raining (water falls)					
243	so:no	river	so:no	Tana river	DN		
244, cf. 129	p'á:ra	termites	p'á:re:t̪e, p'á:ra; p'á:rane	termite	CE; DE		
245	gípe	sand dune	gípe	(ant)hill	CE		
246	mut̪t̪a	anthill	mut̪a, mut̪at̪t̪e	small ant	DE		
247	t̪'uggwa	smoke	t̪'ugʷgʷa, t̪'ogʷgʷa	smoke	DE	t'uggʷa	smoke
248	ŋǀá:da	name of lake east of Witu (cf. Boni [ɲá:dá̄], same meaning)	ŋǀa:ða	area on east of Witu to Pandangu o road			
249	ŋǀiŋkilíʡa, ŋ̊́gilíʡa (Kipini sp. 5 var.)	stars	ʄiŋgiliʕe, (ŋǀkiliʕe, ŋǀkilíʕe)	star	(CE)	ʄiⁿgiliʕe, ʄiⁿgilíʕa	star
250, cf. 223	ʔad̪d̪o	day, sun	ʔad̪o	sun	DE	ʔaddo	sun, day
251	tsa:ka	noon, hot season	tsa:ka	hot season	CE	tsááka	hot season, hottest part of the day
252	ŋgóba	thunder	ŋgóβa	thunderbo lt	CE		
253	mát̪ika	rainy season	mátika	long rains	CE		
254	ʔad̪d̪ókwa	today	ʔad̪ókʷa	today, this day	DE	ʔaddókʷa	today
255	hí:mane	tomorrow	hí:mane	tomorrow	DE	híímane	tomorrow
256	hí:manéʔesu	yesterday				híímanesú[ʔ]u	day after tomorrow
257	ŋǀaba	forest	ŋǀaβa, ŋǀaβutta	forest		nǀaba	forest
258	dʒú:fume	wind	dʒú:fume	wind	CE	dʒúúfune	wind

259	wat̪t̪úkwe	one	wat̪t̪úkʷe	one	DE	vattékʷe (f), vattúkʷe (m)	one
260	ɬi:ma	two	ɬi:ma	two	DE	ɬiíma	two
261	k'aba	three	k'aβa, k'áβâ	three	DE	k'aba	three
262	saʡála	four	saʕála	four	DE	saʕála	four
263	d̪áwat̪t̪e	five	d̪áwat̪t̪e	five	DE	dáwatte	five
264, cf. 276	ɓúet̪e	maize stalk					
265	ʡáwa:ke	sp. palm (fully grown)	ʕáwa:ke	Borassus Palm (or fruit?)	DE	ʔávak'e, ʔávak'a	doum palm
266	ná:d̪it̪e	coconut palm	ná:ðit̪t̪e, ná:ðimamu	coconut	CE		
267	ɓuru	maize	ɓuru	maize	DE	ɓuru	maize
268	mberewere	flower of maize					
269	ʃit̪iŋke	sp. palm	ʃi:t̪iŋke, ʃi:t̪iŋka	Cycad tree	DE, CE		
270	t̪áʔad̪i	fruit of shitinke palm	t̪aʔaði	Cycad fruit	DE		
271	t̪á:mi	grass/thatch	tá:me t̪á:mine, t̪á:mi	leaf of Borassus Palm blade of grass	DE DE	táámine,táámi	grass, blade of grass
272	nt̪ago	grass, thatch					
273	ʎ̥a:bu	leaf	ɬá:βune, ɬá:u	leaf	DE	ɬáábu, ɬáábune	leaf
274	gák'ane	maize tassel					
275	nt̪ét̪e	maize, not dry				(ntée)	(rice after husking)
276, cf. 264	ɓúet̪e	maize stalk					
277	múku ná:d̪it̪e	sp. bush w/thorns					
278	góesa	sp. grass					
279	saméttsa	sp. palm					
280	ʔággwane	sp. tree with thorns	ʔággʷane	sp. thorntree	DE		
281, cf. 145	ŋ̊ǀo:d̪e	sp. tree w/fruit, or its wood	ǀô:ðe	sp. tree			
282	ŋ̊ǀikwe ŋ̊ǀikwa	sp. reddish grass (sg.) sp. reddish grass (plu.)	ǀikʷe	sp. tall grass			
283	ʡé:me	thorn	ʕé:me	thorn	DE	ʕééme, ʕéémi	thorn

<table>
<tr><td>284</td><td>ŋǀóʔo</td><td>sp. palm
(young)</td><td>ŋǀóʔodʒe,
(ŋǀóʔo),
ŋǀóʔekki</td><td>palm
frond</td><td></td><td>nǀóʔo</td><td>palm
frond</td></tr>
<tr><td>285</td><td>ʜít̪t̪ibe</td><td>bird (general
term)</td><td>hét̪i:βe

ɦíd̪d̪iβe</td><td>bird

bird</td><td>AT

DE</td><td>ɦíddibe,
ɦíddiba
ɦíddibe,
ɦíddiba</td><td>bird

bird</td></tr>
<tr><td>286</td><td>ʔábat̪'e</td><td>honey guide</td><td>háβat̪'e</td><td>honey
guide</td><td>DE</td><td></td><td></td></tr>
<tr><td>287</td><td>ʜé:le</td><td>sp. black
bird</td><td>ɦílle

ʕelle</td><td>sp. owl

sp.
vulture</td><td>

CE</td><td></td><td></td></tr>
<tr><td>288</td><td>n̪t̪oke, t̪oke
(Kipini var.)</td><td>sp. black
bird</td><td>toke</td><td>sp. bird
(mlakufi)</td><td>DE</td><td></td><td></td></tr>
<tr><td>289</td><td>mba:re</td><td>cattle egret</td><td>mba:re</td><td>egret</td><td>CE</td><td></td><td></td></tr>
<tr><td>290</td><td>n̪d̪a:re,
ndza:re
(Kipini var.)</td><td>cattle egret</td><td>ndzá:re,
ndzá:re:ma</td><td>cattle egret</td><td>CE</td><td></td><td></td></tr>
<tr><td>291</td><td>ŋkuko,
ŋguko
(Kipini var.)</td><td>chicken</td><td>ŋgúko</td><td>chicken</td><td>DE</td><td>^{n}gúúko,
^{n}gúúku</td><td>cock</td></tr>
<tr><td>292</td><td>ɦáʔali</td><td>chicken
(cock)</td><td>ɦáali</td><td>cockerel</td><td>DE</td><td></td><td></td></tr>
<tr><td>293</td><td>ń̩so</td><td>chick</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>294</td><td>káŋke</td><td>guinea fowl</td><td>káŋke,
(káŋka,
káŋke:ma)</td><td>guinea
fowl</td><td>DE,
(CE)</td><td></td><td></td></tr>
<tr><td>295</td><td>ɦét̪eɦét̪t̪e</td><td>hammerkop</td><td>ɦet̪éɦet̪e</td><td>Hammerk
op</td><td>DE</td><td></td><td></td></tr>
<tr><td>296</td><td>tsílalla</td><td>hawk sp.</td><td>tsílala,
tsílalu:ke</td><td>hawk</td><td>CE</td><td>tsílalla,
tsílalluuke</td><td>hawk</td></tr>
<tr><td>297</td><td>wáŋko</td><td>hornbill</td><td>wáŋkko</td><td>sp.
hornbill</td><td>DE</td><td></td><td></td></tr>
<tr><td>298</td><td>ŋ̊ǀaŋkúru</td><td>sp. red-faced
barbet</td><td>ǀa:k'úru</td><td>big barbet</td><td></td><td></td><td></td></tr>
<tr><td>299</td><td>tʃé:rere</td><td>kingfisher,
woodpecker</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>300</td><td>ŋ̊ǀúpuku</td><td>knob-billed
duck</td><td>ǀúpuŋk'u</td><td>Knob-
billed
duck</td><td></td><td></td><td></td></tr>
<tr><td>301</td><td>ɗo:ri</td><td>ostrich</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>302</td><td>dʒukú:bu</td><td>owl</td><td>dʒú:kúβu,
dʒu:kúβe:ma</td><td>owl</td><td>CE</td><td></td><td></td></tr>
<tr><td>303</td><td>ɓi:t̪a</td><td>pelican</td><td>ɓi:tá,
ɓi:t̪e:ma</td><td></td><td>CE</td><td></td><td></td></tr>
<tr><td>304</td><td>ɗé:ro</td><td>red-headed
woodpecker</td><td>d̪é:re</td><td>woodpeck
er</td><td>DE</td><td></td><td></td></tr>
<tr><td>305</td><td>ŋ̊ǀik'it̪o</td><td>ruddy
weaver bird</td><td>ŋǀik'i:t̪o</td><td>blackneck
ed weaver</td><td></td><td></td><td></td></tr>
</table>

306	ndzigi	scissor-tailed bird	ndzigi dzík'ilifie	red-cheeked Cordonbleu sp. bird (magpie?)	CE DE		
307	tʃoŋgo	small yellow bird					
308	p'op'oroʜódʒa	sp. bird	p'op'oroʕʕo:dʒa	sp. bird	DE		
309	tʃalakakki (tones unknown)	sp. bird	tʃalakakki	kerezende ??	DE		
310	ŋgwale	sp. bird w/red legs	ŋgʷále, ŋgʷále:ma	spurfowl	CE		
311	gólobe	sp. white bird	góloβe, golóβa	Vervet (?) monkey	DE	gólobe, góloba	monkey
312	t̠'úaru	sp. ruddy bird	tʃ'úare	sp. bird (roller?)	DE		
313	ɓa:t̪a	water bird					
314	dza:ʔát̪a	death	dza:ʔe	death	DE	(dzaaʔáta, dzaaʔátuuke)	(wild animal, enemy)
315	mbó:ri	war	mbó:ri	war	DE	ⁿɓóóri	fight, war
316	ndzá:ʔina, dzá:ʔina (Kipini var.)	die	dza:ʔ-	die	DE	dzaaʔ-	to die
318	hákwa	go, to leave	(hakw-)	leave	DE	(ʔakkw-)	(leave, to;, let go, to)
318	fú:runa	I'm full (of food)	fu:r-	be satiated	DE	fuur-	to be satiated
319	ŋ̊|úʔuna	I'm quenched	|u:ʕ-	stop drinking			
320	pak'enna	lay open	pak'k'e:ð-	uncover	DE		
321	ŋ̊|wá:ʔana	massage, knead					
322	dʒuʔenna	quell, douse fire	dʒuʕʕe:ð-	extinguish	DE	dʒuʕeed-	to extinguish (fire)
323	ɓaʔanna	set out to dry	ɓáʔama, ɓaʔámi	hard, dry (pl., m)	DE	ɓaʔad-	to dry (meat)
324	k'éré kámí:na	be hungry	kam-	hold, take	DE	k'ére kam-	to be hungry
325	k'ére	hunger	k'ére	hunger	DE	k'ére	hunger
326	p'úat̪úina	be rotten	p'óʔat̪e.; p'uʔat̪uw-	rotten; be rotten	CE; DE	p'óʔate	rotten
327	k'áʜina	bite	k'afi-	bite	DE	k'aħ-	to bite
328	ɬímpínna nd̪wa:la	blow nose	ɬimpið-	blow out snot	DE		
329	ʜa:ɗinna	boil					

330	faːʡona	break	faːʕ-	to break, smash	DE	faaʕ-	to destroy (impf.)
331	nt̠ikanna	bury	ndiːkaːð-	bury ([nd̠]?)	DE	ⁿd̠igad-	to bury
332	gwat̪'ana	chew	gʷat̪'-	chew	DE	gʷat'-	to chew
333	k'éːrena	chop	k'eːr-	cut	CE	k'eer-	to chop
334	róʔiːa	come	roʔ-	go	DE	roʔ-	to go
335	ɗákana	cook	ɗak-	cook	DE	ɗak-	to cook (perf.)
336	ʡáhoʔóna	cough	ʕőhoʔóna	cough	DE	ʕoʕoʔid	to cough
337	kuppinna	cover	kuːβ-	shut	CE	kupid-	to cover
338	ŋ̊ǀapinna	really cover well	(ŋǀap'ið-); ǀappið-	(smack); cover			
339	dʒakuna	kneel	dʒakkaw-	squat	DE		
340	t̠út̠aʔamát̪ina	crouch					
341	ʔaːʡʡona	cry	ʕaʔ-	cry, bleat	CE	ʕaʕ-	to cry
342	haːbana	cut				haab-	to cut (impf.)
343	maʔuna	drink	maʔaw-	drink	DE	maʔaw-	to drink
344	ʡat̪t̪áana	drop	ʕat̪t̪aɦ-	throw	DE	ʕattaħ-	to throw
345	ʡaguna ʡáːme	eat eat	ʕag- (ʔaːme)	eat (few)	DE	ʕag- ʕaameemit-	to eat to eat continousl y
346	lút̪t̪ukunna	fall	lut̪t̪ukum-	fall	DE	luttokum-	to fall
347	rik'ina	fear	rik'íno; rik'-	fear; obey	DE	rik'-	to be afraid
348	ŋ̊ǀóːkona	fill	ǀoːk-	fill		ǀuuk-	to fill
349	sáʔit̪ona	forget	saɦit̪-	forget	DE	saħid-	to forget
350	héːʔekuna	give	heːʔ-	give	DE	heeʔ-	to give
351	ʔómuʃut̪ina	give birth	ʔumuʃ-	bear (child)	DE		
352	róʔina	go	roʔ-	go	DE	roʔ-	to go
353	kanna	grab, hold, grasp	(kaːneːm-)	(stick tog.)	CE	kam-	to hold
354	ɬéʡekina	have an odor	ɬeːʕ-	smell (intr.)	DE	ɬeeʕ-	to smell
355	ʔet̪inna	hear				ʔeetit-	to hear
356	ʜik'onna	hiccup					
357	paʜuna, t̪aʜuna	hit				paħ-	to beat
358	gubáːlanna	hunt				gubaalid-	to hunt
359	paʜʜina	hit				paħ-	to beat
360	kóːlona	jump				kool-	to fly
361	dzeʔenna	kill				dzeeʔed-	to kill
362	ʜát̪'ana	slaughter				hat-	to slaughter
363	ɬéːʡena	kiss					
364	ʔeledʒo	know				ʔeledʒ-	to know
365	k'iːk'ona	laugh				k'iik'-	to laugh

366	ʜáŋ̊ǀana	lick				ʕaⁿǀ-	to lick
367	ʔúk'ent̪t̪inna	lift					
368	ɓakkéːmina	light				ɓakk-	to light fire
369	kéːkiaʔa	look (at)				kekkeek-	to look around (freq.)
370	ɬaokína	love	ɬaw-	love	DE	ɬaw-	to love, like
371	ŋ̊ǀiɲuʔikina, ŋ̊ǀuɲuʔit̪ina (Kipini var.)	make sucking noise					
372	láːwi	pick fruit				laaw-	to pick up
373	p'úʡʡu	pierce				p'uʕud-; p'uʕ-	to pierce; to prick
374	ɬwaːʜe	pinch				ɬʷaħ-	to pinch
375	ʜeːwona	play, dance					
376	lásana	pull				las-	to pull
377	ʃúːk'ina	pull					
378	ɬwaʡe	pull back husk of maize	ɬwaːʕ-	peel	CE		
379	kokkonna	understand, remember				kokkod-	to remember
380	ɬáʔana	roast, burn	ɬaɦ-	burn up	DE	ɬaħ-	to put on fire
381	poʔʔenna	be burnt (?)				poʔeem-	to be burnt up
382	fuŋ̊ǀinna	root up	funǀuk'-	root up			
383	kwáʡana	run	kʷaʕ-	run (subj. sg)	DE	kʷaʕ-	to run away (to)
384	háːbana	saw	haβɿ-	cut, split (wood)	DE	(haaw-)	(to understand)
385	tʃikatʃikena	saw w/dull knife	tʃikʷatʃikʷ-	saw	DE	tʃikotʃik-	to saw
386	hoːɗona homme	say they said	hoːɗ-	tell	DE	hooɗ-	to say
387	bílabílina	scatter	(ɓil-)	(decorate)	CE		
388	koːranna	scrape					
389	hant̪'inna	scratch	ʔant̪'-	scratch	CE	ʕant'id-	to scratch
390	hát̪'id̪a mono	very close				háát'i	near
391	hoːni	far	hóːni	far	DE	hóóni	far
392	hula gáʡana	sense an odor	hulla; hul-; gaʕ-	smell ; smell; smell (tr.)	DE; CE; CE	(hul-)	(to stink)
393	ʜiːre	shave				ħiir-	to shave
394	sóːʔona	sing	soːʔ-	sing	DE	sooʔ-	to sing

395	ɓonna	sleep (v)	ɓom-	sleep (sg. subj.)	DE	ɓom-	to sleep
396	gwáina	sit	gwaɦ-	sit down, stay	DE	g^{w}aħ-	to stay, to live
397	ŋ̊ǀap’inna	smack	ŋǀap’ið-	smack			
398	makamúmuk a:d̪it̪t̪o	smile				mummukud-	to smile
399	ʜa:ɗinna	sneeze	ɦa:d̪ið-	sneeze	DE		
400	pí:fina	spew out water from mouth	pi:f-	spray out of mouth	CE		
401	but̪úʡuna	spit	ɓut̪:uʕ-	spit	DE	ɓut’uʕ-	to spit (impf.)
402	ɗáɬed̪ona	split	ɗaɬ-	split	DE		
403	sattsina	stand	sa:ɗ-	stand	DE	saaɗ-	to stand
404	ʔággwidʒona	steal	ʔaggwiy-	steal	DE	(ʔaggwid-)	(to weave)
405	súsúʡunna	stoke a fire					
406	gommenna	stoop, bend down	gomme:m-	bend down	CE	gomm-	to bend down
407	kw’a:dʒáʔan a	stretch body	k^{w}a:dʒ-	stretch	CE		
408	nóʔona	suck	noʔ-	suck	CE	noʔ-	to suck
409	ʡunnenna	swallow (v)	ʕonne:ð-	swallow	DE	ʕunneed-	to swallow
410	pe:lanna	sweep	pe:la:ð-	sweep	DE	peelaad-	to sweep
411	k’uk’uʔúna	swell	k’uk’uʔ-	swell	DE		
412	d̪á:kana	swim	d̪a:k-	swim	DE		
413	t̪enna	taste, try	t̪em-	try (DE), look at (CE)			
414	k’o:bokína	want	k’o:β-	want, need	CE	k’oob-	to want
415	ʔút̪unu dzwa:ʜana	wash wound	(ʔút̪unu)				
416	fi:t̪’ina	whistle	fi:t̪’-	whistle, sniff, suck	DE/ CE	(fitina)	(discord, Sw.)
417	kád̪i falona	work	ka:ði fal-	work	DE	(káádi) fal-	(work), to do, to work
418	t̪’ú:buna	wring	t̪’u:ʕ-	press, throttle	CE	t’uub-	to squeeze
419	ha:dʒunna	yawn	ha:yuð-	yawn	DE	hadʒow-	to yawn
420	ŋ̊ǀámut̪t̪e	yellow	ǀámu:t̪e	yellow			
421	kw’éʡuma	white	k’ú:ɦuma	white	DE	k’úúħuma	white
422	rú:mat̪e	tall	rú:mat̪e	long	DE	rúúmate, rúúmatiddze (f), ruumáti (plu)	tall, deep
423	ɬáʜame	sweet, tasty	ɬá:ɦame	sweet	DE	ɬááħame	sweet

424	ʔaminina (tones unknown)	small, little	ʔa:mína	small	DE	ʔaamína	small
425	ʡumat̪'e	short	ʕúmmat̪'e	short	DE		
426	ʎ̥ákane	sharp	ɬákane	sharp (knife)	DE	ɬákane, ɬákani	sharp
427	t̠'iraraʔe	red	ts'íraraʔe	red	DE	ts'íraraʔe, (f=) -ittse, (pl)-i	red
428	ga:no	big	ga:no	big	DE	gaano, gáána	large, big, grown-up
429	k'áʡʡime	many	kaʕime	many	DE	káʕime	many
430	ŋ̊ǀúŋkuma	laziness, lazy	ǀú:kuma	lazy			
431, cf. 184	ŋ̊ǀábat̪e	good smelling	ǀaβa	nice smell (of oil)			
432	wíne	good	wíne, wino	good, clean (m)	DE	víne, vínaddza (f), vino (pl.)	
433	ŋ̊ǀáʔat̪e	dry (thing)					
434	ŋ̊ǀáʔe	dry				ǀaʔ-	to be dry
435	méŋ̊ǀet̪e	carefully	méŋǀet̪e	slowly	DE	mé^{n}ǀate	slowly (sg?)
436	himmat̪t̪e	black	hímmat̪e	black	DE	hímmate	black
437	ŋǀúʔit̪e	bitter, fierce	ŋǀuʔit̪e	st that hurts body			
438	k'árare	bitter	k'ára:re	bitter	DE	k'áraare	bitter
439	mpí:t̪e	bad	mbí:t̪e, (mpit̪e), (mpi:t̪e)	bad, dirty (ms)	DE, (AT), (CE)	^{n}bíite, ^{n}bíidʒadʒdʒi	bad, ugly
440	williʡine (tones unknown)	be cold, get cold	wíliʕine	cold	DE		
441	ŋ̊íɲit̪t̪u	darkness					

The Phonetic Structures of Hadza

Bonny Sands, Ian Maddieson, Peter Ladefoged

Hadza is a language spoken in Tanzania by approximately 800 people. Some researchers classify it as a Khoisan language (Bleek 1931, Greenberg 1966, Ehret 1986), while others maintain that it is a language isolate (Woodburn 1962, Elderkin 1983). Hadza is typologically unusual in being, along with Sandawe and Dahalo, one of the few languages outside of southern Africa to have clicks. A careful and systematic description of the sounds of Hadza is an important first-step in research into the possible linguistic genetic affiliation of the language. A basic phonetic description of any language is relevant for the study of cross-linguistic universals, and a description of Hadza is particularly important for the insights it can provide into the characteristics of some typologically rare sounds such as clicks. This paper discusses some of the articulatory and acoustic characteristics of the clicks in Hadza, as well as the characteristics of all the other sounds. Hadza has a rich consonant inventory, with ejective stops and affricates, and lateral fricatives and affricates. The contrastive states of the glottis for different manners of consonant articulation will also be discussed, as well as the characteristics of the vowel system.

Fieldwork on the Hadza language has been carried out by a number of researchers: D. Bleek (1931, 1956), O. Dempwolff (1916-17), E. Obst (1912), P. Berger (1943), A. Tucker and M. Bryan with J. Woodburn (1977), E.D. Elderkin (1982, 1983), A.J. de Voogt and M. de Meij (de Voogt 1992), and J. Wagner (forthcoming). Despite all this attention, the phonetic characteristics of the language are not well understood. The sound system of Hadza has been described in some detail in a few studies, notably Tucker and Bryan (1977) and de Voogt (1992); but no basic acoustic or instrumental data has been reported. This is significant, especially as there are discrepancies in the phonemic inventories as reported by different researchers.

This study is based on field transcriptions and instrumental analyses. Field work was done in Mangola, near Lake Eyasi in north-central Tanzania, in August 1991 by all three authors. In addition to audio recordings of a fairly complete range of phonetic phenomena, the articulatory characteristics of clicks and lateral affricates were also recorded by the use of palatograms and linguograms Palatal casts were also made to assist in interpreting the palatographic and linguographic data. The first author returned to Mangola for further fieldwork from January to June 1992. The analyses reported here are based on the material recorded during the first field trip, supplemented by observations of some additional words noted during the second period of fieldwork.

All of the consultants for this study resided in the Mangola area and were considered to speak a uniform dialect. Speakers in some areas are considered to be more pronouncedly influenced by Isanzu or Sukuma, both neighboring Bantu languages. The differences between dialects are primarily in the lexicon and not in the sound system, and will not be discussed here. Four women and three men ranging in age from early 20's to early 50's were recorded saying the list of words in the Appendix. All acoustic analysis was conducted on words on this recording.

Consonants

Overview: The consonant inventory of Hadza is shown in Table 1. Words illustrating these sounds are shown in phonemic form in Table 2. The numbers in parentheses refer to the order of words as they appear on the recording which is transcribed in the Appendix. Some words are cited which do not appear on the recording, as they were observed during the second period of fieldwork.

Table 1: Hadza Consonants

	Bilab.	Labio dental	Dental	Alv.	Alv. Pal.	Pal. Alv./ Pal.	Velar	Lab. Velar	Glot.
Plosive	p^h p b			t^h t d			k^h k g	k^hw kw gw	ʔ
Ejective	(p’)						k’	k’w	
Central Oral Click			kǀ		k!				
Lateral Oral Click						kǁ			
Pulmonic Nasal	m			n		ɲ	ŋ	ŋw	
Nasal Cen.Click			ŋ̊ǀ’ ŋǀ		ŋ̊!’ ŋ!				
Nasal Lat. Click						ŋ̊ǁ’ ŋǁ			
Pul. Pre-nas. Stop	mp^h mb			nt^h nd			$ŋk^h$ ŋg		
Pul. Pre-nas. Affr.				nts ndz		ndʒ			
Pul. Cen. Affricate				ts dz		tʃ dʒ			
Pul. Lat. Affricate						tʎ̥			
Ejec.Cen. Affricate				ts’		tʃ’			
Ejec.Lat. Affricate						tʎ̥’			
Central Fricative		f		s		ʃ			
Lateral Fricative				ɬ					
Central Approx.						j		w	ɦ
Lateral Approx.				l					

Place of Articulation: The number and type of contrasts in place of articulation differs depending on the manner of articulation. Plosives contrast five places: bilabial, alveolar, velar, labialized velar and glottal. Ejectives contrast three places: bilabial, velar and labialized velar. Clicks contrast three places; pulmonic nasals five places; fricatives three places; affricates two places and prenasalized stops three places. The terms used for the places of articulation in this overview section are simply indicative of the general phonetic categories involved, which will be made more specific later in this paper.

Plosives: Hadza has a large number of stop consonants. There are voiced, voiceless and aspirated pulmonic stops at bilabial, alveolar, velar and labialized velar places of articulation. There is also a glottal stop.

Ejectives: Ejective stops occur at velar, labialized velar and bilabial places of articulation. Ejective [p'] is a marginal sound occuring in only a few lexical items.

Clicks: There are three click types in Hadza: dental [|], alveolopalatal central [!] and lateral [||]. The lateral click is produced with an alveolar closure and a lateral palatoalveolar release. Each of these types occurs with three click accompaniments: voiceless oral [k|, k!, k||], voiced nasal [ŋ|, ŋ!, ŋ||], and voiceless nasal with glottalization [ŋ̊|', ŋ̊!', ŋ̊||']. A vowel preceding a click with voiceless nasalization will itself be nasalized, i.e. [hãŋ̊!'a-k^{h}o] 'rock'.

Nasals: There are pulmonic nasal consonants at bilabial, alveolar, palatal, velar and labialized velar places of articulation. Voiceless and voiced nasal clicks also occur for the dental, post-alveolar, and lateral click types.

Prenasalized stops and affricates: There are prenasalized stops occuring at bilabial, alveolar and velar places of articulation. Prenasalized affricates occur at dental and palatoalveolar places of articulation. There are no prenasalized lateral affricates or prenasalized labialized velar stops.

Affricates: Voiceless and ejective affricates occur at two places of articulation: alveolar and palatoalveolar. Like the fricatives, the palatoalveolar affricates may have either a central or a lateral articulation. There is a contrast between voiced and voiceless among central alveolar and palatoalveolar affricates. There is no voicing distinction for the lateral affricates. None of the affricates is distinguished by degree of aspiration, unlike the voiceless stops.

Fricatives: Fricatives in Hadza occur at three places of articulation: labiodental, alveolar and postalveolar. Alveolar fricatives can have either central [s] or lateral [ɬ] articulations. All the fricatives are voiceless.

Approximants: Hadza has a labiovelar approximant [w], a lateral alveolar approximant [l], a palatal approximant [j] and a glottal approximant [ɦ]. In intervocalic position, the approximant [l] can appear as [ɾ].

Table 2: Words Illustrating Contrastive Consonants of Hadza

Initial			# on tape	Medial		# on tape
Bilabial:						
p^h	*p^{h}andʒu-p^{h}e*	'sp. plant'	32	*ʔúp^{h}úk^{h}wa*	'leg'	**101**
p	*patáku'ʃé*	'palm of hand'	99	*ɬupá-k^{h}o*	'foam'	81
b	*badá*	'hole'	44	*ŋ||obá-k^{h}o*	'baobab'	138
p'	*p'áʔùwé-*	'to split'				
m	*mák^{h}o*	'clay pot'	85	*sámak^{h}a-p^{h}i*	'three'	**37**
mp^h	*mp^{h}alamaɦo-k^{h}o'*	'slingshot'		*ɦomp^{h}ai-k^{h}o*	'wing'	
mb	*mbalata-k^{h}o*	'cockroach'		*k||amba-bi*	'small intestine'	186

Labiodental:						
f	*fá-*	'to drink'	11	*ts'ifi*	'night'	**46**
Dental:						
k|	*k|útʰi-*	'neck'	177	*k|ak|a*	'large, flat rock'	
ŋ|'	*ŋ|'ats'e-*	'to reheat'	211	*tá ŋ|'e*	'belt'	140
ŋ̊|	*'ŋ̊|aítʰá*	'tongue'	228	*k|ikiríˈŋ̊|a*	'pinky finger'	181
Alveolar:						
tʰ	*tʰasé*	'long'	123	*átʰa'má*	'blood'	**107**
t	*titʃ'i-*	'black'	120	*patáku'ʃé*	'palm of hand'	99
d	*daraŋga*	'flour'		*badá*	'hole	44
n	*nátʰi*	'donkey'		*ʔéna-pʰi*	'grass'	**30**
ntʰ	*ntʰuli-bi*	'beer'		*ʔintʰawe*	'nose'	
nd	*ndagwe-ko*	'notch'		*ŋ!andá*	'agama lizard'	97
l/ɾ	*lalá-kʰo*	'gazelle'	67	*baʔaɾá-kʰo, baʔalá-kʰo*	'honey'	91
ts	*tsipit'í*	'porcupine'	65	*tsetse-*	'to grow old'	
dz	*dzá-*	'come!'	2	*tʎ̥'odzo-*	'to say'	19
ts'	*ts'áke-*	'to steal'	24	*ɦits'á-pʰe*	'fat'	**89**
nts	*ntsá-kʰo*	'star'	47	*tan(t)se-*	'to crack'	
ndz	*ndzopʰa*	'bottle'		*mindza*	'reedbuck'	
s	*sámaka-pʰi*	'three'	37	*pápa'sa*	'hip bone'	109
ɬ	*ɬanó*	'python'	95	*ŋ̥!'ak'iɬá*	'palate'	179
Alveolopalatal:						
k!	*'k!ákú-*	'to jump over'	4	*k!o'k!ó-kʰo*	'back of head'	137
ŋ̥!'	*ŋ̥!'ojé*	'wax'	139	*ɦaŋ̥!'á-kʰo*	'rock'	184
ŋ!	*ŋ!ána-*	'sp. mongoose'	64	*ŋ!ikiŋ!i-*	'to push a lot'	
Palatoalveolar/Palatal:						
ɲ	*ɲau-wa*	'cat'	56	*moɲóda*	'salt'	**92**
tʃ	*tʃatʃa*	'bushbaby'	55	*ʔitʃáme*	'one'	36
dʒ	*dʒándʒai*	'leopard'	167	*gubidʒi-*	'to get s.t. ready'	
tʃ'	*'tʃ'á-kʰo*	'guineafowl'	117	*ɦatʃ'apitʃ'i-kʰo*	'ear'	
ndʒ	*ndʒa*	'reedbuck'		*dʒandʒai*	'leopard'	167
ʃ	*ʃamu-ko*	'Swahili'		*andáʃa*	'caracal'	
tʎ	*tʎ̥ákáte*	'rhino'	183	*kwatʎ̥a*	'shoe'	84
tʎ'	*tʎ̥'áʔa-*	'to sing'	22	*mitʎ̥'aː*	'bone'	**142**
k||	*k||a'pʰá*	'stump'	190	*kak||á-*	'to hunt'	14
ŋ̥||'	*ŋ̥||'ekʰwá*	'sp. root'	188	*kʰaŋ̥||'é-*	'to jump'	203
ŋ̊||	*'ŋ̊||áʔa-*	'to scavenge'	214	*koŋ̊||aɦete*	'man w/2wives'	
j	*jámu-a*	'land'	40	*ʔijátu-bi*	'snakes'	96

Velar:						
k^h	*k^{h}alimo*	'animal'	52	*mak^{h}o-wa*	'clay pot'	85
k	*káŋga*	'sp. mongoose'	57	*ɦaká-*	'to go'	13
g	*gaʃa-bi*	'honey beer'		*damoga-k^{h}o*	'beard'	
k'	*k'apáku-bi*	'jaws'	114	*ts'ik'ó*	'smoke'	43
ŋk^h	*ŋk^{h}ólo-'á-k^{h}o*	'heart'	108	*ts'aŋk^{h}a*	'sp. mongoose'	58
ŋg	*ŋgat^{h}á*	'head ornament'	113	*k!oŋga*	'hare'	**158**
ŋ	*ŋaŋa*	'kind of fruit'	50			
Labialized Velar:						
k^{h}w	*k^{h}wak‖a-*	'to vomit'	224	*uk^{h}wá-k^{h}o*	'arm'	98
kw	*kwaʔí*	'warthog'	68	*ŋ̊‖'ekwa*	'sp. root'	188
gw	*gwanda-k^{h}o*	'shirt'		*ɦagwanda*	'adolescent animal'	
k'w	*k'waʔu-k^{h}o*	'eggshell'		*ɦek'wa-be*	'shell, rind'	
ŋw	*ŋwapo-k^{h}o*	'ditch'				
w	*watʃ'o*	'sp. mongoose'	63	*ʔáwawa*	'bee'	**93**
Glottal:						
ʔ	*ʔá'ɦú*	'skin'	104	*t̥ʎ'óʔa-k^{h}o*	'a skin'	103
ɦ	*ɦaka*	'to go'	13	*k^{h}áɦa*	'to climb'	8

Vowels

Hadza has five contrastive vowel qualities [i, e, a, o, u], as shown in Table 3.

Table 3: Words illustrating the five vowel qualities of Hadza

i	*ŋ‖i-ʔi*	'put poison on f arrow'	210
e	*ŋ‖e-ʔe*	'put poison on m arrow'	209
a	*ŋ‖áʔa-*	'to scavenge'	214
o	*ŋ̊‖'o-ʔo*	'wash, bathe'	226
u	*ŋ‖uʔú-*	'to snore'	

Length is not contrastive, although there are phonetic differences in length which correlate with differences in pitch/accent. Long vowels may occur as the result of the addition of an affix to a word, i.e. /uk^{h}wa-a-k^{h}o/ 'It is an arm.' [uk^{h}waːk^{h}o]. Reduction of intervocalic [ɦ] can also result in a long vowel, i.e. [k^{h}aɦa]/[k^{h}aː] 'to climb'. Final vowels may become voiceless [i̥, e̥, ḁ, o̥, u̥], particularly when preceded by a glottal stop or any other voiceless stop.

Nasalized vowels occur in the environment preceding a nasalized or voiceless nasalized click. The vowels [ĩ, ũ] occur in two recorded lexical items in which their nasality cannot be predicted from the environment. In both of these examples, the nasalized vowel is followed by [ɦ]. It is possible that these words may once have contained nasalized clicks.

Tone and stress

The role of tone and stress in Hadza are not entirely clear. Tucker, Bryan and Woodburn (1977) transcribe stress, high, low and mid (unmarked) tone, although they are careful not to claim that these are contrastive units. They mark five tonal classes for the nouns (in the grid: ___*bàhèà* 'there is'): MMM, MML, MHH, HML, HMH. For the verbs they note four tonal classes (using the first person singular future form as a grid): LH, HL, HH, LL.

Words in this article are transcribed with high tone [´], and stress [']where needed. Low tone is unmarked. The tonal pattern of a word often shows a great deal of variance from utterance to utterance, i.e. [ŋǁ'ekʰwá], [ŋǁ'ékʰwá], [ŋǁ'ēkʰwā] 'sp. root' (188). High toned syllables are typically longer and more stressed than low toned syllables. We have found no minimal or near minimal pairs which contrast mid tone with either high or low tone. In our preliminary analyis, we believe that these facts, as well as the overall behavior of tone and stress can best be accounted for by analysing Hadza as a pitch-accent language. We have identified at least two melodies: LHL and HL; but the interaction of these melodies with additional morphemes bound to the root has not been fully worked out.

Place of articulation

The description of the place of articulation of the consonants was based on field transcriptions combined with the questioning of the consultants about their articulations, and instrumental palatographic records. Because only a few studies, such as Traill [1985], Ladefoged and Traill [1984], Doke [1923] and Beach [1938] have described the place of articulation of clicks with the use of instrumental techniques, these sounds were given particular attention. There has been some disagreement in the description of the Hadza click types. Bleek (1956), and Greenberg (1966) following her example, transcribe a fourth click [ǂ]. Tucker, Bryan and Woodburn (1977) in addition transcribe a bilabial click and a "flapped" version of the [!] click, transcribed [!!].

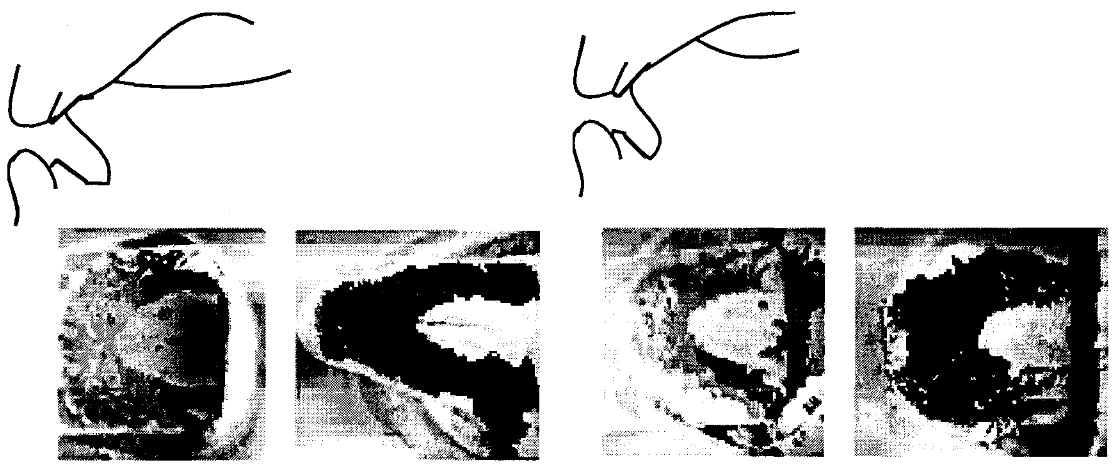
The two instances in which they give an example of a bilabial click are in greetings. Our consultants had aspirated bilabial stops where Tucker, Bryan and Woodburn (1977) transcribe either a bilabial or dental click. Neither click was considered an acceptable substitute for the pulmonic stop, however it was acceptable to precede the greeting with a labio-manual click--a kiss on one's own hand. We will consider later the possibility of a flapped version of the [!] click.

All of the words which Bleek transcribed with the [ǂ] click have been transcribed by Sands (forthcoming) with other sounds such as [!], [ǁ] and [k']. The differences in transcription appear to be due to errors, rather than to any linguistic change which may have occured over the 60 years separating the fieldwork.

Palatograms and linguograms were made for two adult male speakers of Hadza. Each speaker uttered words which contained either a dental, lateral or alveolopalatal click, using techniques described by Ladefoged (1993). The relationship of the place of articulation of clicks with that of other consonants has not been extensively studied. Palatograms and linguograms of the ejective lateral affricate were also made, as this sound had a striking acoustic and articulatory similarity to the lateral click. The speaker uttered a word twice before the contact area was photographed and/or videotaped. The video image was later digitized using a Macintosh computer equipped with a video capture card. For each speaker, a dental impression was made, showing the shape of the roof of the mouth. This was used to create a sagittal view of the fixed structures of the speaker's vocal tract.

The dental clicks [ǀ] can be described as having a laminal coronal closure, extending from the upper teeth to the alveolar ridge. This can be seen in Figure 1, which shows palatograms and linguograms of the front articulation in a dental click, as produced by two speakers. The sagittal views were constructed from the information in the palatograms and linguograms. Although the back click closure cannot be seen on either the palatogram or linguogram, evidence for a closure along the sides of the mouth can be seen. The sides of the tongue made contact with the sides of the palate. This closure, along with the front and back closures are necessary for the creation of a suction chamber which allows for the influx of air characteristic of a click release.

The palatograms of the dental click shown in Figure 1 (and those of the alveolopalatal and lateral clicks in Figures 2 and 3, which we will be discussing later) are markedly different from palatograms of the corresponding clicks in languages spoken in Southern Africa, such as !Xóõ (Traill 1985) and Zulu (Doke 1923). In these other languages the back closure extends further forward, so that the contact of the back of the tongue is visible on the palatograms. The sagittal diagrams in Traill (1985) and in Ladefoged and Traill (1984, 1993), which are based on x-ray cinematography, also show that in the Southern African click languages at the onset of the formation of the click there is usually a smaller enclosed air space than our palatographic records indicate for Hadza. This may be due to the fact that these Hadza speakers have higher palates than those of the !Xóõ speakers, who have shallow palates and virtually no alveolar ridge: We cannot tell how our speakers compare with Doke's Zulu speakers, for whom the shape of the roof of the mouth is not given.



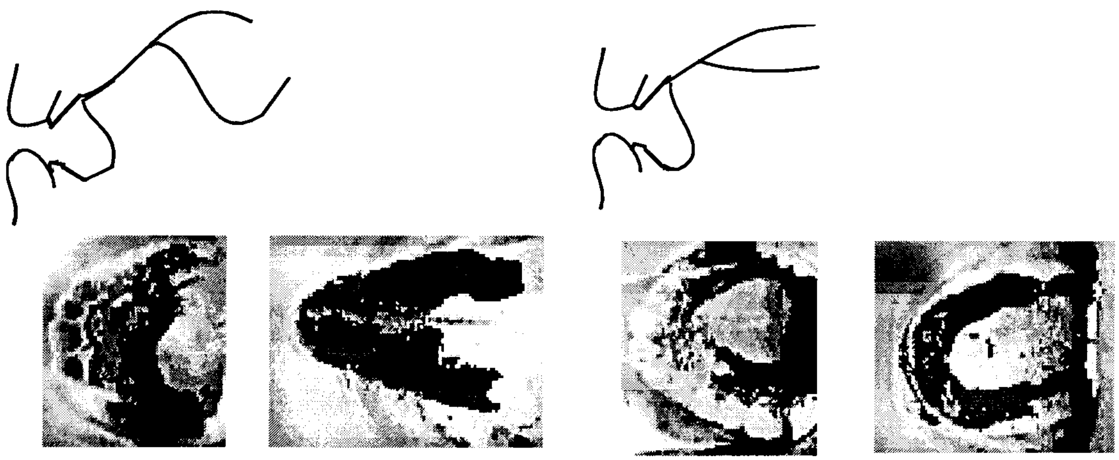
sp.1 sp.2

Figure 1: Palatograms and Linguograms of a dental click in the word 'forget' [ŋǀaha] as spoken by two male Hadza speakers. The sagittal view of each articulation was inferred from the patterns of contact on the tongue and palate.

The alveolopalatal click [!] in Hadza was noted to vary a great deal in terms of how forcefully it was produced by speakers. In some instances, the amplitude of the release was very low, as if the click were produced with very little suction. A notable allophonic variation of the [!] click occured for a number of speakers. In these instances, the tongue tip makes contact with the bottom of the mouth after the release of the front click closure. The release of the front closure and the contact with the bottom of the mouth is one continuous, ballistic movement, with the underside of the tip of the tongue making a percussive sound as it strikes the floor of the mouth. This is the articulation which Tucker, Bryan and Woodburn (1977) characterized as a flapped palato-alveolar click. It is quite clearly a free variant of the unflapped [!] and not a separate phoneme. The within-speaker variation is not reported from any Southern African languages with clicks (Traill, personal communication).

Palatograms, linguograms and sagittal sections of the front articulation of the [!] clicks can be seen in Figure 2. Like the clicks in other languages transcribed [!], the front closure of these clicks in Hadza tends to be made at less anterior place of articulation than the [ǀ] click, and is typically more apical. This is certainly the case for speaker 2, but speaker 1 shows more similarity in his articulations for [ǀ] and [!]. The linguogram for speaker 1 shows front closure contact on the tongue to be similar in length and location for both [ǀ] and [!], but the dental click differs in the shape of the area in the middle of the tongue which did not make contact with the roof of the mouth. In the dental click, this area is tapered toward the front, whereas the alveolopalatal click displays a more rectangular shape for the corresponding area. The linguograms and palatograms indicate that, at the midline, the tongue behind the contact is more sharply lowered for the dental than the alveolopalatal click.

The palatogram of the alveolopalatal click for speaker 1 also shows that contact was made at the back of the front teeth, yet this contact does not extend to the base of these teeth. The blackened area on the front teeth must be the result of a light contact, otherwise we would expect a continuous contact area. The contact is not consistent with a laminal dental articulation. The contact may be the result of the tip of the tongue quickly flipping against the teeth after the front closure contact is released. The length of the front click closure is somewhat longer for speaker 1 than for speaker 2. This is consistent with the idea that speaker 1 articulated the click with a rather forceful release. The apical contact would have been post-alveolar initially, but later alveolar as the cavity behind the closure was enlarged to lower the intraoral pressure.



sp.1 sp.2

Figure 2: Palatograms and Linguograms of the alveolopalatal click in the word ‘cut’ [ŋ!eʔe] as spoken by two male Hadza speakers. In this and subsequent palatograms, the sagittal view of each articulation was inferred from the patterns of contact on the tongue and palate.

The third type of click found in Hadza, the lateral click [ǁ], is especially interesting because of its similarity to the lateral ejective affricate. Acoustically and articulatorily these sounds are very similar. They are both produced with a laminal closure and with a ring-like closure along the sides. For many speakers, the lateral release in these sounds occured far back in the mouth, and could be properly characterized as a lateral palatal release. Several of our field transcriptions show that we transcribed the lateral ejective as [cʎ̥’], or even as [kʟ’]. However, temporarily, because of the location of the central constriction, these sounds are characterized as alveolar obstruents rather than as palatals. Figure 3 shows the lateral click for the two speakers, and figure 4 the lateral ejective. Both laterals for speaker 1 are made with the tip of the tongue down. The laminal contact is on the teeth and alveolar ridge for the click, but only on the alveolar ridge for the ejective. For speaker 2, the tongue tip also appears to be down during both laterals. Contact occured from the bottom of the top front teeth to the back edge of the alveolar ridge. Unfortunately the speaker did not open his mouth sufficiently when the photograph was taken, and his upper teeth prevent us from seeing the backward extent of the contact in the ejective.

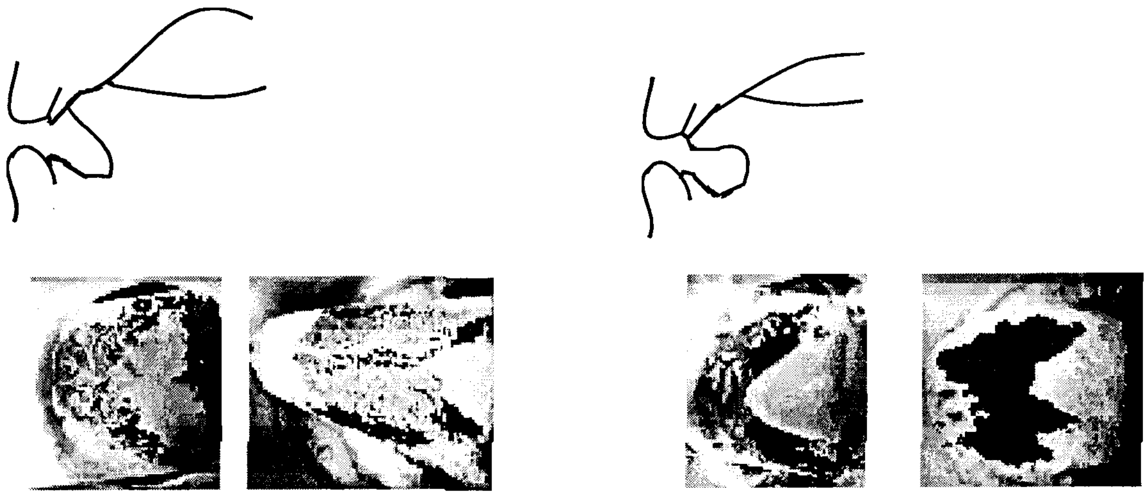


Figure 3: Palatograms and Linguograms of a lateral click in the word 'scavenge' [ŋǁaʔa] as spoken by two male Hadza speakers.

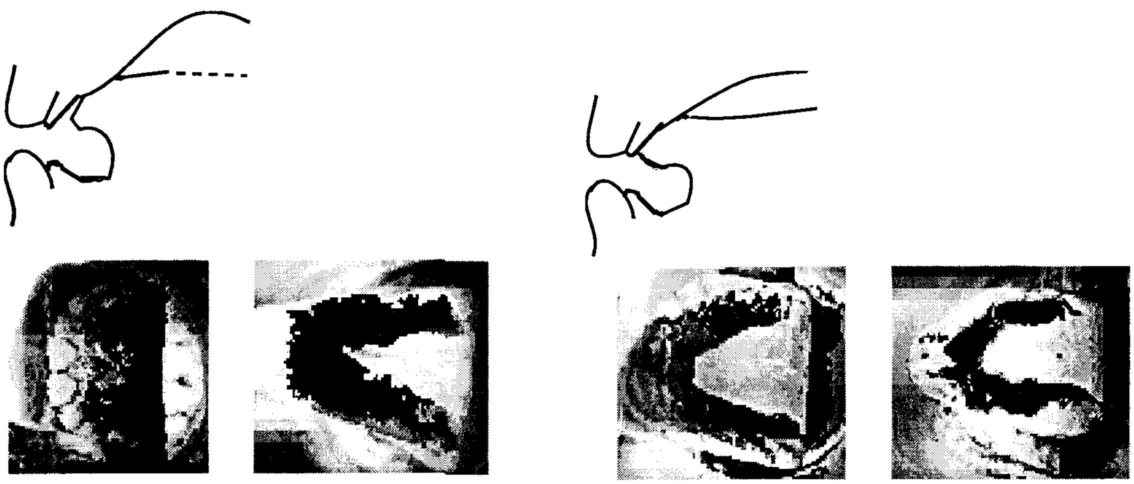


Figure 4: Palatograms and Linguograms of a lateral ejective affricate in the word 'bone' [mitʎ̥'a] as spoken by two male Hadza speakers. The position of the tongue is shown by a dashed line for speaker 1 as the mouth was not open sufficiently and the extent of contact cannot be seen.

Click accompaniment

Each of the three types of clicks, [ǀ, ǁ, !], can have three different accompaniments which can be associated with the velar articulation that occurs in these clicks. The pattern of accompaniments that occurs in Hadza is very different from the patterns that occur in the Khoisan and Nguni languages of Southern Africa. (See Ladefoged and Traill (1993), for a discussion of possible accompaniments to clicks.) In those languages, the first possibility is a plain velar stop [k], giving [kǀ, kǁ, k!]. A waveform of a word in Hadza containing a click with this accompaniment type is shown in Figure 5. We will discuss the degree of aspiration in these clicks in a later section. The second possibility is the voiced velar nasal [ŋǀ, ŋǁ, ŋ!], as shown in Figure 6. The third accompaniment is more complex; it is both nasalized and glottalized. This voiceless nasal accompaniment is transcribed [ŋ̊ǀ', ŋ̊ǁ', ŋ̊!']. The glottalization takes the form of a glottal stop which is formed during the click closure, and released well after the release of the anterior click closure, so that there is a delay before the onset of voicing. The voiceless nasal airflow can be heard as a voiced nasal accompaniment when the click follows a vowel, even across word boundaries. The following vowel is somewhat nasalized. In all environments, the nasalization can be detected by placing a hand in front of the nose of the speaker. Speakers identify both the voiced and the voiceless nasal clicks as having nasal airflow. The waveform in Figure 7 clearly shows that the closure for this click is voiceless. The voiced and voiceless nasal click accompaniments can be difficult to distinguish in intervocalic position, but as Figures 6 and 7 show, the voicing contrast is not neutralized.

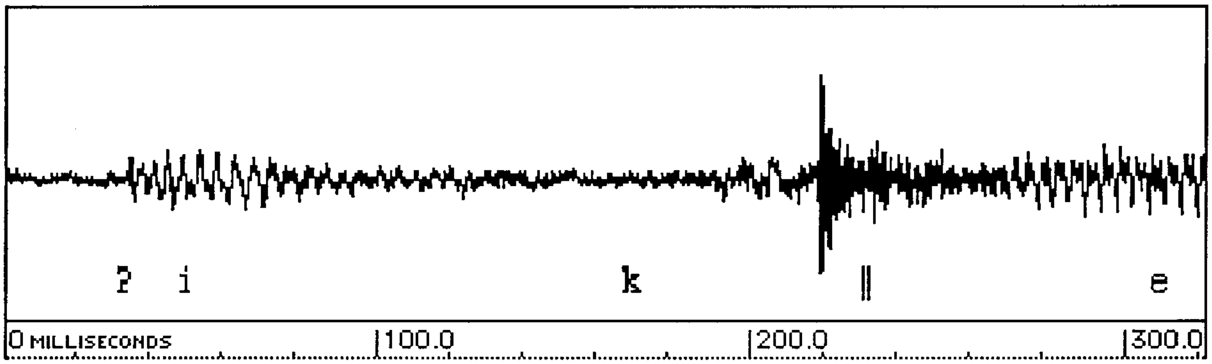


Figure 5: Waveform of voiceless click in intervocalic postition (from /ʔikǁe-ʔe/ 127)

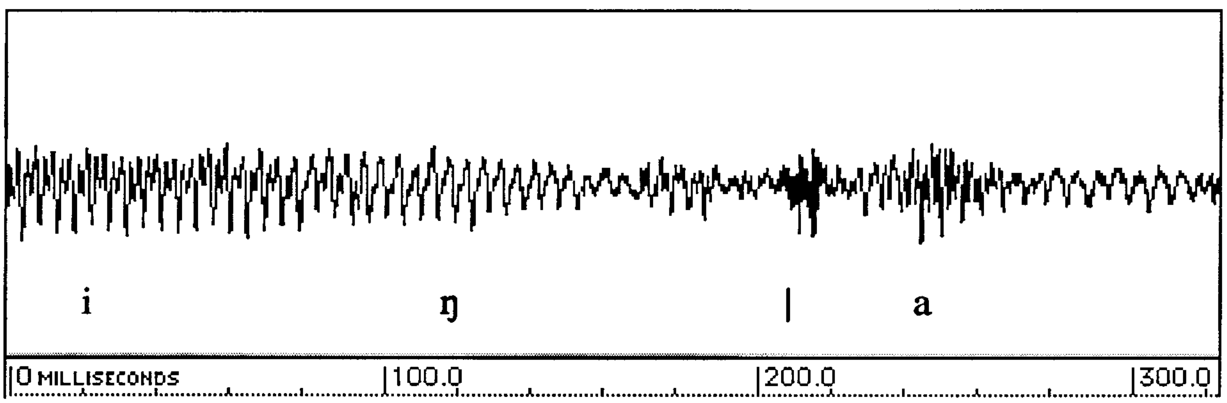


Figure 6: Waveform of voiced nasal click in intervocalic position (from /k|ikiliŋ|a/ 181)

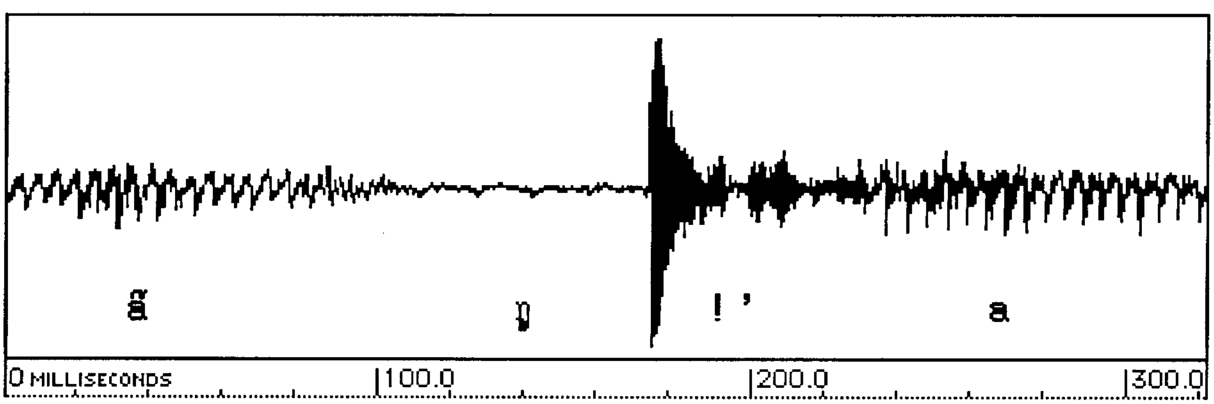


Figure 7: Waveform of voiceless nasal click in intervocalic position (from /haŋ!'a-ko/ 184)

Other researchers have distinguished different sets of accompaniments. Bleek (1958) notes additional click accompaniments [‖kx, ‖k'', g‖], we do not consider these possibilities to be phonologically contrastive. They seem to be simply Bleek's transcriptional variants. Tucker, Bryan & Woodburn (1977) note a nasal click and a nasal compound click allophone of a pausal, or glottalized click. A. de Voogt (1992) transcribes, nasal, aspirated (glottalized), 'simple' glottalized (without delay in voice onset, possibly not glottalized) and glottalized with delayed release. Both sets of researchers fail to note that the "glottalized click with delayed release" or "pausal" click is not merely nasalized intervocalically, but in all environments.

Voice Onset Time

As with place of articulation, there is some disagreement in the literature as to the nature of the contrastive laryngeal states that accompany the consonants. As with all contrasts in the language, there are few minimal pairs to serve as a guide. The contrasts in Voice Onset Time have been particularly difficult for researchers to untangle. Also of interest is the fact that clicks, stops with single articulations, and affricates all pattern differently with respect to phonation type.

In order to investigate VOT, measurements were taken of clicks, pulmonic and ejective stops and affricates in lexical items produced in isolation by 7 speakers of Hadza. Each of these words were said twice, with some lexical roots having additional repetitions. These additional repetitions were made at a different place on the word list and are averaged separately. Measurements were made by looking at simultaneous displays of spectrograms and waveforms on a Kay Elemetrics Computer Speech Lab, with speech digitized at 10 kHz. The duration measured was from the release burst of the consonant to the onset of voicing of the following vowel.

The consonants made with a single place of articulation (bilabial, alveolar, velar) occur phonetically voiced or voiceless aspirated. The voiceless aspirated stops can be further characterized as less aspirated and highly aspirated. The voiceless stops are transcribed as aspirated [p^h, t^h, k^h] and unaspirated [p, t, k], but both stop series are aspirated to some degree. For ease of transcription, we will refer to this contrast as unaspirated vs. aspirated. Measurements of VOT for pulmonic velar stops in 22 separate lexical items were taken. The means and standard deviations of the VOTs for pulmonic velars by word is shown in Figure 8. The first 14 words from the left clearly group together, separate from the 9 rightmost words. The bimodal distribution of VOT can be more clearly seen when the word [ŋk^holo-wa-k^ho], the only prenasalized velar stop in the corpus, is removed. There does not seem to be a two-way contrast in aspiration for the prenasalized stops. The degree of aspiration for prenasalized stops does not correspond to that of the less aspirated or highly aspirated stops. Similar results were found for bilabial and alveolar stops, although smaller data sets were used.

The velar stops additionally contrast an ejective [k'] with the voiceless aspirated [k^h] and unaspirated [k]. This contrast is marginal with the bilabials, but was not found for the alveolars. The VOT measurements for pulmonic velars shown in Figure 8 were compared with measurements of ejective velars in 5 lexical items. VOT was found to be significantly different for /k'/ and /k/ ($p<.0001$) using an ANOVA with speaker and phonation type as variables. There was a tendency for the ejectives to have greater VOTs than the highly aspirated stops, but this difference was not significant.

There is a small but significant difference in VOT between voiceless oral clicks and voiceless nasal clicks with glottalization. The voiceless nasal clicks tend to have longer VOTs. This difference was found to be significant ($p=.0042$) in an ANOVA with speaker and syllable position and accompaniment type as independent variables comparing VOT of [k!] and [ŋ!'] clicks. These click accompaniments are also distinguished by the presence or absence of nasalization on any immediately preceding or following vowel. Differences in VOT between pulmonic and ejective affricates were also found, with the ejective affricates having longer VOTs. These differences tended to be somewhat greater than the differences in the clicks.

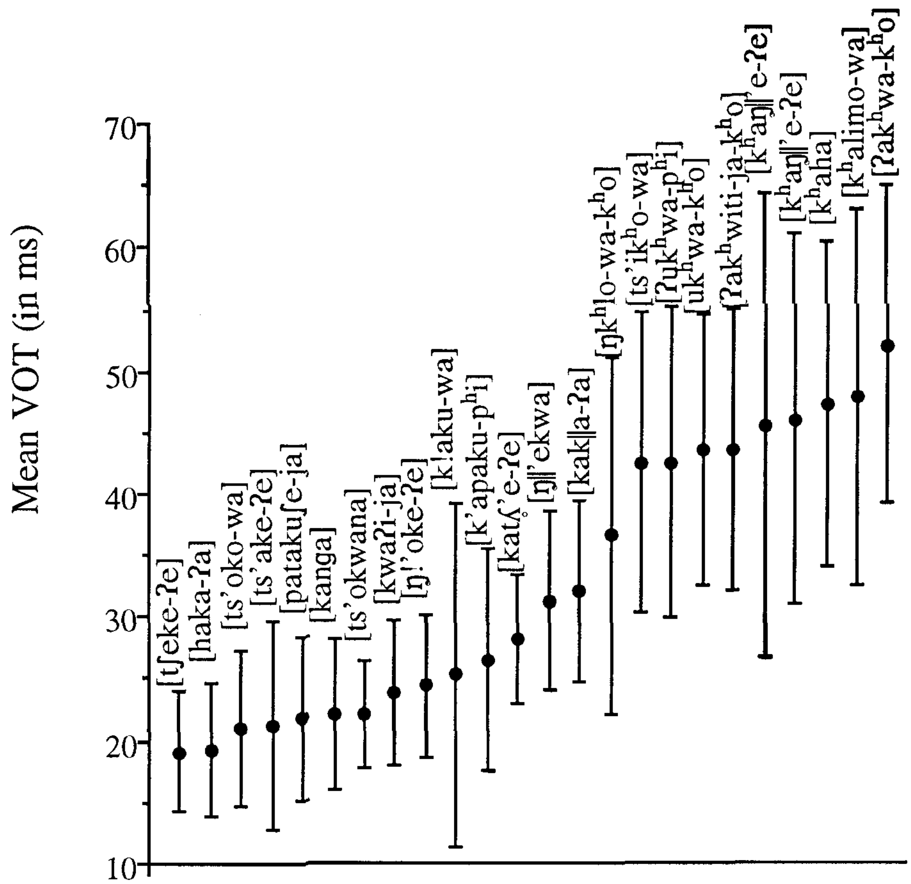


Figure 8: Plot of the mean VOT of a velar stop [kʰ, k] in 22 separate lexical roots. VOT was measured for 7 speakers, producing each item twice. Each mean represents no fewer than 9, and no more than 14 separate tokens.

As in this study, Tucker, Bryan & Woodburn (1977) note that an initial consonant has a very short VOT in a word where the first and second syllables are otherwise the same. During the course of our fieldwork, we noted that this generalization holds for plosives, affricates and clicks. That is, these consonants will have a shorter VOT if they are the initial consonant in the first of two identical syllables than if they occur in a non-identical sequence. VOTs for initial [k!] clicks in words where the first two syllables are identical were found to be significantly different ($p<.0001$) from other [k!] click VOTs, using a paired, two-tailed T-test. The words used in this comparison are shown in Figure 9. As can be seen in this figure, the mean VOTs for the initial clicks in [k!ok!oroma] 'epiglottis' (129) and [k!ok!o-akʰo] 'back of head' (137) tend to be shorter than VOTs for other clicks. A similar comparison with /p, pʰ/ showed the same effect.

In contrast with the pulmonic consonants, we have not observed a distinction between aspirated and unaspirated affricates and clicks. Tucker, Bryan & Woodburn (1977) transcribe an aspirated/unaspirated constrast in the pulmonic affricates and clicks, and de Voogt (1992), transcribes this contrast for the pulmonic affricates, but feels it may be due to allophonic variation.

The lack of a contrast between aspirated and unaspirated voiceless clicks can be seen in Figure 9, a plot of mean VOTs for voiceless oral [k!] clicks. The variation in the means of the three repetition sets of the word [k!eʔe] (10, 196, 216) can be seen to be quite large. As discussed previously, shorter VOTs were found for initial consonants in a word where the first two syllables are identical, such as [k!ok!oroma] ‘epiglottis’ (129) and [k!ok!o-ak^{h}o] 'back of head' (137). The other variation in VOT of clicks occurs when the following syllable contains a nasal. In the data set shown in Figure 9, four words have an initial click with a nasal in the following syllable: [k!uni-p^{h}e] (144), [k!uma-k^{h}o] (145), [k!uma-ʔe] (20) and [k!oŋga-a] (158). Using an ANOVA, with speaker as an independent variable, these clicks had significantly lower VOTs ($p<.0001$) than the other clicks (excluding the initial clicks in [k!ok!oroma] (129) and [k!ok!o-ak^{h}o] (137)) in the data set. Given that both nasalization of the following vowel and a longer VOT are cues to the voiceless nasalized click accompaniment as contrasted with the voiceless oral accompaniment, we might expect a trade-off between these cues to be possible. In the case where a vowel becomes partly nasalized due to the effect of a following nasal consonant, it becomes more difficult for the listener to determine whether a preceding click is oral or nasal based on the cue of the nasalization of a following vowel alone. It is not surprising that voiceless oral clicks which have a nasal in the following consonant would have shorter VOTs than when a nasal does not follow. The VOT cue is exaggerated in the context where the cue of nasalization is more ambiguous.

Similar comparisons for the other clicks and affricates were made. Although smaller data sets were used, similar results were obtained.

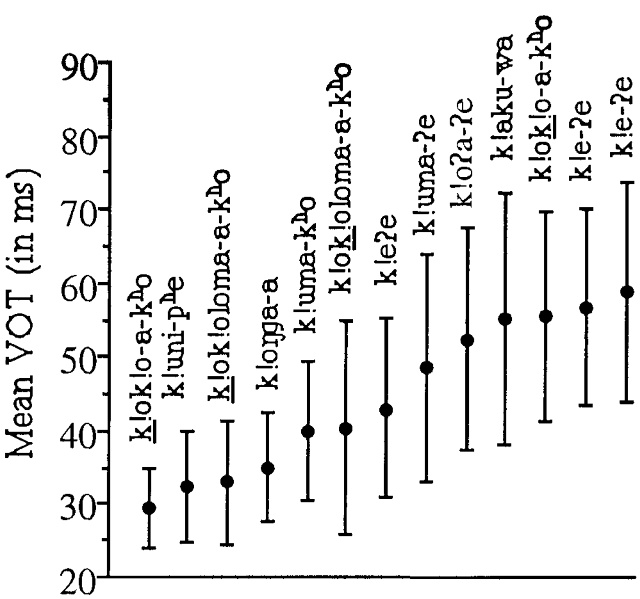


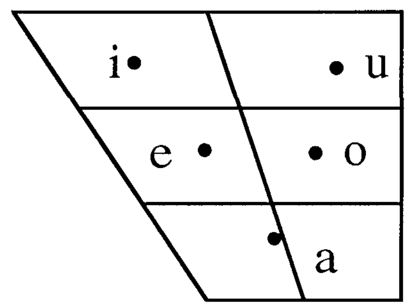
Figure 9: Means and Standard Deviations of VOT of voiceless alveolopalatal clicks for 7 Hadza speakers. Duration is measured from the release of the consonant to the onset of voicing for the following vowel. In words with more than one click, the measured click is underlined.

Vowel Quality

The quality of the five vowels in Hadza was examined by making measurements of the formant frequencies of each of these vowels in a similar environment. The words in Table 4 were used for this purpose. These words were chosen as they represented the nearest possible minimal set available for the vowels, despite the fact that they contain nasal consonants. For each of the seven speakers there were two utterances of each word, and two identical vowels in each word. Formant frequencies were measured in the midpoint of each vowel. Figures 11 and 12 plot the first and second formant frequencies of the four female and three male speakers, respectively. The vowels are plotted on a Bark scale, and the ellipses enclose all data points for a given vowel that are within two standard deviations of the mean of that vowel. The third formant was not plotted due to the large number of missing values across tokens. The range of F1 is considerably greater for the female speakers than for the male speakers. The mid vowels /e, o/ appear to be more centralized for the female speakers.

Table 4: Vowels

i	*ŋǁi-ʔi*	'put poison on f arrow'	210
e	*ŋǁe-ʔe*	'put poison on m arrow'	209
a	*ŋǁáʔa-*	'to scavenge'	214
o	*ŋ̊ǁ'o-ʔo*	'wash, bathe'	226
u	*ŋǁuʔú-*	'to snore'	
ĩ	*ɬiɦe-*	'to blow nose'	
ũ	*saɦũɦe*	'be quiet!'	



(ĩ, ũ)

Figure 10: Vowel Qualities of Hadza

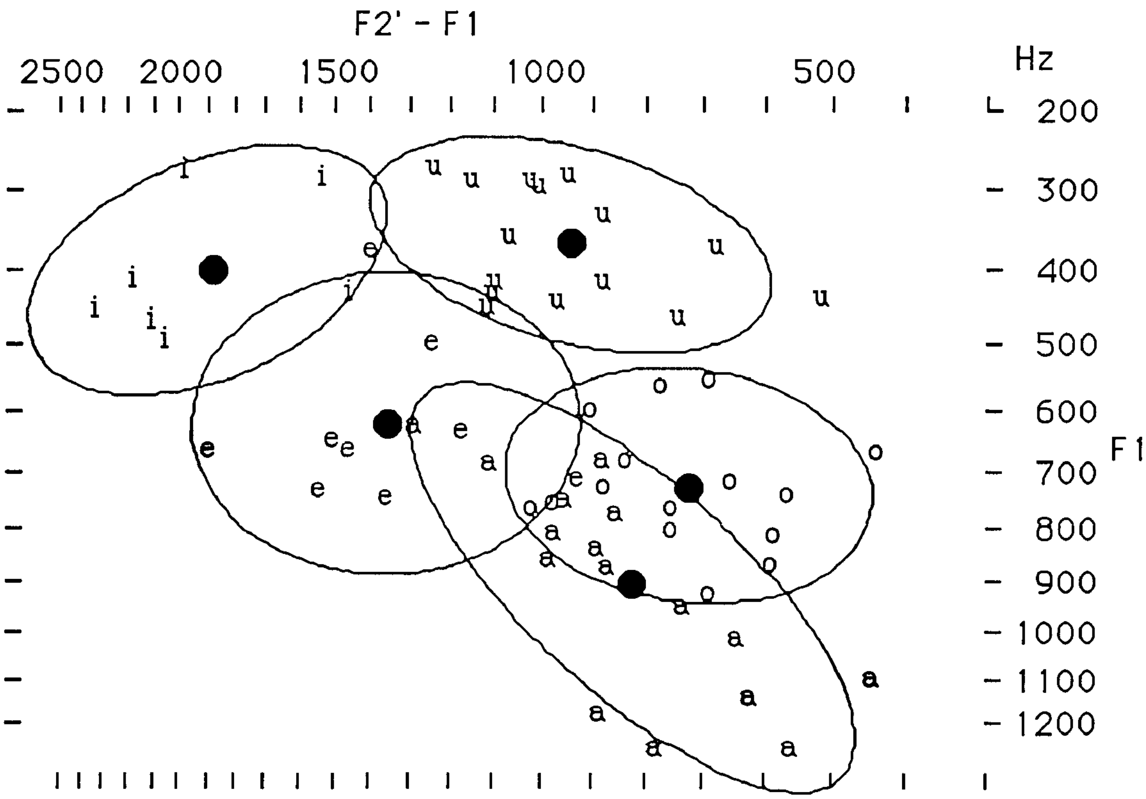


Figure 11: Frequencies of the first and second formants of the vowels [i, e, a, o, u] for four female speakers of Hadza. Formants were measured in the middle of the oral vowels found in the second syllables of the words in Table 4

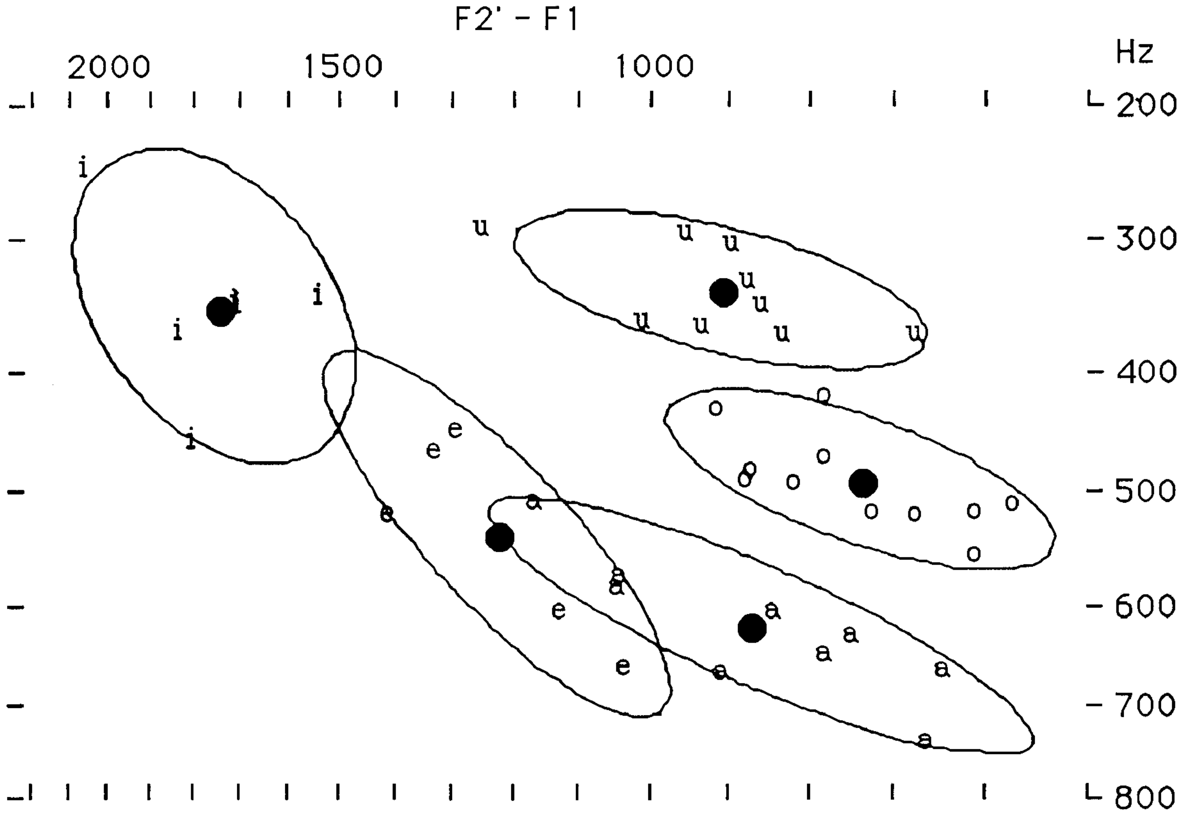


Figure 12: Frequencies of the first and second formants of the vowels [i, e, a, o, u] for three male speakers of Hadza. Formants were measured in the middle of the oral vowels found in the second syllables of the words in Table 4.

Concluding comments

There are three languages spoken in East Africa that have phonetic inventories which include clicks. One of them, Sandawe, is spoken by several thousand people who form a strong community with schools and local government. Another, Dahalo, discusssed in this volume, is clearly a vanishing language, spoken by only a few hundred people scattered among other communities. The situation of Hadza is harder to describe. It appears to have been spoken by a small group for a very long time. Children are learning the language, despite a high frequency of contact with other languages. As linguists, we are glad that this language is viable.

Acknowledgements

We are grateful to all the Hadza speakers who assisted us and shared their linguistic knoweldge with us, most especially Gudo Mahiya, as well as to Nick Blurton-Jones, Herman Batibo, Jeanette Hanby and David Bygott, and to the Tanzania Commission on Science and Technology. At UCLA we also are grateful to Siniša Spajić for his assistance in editing the digitized video images and analyzing the data, and to Stephan Schütze-Coburn for his assistance with the VOT measurements.

REFERENCES:

Beach, D. M. (1938). The Phonetics of the Hottentot Language . Cambridge: W. Heffer & Sons Ltd.

Berger, P. (1943). Überlieferungen der Kindiga. Afrika. 2. 97-122.

Bleek, D. (1931). The Hadzapi or Watindega of Tanganyika territory. Africa (Journal of the International Institute of African Languages and Cultures). 4. 273-286.

-----. (1956). A Bushman Dictionary. New Haven: American Oriental Society.

Dempwolff, O. (1916-17). Beiträge zur Kenntnis der Sprachen in deutsch-Ostafrika. 12. Wörter der Hatzasprache. Zeitschrift für Kolonialsprachen. 7. 319-325.

Doke, C. (1923). Notes on a problem in the mechanism of the Zulu clicks. Bantu Studies, 2(1), 43-45.

Ehret, C. (1986). Proposals on Khoisan reconstruction. Sprache und Geschichte in Afrika. 7.(2). 105-130.

Elderkin, E.D. (1982). On the classification of Hadza. Sprache und geschichte in Afrika. 4. 67-82.

-----.(1983). Tanzanian and Ugandan isolates. In M.Bechhaus-Gerst and R. Voßen (eds.). International Symposium on Languages and History of the Nilotic Peoples. Kölner, Beiträge zur Afrikanistik.

Greenberg, J. (1966). The Languages of Africa. Indiana University.

Ladefoged, P. and A. Traill. (1984). Linguistic phonetic descriptions of clicks. Language, 60(1), 1-20.

Ladefoged, P. and A. Traill. (1993). Clicks and their accompaniments. (forthcoming).

Obst, E. (1912). Von Mkalama ins Land der Wakindiga. Mitteilungen der Geographischen Geschichte. 27. 1-45.

Sands, B. (forthcoming). Hadza Lexicon. UCLA. ms.

Traill, A. (1985). Phonetic and Phonological Studies of !Xóõ Bushman . Hamburg: Helmut Buske Verlag.

Tucker, A., M. Bryan, and J. Woodburn. (1977). The East African click languages: a phonetic comparison. In W.J.G. Möhlig (ed.). Zur Sprachgeschichte und Ethnohistorie in Afrika: Neue Beiträge Afrikanistischer Forschungen. 300-322. Berlin: Dietrich Reimer.

de Voogt, A.J. (1992). Some Phonetic Aspects of Hatsa and Sandawe Clicks. thesis: Leiden University, Leiden.

J. Wagner. (1992, unpublished fieldwork).

Woodburn, J. (1962). The future of the Tindiga. Tanganyika Notes and Records. 59. 268-273.

Appendix

The following word list and comments refer to the tapes of four women and three men recorded in August, 1991.

(tape 1)

#	root	gloss	recorded endings	observed forms
1	bo'tʃó-	'to come'	-ʔo	bɤtʃo-ʔo, bətʃo-ʔo, bʊtʃo-ʔo
2	dzá-	'to come'	-ʔa	dz̥a-ʔa
3	kʰa'ŋǁ'é-	'to jump', 'to spring up' (= 203)	-ʔe	
4	'k!áku-	'to jump over	-wa	
5	'tʃi-	'to run'		
6	te'ŋ!'é-	'to carry on shoulders'	-ʔe	te'ŋ!'ɛ́-
7	'pótʌ'o-we-	'to break'	-ʔe	'pʊ́tʌ'o-ˌwé-
8	kʰáɦa-	'to climb'	-ʔa	káɦa-ʔḁ, káa-ʔ
9	ŋǁ'utʰi-je-	'to cook'	-ʔe	ŋǁ'ʊti-je-ʔe
10	'k!a-	'to cut' (= 196 & 216)	-ʔe	'k!é-
11	fá-	'to drink'	-ʔa	
12	séme-	'to eat'	-ʔe	sɛ́me-ʔ
13	haká-	'to go'	-ʔa	
14	kǁa'kǁá-	'to hunt' (= 202)	-ʔa	kakǁa-ʔa
15	tʃé-	'to take'	-ʔe	
16	ts'aɦa-	'to know'	-ʔe	ts'áɦe-
17	tʰa-ija	'to leave off', 'to stop doing something.'	-ʔe	tʰa-je-
18	tʃeké-	'to put'	-ʔe	
19	tʌ'odzo-	'to say'	-ʔo	
20	k!uma-	'to wrinkle'	-ʔe	
21	k!oʔá-	'to scratch'	-ʔa	
22	tʌ'áʔa-	'to sing'	-ʔa	
23	tsija-	'to sneeze'	-ʔa	
24	ts'áke-	'to steal'	-ʔe	
25	ŋ!eko	'to stir'	-ʔe	ŋ!éke-
26	ŋ!'uʔ-ija-	'to swell'	-ʔa	
27	tʌéɦena	'below, bottom'		tʌɛ́ːna
28	ŋǁ'o-	'to wash' (= 226)	-ʔo	
29	ga'gá-	'grasshopper'	-a	gaː'gá-a
30	ʔéna-	'grass'	-pʰi	
31	ɦáts'apʰi-	'leaves'	-pʰi	
32	pʰándʒu-	'sp. plant'	-pʰe	
33	pʰisé-	'thorn'	-ja	
34	ts'ití-	'tree'	-ja	
35	bone-	'four'	-pʰe	bonɛ-pʰe, bole-pʰe
36	ʔitʃáme-	'one'	-ja	
37	sámaka-	'three'	-pʰi	
38	píje-	'two'	-pʰe	pʰíjɛ-pʰe
39	ɦék'wa-	'bark, shell, rind, crust'	-pʰe	
40	jámu'ʔá-	'country, land'	-a	
41	tʌ'alá-	'dust'	-a-kʰo	
42	ts'okó-	'fire'	-wa	
43	ts'ik'ó-	'smoke' (= 86)	-wa	
44	badá-	'hole'	-a	

45	'sétʰa-	'moon'	-a	seːtʰaː
46	ts'ifi-	'night'	-ja	
47	ntsa-	'star'	-a-kʰo	tsa-kʰo
48	iʃó-	'sun'	-wa-kʰo	
49	atʰí-	'water'	-pʰi	
50	ŋaŋa	'sp. fruit'		
51	maná-	'meat'	-a-kʰo	
52	kʰarimo-	'animal'		kʰalimo, -wḁ
53	ɦáts'e-	'hunger'	-ja	
54	nak'oma-	'buffalo'	-a	
55	tʃatʃa-	'lesser bushbaby'	-a	
56	ɲáu-	'cat'	-wa-kʰo	
57	káŋga-	'sp. mongoose'	-a	kaːŋgaː
58	ts'aŋkʰá-	'banded mongoose'	-a	
59	ts'okwana-	'giraffe'	-a	
60	ts'iŋgaʔú	'oryx, sable or roan antelope'	-wa	
61	ˌʔúndaʔúnda-	'hedgehog'		
62	we'ts'ái-	'hippo'	-ja-kʰo	
63	watʃ'o-	'sp. mongoose'	-wa	
64	ŋ!ána-	'kudu'	-a	
65	tsipiti-	'porcupine'	-ja	
66	jóndo-	'rat'	-wa	
67	lalá-	'gazelle'	-a	laːlaːkʰo
68	kwaʔí-	'warthog'	-ja	
69	mbugida-	'wild dog'	-a	
70	dóŋgo-	'zebra'	-wa-kʰo	
71	ʔolá-	'child'	-a	
72	báwa-	'father'		bawə
73	'ɬemé-	'man'	-ja	'ɬɛme-ja
74	ʔákʰwiti-	'woman'	-ja-kʰo	
75	ɦóts'o-	'ashes'	-wa-kʰo	
76	loʔo-	'horn'	-pʰe	
77	mbo'góʃi-	'bag'	-ja-kʰo	bogoʃi-, bʊgʊʃi-
78	ndzópá-	'bottle'	-a	nzópá-
79	kʰó'ʔó-	'bow'	-wa	
80	'kʰómati-	'eland'	-ja	
81	ɬupʰá-	'foam'	-a-kʰo	
82	k‖ek‖etʃe-	'woman's loincloth'	-ja	ŋ‖'eŋ‖'etʃe
83	logo-	'shield'	-wa	
84	kwatʎa-	'shoe'	-a-kʰo	kwatʎ'a-
85	mákʰo-	'clay pot'	-wa	
86	ts'í'k'o-	'smoke' (= 43)	-wa	
87	utʰumé-	'spear'	-ja	
88	ɦek'wá-	'shell, bark, rind' (= 39)	-a-kʰo	
89	ɦits'á-	'fat, oil'	-pʰe	
90	semé-	'food'	-ja	sɛme-
91	baʔalá-	'honey'	-ko	baʔalá-ko
92	moɲóda-	'salt'	-a	moɲoːdaː
93	ʔawawa	'bee'	-a	
94	tʃ'áɦi-	'maggot'	-pʰi	
95	ɬanó-	'python'	-wa	
96	ʔijátu-	'snake'	-pʰi	

97	ŋ!'anda-	'agama lizard'	-a	
98	ʔu'kʰwa-	'arm'	-ko	
	ʔu'kʰwa-	'hand'	-pʰi	
99	patákuʃé-	'palm of hand', 'sole of foot'	-ja	
100	guɾiŋgu'ɾí-	'kneecap'	-ja-kʰo	guɾuŋguɾi-
101	ʔúpʰúkʰwa-	'leg'	-a	
102	ɦáts'áts'e-	'lower leg bones'	-ja	
103	tʎ'óʔa-	'cloth tied around shoulder'	-a-kʰo	
104	ʔá'ɦú-	'skin'	-wa	
105	ts'á'ɦó-	'tail'	-wa	
106	ɦompʰái-	'wing'	-ja-kʰo	
107	ʔátʰa'má-	'blood'	-a	
108	ŋkólo-	'heart'	-wa-kʰo	
109	pápasa-	'innominate bone'	-a	
110	ho'tʃ'ó-	'lung'	-pʰe	
111	ʔatʃú-	'sinew'	-pʰi	
112	ʔakʰwá-	'eye'	-a-kʰo	
113	ŋkʰatʰa-	'strand of beads worn on head'	-a	ŋkatʰa
114	k'apáku-	'mandible'	-pʰi	
115	'ʔáwaniká-	'mouth'	-a	
116	ʔáɦá-	'tooth'	-pʰe	
117	'tʃ'á-	'guineafowl'	-a-kʰo	
			-pʰe	
118	ˌwáʔi'ná-ma	'all'	-ma	
119	pákapaʔá-	'big'	-a	
120	titʃ'i-	'black'	-jé-ja	
121	ts'útʃi-	'wind'	-pʰi	
122	ŋ‖a'tʎ'á-	'to be cold'	-ne-ja	
123	tʰasé-	'tall, long'	-ja	
124	petʎ'ái-	'white'	-ja	
125	k'alaɦái	'sp. fruit'		
(Tape 2)				
126	ts'ukú-	'firewood'	-pʰi	
127	ʔi'k‖á-	'to close'	-ʔe	ʔi'k‖é-
128	hi'ŋ!'é-	'to come out of', 'to exude', 'to give out'	-ʔe	
129	k!ok!óroma-	'epiglottis'	-a, -a-kʰo	
130	ŋ‖'ekéjo-	'ankle'	-wa	
131	huŋ‖'ú-	'anthill'	-wa	
132	k‖ána-	'arrow, female'	-a-kʰo	
133	k‖ána-	'arrow, male'	-a	
134	'k‖ákʰá-	'arrowstand'	-pʰi	
135	néʔe-	'baboon'	-ja-kʰo	
136	ŋ!aleʔa-	'red flesh which sticks out of the anus, or red area on a baboon		
137	k!ok!ó-	'back of head'	-a-kʰo	
138	ŋ‖obá-	'baobab'	-a-kʰo	
139	ŋ!ojé-	'beeswax'	-ja	
140	taŋ‖'e-	'belt, rope'	-ja	
141	ˌŋ‖áwé'tʃ'e-ne	'blue, green'		
142	mitʎ'á-	'bone'	-a	
143	tʎ'áɦa-	'bushpig'	-a	
144	k!úni-	'calf muscle'	-pʰe	
145	k!uma-	'club'	-a-kʰo	

146	ŋgets'ea-	'forehead'	-p^{h}e	
147	kǁátʃ'o-	'fontanelle' (same root as 'frog' 156)	-wa-k^{h}o	
148	kǁáʔano-	'dog'	-wa	
149	tʎ'ápo-	'dove', 'gull'	-wa-k^{h}o	
150	tʎ'áˈlá-	'dust'	-a-k^{h}o	
151	ŋ!'oko	'to pierce'	-ʔ	ŋ!'uki-
152	ŋǀ'úkúˈmaje-	'elbow'	-ja-k^{h}o	
153	ŋǁ'inˈtʃino-	'fang'	-p^{h}i	ŋǁ'indʒino-
154	ŋǀ'ama-	'fish'	-a	
155	kǀaˈkǀá-	'flat rock'	-a	
156	kǁátʃ'o-	'frog' (same root as 'fontanelle' 147)		
157	ˈɦátʎ'é-	'hair'	-p^{h}e	
158	k!óŋˈga-	'hare'	-a	
159	tʎóma-	'head'	-a-k^{h}o	
160	ʔets'á-, ŋǀ'ets'á-	'house'	-a-k^{h}o	
161	ŋ!uŋguwe-	'hundred'		
162	tʎáʃo-	'tree hyrax'	-wa	
163	ˈŋ!'úkunˌdʒu-	'kidney'	-wa	
164	!'ama-	'klipspringer'	-a-k^{h}o	
165	ʔitʎá-	'knife'	-a-k^{h}o	ʔitʎ'a-
166	ŋ!'elé-	'hartebeest'	-a-k^{h}o	
167	dʒándʒai	'leopard'		
168	ŋ!é-	'leopard'	-ja	
169	ŋǁ'e-	'liver'	-ja-k^{h}o	
170	ŋǀ'áˈmáts'i-	'louse'	-ja-k^{h}o	
171	kǁáʔaˈkǁaʔa-	'middle'		
172	huŋǁ'uˈk'ó-	'molar tooth'	-wa-k^{h}o	haŋǁ'uˈk'ó-, huŋǁ'oˈk'ó-
173	kǁuwi-	'mosquito'	-ja-k^{h}o	
174	ŋǁ'utʎ'e-	'mountain, hill'	-ja	[ŋǁ'ũtʎ'e-]
175	toŋ!'oko-	'mud'	-wa-k^{h}o	toŋ!'ok'o-
176	ŋ!'uʃu-	'navel'	-wa-k^{h}o	
177	kǀúti-	'neck'	-ja	
178	ŋ!oˈmo-	'half'	-ja	ŋ!umo-ja; ŋ!ʊmo-ja
179	ŋ!'akiɬá-	'palate'	-a	ŋ!'ak'iɬá-
180	ˈtʎ'óŋk^{h}o-	'tawny eagle'	-wa	
181	kǀikiliŋǀa-	'pinky finger'	-a	kǀikiŋǀa-
182	kǁaˌtak'áˈnó	'rainbow'		
183	tʎákáte-	'rhino'	-ja	tʎ'akate
184	ɦaŋ!'á-	'rock'	-a-k^{h}o	
185	ŋǀ'its'é-	'short'	-ja	ŋǀ'ɪts'e-
186	kǁamba-	'small intestine'	-p^{h}i	ŋǁ'amba-p^{h}i kǁampa-p^{h}i
187	ŋǀáláka-	'snail'	-a	
188	ŋǁ'ekwá-	'sp. root'	-a	
189	púˈkǀ'é-	'spleen'	-ja	
190	kǁaˈpá-	'stump'	-a	ŋǁ'apa-
191	kǁáts'i-	'sweat'	-ja	
192	ŋǁ'uˈk'wá-	'larynx'	-a-k^{h}o	
193	ŋǀ'ets'é-	'tick'	-ja	
194	ŋǀ'óso-	'to be full'	-ʔo	

195	ka'tʎ'é-	'to bite'	-ʔ	
196	k!a-	'to cut' (=10, 216)	-ʔe	'k!é-
197	taŋ‖'i-	'to die'	-ʔi̥	
198	ŋ!óʔo-	'to enter'	-ʔ	
199	'ŋ|áɦa-	'to forget'	-ʔ	
200	'ŋ‖áʔe-	'to hear'	-ʔ	
201	ŋ‖'aka-	'to hit with an arrow', ('to shoot at, to hit')	-ʔ	ŋ‖'ake-
202	ka'k‖á-	'to hunt' (=14)	-ʔḁ	
203	kʰa'ŋ‖'é-	'to jump', 'to spring up' (=3)	-ʔe	
204	'k‖ó-we-	'to kill'	-ʔ	
205	ts'uʔa-	'to remove s.t.'	-ʔe̥	ts'u-we-
206	ŋ|e'ɦé-	'to whistle' (=227)	-ʔe	
207	'ŋ!í-je-	'to push'	-ʔ	
208	ŋ|'uts'u-we-	'to push'		
209	ŋ‖e-	'to put poison on a (male) arrow'	ŋ‖e-ʔe	
210	ŋ‖e-	'to put poison on a (female) arrow'	ŋ‖i-ʔi	
211	ŋ|'ats'a-	'to reheat'	ŋ|'ats'e-ʔe	
212	'máŋ!'i-	'to circle around'	-ʔi	
213	k‖aŋgála-	'to pass legs under, to be lying down'	-ɦi	
214	'ŋ‖áʔa-	'to scavenge'	-ʔ	
215	k|i-jé-	'to see'	-ʔ	
216	'k!a-	'to cut' (=10,196)	-ʔe	'k!é-
217	'ŋ!'óko-	'to slap'	-ʔe	'ŋ!'óke-
218	ʔase-	'to sleep'	-ʔe	
219	'k‖úpi-	'to sleep'		
220	'ŋ‖úʔu-	'to snore'	-ʔ	
221	ŋ|'a'k'wé-	'to swallow'	-ʔ	
222	utʎ'ú-we-	'to uproot (roots)'	-ʔ	
223	'ŋ!'ó-we-	'to uproot'	-ʔ	
224	'kʰwák‖a-	'to vomit'	-ʔa	
225	'ŋ|íts'i-	'to wait for', 'wait!'	-ʔi	
226	ŋ‖'o-	'to wash', 'to bathe' (=28)	-ʔo	
227	ŋ|e'ɦé-	'to whistle' (=206)		
228	'ŋ|áta-	'tongue'	-a	
229	ŋ‖ímó-	'van der Decken's hornbill'	-a-kʰo	
230	ŋ!'oŋgojó-	'area of body encompassing buttocks, hips, pelvis and tail'	-kʰo	
231	ŋ|'á'mé-	'white hair'	-ja-kʰo	
232	'tʎ'ómásá	'pipe'	-a	

Phonetics of Toda

Michael Shalev, Peter Ladefoged and Peri Bhaskararao

Toda, a Dravidian language, has some of the most interesting phonetic properties in the world. It has a wider variety of vowel qualities than any other Dravidian language, and it uses both individual consonant sounds and sets of consonant contrasts that are not attested in any other known language. This paper will present a detailed account of the vowel system, and a preliminary investigation of some of the consonants.

The most complete description of the language is Emeneau's monumental work *Toda Grammar and Texts* (Emeneau 1984), which is based on a "collection of linguistic data made in a total of about a year of contact with the Todas in the three year period from mid-1935 to mid-1938." (Emeneau 1984:1). As far as we can see, there has been virtually no other linguistic fieldwork with Toda speakers since that time. All later linguistic publications (e.g. Agesthialingom and Sakthivel 1970, Sakthivel 1976) have relied on Emeneau's work, and have not been independently checked with speakers of the language. The present paper is based on newly collected data, including recordings from 13 speakers (6 men and 7 women). We, too, have profited enormously from Emeneau's work. We have followed his phonological interpretations, and benefited from his shrewd phonetic insights. We found that the Toda language of today differs in only a few, comparatively small, respects from that described by Emeneau on the basis of his fieldwork over half a century ago.

Toda's phonetic idiosyncrasies have arisen as a result of its comparatively long isolation. Toda is spoken by about 1,000 people in the Nilgiri Hills in Southern India. The Nilgiri Plateau was almost totally cut off from its surroundings until the nineteenth century. Today, five Dravidian languages are spoken there: Toda, Kota, Badaga, Irula and Tamil. Toda is most closely related to Kota, which conforms much more to the norm for Dravidian languages. In contrast, McAlpin (1981) notes that Toda is highly aberrant in comparison to the areal norm, and he does not consider it even to look like a Dravidian language. Emeneau (1984) notes that the Todas' ethnological aberrancy both in customs and appearance has sparked much anthropological interest in them as a group. Emeneau classifies Toda as a member of the southern subgroup of Dravidian, so that, along with Kota, it is closely related to Tamil. Emeneau suggests that Toda's aberrancy lies in its having a disproportionately large number of rules that are not shared by other Dravidian languages (except to a small extent by Kota).

The inventory of Toda sounds is shown in Figures 1 and 2. The phonological contrasts specifiable by the items in these charts are exactly the same as those given by Emeneau (1984), although the phonetic symbols are different, largely because they adhere more closely to IPA principles (International Phonetic Association, 1989).

	Broad phonetic transcription			Narrow phonetic transcription		
Primary qualities	**i, iː**		**u, uː**	**i, iː**		**u, uː**
	e, eː		**o, oː**	**e**		**o, oː**
					æː	
		a, aː			**ɑ, ɑː**	
Secondary qualities	**y, yː**		**ɯ, ɯː**	**y, yː**		**ɨ, ɨː**
	ø, øː				**ɵ, ɵː**	

Figure 1. Vowel diagrams using broad and narrow phonetic transcriptions of Toda vowels.

The Toda vowel inventory is composed of sixteen vowels, which may be considered to be, in a broad phonetic transcription, long and short versions of the five primary vowels **i e a o u**, and long and short versions of the front rounded vowels **y ø**, and of the unrounded vowel **ɯ**. We will use a more narrow transcription, as shown on the right hand side of Figure 1, which we will *justify in the next section. The following set of near minimal pairs illustrates the contrasts among* these vowels.

iʈ	'flour'	**i:t̠**	'spear'
et̠	'where'	**æ:t̪**	'to take'
pot̪	'ten'	**mo:ʈ̪**	'speech'
puʈ	'stirring stick'	**pu:ʈ**	'eighteen'
pɑθ	'eagle'	**pɑ:t̠**	'cockroach'
nyts	'broken rice'	**py:l**	'summer'
ɵt̪	'to climb'	**ɵ:s̪t̪**	'to boil (e.g. vegetables)'
pis̠	'bow'	**pi:t̪**	'meat curry'

An overview of the Toda consonant system is provided by the chart in Figure 2. The row and column headings should be taken only as rough labels for the sounds, which will be described in more detail below. IPA symbols illustrating the six places of articulation are **p b t̪ d̪ t d ʈ ɖ tʃ dʒ k g**. Dravidian scholars often use the symbols **p b t d t̠ d̠ t5 d5 tʃ dʒ k g**. There is thus a source of potential confusion in that the IPA symbols **t d** without diacritics are used for alveolar articulations, whilst in the Dravidian tradition unmarked **t d** represent dental articulations. In order to make this paper useful to both Dravidian scholars and a wider audience, *we will use over-differentiated symbols, representing both dentals and alveolars with diacritics, as* shown in the chart. Toda adds to the basic Dravidian inventory by having stops which Emeneau transcribes as **c j** which contrast, in his transcription, with **č ǰ**. We will transcribe these stops as **ts dz** (contrasting, in an IPA transcription, with **tʃ dʒ**). Words illustrating consonant contrasts will be given later.

	labial	**dental**	**denti-alveolar**	**alveolar**	**retroflex**	**palato-alveolar**	**velar**
stop (and affricate)	p b	t̪ d̪	ts dz	t̠ d̠	ʈ ɖ	tʃ dʒ	k g
nasal	m			n	ɳ		
fricative	f	θ	s̪	s̠	ʂ	ʃ	x (ɣ)
trill		r̪		r̠	ɽ		
approximant						j	w
lateral			ɬ l	ɭ̊ ɭ			

Figure 2. Toda consonant chart.

Vowels

The following description of the Toda vowel system is based on an acoustic analysis of data from the six male speakers. We hope that we will be able to complete a similar analysis of the seven female speakers in the near future; but we do not wish to delay publication of this report any further. The six male speakers consisted of two groups of three, each group coming from a different *mund* (tribal location). Their two villages were a few miles apart, both being in the immediate outskirts of Ootacamundalam.

Each group recorded the list of words given in the appendix, saying the words first in isolation and then in a frame sentence. In each group one of the men said the word or sentence first, and then the other two men repeated it after him. This method of elicitation has a tendency to standardize the pronunciation of the tokens in a way that has both advantages and disadvantages. On the one hand, it is beneficial when examining the acoustic properties of tokens of the language, as it minimizes personal differences. On the other hand, it may introduce artifacts. We found differences between the two groups that could be due to slight differences in accent between the two villages, but may have been simply due to a 'follow the leader' effect. For example, the differences in duration of individual vowels were not the same for the two groups; this effect was statistically highly significant ($p = 0.0001$). In both groups the second and third informants may have been taking their cues for an appropriate vowel duration from the first informant, so that the effects in terms of vowel duration are evident across tokens.

The formant frequencies of the sixteen Toda vowels were measured from tape recordings of tokens uttered in isolation by the six male speakers. The measurements were made on a Kay Elemetrics Computerized Speech Lab (CSL), using Linear Predictive Coding (LPC) analysis. Table 1 shows the mean formant frequencies of the different vowel qualities — the eight short vowels and **æː** — for the data set: **kits** 'link hands', **t̪it̠** 'to wear a cloth', **is̠** 'rat', **nyts** 'broken rice', **pɨs̠** 'bow', **ɨts** 'to reach', **puʈ** 'stirring stick', **ut̪** 'anthill', **t̪et̠** 'wrap garment around waist', **et̠** 'where', **ɵt̪** 'to climb', **ɵʈ** 'eight', **pot̪** 'ten', **motʃ** 'cot', **æːt̪** 'to take', **pæːt̪** 'delivery', **t̪æːt̠** 'to fold', **pɑs̠** 'tooth', and **ɑt̪k** 'that much'.

Table 1. Mean formant frequencies of Toda vowel qualities (Hz)

Vowel	F1	F2	F3	F2'
i	363	1962	2569	2182
y	373	1836	2628	2093
ɨ	412	1485	2556	1753
u	368	1040	2364	1263
e	471	1864	2680	2121
ɵ	461	1530	2439	1776
o	537	1221	2694	1455
æː	684	1679	2686	1929
ɑ	673	1366	2727	1596

In accordance with Ladefoged (1993) among others, these formant frequencies may be regarded as determining vowel quality in terms of *height* and *backness*, taking **F1** as a measure of height, and either **F2'** or **F2'-F1** as a measure of backness. F2' is a weighted average of F2, F3 and F1, and is computed as follows (Fant, 1973:52):

$$F2' = F2 + 0.5\ (F3 - F2)\frac{F2 - F1}{F3 - F1}\ .$$

If the measured formant frequencies are converted from Hertz to Bark, then equal intervals in Bark may be taken as equivalent to equal perceptual intervals. A suitable formula for this purpose is:

$\mathbf{B} = 13 \arctan(0.76f) + 3.5 \arctan(f/7.5)^2$, where **B** is the critical band value in Bark and f is the frequency in kHz (Syrdal & Gopal, 1986:1088). Figure 3 shows the mean vowel qualities plotted in a way that agrees with traditional IPA descriptions, with F1 representing *Height* and and F2'-F1 representing *Backness*.

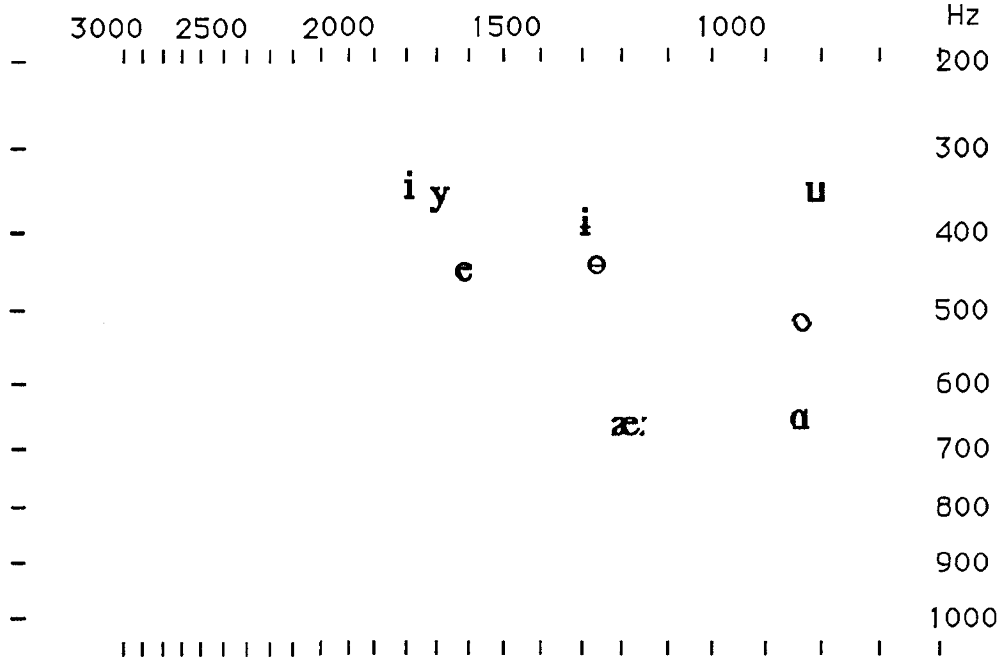


Figure 3. Mean F1 & F2' - F1 formant frequencies of Toda Vowel Qualities

The acoustic data, taken together with our direct observations, provide a basis for the following description of the vowel qualities:

i is a high, front unrounded vowel, lower than cardinal **i**, with conspicuous lip retraction.

y is a high, front vowel, with some lip rounding.

ɨ is a high-mid, central, unrounded vowel. Emeneau transcribes it as ï, and describes it as a high back unrounded vowel. In the case of our speakers, however, the acoustic analysis supports a central classification of both short and long variants of this vowel as central, unrounded vowels, **ɨ** and **ɨː**. Figure 3 shows **ɨ** to be much closer to **ɵ** than to **u**, or **o**.

u is a high, back, rounded vowel, not quite so high as cardinal **u**, with conspicuous lip rounding.

e is a short, front, mid, unrounded vowel.

ɵ is a central, mid rounded vowel. It is distinctly further back than the front rounded vowel **ø**, found in European languages such as French and German.

o is a back, mid rounded vowel. Emeneau describes it as showing strong lip-rounding, but we consider it to be no more rounded than usual for a cardinal vowel in this area.

æː is a long, front, lowered mid, unrounded vowel. Emeneau represents this vowel as **eː**, noting the large allophonic variation in mid vowels. Acoustic analysis shows this vowel to have a mean F1 of 684 Hz, which is higher than that found by Disner (1983) for **ɛː** both in Italian (470-630 Hz) and Yoruba (450-700 Hz), and lower than that of **ɑ** in Toda. The auditory impression of this vowel is much closer to **æː** than to **ɛː**.

ɑ is a back, somewhat raised low, unrounded vowel.

Intrinsic vowel length

High vowels are observed to be intrinsically shorter than low vowels in many languages, including: English, German, Danish, Swedish, Thai, Lappish, and Spanish (Lehiste, 1960). Table 2 shows mean durations of Toda vowels arranged by vowel height, for the data set:

kits ‘link hands’, **t̪it̠** ‘to wear a cloth’, **iʈ** ‘flour’, **nyts** ‘broken rice’, **its** ‘to reach’, **puʈ** ‘stirring stick’, **up** ‘salt’, **ut̪** ‘anthill’, **t̪et̠** ‘wrap garment around waist’, **et̠** ‘where’, **peʈ** ‘cane’, **ɵʈ** ‘eight’, **ɵt̪** ‘to climb’, **pot̪** ‘ten’, **oʈ** ‘to pour’, **kots** ‘to bite’, **t̪ɑʈ** ‘churning vessel’, and **ɑt̪k** ‘that much’. This set includes tokens with all of the short vowels, and with the context of the following consonant limited to voiceless coronal stops and affricates.

Table 2. Mean durations (ms) of high, mid, and low vowels in Toda

Vowel	High	Mid	Low
i	60		
y	49		
ɨ	70		
u	61		
e		75	
ɵ		85	
o		68	
ɑ			76
mean	**60**	**76**	**76**

These data show an apparent variation from the expected results: as is the norm, high vowels are shortest, followed by mid and low vowels which appear to have the same mean duration. Even the duration difference between high and low vowels is not significant ($p = 0.1150$), perhaps due to the large standard error of the **ɑ** tokens (6.129 versus 3.328 and 3.452 for the high and mid vowels, respectively), while the difference between high and mid vowels is almost significant ($p = 0.0641$). Since mid and low vowels are virtually identical in duration ($p = 0.9130$), the contrast in intrinsic length in Toda appears to be between high and non-high vowels. Combining mid and low vowels into one group, we see that high vowels are significantly shorter than non-high vowels ($p = 0.0090$), with high vowels having a mean duration of 60 ms, and non-high vowels having a mean duration of 76 ms.

Phonemic vowel length

In languages which distinguish between short and long vowels, there appears to be no single durational ratio by which this phonological distinction is implemented phonetically. The ratio of short to long vowel durations (V/Vː) "is close to 50%, but may vary a great deal" (Lehiste, 1960:34). Table 3 shows mean durations of vowels in the data set: **iʈ** ‘flour’, **kits** ‘link hands’ ~ **iːt̠** ‘spear’; **nyts** ‘broken rice’ ~ **pyːʈ** ‘summer’; **its** ‘to reach’ ~ **pit̪** ‘flesh’, **witʃ** ‘large lizard’; **puʈ** ‘stirring stick’ ~ **puːʈ** ‘eighteen’, **ut̪** ‘anthill’ ~ **uːts** ‘throw away water’; **et̠** ‘where’ ~ **æːt̪** ‘to take’, **t̪et̠** ‘wrap garment around waist’ ~ **t̪æːt̠** ‘to fold’, **pæːt̪** ‘delivery’; **ɵt̠** to climb’ ~ **ɵːs̪t̪** ‘to boil’; **pot̪** ‘ten’, **kots** ‘brass vessel’ ~ **moːt̪** ‘language’, **moːt̠** ‘change direction’; **t̪ɑʈ** ‘churning vessel’ ~ **pɑːt̠** ‘cockroach’.

While we do not consider **e** and **æː** to be short and long variants of the same quality (see above) they are included in the data set because they constitute a pair of a short and a long vowel, and as such should show the same durations as the other short and long vowels that are phonemically distinct from each other, and do not vary in quality. The mean duration of short vowels is 68 ms and that of long vowels 139 ms. The *short:long* ratio is therefore 1:2.04, or slightly more than **1:2.**

Table 3. Mean durations of short versus long vowels in Toda (ms)

Vowel	Duration	Vowel	Duration
i	67	iː	122
y	49	yː	161
ɨ	70	ɨː	123
u	60	uː	148
e	75		
ɵ	85	ɵː	124
o	64	oː	126
		æː	151
ɑ	74	ɑː	156
V	68	**V:**	139
V/V:	0.49	**V:/V**	2.04

Vowel quality and phonological length

There is a tendency for the short peripheral vowels **i**, **u**, and **ɑ**, to be more centralized than their long counterparts **iː**, **uː**, and **ɑː**, as may be seen from the vowel chart in Figure 4. This may be due to achievement of a more complete articulation of the long vowels in comparison to that of the short vowels. Lehiste (1960:30-31) found the same tendency for vowels in disyllabic Czech words although there the nonperipheral vowels also behaved in a regular fashion (the short vowels were more centralized), while in Toda as we see, this is not the case.

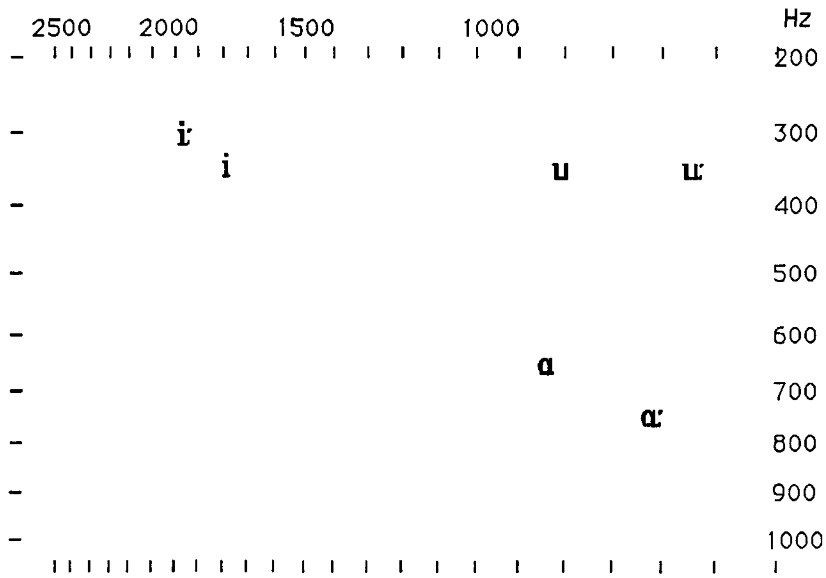


Figure 4. Mean F1 & F2' - F1 formant frequencies of peripheral vowels

Figure 5 plots the non-peripheral vowels: **y**, **yː**, **ɨ**, **ɨː**, **e**, **æː**, **ɵ**, **ɵː**, **o**, and **oː**, for the data set: **nyts** 'broken rice' ~ **pyːʈ** 'summer'; **its** 'to reach' ~ **pit̪** 'flesh', **witʃ** 'large lizard'; **et̲** 'where' ~ **æːt̪** 'to take', **t̪et̲** 'wrap garment around waist' ~ **t̪æːt̲** 'to fold', **pæːt̪** 'delivery'; **ɵt̲** 'to climb' ~ **ɵːʂt̪** 'to boil'; **pot̪** 'ten', **kots** 'brass vessel' ~ **moːt̪** 'language', **moːt̲** 'change direction'. In contrast to the peripheral vowels, we see no regular tendency distinguishing short from long variants here. Only **o** and **oː** differ significantly in their F1, and this is perhaps due to the nasal stop preceding **oː** in the words for 'language' and 'change direction', in comparison to the preceding oral stop in 'ten'. **e** and **æ:** are not considered here due to their being different vowel qualities.

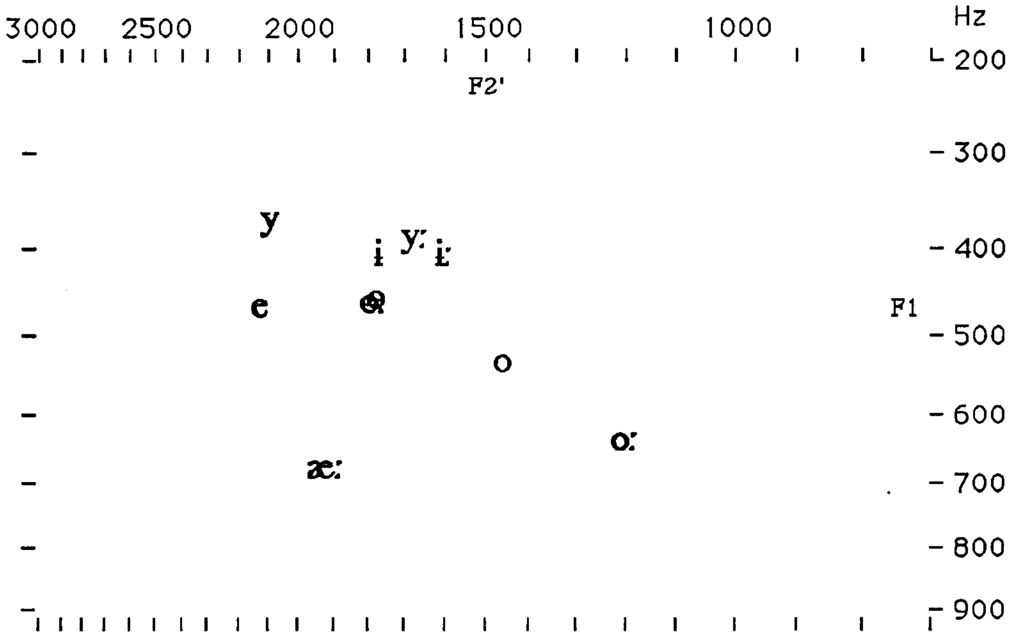


Figure 5. Mean F1 & F2' formant frequencies of non-peripheral vowels

Vowel quality allophony

The voicing of the following consonant triggers allophonic variation of vowels in Toda. Emeneau characterizes the allophonic difference in vowel quality as the degree of *closeness* or vowel height. If the following consonant is voiced, a higher vowel allophone occurs; if it is voiceless, or word final, then a lower allophone occurs. Emeneau states that the mid vowels show more allophonic variation than do the high and the low vowels. Our data do not permit us to examine this question for each vowel. We have few exactly matched pairs, where the only contrast is one of voicing in the final consonant. The data set we used in our comparison is: **kits** 'link hands', **t̠it̠** 'wear a cloth' ~ **kidʒ** 'ankle bells'; **i:t̠** 'spear' ~ **pi:g** 'knot in hair'; **nyts** 'broken rice' ~ **ydz** 'five'; **py:ɬ** 'summer' ~ **y:r̠** 'fingernail'; **its** 'to reach' ~ **nidz** 'heart', **wid̠** 'to pry open'; **ut̠** 'anthill' ~ **ud̠** 'black gram'; **pu:ts** 'dress someone', **pu:tʃ** 'blue-grey' ~ **pu:dʒ** 'fifteen'; **et̠** 'where' ~ **ed̠** 'on what day'; **ɵ:ʂt̠** 'to boil' ~ **pɵ:r̠** 'name, root'; **mots** 'dusk' ~ **modz** 'buttermilk'; **mo:t̠** 'language' ~ **mo:dz** 'to hide (news)', **o:dz** 'to have fear'; **æ:t̠** 'to take' ~ **æ:d̠,** 'why' **pæ:t̠** 'delivery' ~ **pæ:d̠** 'to be born'; **ɑt̠k** 'that much' ~ **ɑd̠** 'that day'.

Table 4 shows the means of F1 and F2' for the vowel midpoints in each of two contexts: preceding voiced or voiceless consonants. The rightmost columns represent the difference in formant frequencies, computed here with the formant in the voiced context being subtracted from that in the voiceless context. We perform this subtraction to try and determine whether the relation between the two formant frequencies is regular. Statistical analysis shows the voicing of the following consonant (context) to have a significant effect on the vowels' F1 (*height*): ($p = 0.008$), but this effect varies from vowel to vowel, as the interaction of vowel identity and context ($p = 0.0114$) shows. Context has only a marginal effect on the vowels' F2' (*backness*): ($p = 0.0623$), although here again vowels do not behave uniformly – there is a significant interaction between vowels and context ($p = 0.0006$). Using the other definition of *backness* used here (F2'–F1), context does have a significant effect on vowel backness ($p = 0.0009$), but once again vowel identity and context interact significantly ($p = 0.0001$), showing that the effect of context on backness is not uniform across vowels. The variations in F2' apparent in Table 4 may perhaps be associated with other contextual effects, due to the size and composition of the corpus analyzed here. For example, the lowering of F2' of [**ɵː**] before the voiceless consonant may be due to this consonant being retroflex. These results, however, do not support Emeneau's claim for more allophony in mid vowels than in high and low vowels.

Table 4. Means of F1 & F2' formant frequencies (Hz) preceding voiceless and voiced consonants

Vowel	_ [-vce]	_ [-vce]	_[+vce]	_[+vce]		
	F1	F2'	F1	F2'	ΔF1	ΔF2'
i	367	2060	267	2434	101	-374
iː	301	2364	338	2308	-37	56
y	373	2088	312	2169	61	-81
yː	391	1680	370	1916	21	-236
ɨ	384	1870	438	1648	-53	221
u	373	1419	378	1192	-5	227
uː	377	1129	375	1180	2	-51
e	485	2145	469	2116	16	29
ɵː	465	1788	448	1741	17	46
o	587	1329	445	1487	141	-158
oː	641	1187	468	1312	173	-125
æː	722	1882	631	1923	91	-41
a	698	1658	647	1762	51	-105

Allophonic vowel duration

Vowels are longer before voiced consonants than before voiceless consonants in many languages (see Maddieson & Gandour, 1977 for references given there). We will examine whether this observation is true in regard to Toda as well. In measuring vowel duration, the duration from the onset of voicing (excluding the preceding consonant's release) was chosen, to allow generalization across syllable types (CVC and VC). The data set used here is: **kits** 'link hands', **t̠it̠** 'wear a cloth' ~ **kidʒ** 'ankle bells'; **iːt̠** 'spear' ~ **piːg** 'knot in hair'; **nyts** 'broken rice' ~ **ydz** 'five'; **pyːʈ** 'summer' ~ **yːr̥** 'fingernail'; **its** 'to reach' ~ **nidz** 'heart', **wid̪** 'to pry open'; **ut̪** 'anthill' ~ **ud̠** 'black gram'; **puːts** 'dress someone',**puːtʃ** 'blue-grey' ~ **puːdʒ** 'fifteen'; **et̠** 'where' ~ **ed̠** 'on what day'; **ɵːs̪t̠** 'to boil' ~ **pɵːr̥** 'name,root'; **mots** 'dusk' ~ **modz** 'buttermilk'; **moːt̪** 'language' ~ **moːdz** 'to hide (news)', **oːdz** 'to have fear'; **æːt̠** 'to take' ~ **æːd̠** 'why', **pæːt̪** 'delivery' ~ **pæːd̪** 'to be born'; **ɑt̪k** 'that much' ~ **ɑd̠** 'that day'.

Table 5 shows that the mean allophonic difference in vowel duration is 23 ms for short vowels and 49 ms for long vowels, showing that for these data, the mean duration difference for long vowels is double that for short vowels, just as long vowels are twice as long as short vowels. This suggests that so far as these data are representative of Toda, the V/Vː ratio of 50% is maintained even when other factors (the voicing value of the following consonant) affect vowel duration, which indicates that this ratio is the cue for quantity. Statistical analysis reveals that although context *does* affect vowel duration (p = 0.0001), this effect differs from vowel to vowel, as the interaction between vowel identity and context shows (p = 0.0114). This is clear in table 2–4: we see the allophonic duration difference for **oː** is practically nonexistent (1 ms), as is that for **ɨ** (5 ms), in comparison with the large duration differences apparent for **yː** (87 ms) and for **ɵː** (90 ms). Here again, the possibility exists that these results are affected by the specific tokens analyzed.

Table 5. Mean durations of Toda vowels (ms) preceding voiceless and voiced consonants

Vowel	_ [-vce]	_ [+vce]	[+vce] – [-vce]
i	51	69	17
iː	131	155	24
y	49	100	51
yː	161	247	87
ɨ	70	76	5
u	75	103	28
uː	121	178	57
e	81	101	20
ɵː	124	214	90
o	66	77	11
oː	118	119	1
æː	152	193	41
ɑ	78	107	30
V	67	90	23
V:	135	184	49

The Pattern of Toda vowels

We are now in a position to consider Toda vowels in relation to the general pattern of vowels in the world's languages. Table 6 shows the percentage of languages listed in *Patterns of Sounds* (Maddieson, 1984), and the *UCLA Phonological Segment Inventory Database v1.1 (UPSID)* (Maddieson & Precoda, 1991) which include vowels having the characteristics of the vowels appearing in the Toda vowel inventory. It may be seen that a number of the vowels found in Toda are very uncommon.

Table 6. Frequency of languages having vowels with the characteristics of Toda vowels

	Features	P of S	UPSID
i	high, front, unrounded	85.5%	87.1%
iː	long, high, front, unrounded	12.9%	8.9%
y	high, front, rounded	6.6%	5.3%
yː	long, high, front, rounded	1.6%	0.9%
ɨ	high, central, unrounded	12.6%	13.5%
ɨː	long, high, central, unrounded	0.6%	0.7%
e	mid, front, unrounded	35.7%	37.5%
æː	long, raised low, front, unrounded	3.2%	1.8%
ɵ	mid, central, rounded	1.6%	1.1%
ɵː	long, mid, central, rounded	0.3%	n/a
ɑ	raised low, back, unrounded	6.9%	5.5%
ɑː	long, raised low, back, unrounded	2.2%	1.8%
o	mid, back, rounded	42.0%	40.1%
oː	long, mid, back, rounded	4.1%	2.7%
u	high, back, rounded	80.1%	81.8%
uː	long, high, back, rounded	11.7%	8.0%

The vowels inventories of the seven Dravidian languages surveyed in *Patterns of Sounds* and *UPSID* are shown in Figure 6. The five vowels of Proto-Dravidian (McAlpin, 1981) are well represented in Toda as well as in these languages, although only four of these languages have two series of vowels (one short and the other either long or nasalized). The four additional vowels in Toda: **y**, **ɨ**, **ɵ** and **æː**, do not form a coherent series in terms of a common feature. Telugu and Tulu both have a fourth, low–mid front long vowel – either **æː** or **ɛː**, so the presence of a long low-mid vowel in Toda is not unattested in Dravidian languages. But **y**, **ɨ**, **ɵ** do not appear in any of the other Dravidian languages. Finally, as we see in Table 6, their frequency of appearance in any of the languages surveyed in *Patterns of Sounds* and in *UPSID* is quite low.

Telugu	i u / **e** o / **ɛ**	iː uː / eː oː / æː / **aː**
Tulu	i ɯ u / e o / ɛ / ɛ̰	iː uː / eː oː / ɛː / aː
Kurukh	i u / e o / **a**	ĩ ũ / ẽ õ / ã
Brahui	i u / e o / a	iː **uː** / eː oː / ɑː
Malayalam	i u / e o / a	
Koya	i u / e / ɛ̰ ɑ	
Kota	i u / e o / a	

Figure 6. Representations of the vowel systems of 7 Dravidian languages.

The three 'additional' vowel qualities: y, ɵ, and **ɨ**, differ in their distribution pattern from the other vowels: they do not appear in final position. Sakthivel (1976) cites a single example of **ɨː** in final position: **pæːʈθwiː** 'mund of the **noːʂ** clan' (p. 53,110). Emeneau (p.c.) concurs with Sakthivel on this word. This distributional evidence supports the distinction drawn here between the 'basic' five vowel qualities of Proto-Dravidian: **i**, **e**, **a**, **o**, **u**, which are common among the Dravidian languages mentioned here, and **y**, **ɵ**, and **ɨ**, which appear to be unique to Toda.

Dispersion Theory (Lindblom, 1990) predicts that an additional vowel would appear in the region that allows maximal dispersion — the high vowel series. Toda has one low vowel, one raised low vowel, three mid vowels and four high vowels. Lindblom also predicts that in a system with nine different vowel qualities, there will be four values of vowel height ('sonority') and three values of backness ('chromaticity'). Toda does indeed have four values of vowel height, and three values of backness.

Phonological descriptions of vowel systems generally assume either binary values for backness and roundness; or consider roundness to be a privative feature. Few languages have vowel inventories that require more than two values of backness for a phonological description. Examples include Nimboran and Norwegian (see Ladefoged and Maddieson 1990 for references given there). In both these languages, there are three high vowels that do not contrast in rounding, but do differ in backness. As we have seen, in Toda there is a three way contrast in backness both in the high and the mid vowel series, although in both cases the backness contrast co-occurs with a variation in rounding.

A feature specification of Toda vowels is given in Table 7. By this classification Toda has four high vowels, three mid vowels, and two low vowels. As we have noted, this corresponds with a vowel dispersion theory such as that proposed by Lindblom (1990), which views the vowel space as a trapezoid in which the upper base is longer than the lower base. The dimensions of the traditional IPA vowel space also reflect this asymmetrically shaped vowel space: it is a right angle trapezoid, in which the horizontal dimension of high vowels is twice that of the low vowels. A dispersion theory would predict that more high than low vowels would occur in a given inventory.

Table 7. Features of Toda vowels

	Front		Central		Back	
	[-round]	[+round]	[-round]	[+round]	[-round]	[+round]
High	**pisk** urine	**ydz** five	**pɨs̱** bow			**puf** worm
Mid	**pem** hip			**ɵṯ** to climb		**poṯ** ten
Low	**pæːṯ** delivery				**pɑs̱** tooth	

Table 8 presents a minimal feature description of the vowels of Toda. This is the description that Emeneau uses, except for the specification of **æː**. Concerning **æː** (or **eː**, as Emeneau represents it), we see that from a historical point of view there is reason to represent it as **eː**, a vowel present in the vowel inventory proposed for Proto-Dravidian (McAlpin, 1981). But Maddieson (1984) and Maddieson & Precoda (1991) show four of the seven Dravidian languages in their sample as having four vowel heights. As we have seen, **æː** has the same F1 as **ɑ**, and is clearly distinct from **e**.

Table 8. A minimal feature specification of Toda vowels

	Front		Back	
	[-round]	[+round]	[-round]	[+round]
High	**i**	**y**	**ɨ**	**u**
Mid	**e**	**ɵ**		**o**
Low	**æː**		**ɑ**	

Emeneau's interpretation is based on auditory impressions of the vowel qualities, but he adopts a parsimonious classification of the vowel inventory using (a minimum) of phonological features. Emeneau classifies **ɨ** (his ï) as *back* rather than *central* since: "the acoustic effect is that of back, rather than central, vowels" (1984:6), and *not* due to the phonological patterning of the vowels. As we have seen, we disagree with this description, which yields a vowel inventory in which there are front high rounded and unrounded vowels, and back high rounded and unrounded vowels, as well as front mid rounded and unrounded vowels, and a mid back rounded vowel. The central vowels can be distinguished by the minimal feature set due to redundancies in the distribution of rounding: no central vowel has the same value of rounding as both a front and a back vowel of the same height. Since Emeneau provides no evidence showing phonological patterning of this classification of [–back] versus [+back] vowels, we present the feature classification as in Table 7, rather than the minimal feature classification as in Table 8, as the preferable representation of the vowel inventory.

Consonants

This section will focus on certain series of consonants in Toda which are interesting both from a phonetic and from a phonological point of view. We will first report some preliminary work on the stops and affricates, and then consider the sibilant and non-sibilant fricatives. The Toda consonant inventory contains other segments that are of interest, notably its rhotic series **r̪**, **r̠**, and **ɽ**. There is also a contrast between voiced and voiceless retroflex laterals **ɭ** and **ɭ̊**, which is unique among the world's languages.

Toda differs from most languages in that the range of consonants that can appear in word-medial and word-final position is less constrained than the range of consonants that can appear in word-initial or syllable-initial position. Many languages, such as Japanese and Finnish, show syllable-final, or word-final, constraints on the consonants that can appear in these positions, as opposed to few or no constraints on the consonants that can appear in word-initial, or syllable-initial, position. The set of consonants that can appear word-initially in Toda is discussed later.

The phonological contrasts among Toda consonants were shown in Figure 2 above. While Sakthivel (1976) also classifies the voiced sibilants: **z̪** and **z̠** as phonemic consonants, we follow Emeneau (1984) who views them as allophones, while retaining the voiced sibilants **ʒ** and **ʐ** as "invariable voiced sibilants" (p. 16), and thus phonemic. Sakthivel lists only the following near-minimal pairs in which voiced and voiceless dental sibilants contrast:

kɵːs̪	'Eugenia arnottliana'	**kɵːz̪**	'Litsaea wightiana'
pis̠pini	'carried-I'	**tiz̪pini**	'ate-I'

Emeneau (p.c.) notes that the word for 'Eugenia arnottliana' is **kɵs̠**, the word for 'Litsaea wightiana' is **kæːdz**, the word for 'I carried' is **pis̠pini** , and the word for 'I ate' is **t̪iz̠pini.** In the revised Dravidian Etymological Dictionary (Burrow and Emeneau,1984, henceforth DEDR) the latter two words are ***pit̠-s̠-pini** (DEDR 4565) and ***t̪id̠-s̠-pini** (DEDR 3263). Thus there is no evidence for voicing contrast between dental sibilants, as the sibilants in question are alveolar (variable voiceless) sibilants. The word for 'I ate' was recorded and for these speakers, the sibilant is clearly voiceless. Sakthivel does not appear to be as reliable or as meticulous a source as Emeneau; the accuracy and originality of his work is called into question in a book review of his 1976 book by Hockings (1976). Accordingly, we will follow Emeneau's classification of the consonants. Emeneau (1984:14), divides the consonants of Toda into the following sets:

invariable voiceless consonants:	**p t̪ ts t̠ tʃ ʈ k**
invariable voiced consonants:	**b d̪ dz d̠, dʒ ɖ g l ɭ ʒ ʐ r w**
variable voiceless consonants:	**f θ x r̠ ɽ ɬ ɭ̊ s̪ s̠ ʃ ʂ**
variable voiced consonants:	**m n ɳ y**

The invariable sets have no allophones of the opposite voicing, while the variable sets show intervocalic voicing of the voiceless consonants, and voiceless allophones of the voiced consonants in final position and following invariable voiceless consonants. For a more detailed discussion of this phenomenon, see Emeneau (1984:16-17).

Consonant distribution

All consonants may appear word-medially and word-finally, but the set of consonants that may appear word-initially in Toda is constrained to a small subset of each manner:

stops:	**p t̪ k**
fricatives:	**f s̪**
nasals:	**m n**
rhotics:	**r̪**
laterals:	**l̪**
approximants:	**j w**

These distributional facts support a claim for this set of consonants being the basic consonants of Toda. We summarize this basic consonant set in Table 9.

Table 9. The basic set of consonants that can occur in both initial and final position in Toda.

Manner	Labial	Dental	Alveolar	Palatal	Velar
Stop	p	t̪			k
Nasal	*m*		*n*		
Trill		r̪			
Fricative	f	s̪			
Approximant				j	w
Lateral			l		

For stop and fricative series that have three or four coronal places of articulation, the unmarked coronal place of articulation is dental, in that this place of articulation has no distributional constraints. The unmarked value for voicing is voiceless, as is observed in many other languages of the world (Maddieson, 1984). Dravidianist notation also supports a dental place of articulation as the unmarked coronal, with dental articulations having no additional diacritics, while alveolar places of articulation are noted with an additional diacritic. This preference for dental over alveolar consonants is not universal, as in general, more languages of the world have alveolar as their only coronal series (Maddieson, p.c.).

The richness of the consonant set that may appear in medial- and final-position, or more generally, post-vocalically, suggests that this may be due to historical context effects of lost final vowels, or perhaps a merging of more than one consonant, following the loss of a vowel or the collapse of a syllable.

Both Sakthivel (1976) and Emeneau (1984), list consonant clusters of varying length as appearing in medial and final position, but limit initial clusters to: **t̪w-**, **kw-**, **sw-**, **nw-**, and **pj-**, **t̪j-**, **kj-**, or more generally, labialized and palatalized consonants chosen from the minimal consonant set that appears in Table 9. Emeneau points out the lack of examples of labial consonants preceding **w** in initial clusters; this may be explainable as a constraint on initial adjacent consonants having the phonological feature [+labial], reminiscent of the Obligatory Contour Principle (Leben, 1973, Goldsmith, 1979). An alternative approach is proposed by Kawasaki & Ohala (1984): they explain this as "a strong tendency among languages to avoid sequences of the sort... *labial(ized) consonants+[w] or high back rounded vowels" (1984:122)* and consider it to result from the fact "that they create minimal modulations in amplitude, periodicity, and spectrum" (1984:123) and as such will not be accurately identified.

Consonant Series

If we collapse the voiceless stops and affricates into a single series, the voiced stops and affricates into another, and the non-sibilant and sibilant fricatives into a third, they form three complete series of consonants, each with seven places of articulation, as shown in Table 10. The great majority of the these consonants can be considered to be coronal. Both the voiced and the voiceless stop/affricate series have five coronal stops/affricates, and the fricative series has four voiceless and two voiced coronal sibilants, and one interdental non-sibilant fricative.

In *The Sound Pattern of English* (Chomsky & Halle, 1968), a feature [±distributed] is proposed to distinguish between consonants in languages with series such as: **/p, t̪, t̠, ʈ, k/**, or in our case: **/f, θ, s̪, s̠, ʃ, ʂ, x/**. Toda is mentioned as one such language (p. 312). [±distributed] is proposed to replace the apical/dental distinction, so as to include the distinction between bilabial and labiodental consonants, in which the tongue is not a common primary articulator. This feature refers to the constriction length *in the direction of the airflow*: [+distr] characterizes a longer constriction than [–distr]. A description of Toda fricatives (and we assume of stops/affricates) using SPE features would be as follows:

Consonant	anterior	coronal	distributed	strident	high	back
f	+	–	–	+		
θ	+	+	–	–		
s̪	+	+	+	+	–	–
s̠	+	+	–	+	–	–
ʃ	–	+	+	+	+	–
ʂ	–	+	–	+	–	–
x	–	–	–		+	+

A hypothesis for future study is that stop articulation in word final position will be more precise than in word initial position. More variation in the place of articulation (dental-alveolar) is anticipated initially because only one coronal articulation occurs in this position in the word. In final position, where there are four possible coronals, it will presumably be necessary to make each of these sounds precisely, so that the distinctions among them are maintained. While this hypothesis is interesting not only in regard to Toda, it lies beyond the scope of this study.

Table 10. Toda consonant inventory with collapsed consonant series

	Labial	Dental	Dent-Alv.	Alveolar	Retroflex	Palato-Alv.	Velar
Voiceless stop	p	t̪	ts	t̠	ʈ	tʃ	k
Voiced stop	b	d̪	dz	d̠	ɖ	dʒ	g
Fricative	f	θ	s̪	s̠	ʂ ʐ	ʃ ʒ	x (ɣ)
Approximant						j	w
Lateral				ɬ l	ꞎ ɭ		
Nasal	m			n	ɳ		
Trill			r̪	r̠	ɽ		

Stops and affricates

We will consider stops and affricates together as this distinction can be made only in an arbitrary way. There is no doubt that **ts** and **tʃ** are affricates; but there is also a great deal of friction in the release of the voiceless apical alveolar stop **t̠** . If we consider the affricates **ts** and **tʃ** to have different places of articulation from each other and from any of the other stops (as is clearly true for **tʃ**, and is at least arguably so for **ts**, particularly considering that t̠ is also very affricated), then Toda has 7 places of articulation for stops as well as for fricatives. Dravidian languages typically have five places of articulation. Malayalam is exceptional in that it has six, which are usually described as bilabial, dental, alveolar, retroflex, palato-alveolar (or palatal) and velar. Toda is even more exceptional in that it has a seventh member of this series, the affricate **ts**. Words illustrating the complete set of Toda stops and affricates are given below. As only three of these stops, **p t̪ k**, contrast in initial position, the Toda contrasts are exemplified in final position.

kɑp	'black dot' (between eyes or on chin)	**kɑb**	'sugar cane'
pot̪	'ten'	**mod̪**	'churning stick'
pɑːt̠	'cockroach'	**mod̠**	'village with dairy'
mots	'dusk'	**podz**	'cotton blossom'
motʃ	'cot'	**kidʒ**	'bells tied to the ankles while dancing'
t̪ɑʈ	'churning vessel'	**mɑɖ**	'head'
wɑːk	'small bamboo vessel'	**toːg**	'to support burden'

We will focus our attention on the coronal stops **t̪ d̪ ts dz t̠ d̠ ʈ ɖ tʃ dʒ**, as there is little new that can be said about the bilabial and velar stops **p b k g**. Among the coronals, many additional examples illustrating differences among the dental, alveolar and retroflex stops, could have been given, but minimal pairs showing the difference between **ts tʃ** and between **dz dʒ** are comparatively rare; Emeneau (1984) does not provide *any* contrasting examples in his discussion of these sounds. In the neighboring language, Kota, he notes that **ts** and **tʃ** are free variants (Emeneau 1944). We found that in Toda there was also a great deal of variation. For example, on one occasion a group of speakers might pronounce the word for 'cricket' (insect) as **ɑts**, and five minutes later, when being asked to make a recording, as **ɑtʃ**. Similarly, one in a group of three speakers might say **kidz** 'bells tied to the ankles while dancing', and the other two say **kidʒ**. There are, however, a few pairs that seem stable, at least for the voiceless series, such as **mots** 'dusk' and **motʃ** 'cot'.

We investigated the articulatory characteristics of the Toda stops using palatographic analysis techniques which are fully described elsewhere (Dart 1991, Ladefoged 1993). The basic procedure was to paint the tip, blade and front of the tongue with a mixture of edible oil and finely powdered charcoal. The speaker then said the word to be investigated, which was always carefully chosen so that it had only one coronal consonant. The pronunciation was recorded using a video camera, which was also used to photograph a mirror placed in the mouth so as to show the area of the upper part of the vocal tract that had been contacted. The procedure was then reversed, painting the upper part of the vocal tract, and observing which part of the tongue had made the contact. Dental impressions were made of each speaker's upper teeth and palate, so that the photographs could be reproduced at life size, and sagittal sections of part of the upper surface of the vocal tract could be drawn. We have records of five speakers (three men and two women) made in this way.

All five speakers made the dental stops with the tip and blade of the tongue making contact with the upper front teeth and the alveolar ridge, as exemplified in the upper pair of photographs in Figure 7. Often, as in the case of the speaker illustrated, a considerable part of the alveolar ridge was involved. It would seem appropriate to classify these sounds as laminal denti-alveolar stops. The middle pair of photograhs in the figure illustrate the alveolar stops. Note that the dark areas on the backs of the subject's front teeth in the lefthand photograph (and in subsequent palatograms) are stains and not the result of tongue contacts. These stops were always made with the tip of the tongue contacting the middle of the alveolar ridge, making them clearly apical alveolars. The retroflex stops, exemplified in the lower pair of photographs, also present no problems. There is very little contact visible on the tip of the tongue, as, instead of the tip itself, it was the underside of the tip that was involved, making these sounds sublaminal retroflex stops. The contact was between the curled back tip of the tongue and the roof of the mouth well behind the alveolar ridge, in the region of the hard palate.

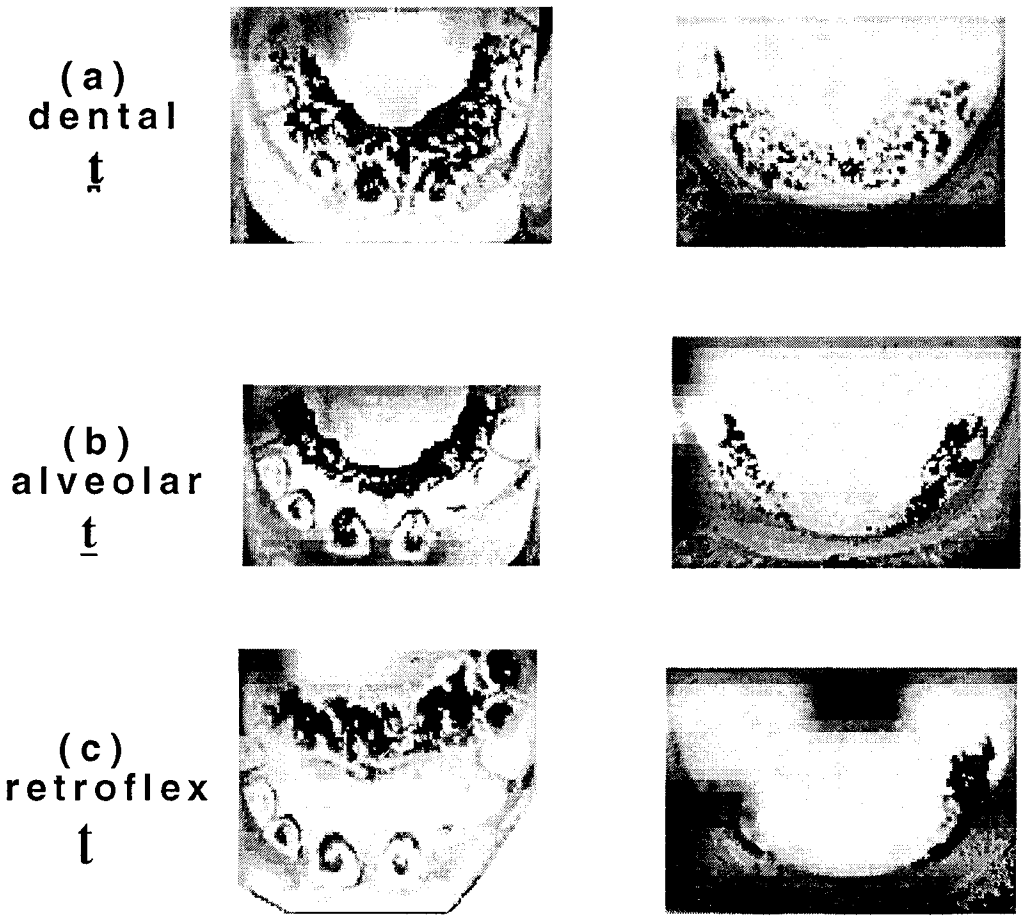


Figure 7. Palatographic and linguographic records of **pot̪** 'ten', **paːt̠** 'cockroach', **t̪aʈ** 'churning vessel' as spoken by Tes (speaker 1). The photographs have been retouched so as to remove highlights. The dark areas on the backs of the subject's front teeth are stains and not charcoal deposits.

Figure 8 shows the contrast between the minimal pair **mots** 'dusk' and **motʃ** 'cot' for Tes (speaker 1), and for one of our female speakers. In discussing these sounds, Emeneau notes that **ts** and **dz** are post-dental and *not* dental. But the contact in **ts** for the first speaker is actually very similar to that in his dental t̪ in Figure 7. The same is true for the second speaker and for the other three speakers for whom we have palatographic records (not shown here). We would classify both **ts** and **t̪** as laminal denti-alveolars for all these speakers, the distinction being that **ts** is affricated. For both speakers in Figure 8 the contact on the roof of the mouth must have been very light in the tokens of **mots** before the photograph of the tongue was taken, as for both of them there is virtually no marking in the center of the blade of the tongue. The records for **tʃ** are similar in that respect. Again for both of the speakers in Figure 8 there is only only very weak contact on the tongue in the center line. **tʃ** differs from **ts** in the extent of the contact on the hard palate behind the alveolar ridge and the greater part of the blade of the tongue that is involved. It is interesting that both speakers have far greater contact on one side of the mouth; we did not observe this with our other speakers. For both speakers illustrated (and for our other speakers) the tip of the tongue was behind the lower front teeth, and we infer that the front of the body of the tongue must have been raised as in a palatalized sound. We will consider another articulation of this kind when we discuss the fricative **ʃ**.

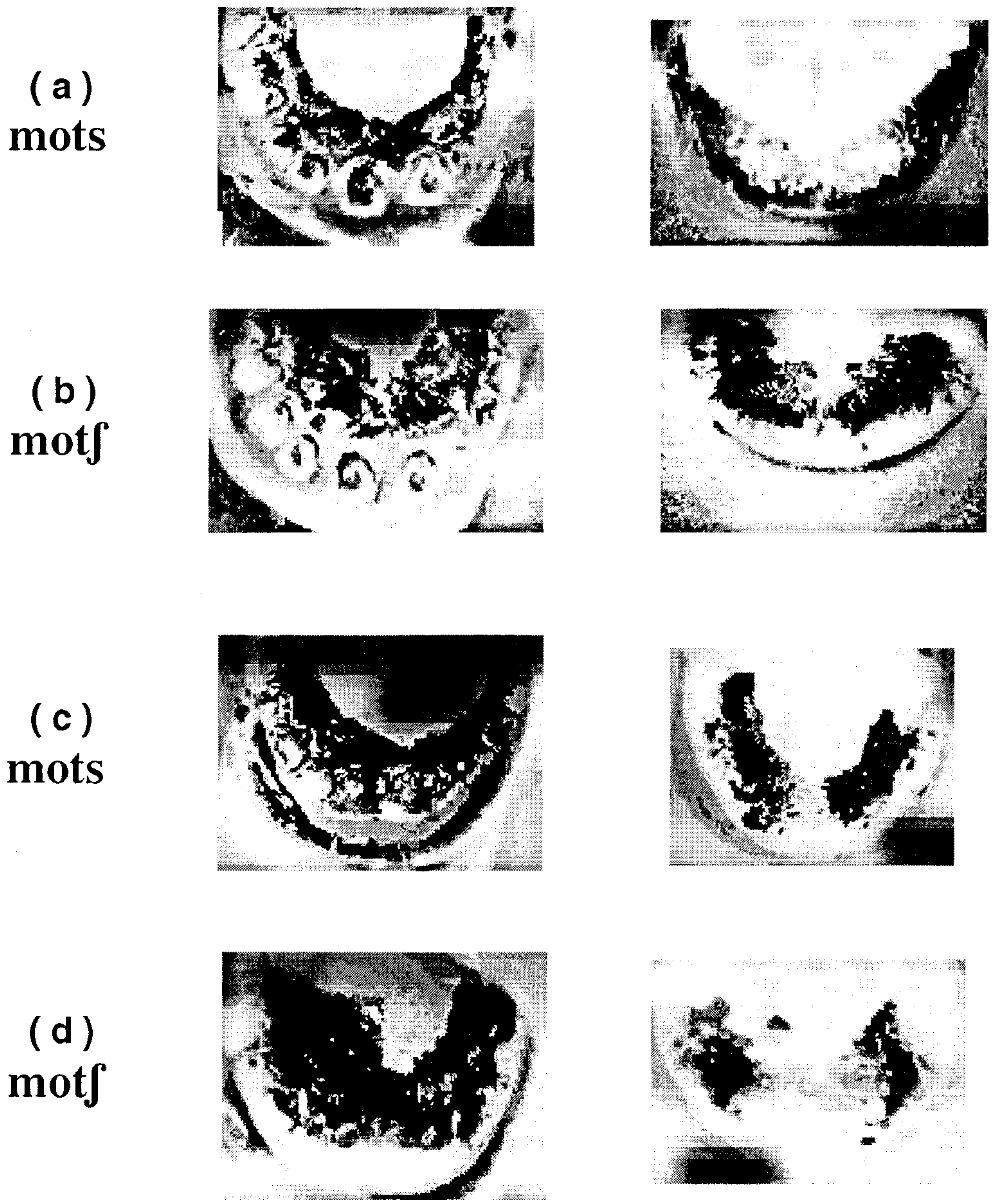


Figure 8. Palatographic and linguographic records of **mots** ‘dusk’ and **motʃ** ‘cot’ for Tes (speaker 1), and for Jayamisi, a female speaker.

Voice onset time

Voice Onset Time (VOT) is here taken to be the interval between the release of a stop closure and the start of voicing in the following vowel; thus in this study the fricative portion of a release is counted as part of the voiceless interval to be measured. All the Toda voiced stops had voicing virtually uninterrupted throughout the closure and on into any following voiced sound; these stops will not be considered further in this section. In this preliminary study of the stop consonants, we wanted to examine the effects of different places of articulation. As many of the Toda stops contrast only in final position, we used a data set in which each word was placed in a frame sentence such that the word following the test word began with a vowel. The words were said in the frame **kɑːḍ** __ **es̱s̱i** ‘Kad [a Toda name] said ___’. The data set used was **kɑp** ‘black dot', **pot̪** 'ten', **pɑːt̠** 'cockroach', **mots** 'dusk', **motʃ** 'cot', **t̪ɑʈ** 'churning vessel', **wɑːk** 'small bamboo vessel', **nep** 'live coal', **pæːt̪** 'delivery', **et̪** 'how many', **t̪æːt̠** 'to fold', **et̠** 'where', **pæːʈ** 'to hunt', **tek** 'small basket used for dairy'. VOT was measured from tape recordings of the six male speakers. One speaker did not produce a measurable record of **nep** 'live coal' and three speakers did not say **motʃ** 'cot'. Measurements were made on a Kay Elemetrics Computerized Speech Lab (CSL), using spectrographic and waveform displays. T tests were used to determine the significance of the difference in mean VOT between each of the places of articulation, except for the palato-alveolars, for which there was insufficient data. Results are as shown in Table 11.

Table 11. VOT (in ms) of Toda voiceless stops. The p values (significance) resulting from t-tests are shown in the lower part of the table.

	bilabial	dental	alveolar	retroflex	velar	dental affricate	palatoalveolar affricate
Speakers	5	6	6	6	6	6	3
Words	2	3	3	2	2	2	1
Tokens	10	18	18	12	12	12	3
Mean (ms)	17	22	25	12	28	52	43

Significant differences:

↑---------- .02 ----------↑ (bilabial – dental)

↑-----------------.04 -----------------↑ (bilabial – alveolar)

↑---------------------------.012 --------------------↑ (bilabial – velar)

↑---------.0006 --------- ↑ (dental – retroflex)

↑---.0003 --↑ (alveolar – retroflex)

↑---.0001---↑ (retroflex – velar)

The dental affricate is very significantly different from each of the other stops

It may be seen that Toda follows a universal tendency, in that the bilabial stops have a shorter VOT than the velar stops, and the dental and alveolar stops are intermediate (although not significantly different either from each other, or from the velar stops). What is more noteworthy is that the retroflex stops have a significantly shorter VOT than any of the other stops. Thus Toda sheds a little more darkness on our lack of explanations for the well known differences in VOT. It cannot be associated simply with the volume of air in the cavity behind the closure, in that the retroflex stops, even with their presumed lower jaw position, cannot have a considerably larger cavity than the bilabial stops. Nor can it be the length of the closure, as the VOT of the apical alveolar stops, which also have a short contact length, is intermediate between the laminal dental stops and the velar stops (although none of these tendencies is significant). Our best suggestion at the moment is that the extremely short VOT for the retroflex stops is associated with the forward movement of the tongue tip and blade that occurs during the release of a retroflex stop. This movement draws air upwards from the pharynx, effectively lowering the pressure above the vocal cords. But more experimentation and aerodynamic recording is needed before we can truly explain these phenomena.

Acoustic Characteristics of Toda Stops

Our characterization of the acoustic structure of Toda stops is still in a preliminary stage. We made a number of observations based on our examination of the spectrograms made in the course of measuring the voice onset times reported above. These observations have yet to be tested by a full statistical analysis of all our data. It would also be appropriate (and not too difficult logistically) to conduct perceptual tests of these preliminary findings, using synthetic speech stimuli. We hope that this will one day be done.

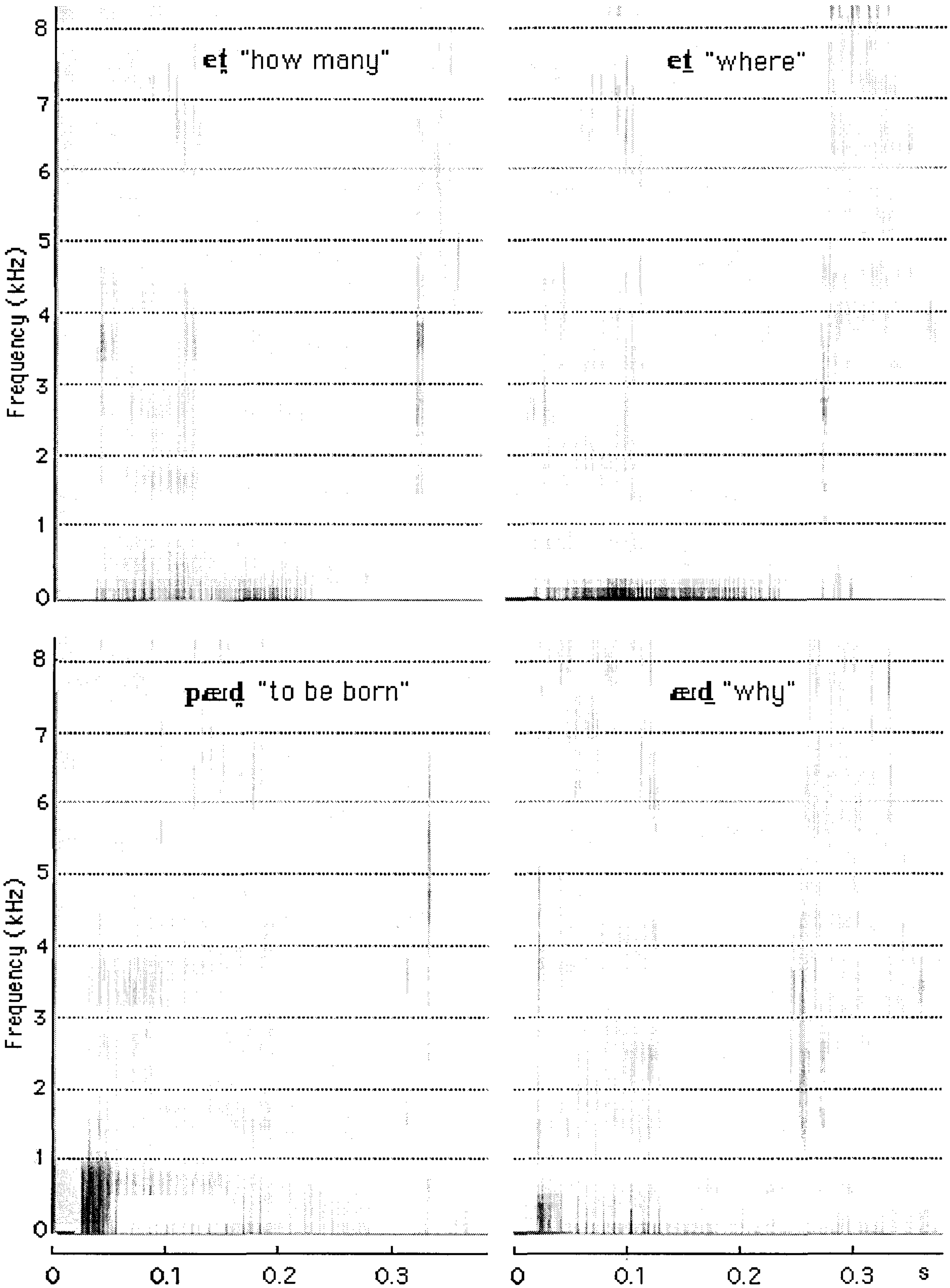


Figure 9. Spectrograms of **et̪** 'how many', **et̲** 'where', **pæːd̪** 'to be born', and **æːd̲** 'why' (speaker Tes).

The bilabial stops are similar to those in other languages: they have a comparatively weak burst, with energy centered around 1200 Hz. The dental and alveolar stops differ in that the burst has greater intensity and a longer duration in the alveolar stop than in the dental stop, as can be seen in the spectrograms in Figure 9, illustrating the difference between **et̪** 'where' and **et̠** 'how many'. This is somewhat surprising in that the palatographic data shows that the dental stops have a larger area of articulatory contact and thus might be expected to be more affricated. Presumably the shape of the constriction in the alveolar stops produces a more efficient noise source. There are no obvious differences in the loci of these stops; the formant transitions going into the dental and alveolar stops are very similar. However, the second and third formants have marginally higher frequencies during the vowel before the alveolar consonant than before the dental consonant. This is consistent across the majority of the speakers we have examined. It is indicative of a higher tongue body during the alveolar stops.

In **ts** most of the fricative energy is above 4000 Hz, giving an auditory impression similar to English **ts**. For the palato-alveolar affricate **tʃ**, the second and third formants come together, indicating an articulation in which the body of the tongue may be raised well behind the alveolar ridge . The lowest center of energy in the burst is around 3400 Hz. There is a large amount of fricative energy, extending to above 8000 Hz.

The retroflex stop **ʈ** is marked by the third formant descending, sometimes to as low as 2000 Hz. The burst has energy at about 1800 Hz. It is usually a short, sharp burst, for some speakers simply a spike with transient energy extending up to above 8000 Hz.

The velar stops have a burst centered at about 1700 Hz. There is less evidence of the second and third formants coming together than in English velar stops, perhaps indicating that the Toda stops have a more retracted articulation. Most of the speakers have some sort of double burst at the release of this stop.

We should emphasize that the preceding comments on the acoustic characteristics of the stop consonants should be regarded with caution. They are based on observation of at least 6 tokens of each type. But rigorous measurements have not yet been made, and we do not know the statistical validity of the numbers cited.

Fricative Articulations

Fricative sounds result from a turbulent airstream occurring in the vocal tract. The turbulence may be generated at the constriction itself, or it may be due to the high velocity jet of air formed at a narrow constriction going on to strike the edge of some obstruction such as the teeth. Following Holder (1669) and Shadle (1985) among others, we will distinguish between two classes of fricatives: sibilant and non-sibilant, or obstacle and non-obstacle, fricatives. Sibilant, or obstacle, fricatives are produced when the jet of air hits an obstacle (such as the teeth) situated downstream of the point of constriction.

As we saw in Table 2, Toda has seven voiceless and two voiced fricatives. There are three voiceless non-sibilants: **f**, **θ**, and **x**, four voiceless sibilants: **s̪**, **s̠**, **ʂ**, **ʃ**, and two voiced sibilants: **ʐ**, **ʒ**. The status of the voiced fricatives is complex. **ʂ** and **ʐ**, and **ʃ** and **ʒ** contrast in voicing, as the following tokens show: **kwiːʂ** 'pocket in outside edge of cloak' and **twiːʐ** 'in the pen', and **piːʃ** 'you (pl.) went' and **tiːʒ** 'at the **tiː** dairy'; but such contrasts are limited in the vowels that can precede them and are extremely rare. To the best of our knowledge, no actual minimal pairs exist in Toda. As Emeneau (1984:16-17) deals with the classification of these fricatives at some length, we will not address this issue here. We will simply examine the series of voiceless fricatives: **f**, **θ**, **s̪**, **s̠**, **ʃ**, **ʂ**, **x**, paying most attention to the sibilant fricatives. There is little that need be said about the non-sibilant fricatives: **f** is always labiodental (although we have no quantitative data, it seems

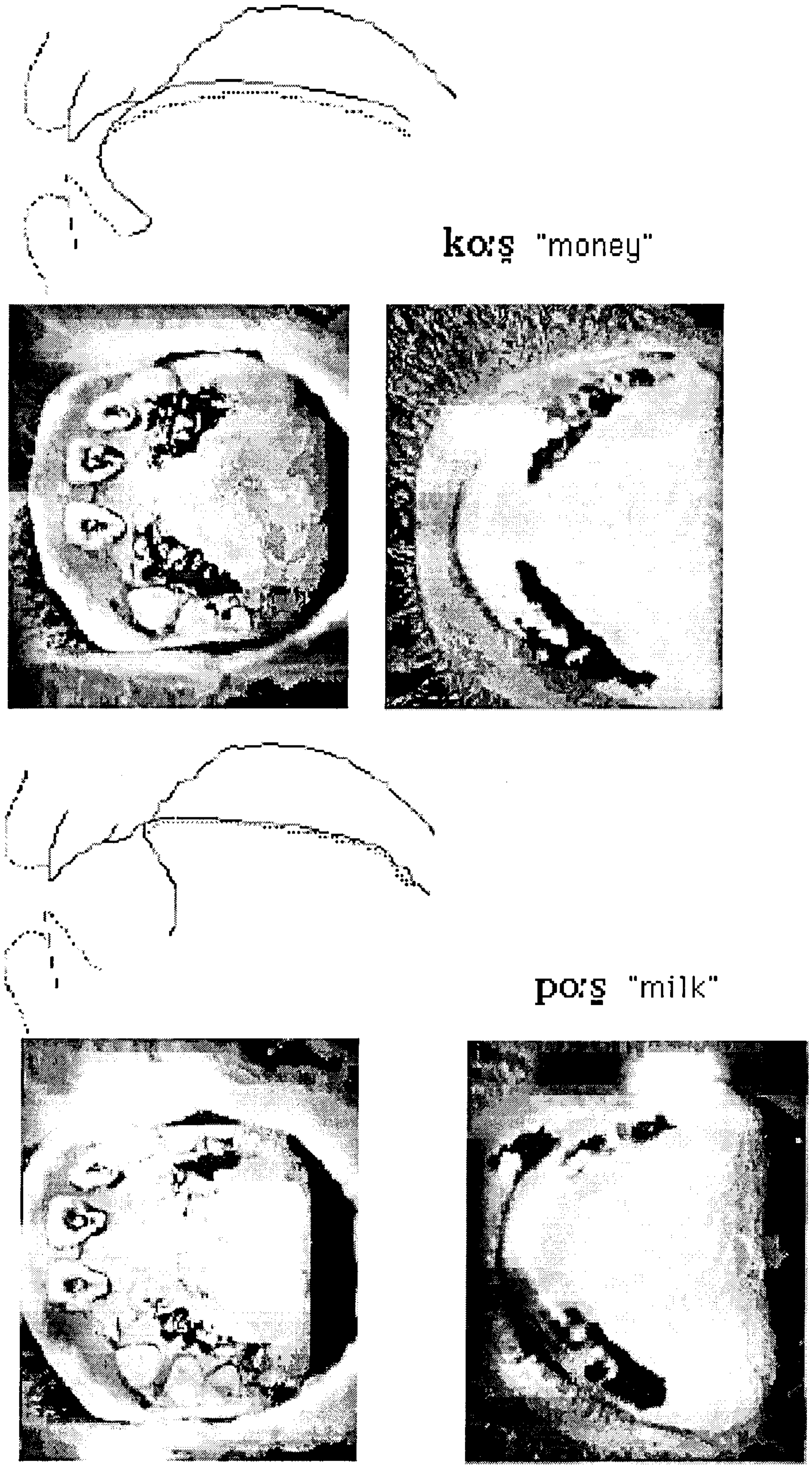


Figure 10. Palatographic and linguographic records of **koːs̪** ‘money’, **poːs̪** ‘milk’ (speaker Tes).

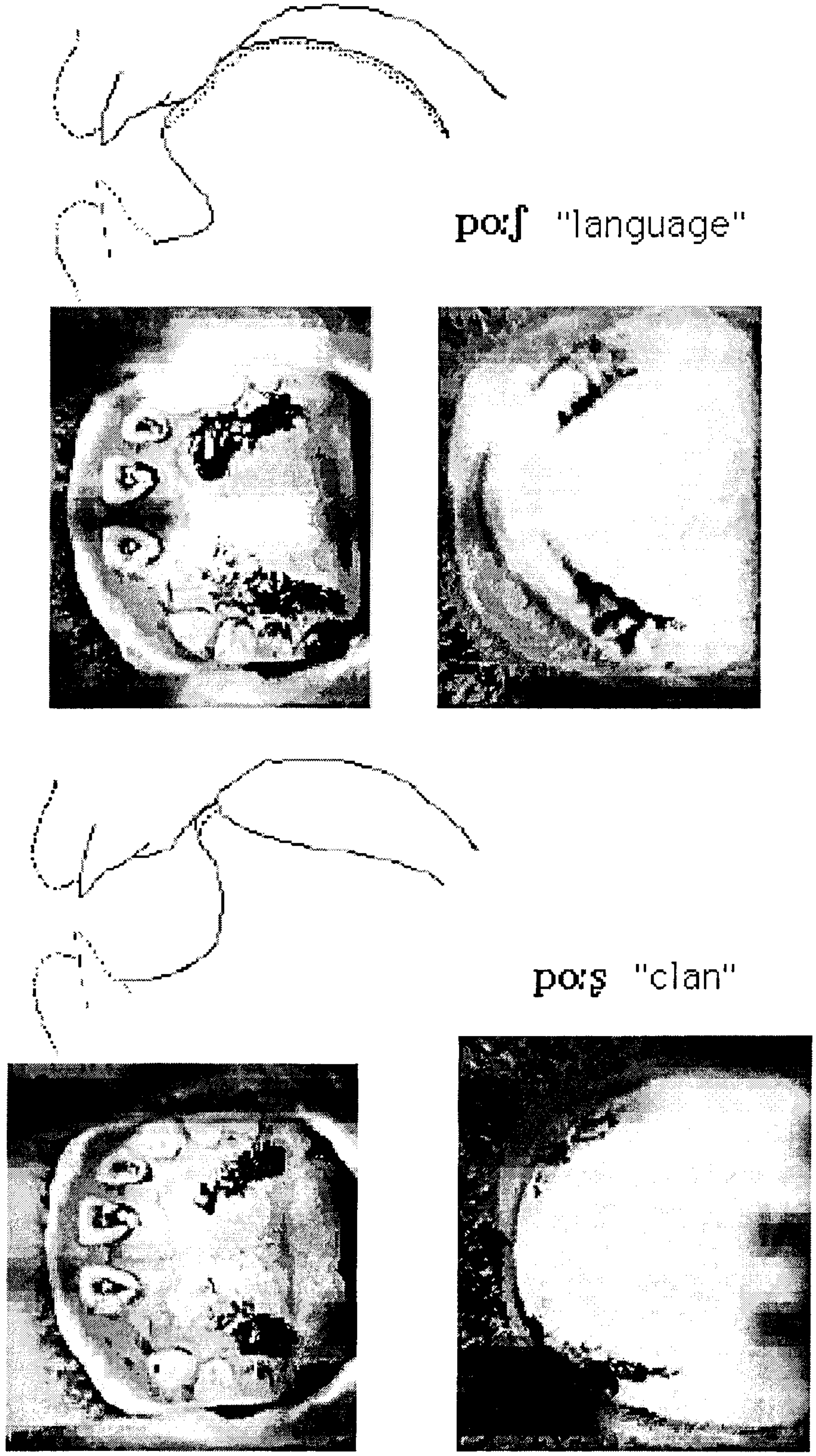


Figure 11. Palatographic and linguographic records of **po:ʃ** 'language', **po:ʂ** 'name of a mund' (speaker Tes)

to us that many of the Toda have comparatively large frontal incisors); **θ** is laminal dental or interdental (we have observed both possibilities, but our data do not allow us to quantify the relative frequency of each of them); and **x** is a velar fricative.

For our investigation of sibilants, we used the following words: **koːs̪** 'money', **poːs̱** 'milk', **poːʃ** 'language', and **poːʂ** 'name of a mund', which are a near minimal set. Palatographic techniques (described above) were used to determine both the exact place of articulation as well as the areas of the tongue that make contact with the palate. Figures 10 and 11 show the palatograms, linguograms and sagittal reconstructions of the four sibilants from one speaker (Tes). The photographs provide evidence only of the actual contacts made by the tongue against the upper surface of the vocal tract. The shapes of the center of the tongue shown in the diagrams are deduced on the basis of the known positions of the sides of the tongue, the casts giving the shape of the palate, and inferences from the acoustic characteristics of the sounds. It should always be remembered that only the place of contact and the parts of the tongue making contact are certain.

s̪ is described by Emeneau as being post-dental (pre-alveolar), with the tongue shaped like an English s. The palatographic evidence substantiates this description of the place of articulation, and shows that, as in English, there is a narrow groove through which the airstream passes. But, as Dart (1991) has shown, although English s is pronounced in a variety of different ways, the most common articulation is apical. Toda s̪ is laminal for all three subjects for whom we have satisfactory records. It is also noteworthy that, although (as we have seen) the stop t̪ is definitely dental, the fricative s̪ is more denti-alveolar.

s̱ is an alveolar sibilant, and further in contrast to s̪ it is apical. It also has a wider channel for the airstream. Toda thus follows the general pattern in the languages of the world, in that dental consonants are usually laminal, while alveolar consonants are typically apical (Ladefoged and Maddieson, 1986, Stevens, Keyser, and Kawasaki, 1986).

ʃ is a laminal palato-alveolar sibilant, with more contact of the tongue with the palate than either of the preceding sibilants. The tongue is domed up towards the roof of the mouth, so that this might be regarded as a palatalized consonant.

ʂ is a post-alveolar retroflexed sibilant, with apical or sublaminal lingual articulation. As we will see, the acoustics of this sound, particularly the rapid fall of the third and fourth formants, point to the tongue producing a sublaminal cavity in this sound (c.f. Stevens, 1989).

Acoustic analysis of fricatives

One of the interesting questions raised by a language having series of consonants with so many places of articulation is how speakers of such a language distinguish these segments in acoustic terms. In this section we will examine the fricative series of Toda, and attempt to answer this question for this set of sounds. Providing an acoustic description of fricatives is problematic in that it is not clear how this should be done. Vowels can be appropriately described in terms of their formant frequencies, as we saw in a previous section. But there is no general agreement on which acoustic attributes should be used to characterize differences among fricatives. We will take it that fricatives are distinguished by their overall intensity, and by the particular regions of high intensity in their spectra. They are also distinguished by the formant transitions — the movements of the formants into or out of the adjacent sounds.

Tokens of words containing sibilants were digitized onto a Kay CSL, sampling at 20480 Hz with 16 bit quantization. Spectrograms and Fourier Fast Transform (FFT) spectra were made as exemplified in Figures 12 - 15. The spectrograms show the acoustic energy up to 8320 Hz, and the spectra show acoustic energy up to 10240 Hz. The spectra are the means of 17 spectra, each with a duration of 5 ms, made during the central part of the fricative.

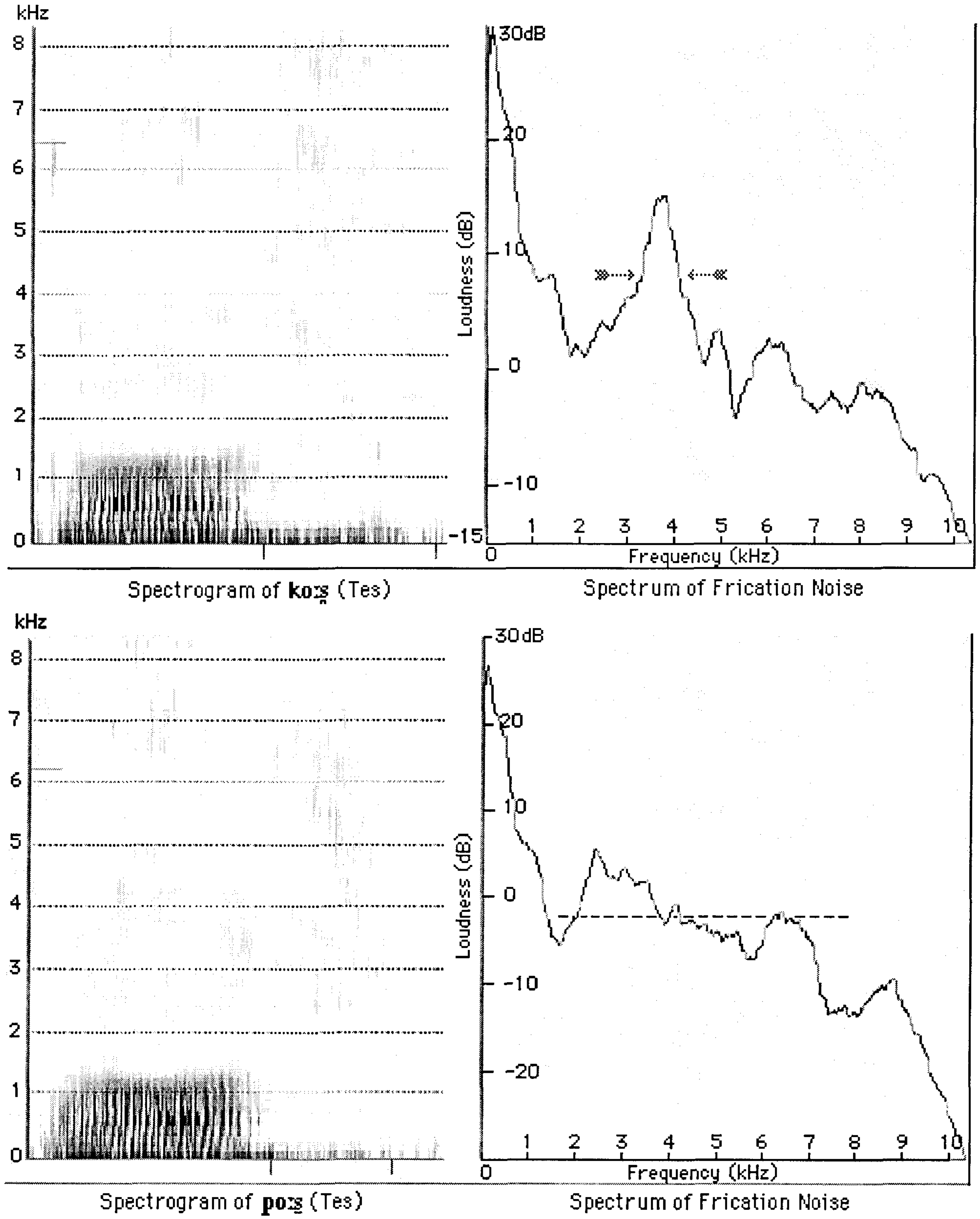


Figure 12. Spectrograms and FFT Spectra of **ko:s̪** 'money' and **pos̪** 'milk'

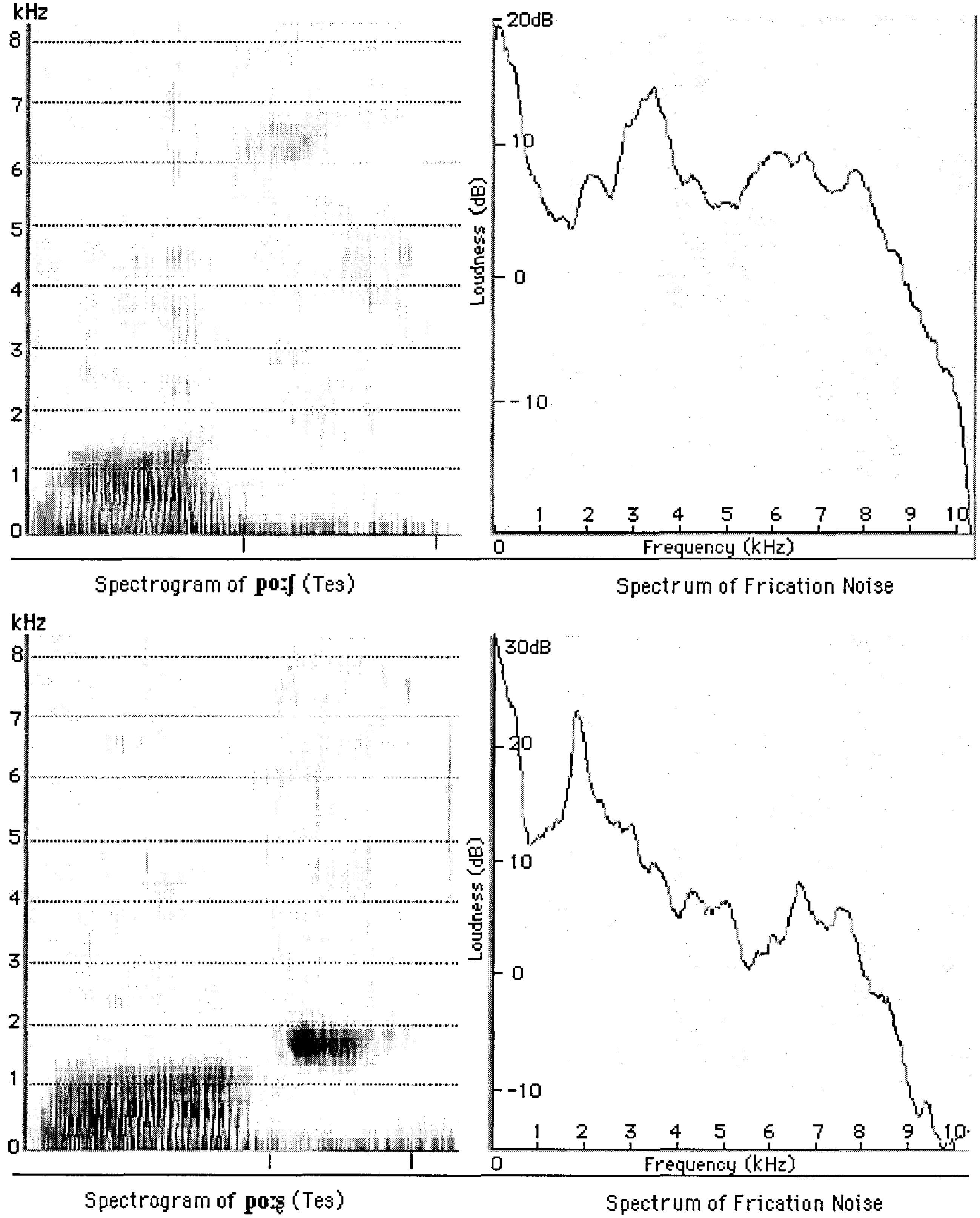


Figure 13. Spectrograms and FFT Spectra of **po:ʃ** ‘language’ and **poʂ** ‘name of a mund’

Table 12. Observations on spectral peaks and ranges (in Hz) of word-final sibilants following oː.

Speaker	Word	1st peak	Last peak	Most prominent	Comments	Frequencies at same dB
Tes	koːs̻	3000	8000	1st	3000–4000	3000–6200
Putkudn	koːs̻	2500	8400	N/A	No prominent peak	2500–9300
Putdil	koːs̻	3100	7300	N/A	No prominent peak	3100–7800
Po:thili	koːs̻	2700	8050	1st	3000–3650	2700–8450
KoTa:Dwit	koːs̻	3000	9500	3rd	5300–6900	3000–9500
SindiSk	koːs̻	3000	8500	N/A	No prominent peak	3000–9200
Tes	poːs̱	1600	8300	1st	2150–2500 Zero at 7000–7700	1600–6700
Putkudn	poːs̱	2200	7900	1st	2350–4000 No pronounced zero	2200–7950
Putdil	poːs̱	2000	8800	1st	2150–3650 Zero at 7000–7400	2000–6100
Po:thili	poːs̱	2900	8200	N/A	No prominent peak	2900–6500
KoTa:Dwit	poːs̱	1850	9500	1st	2200–3250	1850–8400
SindiSk	poːs̱	2600	9200	1st	2750–3600	2600–7800
Tes	poːʃ	1600	7400	2nd	2500–3200 Zero at 4400–5000	1600–7900
Putkudn	poːʃ	1450	8000	2nd	20503350 Zero at 4000–4700	1450–8100
Putdil	poːʃ	1300	6000	2nd	2600–3000 Zero at 3700–4300	1300–6800
Po:thili	poːʃ	1200	8500	1st	1750–2450	1200–7100
KoTa:Dwit	poːʃ	2400	8400	1st	2800–3600	2400–7750
SindiSk	poːʃ	2400	6750	1st	2400–4000	2400–7300
Tes	poːʂ	1400	7300	1st	1450–2300	1400–2400
Putkudn	poːʂ	1650	8200	1st	1950–3050	1650–7050
Putdil	poːʂ	1500	6800	1st	1750–2950	1500–6000
Po:thili	poːʂ	1350	7800	2nd	1st is a step up to 2nd	1350–4800
KoTa:Dwit	poːʂ	1300	1800	2nd	1st is a step up to 2nd	1300–5600
SindiSk	poːʂ	1400	9050	1st	not a large difference	1300–7500

Table 12 describes the FFT spectra of word-final sibilants appearing in the tokens: **koːs̻** 'money', **poːs̱** 'milk', **poːʃ** 'language', **poːʂ** 'name of a mund' for six speakers. The tokens analyzed form a near minimal set, in that all of the sibilants are in final position in monosyllabic words following the same vowel: **oː**. Tokens in which the fricative follows a front vowel were also analyzed, to ensure that the proposed analysis is not limited to the environment of following a mid back rounded vowel. Table 13 describes FFT spectra of tokens containing three of the four sibilants: s̻, s̱, ʃ, in which they follow front vowels in **pæːs̻** 'to thatch', **næːs̱** 'shade', and **t̪eʂ** 'gun'. One of these vowels is a long low front vowel, and the other is a short high mid vowel.

Table 13. Observations on spectral peaks and ranges (in Hz) of word-final sibilants following front vowels.

Speaker	Token	1st peak	Last peak	Most prominent	Comments	Bandwidth at same dB
Tes	pæːs̪	2900	8500	N/A	No prominent peak	2900–8200
Putkudn	pæːs̪	3200	9000	N/A	No prominent peak	3200–8700
Putdil	pæːs̪	2800	8200	N/A	No prominent peak	2800–8700
Tes	næːs̠	1100	9400	2nd	2800–3900 Zero at 4100–5700	1100–9300
Putkudn	næːs̠	2000	9000	2nd	3200–4500 Zero at 5000–5500	2000–7200
Putdil	næːs̠	2500	8300	1st	2900–4250	2500–6500
Po:thili	næːs̠	1100	8050	2nd	3200–4200	1100–7000
KoTa:Dwit	næːs̠	1150	8300	2nd	3300–5000	1150–8500
SindiSk	næːs̠	2600	8300	1st	3200–4250	2600–7550
Tes	t̪eʂ	1000	9000	2nd	2600–3700	1000–7000
Putkudn	t̪eʂ	1300	7400	2nd	2500–4200	1300–8500
Putdil	t̪eʂ	1000	7500	2nd	2500–3500	1000–7750
Po:thili	t̪eʂ	1000	8450	N/A	Zero at 4150–4600	1000–8500
KoTa:Dwit	t̪eʂ	1400	8400	1st	1900–4350	1400–8600
SindiSk	t̪eʂ	1200	8600	1st	18002900	1200–8750

The first two columns in each of these tables give the names of the speakers (used with their permission) and the token analyzed. The first speaker (Tes) is the speaker for whom palatographic and linguographic evidence for the articulation of these sibilants was given above. The third and fourth columns specify the locations of peaks in the spectrum. Designating peaks in an irregular spectrum is always somewhat arbitrary. We defined the first peak as the first local maximum point in the FFT spectrum after a rise of at least 10 dB. The last peak was said to be the local maximum before the last fall of at least 10 dB. The fifth column shows the most prominent peak, in cases where one peak is notably higher than its surroundings. The sixth column includes information on the frequency range around the most prominent peak. This range was measured at the lowest points at which the peak was distinct from its surroundings (see arrows in the FFT spectrum of **koːs̪** in Figure 12). This column also includes comments on local minima (the zeroes) that recurred in the spectra of some tokens. The seventh column shows the frequency range measured from the starting point of the first peak to the highest frequency which was at the same intensity level (see dotted line in FFT spectrum of **poːs̠** in Figure 12).

We will begin our discussion of the acoustic characteristics of the sibilant fricatives by considering the difference between ʂ and s̠. As we see in column three of Table 12, for five of the six speakers the frequency of the first peak in ʂ is higher than that of the first peak in s̠, with differences between the peaks ranging from 300 Hz to 1400 Hz, and a mean difference of 870 Hz. As column six shows, the highest frequency at which the intensity is at the same level as the first peak is also higher for ʂ than for s̠ for five of the six speakers (the exceptions are not the same speaker). Here the differences range between 1100 Hz and 1950 Hz, with a mean difference of 1500 Hz. The frequency of the most prominent peak is not a reliable metric here as there is no prominent peak (as defined here) in three out of the six tokens of **koːs̪**, and in one of the six tokens of **poːs̠**.

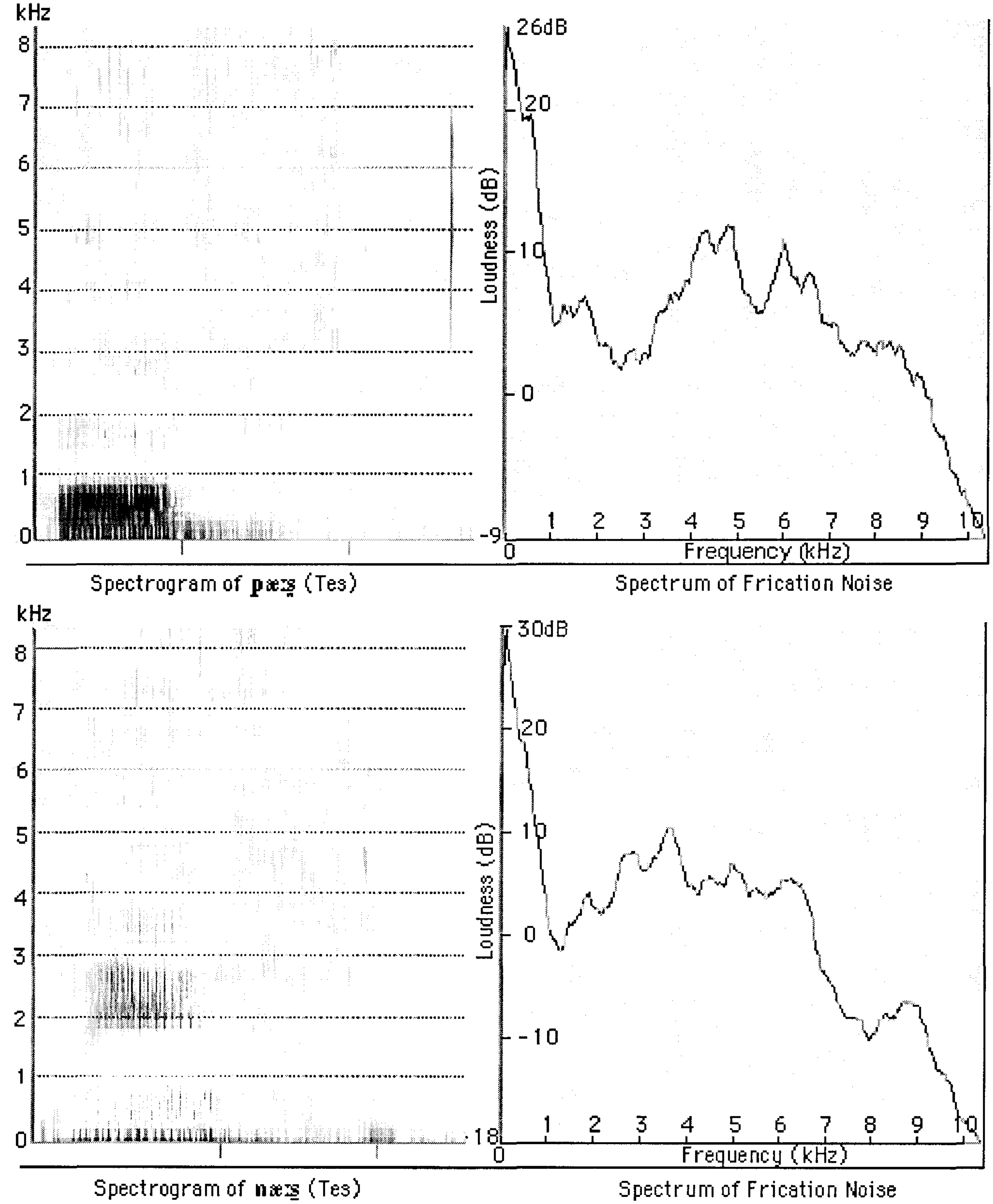


Figure 14. Spectrograms and FFT spectra of **pæːs̪** 'to thatch' and **næːs̪** 'shade' (speaker Tes)

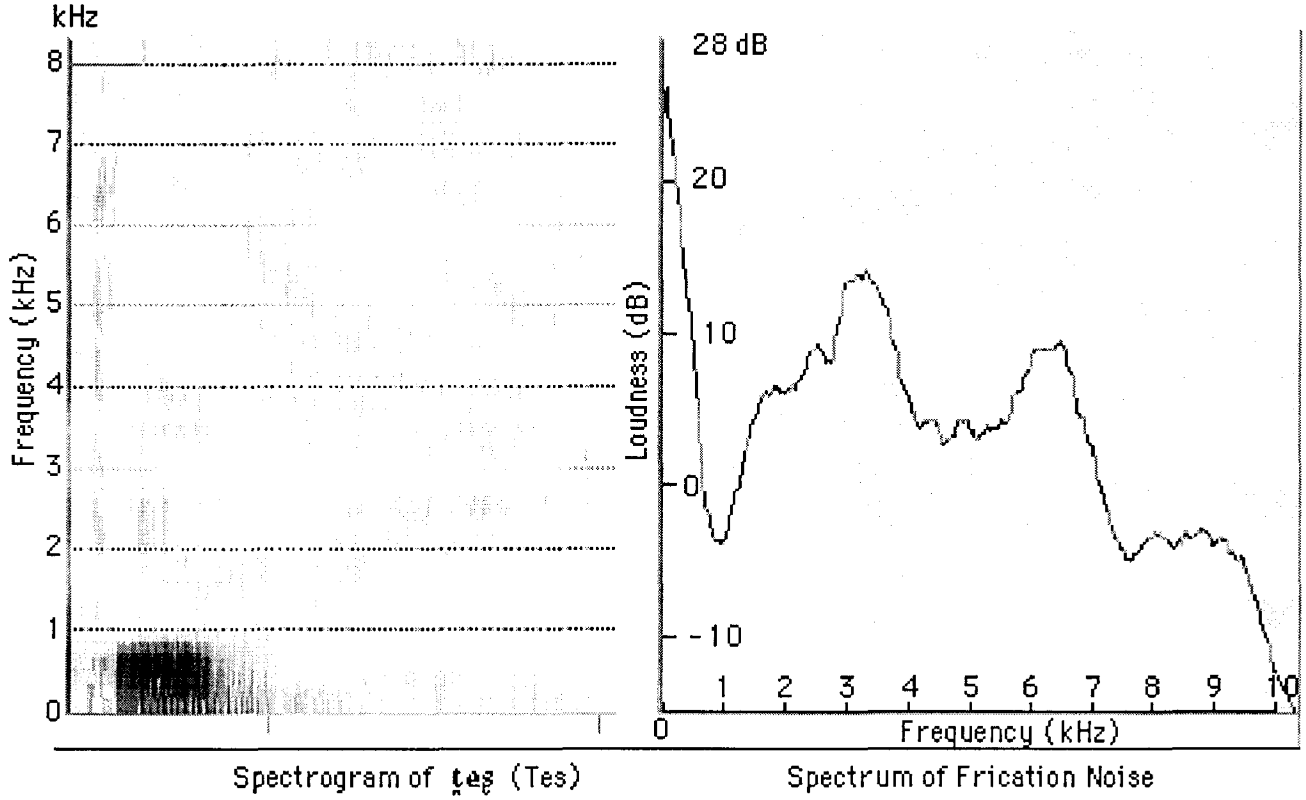


Figure 15. Spectrogram and FFT spectrum of **teʂ** 'gun' (speaker Tes)

Similar observations apply in part to the spectra of **pæ:ʂ** and **næ:s̠**, summarized in Table 13. For the three speakers that uttered both words, all of the first peaks in **pæ:ʂ** are higher in frequency than in **næ:s̠**: the differences range from 300 Hz to 1800 Hz, with a mean difference of 1100 Hz. Accordingly we may say that this pattern appears to hold across vowels, as for each of these three speakers the first peaks in tokens of both words containing ʂ occur at higher frequencies than the first peaks of both words containing s̠.

In regard to the frequency range between the first peak and the highest frequency at the same level of intensity, the results for the tokens in Table 13 are not as consistent. For one speaker, the frequency range is greater for s̠ than for ʂ, while for the other two the relationship is the same as in Table 12. Here again only the three speakers that uttered both words were compared. It is important to note that the same speaker (Tes) has the frequency range of ʂ being narrower than that of s̠ in all tokens, hence this might be due to the idiosyncratic pronunciation or vocal tract characteristics of this speaker. Nearly all the Toda speakers we recorded had poor dentition, and his was certainly aberrant. This leaves open the possibility that the observations made for the tokens in which the sibilant followed the long mid back round vowel hold true for sibilants following the mid front unrounded vowel as well.

In general, it appears to be the norm for the energy in ʂ to be distributed over a higher frequency range than that of s̠. There are various articulatory correlates of this difference. Dart (1991:63) notes that laminal articulations will have a higher F2 than apical articulations, because of the higher tongue body position of forming a constriction in the front of the mouth. The laminal articulation will also have a higher F3 and F4 than the apical articulation due to the sublaminal cavity formed by the apical articulation. The reason for this is that the front cavity's resonance, while relatively high due to its small size, could nonetheless be lower than the third resonance of the oral tract's main cavity. Dart found evidence supporting this in languages that contrast laminal and apical articulations, 'O'odham, an Amerindian language, and Malayalam, a Dravidian language. We may take the lower first peak of the frication noise to be associated with F3, which is higher in denti-alveolar/laminal s̪ (~2883 Hz) than in alveolar/apical s̠ (~2192Hz). As there are no other readily distinguishable resonances in the spectrograms, it is not clear if these two sibilants differ in regard to F2 and F4.

We propose that one acoustic cue for discriminating between s̪ and s̠ may be the frequency of the first peak of energy in the frication noise: for s̪ this peak is at a higher frequency than for s̠. The frequency range of the energy in the each of these sibilants supplies possible additional cues: the frequency range of energy (as defined in column six) in s̪ is situated higher than that of s̠; and it is also broader on the average than that of s̠ (the means are 5692 Hz versus 5050 Hz, respectively). The broader frequency range of frication noise in s̪ might be the result of the narrower tongue groove than that involved in the articulation of s̪ (as we see in figure 3–1) causing more turbulence.

Next we will consider the difference between ʃ and ʂ. Table 12 shows the mean peak energy in ʃ to be between 2470 Hz and 3430 Hz for five of the tokens, while for one speaker (Po:thili) the peak energy is between 1750 Hz and 2450 Hz. Prominent peaks (as defined here) appeared in only three of the spectra of ʂ; for these tokens the mean peak energy is between 1717 Hz and 2767 Hz, which is well below the mean peak energy for ʃ. This difference might be due to the size of the sublaminal cavity created by raising the front of the tongue: the retroflexion may create a larger sublaminal cavity than the articulation of palato-alveolar ʃ does, which enhances the frication noise that corresponds to F3 at a lower frequency. The spectra of ʃ for three speakers show zeroes between 4033 Hz and 4667 Hz (means), while none of the spectra of ʂ show significant zeroes. It seems likely that the frequency range of the highest energy intensity may serve as an acoustic cue for discriminating between ʃ and ʂ, and that the zeroes that appear in half of the tokens of ʃ may serve as an additional acoustic cue.

The spectrograms of the tokens of **poːʃ** (Figure 13) show vowel transitions from the vowel formants into the fricative noise in which F2 & F3 converge to around 2000 Hz for all six speakers (in four of the spectrograms F2 & F3 converge at 2000 Hz; in one spectrogram they converge at 2300 Hz, and in the sixth they converge at 1900 Hz). The rapid descent of F3 and F4 that is evident in the vowel transitions in **poːʂ** is not evident in the spectrogram of **teʂ**, and may be specific to cases in which the preceding vowel is a mid back rounded vowel. The strong band of energy at 3000 Hz that is evident in the spectrogram of **poːʂ** (Figure 13) is also evident in the spectrogram of **teʂ**. While the acoustic cues proposed above might be sufficient for discriminating between ʃ and ʂ, the vowel transitions serve as an additional, perhaps redundant, acoustic cue.

s̪, s̠ and ʃ, ʂ are distinguished by ʃ and ʂ having much more energy in their frication noise (relative to the preceding vowel) than either s̪ or s̠, and in this high energy being spread much more homogeneously over a broad frequency range. The frequency at which the first peak appears may serve as another cue to distinguish between the two pairs. The mean frequencies for the first peaks in these sibilants are:

Following **oː**	Following **æː** or **e**
s̪ – 2883 Hz	s̪ – 2967 Hz (only 3 speakers)
s̠ – 2192 Hz	s̠ – 1742 Hz
ʃ – 1725 Hz	
ʂ – 1433 Hz	ʂ – 1150 Hz

While the means are different for tokens following different vowels, a general tendency in the frequency of the first peak (the lowest frequency after an increase of at least 10 dB) is apparent. The laminal denti-alveolar s̪ has the highest frequency (2883 Hz or 2967 Hz), followed by the apical alveolar s̠ (2192 Hz or 1742 Hz), then a substantial drop to that of the laminal palato-alveolar ʃ (1725 Hz — no example after a front vowel), and finally a further drop to that of the sublaminal post-alveolar retroflexed s̢ (1433 Hz or 1150 Hz).

The distinction between s̺, s̠ and ʃ in Toda may be made on the basis of the different frequency range of the frication noise, specifically, the lower boundary of the frication noise. As we see above, in s̺ the first peak is at 2883 Hz, and in s̠ it is at 2192 Hz, while in ʃ it is at 1725 Hz. In experiments on the perception and discrimination of fricatives in American English, Harris (1958) showed that differences of this magnitude are sufficient to distinguish between s and ʃ.

An Alternative Approach

Jassem (1979:82), characterizes the spectra of Polish fricatives with the binary features: Strident/Mellow, Compact/Diffuse, and Grave/Acute, originally proposed by Fant (1960). Jakobson, Fant & Halle define these feature values as follows:

<u>Strident</u>: sounds that have irregular waveforms... in the spectrogram such sounds are represented by a random distribution of black areas;
<u>Mellow</u>: sounds with more regular waveforms... [which] have spectrograms in which the black areas may form horizontal or vertical striations;
<u>Compact</u>: sounds characterized by the relative predominance of one centrally located formant;
<u>Diffuse</u>: sounds in which one or more non-central formants... predominate;
<u>Grave</u>: sounds in which the lower side of the spectrum predominates;
<u>Acute</u>: sounds in which the upper side [of the spectrum] predominates.

Jassem proposes that three of these feature values have direct phonetic realizations:
[+mellow] is characterized by the high energy peak being between 1500-3200 Hz;
[+compact] is characterized by the high energy peak being between 3200-5000 Hz;
[+acute] is characterized by a high energy peak between 5000 and 8000 Hz.

While [+mellow] and [+compact] are mutually exclusive, [+acute] can occur either with one of the two, or alone, when the highest energy peak is between 5000 Hz–8000 Hz. Table 14 characterizes the data in Tables 12 and 13 following Jassem, and we will see whether this scheme can distinguish between fricatives in Toda. The data on non-sibilant fricatives were not included in the analysis of sibilants, but were collected using the same method as the rest of the sibilants, and are given here to show what Jassem's analysis must account for.

As we see in Table 14, the criteria proposed by Jassem (1979) are not sufficient to capture the acoustic distinctions between these fricatives. The large amplitude difference evident in the spectra of sibilant and non–sibilant fricatives, which Jakobson, Fant and Halle (1949) capture by the *Strident/Mellow* distinction cannot be encoded using only three of the six possible values of these features. Using Jassem's definition of *Mellow*, we may characterize a spectrum in which there are only small energy peaks, the largest of which is below 3200 Hz as *Mellow*, while according to Jakobson, Fant & Halle's features, it would be defined as *Strident*.. In addition, Jassem's definition of *Acute* requires only that there be a high energy peak between 5000–8000 Hz regardless of the energy below 5000 Hz, while Jakobson, Fant & Halle's definition of *Acute* requires that the upper side of the spectrum predominates. Crucially, applying these criteria to the data does not yield any generalizations which distinguish between different fricatives. The only values that are consistently assigned to segments across speakers are:[–mellow, +acute] to **θ**, [+acute] to **θ**, **ʃ**, s̢, and **x**, [–mellow] to **θ**, and s̺. If we consider the distinctions between fricatives for each speaker, the situation is not much better: for Tes, s̠ and ʃ are the same, as are s̠ and **x**; for Putkudn, **f**, s̠, and s̢ are the same; for Putdil, **θ** and s̺ are the same; for Po:thili, **f**, s̺, and **x** are the same, as are **θ**, s̠ and s̢; for KoTa:Ḍwit, **θ**, **ʃ** and **x** are also the same, as are s̠ and s̢; and finally for SindiSk, **f** and **θ** are the same, as are s̺, s̢, and **x**.

Table 14. Characterization of fricative spectra following Jassem (1979)

Feature	Speaker	f	θ	s̪	s̱	ʃ	ʂ	x
mellow	Tes	–	–	–	+	–	+	–
compact	Tes	+	–	+	–	+	–	+
acute	Tes	–	+	+	+	+	+	+
mellow	Putkudn	+	–	–	+	–	+	–
compact	Putkudn	–	+	+	–	+	–	+
acute	Putkudn	+	+	–	+	+	+	+
mellow	Putdil	–	–	–	+	–	+	+
compact	Putdil	–	–	–	–	+	–	–
acute	Putdil	–	+	+	–	+	+	+
mellow	Po:thili	–	–	–	–	+	–	–
compact	Po:thili	+	–	+	–	–	–	+
acute	Po:thili	+	+	+	+	+	+	+
mellow	KoTa:Dwit	+	–	–	+	–	+	–
compact	KoTa:Dwit	–	+	–	–	+	–	+
acute	KoTa:Dwit	–	+	+	+	+	+	+
mellow	SindiSk	–	–	+	–	–	+	+
compact	SindiSk	–	–	–	–	+	–	–
acute	SindiSk	+	+	+	+	–	+	+

We hope to have shown that while Jassem's approach might have been sufficient to characterize the fricative series of Polish, the attempt we have made to extend his analysis to the fricative series of Toda does not meet with success. The acoustic analysis and spectrogram characterizations suggested in the preceding sections, offer some tentative generalizations that appear to be valid across speakers, for the data considered here. Nonetheless, for a definitive answer to be given to this question, further data must be collected and analyzed.

The evidence presented in this paper appears to support Dart's (1991) claim that in languages which contrast laminal and apical articulations ('O'odham, Malayalam in her sample), certain acoustic properties characterize and may serve as acoustic cues for distinguishing between these places of articulation, specifically a higher F2 for laminal articulations versus apical articulations, and a narrower bandwidth of fricative noise in the apico–alveolar sibilant versus the wider bandwidth of the laminal denti–alveolar sibilant. We believe we have shown that the sibilants themselves in Toda may contain sufficient acoustic cues to allow speaker/hearers to distinguish between them, although their post–vocalic distribution allows for other, perhaps redundant, acoustic cues such as vowel formant transitions. Further work on the stop/affricate series in Toda will help establish whether the generalizations proposed here for fricatives are true for the stop/affricate series as well. As we noted at the beginning, Toda is one of the most interesting languages in the world from a phonetician's point of view. There is much that is yet to be done.

Acknowledgements

This paper is for the Toda; may it help preserve some knowledge concerning their wonderful language. We would especially like to thank Vasamalli and K. Pothili, and Piljin Wiedeman for their help. The linguistic basis for the paper is Murray Emeneau's great work; we would like to thank him, too, for his continuing interest and helpful comments on our research.

References

Agesthialingom, S., & S. Sakthivel (1970) Toda phonology. *Journal of the Annamalai University, Part A – Humanities,* **27**, 131-168.

Burrow, T. & M.B. Emeneau (1984) *A Dravidian Etymological Dictionary.* 2nd edn. Oxford: University Press.

Catford, J.C. (1977) *Fundamental Problems in Phonetics.* Bloomington: Indiana University Press

Chomsky, N. & M. Halle (1968). *The Sound Pattern of English.* New York: Harper & Row.

Dart, S.N. (1991) Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics,* **79**.

Disner, S.F. (1983) Vowel Quality: the relation between universal and language-specific factors, *UCLA Working Papers in Phonetics,* **58**.

Emeneau, M.B. (1984) *Toda Grammar and Texts.* Philadelphia: American Philosophical Society.

Emeneau, M. B. (1944). *Kota Texts, I.* University of California Publications in Linguistics, Berkeley and Los Angeles: University of California Press.

Fant, G. (1960). *Acoustic Theory of Speech Production.* The Hague: Mouton.

Goldsmith, J. (1979) *Autosegmental Phonology.* New York: Garland Press.

Holder, William. (1669). *The Elements of Speech.* London.

Hockings, P, (1976) Review of S. Sakthivel: 'The Phonology of Toda with Vocabulary.' *Language,* **55**, 2, 489-490.

Jakobson, R., G. Fant, & M. Halle. (1965). *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates.* Cambridge, Mass.: MIT Press.

Jassem, W. (1979). Classification of fricative spectra using statistical discriminant functions. In Lindblom, B. & S. Öhman (eds.) *Frontiers of Speech Communication Research.* London: Academic Press.

Kawasaki, H. & J. Ohala (1984). Prosodic phonology and phonetics. In C.J Ewen, & J.M. Anderson (eds.), *Phonology 1*, Cambridge: Cambridge University Press, 113-128.

Kuipers, A.H. (1960) *Phoneme and Morpheme in Kabardian.* The Hague: Mouton.

Ladefoged, P. (1993) Linguistic phonetic fieldwork: a practical guide. *UCLA Working papers in Phonetics* 84, 1-24.

Ladefoged, P. & I. Maddieson (1986). (Some of) the sounds of the world's languages, *UCLA Working Papers in Phonetics,* **64**.

Ladefoged, P. & I. Maddieson (1990). Vowels of the world's languages, *Journal of Phonetics,* **18**, 93-122.

Leben, W. (1973). *Suprasegmental Phonology.* Ph.D. dissertation, MIT.

Lehiste, I. (1960). *Suprasegmentals*. Cambridge, Mass: MIT Press.

Lindblom, B. (1990). Models of phonetic variation and selection, *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)*, **XI**, 65-100.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Maddieson, I. & J. Gandour (1977) Vowel length before aspirated consonants, *Indian Linguistics*, **38**, 6-11.

Maddieson, I. & K. Precoda (1991) *UCLA Phonological Segment Inventory Database version 1.1*

McAlpin, D.W. (1981) *Proto-Elamo-Dravidian: The evidence and its implications*. Philadelphia: The American Philosophical Society.

Sakthivel, S. (1976) *Phonology of Toda with vocabulary*. Chidambaram: Sri Velan Press.

Shadle, C.H. (1985) The acoustics of fricative consonants. Ph.D. dissertation, MIT.

Syrdal A.K. & H.S. Gopal (1986) A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of theAcoustical Society of America,* **79** (4), 1086-100.

Appendix

Recordings of the 6 Toda men and 6 Toda women were made on various occasions in February 1992 using the following words (not in this alphabetic order). In approximately half the recordings the words were said as citation forms, in others they were spoken within a frame such as: **kaːd** ___ **es̠s̠i** 'Kad said ___' (Kad is a Toda name) . Copies of the recordings and further particulars are available on request.

aːc	crawling insect, cricket, cockroach
aːn	elephant
aːs̠	house
aːt̠	to dry up
ad̠	that day
amn	smallpox
atk	*that many*
atwok	at that time
aꞎ	rice put in ghee to clarify it
eːd̠	why
eːf	fever
eːr̠	male buffalo, bison
eːt	to receive, take
ed̠	on what day
ely	boundary
eɭ	a stick used in a native game
er	throw
er̠s-	charge with horns, sacrifice
es̠	leaf
et	how many
et̠	where
fuːc	to put garment on someone else
iːɳ	cicada
iːɳy	now
iːt̠	spear
iθ	this one
is̠	rat
iʈ	flour
kaɽ	pen for calves
kaːk	crow
kaːty	to make to fall
kaːw	forked stick
kab	sugar cane
kal	bead
kap	black dot (between eyes or on chin)
kar̠	juice, sap
kas̠	stone
kar	border of cloth
kaɬ	study
keːj	tree sp.
keb	metal vessel
kedy	to push a stick into mud
kep	mud wall
ket̠	vessel for collecting milk
kic	linking hands in dancing
kĭj	bells tied to the ankles for dancing
kiɭy	parrot
kiꞎy	to laddle out
koːrup	salt giving ceremony
koːs	money
koc	brass vessel
koc	to bite
kof	*to be turned upside down*
kofy	to hide
korɯɖ	cutting instrument
koꞎ	toddy
kud̠	bundle of firewood or grass
kud̠y	pit
kuʈ	small
kuʈy	handcuffs
kwaːw	carrion
kwɯːs̠	stick
kwɯːs̠y	a type of plant
kəːg	black thread
kɯtil	old
kɯɖy	a measure
mad	medicine
maɖ	head
meːɽ	to drive buffalo on migration
meꞎk	mouthful
miːn	star
midy	upstairs
miʈoxy	sweet meet
moːj	to hide (news)
moːt	words, speech
moːt̠	change direction
moɳ	earth
moɳy	sacred bell
moc	cot
mod	churning stick
mod̠	village with dairy
moj	buttermilk
mupoθ	thirty
məː	rain
naːm	saffron
nan	leaf-shoot
neːs̠	shade
nep	live coal
nes̠	rice paddy
nes̠aːɽ	buffalo name
nes̠oːf	moonlight
nex	carduus pycnocophalus, a thistle

no:ɳ	shame
no:x	buffalo heifer
no:xi:r	Nagore (place)
no:xorof	cobra
nobi	to believe, trust
nobky	belief
noj	poison
nɤc	broken rice
nəʐpini	I made rope
nɯ	shoulder
nɯb	to pry open
nɯʒpini	I swam
o:fy	breath
o:r	six
o:ry	who
ø:ʂt	to boil vegetables
o:wuk	elder sister
ofs	to calm down
ofɯl	puffed rice
oʈ	to pour
øʈ	eight
øt	to climb
pa:t̠	cockroach
pa:w	17
paɭ	bangle
paθ	eagle
pas̠	tooth
paꞎ	valley
pe:d-	to be born
pe:s	to thatch
pe:ʈ	to hunt
pe:t	delivery
pem	hip
pep	coagulant agent put in milk
peʃf	to form curds
peʈ	sp. wood reddish in color
pi:g	a knot in the hair
piku:	navel
pily	Toda sorcery
pisk	urine
po:b	snake
po:f	swelling
po:n	sky
po:p	boil, carbuncle
po:θy	moss
pø:r	name, root
pø:r̠	cliff
po:ʃ	language
po:s̠	milk
po:ʂ	clan name
po:x	blood
poj	cotton blossom
pomy	toy
pony	dew
pot	ten
pot	ten
poʈʃpini	I slept
poɖ	large fly sp.
poɬ o:ꞎ	chief priet of the Ti dairy
poɬy	sacred dairy
pøꞎt	white
puɽ	beating
pu:č	blue-gray
pu:f	flower
pu:my	earth
pu:ʈ	eighteen
puf	worm
pum	fruit
puʈ	stirring stick
pɤ:ꞎ	summer
pɯ:t	flesh, meat, muscle
pɯl	price
pɯɭ	crevice in wall
pɯs̠	bow
pɯsspini	I will beat
pɯt̠s̠pini	I carried
pɯtyʃpini	I sowed
pɯꞎ	bird
seɖ	ball
sim	lion
sin	gold, ear-ring
so:l	furrow
sə:ʐpini	I arrived
ta:m	Thursday
taʈ	churning vessel
te:t̠	to fold
teb	copper vessel
teg	to fasten loin cloth
tek	a small basket used for dairy
teʂ	gun
tet̠-	to wrap garment tightly round waist
to:g	to support burden
to:k	to shout with anger
to:θ	powdery, soft
tog	to be humbled
tok	species of thorny bush
tu:r̠	bunch of leaves
tu:r	knife
tɤb	to sneeze
tɤby	honey making insect
u:f	name of a buffalo
ud	black gram
uf	back of trunk of body
umy	husk of gram
up	salt
upum	much
upy	to agree
ut	ant hill

viːs	to swing arm
waːk	small bamboo vessel
wa̱rxy-	to sleep
weːɭ	saliva
wɯḏ	one
wɯḏy	bell shape
ʏːr	finger nail
ʏj	five
ʏry	sweat
yuːkuḏy-	be in comfortable circumstances
ɯb	needle
ɯḏ	to-day
ɯtk	this many
ɯty	small in quantity
ɯẕpini	*I said*

Phonetic structures of Khonoma Angami

Barbara Blankenship, Peter Ladefoged, Peri Bhaskararao and Nichumeno Chase

Angami is a Tibeto-Burman language spoken in the Naga Hills in the northeastern parts of India. There are several dialects, the most prominent being Chokri, Khonoma, and Kohima. The last is considered the standard variety for publications and is taught in schools. Although there is no Angami written literature, there are published translations of textbooks and religious writings (Ravindran 1974). Published descriptions of Kohima dialect include Burling (1960), Ravindran (1974) and Giridhar (1980). The various dialects are mutually intelligible but differ in tonal and segment inventory. This study describes the phonetic inventory of one of the smaller Angami dialects, Khonoma, which is spoken by about 4000 people in the extreme west of the Angami region. Further information about the phonology and grammar of Khonoma Angami can be found in Chase (forthcoming).

The data for this study are recordings of two female and four male adult native speakers, made in February 1992 at the Linguistics Department of Deccan College, Pune, India. All of the speakers were students at institutions in the neighborhood of Pune. The primary corpus is a list of words spoken in isolation and in the frame *a*^{2} ___ *si*^{3}*ʃu*^{2}*to*^{3} , "I write ___", which was recorded by all six speakers. One of the female speakers (NC) recorded the complete set of material on two separate occasions, giving us seven sets of recordings in all. There are also palatographic samples and airflow data for selected phonemes, some of which have been previously reported in Bhaskararao and Ladefoged (1991).

Vowels

Khonoma Angami has 6 vowels, which we shall transcribe as [i,e,a,o,u,ə]. Each of them can occur on any one of the four tones. The vowels occur only in open syllables, and there are no contrasting lengths. Although diphthongs do occur, they are very infrequent, and will not be discussed in this paper

The following sets of words illustrate contrasting vowel qualities.

i	*pi*^{2}	'mushroom'	*mi*^{1}	'fire'
e	*pe*^{3}	'to shiver'	*me*^{2}	'always'
a	*ta*^{3}	'to chew'	*ma*^{3}	'price'
o	*po*^{4}	'one'	*mo*^{3}	'to moo'
u	*pu*^{3}	'to bloom'	*mu*^{4}	'sweet'
ə	*pfə*^{4}	'to carry'	*mə*^{4}	'to refuse'

The seven sets of recordings of all these words were analyzed, using the Kay CSL instrumentation, and the frequencies of the first three formants were determined. The formants were plotted on a graph of F1 and F2', a weighted average of F2 and F3, calculated by using the formula given by Fant (1973: 52):

$$F2' = F2 + \frac{(F3 - F2)(F2 - F1)}{2(F3 - F1)}$$

Figure 1 shows the F1 and (F2'-F1) Hz values plotted on a Bark scale for each vowel spoken by the male speakers. Ellipses enclose all points within two standard deviations of the mean. Values for the two female speakers are shown in Figure 2.

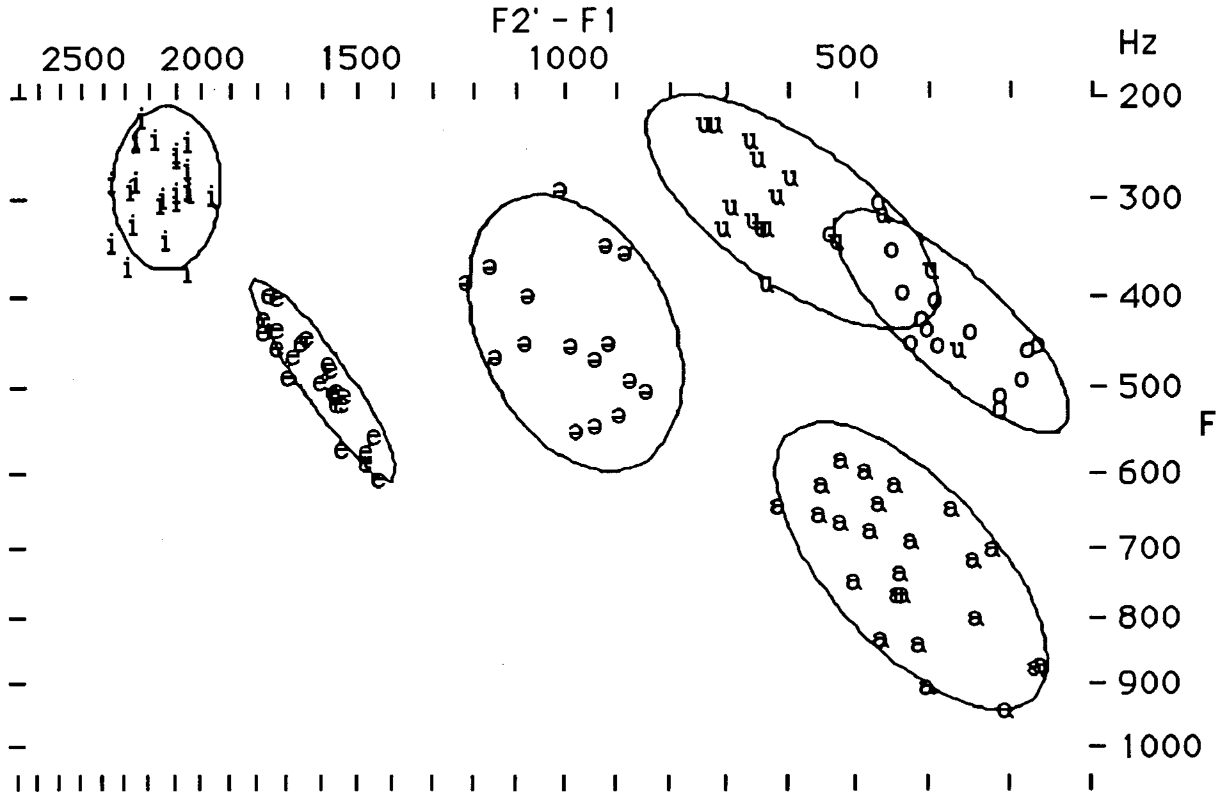


Figure 1. Angami vowels, male speakers (frequencies in Hz on a Bark scale).

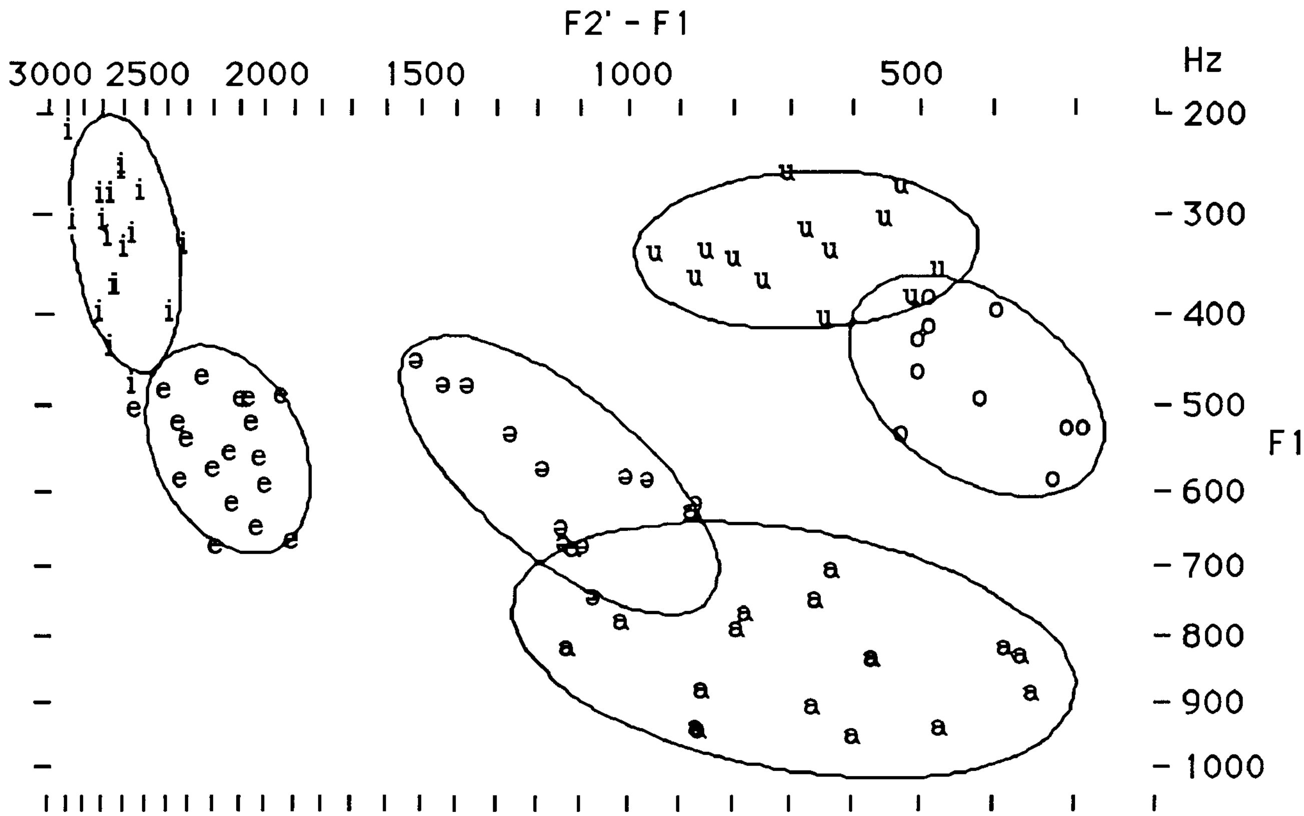


Figure 2. Angami vowels, female speakers (frequencies in Hz on a Bark scale).

The mean formant frequencies of the six vowels as spoken by the six speakers are as shown in Table 1. It is apparent from the plots and the mean formant frequencies that /i/ is a high front unrounded vowel close to cardinal [i]; /e/ is a mid-front unrounded vowel between cardinal [e] and [ɛ]; /a/ is a low central unrounded vowel between cardinal [a] and [ɑ]; /o/ is a mid-high back rounded vowel close to cardinal [o]; /u/ is a high rounded vowel near cardinal [u], but forward toward [ʉ]; and [ə] is a mid-central vowel that covers a range from the canonical [ə] upward toward [ɨ].

Table 1. The frequencies of the first three formants of the Angami vowels.

Vowel	Male speakers			Female speakers		
	F1	F2	F3	F1	F2	F3
i	301	2403	3028	347	2703	3238
e	520	2054	2776	577	2419	3193
a	749	1187	2640	841	1365	3082
o	470	864	2618	508	814	2725
u	331	941	2374	354	855	2700
ə	457	1438	2584	603	1482	2982

Vowels vary allophonically depending on the preceding consonant. Figure 3 gives a sample for the vowel [u], which is pronounced farther forward after [s] than after the bilabials [p] and [m]. Since a complete set of consonants before [u] was not available, we do not know the full range of contexts for this process.

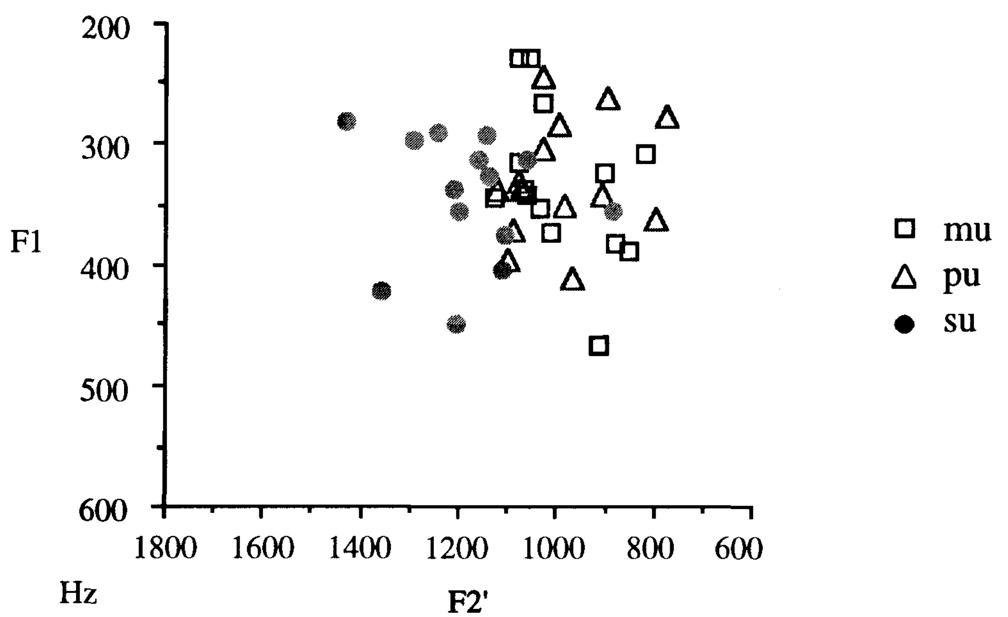


Figure 3. Distribution of [u] vowels in the contexts [mu], [pu], and [su], as spoken in isolation and in a frame by all six speakers

Tones

Khonoma Angami has four tones. (Five tones are reported for the Kohima dialect by Burling 1960 and Ravindran 1974, although they do not characterize each of the tones in the same way.) Throughout this paper the highest tone is indicated by superscript ^{1} and the lowest by superscript ^{4}. The four tones are exemplified in the following minimal sets:

g^we^1	'to bud'	*ke*^{3}*ɻi*^{1}	'to twist'
g^we^2	'to occupy'	*ke*^{3}*ɻi*^{2}	'to marry'
g^we^3	'to be thin'	*ke*^{3}*ɻi*^{3}	'to be ill at ease'
g^we^4	'physique'	*ke*^{3}*ɻi*^{4}	'to mix'
su^1	'to wash face'	ke^3ba^1	'snare'
su^2	'in place of'	ke^3ba^2	'time'
su^3	'to block (as of view)'	ke^3ba^3	'to place on top of each other'
su^4	'deep'	ke^3ba^4	'to play, mess about in mud'

The Khonoma Angami tones are illustrated in Figure 4. The four words (the fourth set shown above) were each cut out from the frame in which they were spoken. The arrows indicate the onset of the fifth harmonic in the second syllable of each of the four words. As the first syllable in each of the word is on tone 3, the tone rises during the second syllable of each of the first two words. In the fourth word, in which the second syllable has a low tone, there is a breathy voice quality, which makes it difficult to see the harmonics.

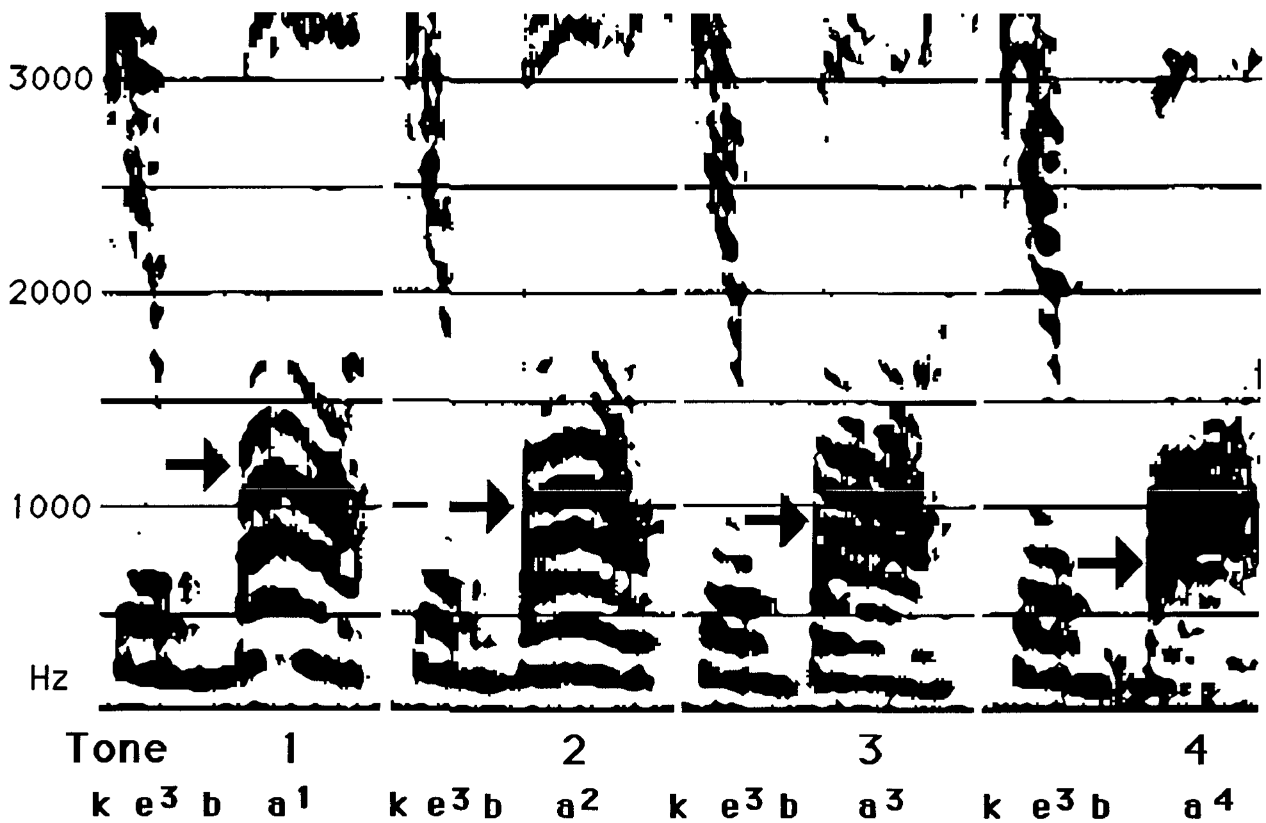


Figure 4. Narrow band spectrograms illustrating the four Angami tones.

Table 2 gives the average fundamental frequencies corresponding to the four tones for male and female speakers. The pitches are reliably different (significant at the .0001 level in analysis of variance), both by gender and by individual speaker.

Table 2. Mean fundamental frequencies of tones 1- 4 in Hz.

	male	female
Tone 1	165	287
Tone 2	141	230
Tone 3	127	206
Tone 4	110	173

Angami consonants

The following set of words illustrates the consonants, using (as in previously cited parts of the list) the conventional Angami orthography, with the addition of tone markings and using [ə] in the place of 'ü'. As we will see, not all these sounds are phonologically contrastive.

pʰ	*phe*^{4}	'to uproot'
p	*pe*^{3}	'to shiver'
b	*be*^{3}	'to cut'
m̥ʰ	*mhe*^{4}	'to blow'
m	*me*^{1}	'mouth'
v	*va*^{3}	'to illuminate'
ʍ	*whi*^{4}	'to roam'
w	*we*^{2}	'we'
pf	*pfə*^{4}	'to carry'
pʰɻ	*phrə*^{3}	'to read'
pɻ	*prə*^{2}	'hail stone'

tʰ	*the*^{3}	'to squeeze'
t	*te*^{2}	'to catch'
d	*de*^{4}	'to hem'
n̥ʰ	*nhe*^{4}	'to blow the nose'
n	*ne*^{1}	'to push'
s	*se*^{4}	'to deceive'
z	*ze*^{4}	'to sleep'
ʃ	*sha*^{1}	'road'
ʝ	*zha*^{1}	'big'
ɹ̥	*rha*^{3}	'to afford'
ɻ	*ra*^{3}	'village'
ɹ̥ə	*rhə*^{3}	'to plan'
ɻə	*rə*^{3}	'to aim at'
l̥ʰ	*lha*^{3}	'to rummage'
l	*la*^{1}	'for'
ts	*tsa*^{1}	'little'

kʰ	*kha*^{2}	'to prohibit'
k	*ka*^{2}	'to spoon out'
g	*ga*^{2}	'to winnow'
ɲ̊ʰ	*nyhie*^{4}	'to paste onto a wall'
ɲ	*nyie*^{3}	'thousand'
ŋ	*ngu*^{2}	'to see'
h	*ha*^{3}	'to breathe'
j̊	*yha*^{4}	'to dig out'
j	*ya*^{1}	'to shade'
tʃ	*ca*^{1}	'white'
kʰɻ	*khru*^{2}	'to bury'
kɻ	*kru*^{1}	'to flow'
kʍ	*khwe*^{4}	'cloth'
kʷ	*ke*^{3}*kwe*^{2}	'to tidy'
gʷ	*gwe*^{3}	'to be thin'
kʰf	*khfə*^{4}	'bitter'
kv	*kvə*^{3}	'to step up'
gv	*gvə*^{2}	'to bear fruit'

The consonant phonemes are as in the chart below. In the Khonoma dialect there are three-way contrasts in voice onset time among the plosives, but not among the affricates. Other dialects of Angami have a three way contrast at least among the affricates [tʃʰ, tʃ, dʒ] . All the other consonants also contrast in voice onset time, with the exception of /v,ŋ,h/. [f] does occur, but only as an allophone of [p]. [h] is in free variation with [x]. The retroflex [ɻ] is laminal before high vowels, but sublaminal before other vowels. The syllable [ɻə] can optionally be pronounced as a syllabic [ɻ], as it is in about half the instances in our recordings. This rhotic vowel has average formant values at 475, 1542, 1997, and 3150 Hz. Its location in the F1 x F2 vowel space is similar to that of [ə], but acoustically it is differentiated by the lowering of F3 and F4 that is characteristic of retroflexion.

Khonoma Angami consonant phonemes

	Bilabial	Labio-dent.	Alveolar	Postalv.	Retroflex	Palatal	Velar	Labial velar	Glottal
Plosive	pʰ p b		tʰ t d				kʰ k g	kʍ kʷ gʷ	
Affricate			ts	tʃ					
Nasal	m̥ʰ m		n̥ʰ n			ɲ̊ʰ ɲ	ŋ		
Fricative		v	s z	ʃ ʒ					h
Approx.					ɻ̊ ɻ	j̊ j		ʍ w	
Lateral			l̥ʰ l						

In addition to whatever affricates may occur, all dialects of Angami have consonant clusters involving [ɻ]. There are also sequences of consonants that result from a labiodentalization process which is recognized in the orthography. Labial and labialized consonants have allophones whose distribution is summarized in Table 3. For such cases, the consonant chart gives the allophone with the broadest distribution. An additional allophone, [bv], appears in the literature on Kohima Angami (e.g. [bvə] 'swollen', in Ravindran 1974: 30), but it does not occur in Khonoma Angami.

Table 3. Allophonic distribution of consonants.

Phoneme or cluster	Realization before [ə]	Realization elsewhere
/p/	[f]	[p]
/m̥ʰ/	[ɱ̊ʰ]	[m̥ʰ]
/m/	[ɱ]	[m]
/kʰw/	[kʰf]	[kʰw]
/kw/	[kv]	[kw]
/gw/	[gv]	[gw]

Voice Onset Time

As we have noted, Angami stops display a three-way distinction in the timing of the onset of voicing. The voiced stops, like those in most of the languages spoken in India, are usually voiced throughout the closure. In our data this means that they may be said to be pre-voiced for an average of 160 ms prior to the stop burst, the voicing then continuing into the following vowel. Voiceless unaspirated stops have an average voiceless interval of 15 ms between the stop burst and the onset of the following vowel. For aspirated stops, the voiceless interval averages 82 ms. A breakdown of voice onset times for the separate places of articulation is shown in Table 4. The VOT durations for the voiceless aspirated and unaspirated stops, both in isolated words and in a frame sentence, are similar to those recorded by Lisker and Abramson (1964) for other languages that make a three-way VOT contrast. The duration of voicing in Angami voiced stops, however, is noticeably longer than in the languages recorded by Lisker and Abramson: about 50% longer in isolated words and 125% longer within a frame sentence.

Table 4. Voice onset times for Angami stop consonants (measured in ms from the stop release).

Place of articulation:	Bilabial	Dental	Velar	Average
In isolated word:				
Voiced	-149	-152	-118	-140
Voiceless unaspirated	10	9	20	13
Voiceless aspirated	83	55	91	77
In frame sentence:				
Voiced	-172	-201	-169	-180
Voiceless unaspirated	10	14	29	17
Voiceless aspirated	83	70	108	87

Voiceless nasals

The voiceless nasals of Angami differ from those reported for other South and Southeast Asian languages in that they are aspirated. They are not like the voiceless nasals in Burmese, for example, which consist of a voiceless portion followed by a short voiced portion just before the release of the articulatory stricture into the following vowel. The nasal continues for a short time into the vowel before the velic stricture is closed. In Angami the voiceless nasals remain voiceless throughout the nasal articulation and even beyond the release; voicing of the following vowel begins well after the articulatory stricture has been released. The timing of glottal vibration (voicing) relative to the velic and articulatory strictures is shown schematically in Figure 5. Although the figure shows distinct events, the actual articulation and acoustic signal are blended at the segment boundaries: the Angami aspiration and vowel may be partly nasalized; the vowel may be only partly voiced at the beginning.

The voiceless nasals in Angami sound very different from those in Burmese. As noted above, there is no voiced portion towards the end of the voiceless nasal consonant, but instead, before the voicing for the vowel begins, the oral occlusion is released while air is still flowing out through the nose. The auditory impression is that there is an epenthetic voiceless plosive after the voiceless nasal and before the vowel. But this is an incorrect description as a plosive involves a complete stoppage of the air, which is then released orally. In Angami voiceless nasals there is continuous nasal airflow, which even persists into the following vowel.

```
                                  NASAL              VOWEL
Burmese
VELIC STRICTURE              |-------open---------|------closed-------|
ARTICULATORY STRICTURE       |----closed-----|----------open----------|
GLOTTIS                      |--voiceless--|----------voicing--- -----|

Angami
VELIC STRICTURE              |-------open---------|------closed-------|
ARTICULATORY STRICTURE       |----closed-----|----------open----------|
GLOTTIS                      |----voiceless------|------voicing-------|
                                              **
```

Figure 5. Overlap of glottal vibration (voicing) with velum opening and the release of the articulatory stricture in Burmese and Angami voiceless nasals. ** indicates the aspirated portion.

The structure of these unusual voiceless nasals may be seen in the aerodynamic records in Figure 6 (from Bhaskararao and Ladefoged, 1991), which show examples of each of the three voiceless nasals extracted from the frame sentence, as pronounced by one speaker. Significant moments in time are marked with arrows in the top example, At time (1) the articulators (in this case, the lips) close, and after a few vibrations of the vocal cords voicing ceases. The line indicating the oral airflow slopes slightly upwards from this point, probably because the lips are being pushed forward into the mouthpiece. At time (2) the articulators open and there is a rapid flow of air from the mouth. At the same time the airflow from the nose drops, but the velum is still lowered so that there is still a considerable flow of air through the nose. At time (3) voicing starts, probably with somewhat breathy vibrations, as there is a high rate of airflow through the mouth, as well as through the nose. If we take it that the vowel begins at this point, then we must consider at least the first part of it to be nasalized.

A similar sequence of events may be seen in the records for the other two voiceless nasals in this language. The oral airflow on the release of the alveolar closure (in the middle set of records) is particularly strong. It even causes some artifacts on the audio record which was made via a microphone held just outside the oral mask. It is hardly surprising that this sharp increase in oral airflow in the voiceless alveolar nasal gives rise to the auditory impression of there being an epenthetic voiceless alveolar plosive. But, although the nasal airflow drops at this moment, it still remains at about 500 ml/s, so this cannot be described as a regular plosive. The voiceless palatal nasal at the bottom of the figure shows a far less sharp release of the oral air. The nasal airflow also drops more slowly, and we may conclude from this and the comparatively smaller oral airflow that there is an oral obstruction comparable to a voiceless palatal fricative accompanying this voiceless nasal.

These patterns were consistent across all repetitions for all of the six speakers of Khonoma Angami that we recorded (as well as for a further three speakers of Kohima Angami who were recorded). Oral airflow began a little over half way through the voiceless section. Unlike the Burmese and Mizo sounds described by Bhaskararao and Ladefoged (1991), in which there was always some voicing during the last part of the nasal, in Angami there was never any voicing during the nasal. There were notable differences among the Angami voiceless nasals. The general pattern of these differences was similar to that exemplified in Figure 6. The oral airflow was usually greatest during the alveolar nasal, and almost as great during the bilabial nasal. The palatal nasal usually had both a slower increase of nasal airflow and a slower decrease.

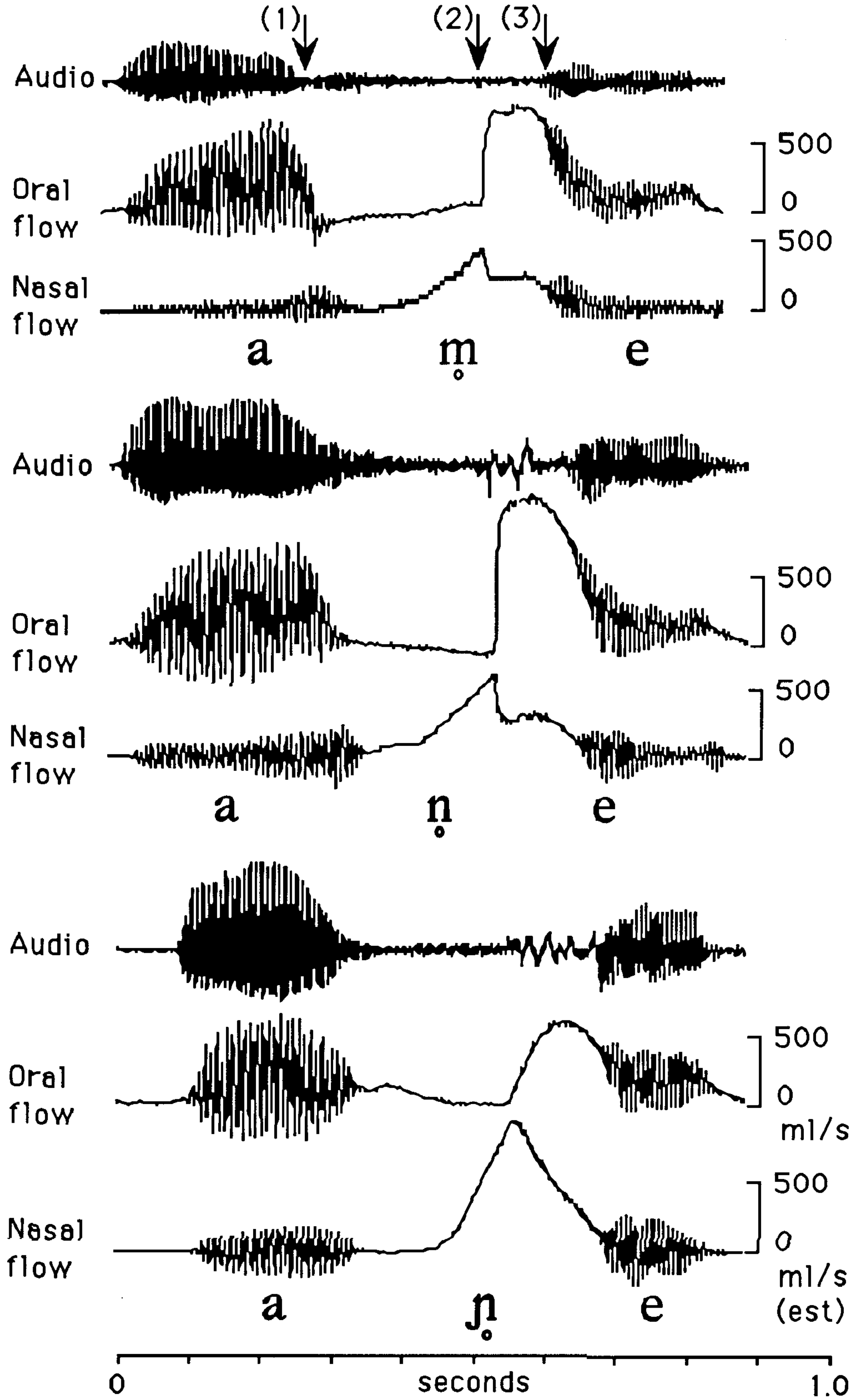


Figure 6. Angami aerodynamic records. See text for explanation.

We measured the durations of the different parts of the voiceless nasals for the six speakers of the Khonoma dialect of Angami in two different ways. Table 5 shows measurements made on the aerodynamic records. In this set of measurements we were mainly concerned with relating the nasalized aspiration – that is, the voiceless part of the sound after the release of the oral stricture and before the commencement of voicing within the vowel – and the purely nasal consonantal part – that is, the part in which there was an oral closure but nasal flow. The consonant was not the only part with nasal airflow; nasalization invariably occurred during the preceding vowel. Nor, when it occurred after a voiced sound, was the nasal consonant entirely voiceless; there was always voicing extending into the consonant from the preceding vowel. The measurements in Table 5 reflect a certain amount of variation in the durations due to overall differences in speed among the speakers, the different degrees of emphasis which the speakers placed on the word, and their varying degree of accommodation to the experimental task. Nevertheless it is clear that in by far the majority of cases the duration of the nasalized aspiration is substantial .

Table 5. Durations in ms of (1) the nasal consonants and (2) the nasalized aspiration for six speakers of Khonoma Angami. Several sets of data were obtained for some of these speakers.

Speaker	bilabial (1)	bilabial (2)	dental (1)	dental (2)	palatal (1)	palatal (2)	mean (1)	mean (2)
1 (a)	150	162	91	117	97	230	113	170
(b)	96	130	159	101	141	182	132	138
(c)	137	121	189	137	146	177	157	145
2 (a)	135	119	105	207	155	139	132	155
(b)	110	93	143	79	87	131	113	101
(c)	105	145	175	124	115	158	132	142
3 (a)	206	45	260	85	213	185	226	105
(b)	253	91	302	95	282	88	279	91
4	248	89	227	107	221	126	232	107
5	73	44	184	22	139	26	132	31
6	146	57	182	49	116	102	318	69
mean	151	100	183	102	156	140	179	114

We also made measurements of the audio recordings of Khonoma Angami, using the Kay CSL analysis system. Figure 7 shows the waveform and spectrogram for the word [m̥e^{4}] (to blow) as said within the frame sentence. Records of this kind enabled us to divide the nasal consonants into two parts, which were measured separately. The first, voiced, part extended from the time marked by the first arrow in the figure to the time marked by the second arrow; the second, voiceless, part was from the time marked by the second arrow in the figure to the time marked by the fourth arrow. The third arrow in the figure shows the moment of oral release within the consonant. The time between this point and the onset of voicing (the fourth arrow) corresponds to the aspiration, the voiceless part after the release of the oral closure, which was also measured. These acoustic records did not allow us to determine whether the aspiration was nasalized (as we could see it was from the aerodynamic records).

Table 6. Average duration in milliseconds and percentage of total duration for each portion of the voiceless nasal in VCV context.

	Voiced nasal	Voiceless nasal	Aspiration	Total
[m̥]	84 (41%)	73 (36%)	48 (23%)	205 ms
[n̥]	115 (55%)	54 (25%)	44 (20%)	213 ms
[ɲ̊]	94 (46%)	57 (28%)	54 (26%)	205 ms
Mean	98 ms (48%)	61 ms (29%)	49 ms (23%)	208 ms

The average durations of the voiced and voiceless portions of the nasal consonant and of the aspiration are shown in Table 6, which also indicates the percentage of time occupied by each portion. The voiced portion is about half of the entire segment, the voiceless nasal somewhat more than one fourth, and the aspiration somewhat less than one fourth. In these measurements, therefore, the aspiration is a somewhat smaller proportion of the total than in the aerodynamic measurements, in which it accounted for about 40% of the entire segment. We are not sure of the reasons for this discrepancy. Obviously the complexity of the procedure for recording aerodynamic data (described in more detail in Ladefoged 1993) accounts for some of it, in that it may have induced a more hyper-articulated style of speech. But some of the variation may be due to measurement procedures. What is comparatively easy to determine on acoustic records — onset and offset of voicing — is hard to see precisely on aerodynamic records; and what is easy to see on aerodynamic records — the onset of oral airflow when there is also nasal airflow — is often lost in the noise of the voiceless nasal on the acoustic records.

We also compared the durations of voiced and voiceless nasals. Voiceless nasals average 7-10% longer than their voiced counterparts. The difference, while interesting, is not great enough to support a theory that the aspirated voiceless nasal is really a cluster of nasal plus [h]. There is also no systematic phonological support for such a theory. Dantsuji (1984) discusses a similar finding for Burmese unaspirated voiceless nasals.

Acoustic correlates of place of articulation in voiceless nasals

When a word beginning with a voiceless nasal is preceded by a vowel, formant transitions from the vowel into the nasal give a distinct cue as to the place of articulation. But when the same word is spoken in isolation, there are no place cues at the beginning of the nasal. We are then left with the problem of determining how the voiceless nasals are distinguished from each other in these circumstances. They are all very similar sounds, but tests on speakers of the language left us in no doubt that they could distinguish between them.

Several aspects of the acoustic signal could serve as cues to place of articulation, among them:

1. the spectral pattern during the nasal portion
2. the frequencies of peaks in the spectrum at the time the nasal is released
3. the frequencies of peaks in the spectrum during the voiceless (aspiration) portion
4. timing of voice onset after the nasal release

We will consider each of these factors in turn. Each of them was measured in the words [m̥e^{4}] [n̥e^{4}], and [ɲ̊e^{4}] spoken in isolation by the six speakers.

The LPC (Linear Predictive Coding) spectrum of the kind used for analysis of vowels is inappropriate for the analysis of the nasal signal, in that it is an all pole model which cannot take the nasal zeroes into account. Instead an FFT (Fast Fourier Transform) of 128 points was used to determine the spectrum. The first pole was uniform at about 300 Hz for all nasals (as explained in Fujimura 1962). Therefore the frequency of the second pole was used as an indicator of spectral difference related to place of articulation. All factors were analyzed by means of T-tests between the pairs of consonant types.

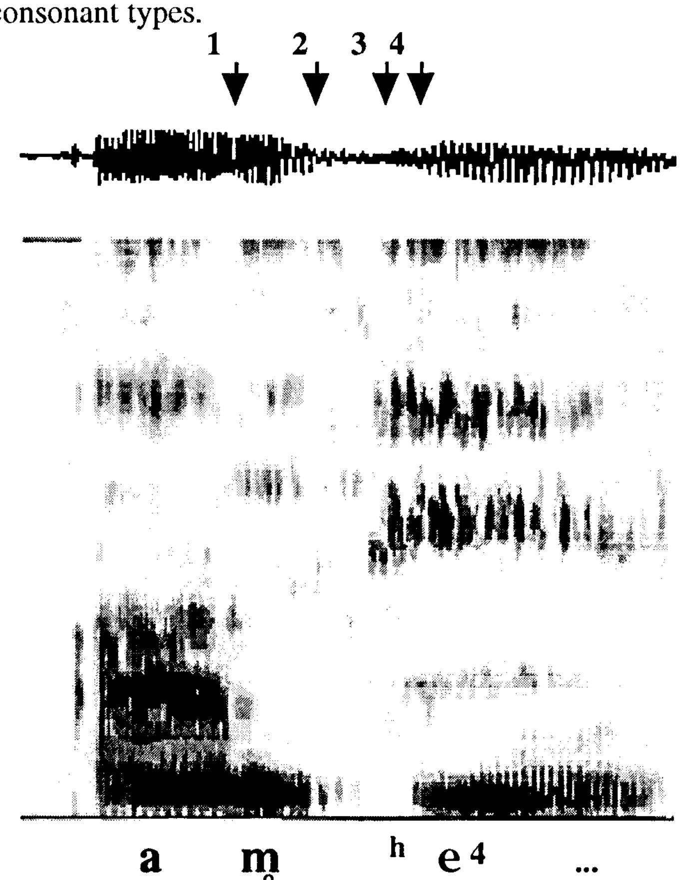


Figure 7. Waveform and spectrogram of the word [m̥e^4] spoken in a frame, only the preceding word being shown here. The arrows indicate: (1) onset of oral closure, (2) end of voicing of the previous vowel, (3) end of oral closure, and (4) onset of voicing of the following vowel.

Our findings with respect to the four points listed above were as follows:

1. There was no evidence that spectral patterns during the nasal closure varied systematically with place of articulation.

2. The mean frequency of the second pole of [ɲ̊] at the nasal release was significantly different (p<.001) from those of [m̥] and [n̥]. The means of [m̥] and [n̥] differed only at the p=.05 level. The mean frequencies were:

[m̥e^4]	1992 Hz
[n̥e^4]	2026 Hz
[ɲ̊e^4]	2692 Hz

3. The mean frequency of the second pole of [ɲ̊] at 30 ms after the nasal release was also significantly different (p<.001) from those of of [m̥] and [n̥]. The means of [m̥] and [n̥] were not significantly different. The mean frequencies were:

[m̥e⁴]	2129 Hz
[n̥e⁴]	2157 Hz
[ɲ̊e⁴]	2769 Hz

Comparisons with the frequencies of the poles at the time of the release (reported in (2) above) show that the second pole of both [m̥] and [n̥] increases by about 140 Hz over the first 30 ms of the aspiration, but that of [ɲ̊] increases by about 180 Hz. The transitions, as well as the frequency at the time of release, distinguishes the palatal place of articulation.

4. The voice onset time of [m̥] was different from that of [n̥] and [ɲ̊] at the p<.08 level. The differences between [n̥] and [ɲ̊] were less significant. The mean onset times were:

[m̥e⁴]	69 ms
[n̥e⁴]	84 ms
[ɲ̊e⁴]	92 ms

Since the data set is small (seven tokens of each consonant), the differences fail to reach a convincing level of probability. They can be taken only as indicative of trends that might be seen in a larger set of data. The palatal nasal can be distinguished from the others by its higher F2 frequency at the nasal release and during the aspiration, and the bilabial can be distinguished by its shorter voice onset time. No single factor among those tested appears adequate for distinguishing all three consonants, although the frequency of F2 at the nasal release might adequately differentiate them if there were more data.

Voiceless approximants

Angami distinguishes between voiced and voiceless approximants [l]/[l̥], [ɹ]/[ɹ̥], [j]/[j̊], and [w]/[ʍ.]. It is difficult to determine whether the voiceless approximants are aspirated like the voiceless nasals, since most of the approximant tokens do not display the instantaneous release which is characteristic of nasals and which demarcates the nasal portion before the release from the aspirated portion after the release. Among the approximants, only the laterals usually have a definite release followed by a voiceless non-lateral portion, similar to the structure of the nasals.

The voiceless approximants have a great deal of high-frequency noise, similar to the [x] sound that is a variant of Angami [h]. Thus the [l̥] could be characterized as a lateral [h], the [ɹ̥] as a rhotacized [h], the [j̊] as a palatalized [h], and the [ʍ] as a labialized [h]. Although such a characterization is an extremely useful device for teaching non-Angami speakers how to make these sounds, there is no systematic evidence to favor their interpretation as [h] plus secondary articulation, instead of approximants. In fact, for [l̥] tokens that have a release, the high-frequency noise is not present during the voiceless lateral portion, but only during the portion after the release.

References

Bhaskararao, P. and Ladefoged, Peter. 1991. Two types of voiceless nasals. *Journal of the International Phonetic Association* 21.2 80-88. [actually published in 1992].

Burling, Robbins. 1960. Angami Naga phonemics and word list. *Indian Linguistics* 21: 51-60.

Chase, Nichumeno. Forthcoming. Descriptive Analysis of Khwünomia Dialect. University of Poona Ph.D. dissertation.

Dantsuji, Masatake. 1984. A study on voiceless nasals in Burmese. *Studia Phonologica* XVIII.

Dantsuji, Masatake. 1987. An acoustic study of the distinction of place of articulation for voiceless nasals in Burmese. *Folia Linguistica* XXI: 281-291

Fant, Gunnar. 1973. *Speech sounds and features.* MIT Currrent Studies in Linguistics 4. Cambridge, MA: MIT Press.

Fujimura, Osamu. 1962. Analysis of nasal consonants. *Journal of the Acoustical Society of America* 34 (12): 1865-1875.

Giridhar, P.P. 1980. *Angami grammar.* CIIL Grammar Series 6. Mysore: Central Institute of Indian Languages.

Ladefoged, Peter. (1993) Linguistic phonetic fieldwork. This volume

Lisker, Leigh, and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20 (3): 384-422.

Ravindran, N. 1974. *Angami Phonetic Reader.* CIIL Phonetic Reader Series - 10. Mysore: Central Institute of Indian Languages.

Recasens, Daniel. 1983. Place cues for nasal consonants with special reference to Catalan. *Journal of the Acoustical Society of America* 73 (4): 1346-1353.

Appendix

Transcription of recordings

Recording of 20 February 1992.
Speakers: Nichəmeno Chase (f), Ketouseno Vəprə (f), Sanyə Iralu (m), Vitsilhuto Dolie (m).

Recording of 25 February 1992.
Speakers: Nichəmeno Chase (again), Avise Terhəja, (m) Petevizo Terhəja (m).

Each word is spoken in isolation, then the list is repeated with each word in the frame: a^{2} ____ si^{3}ʃu^{2}to^{3} "I write ____". (Chase is the only subject who is reading the word list; the others repeat after her.)

		tʃa^{1}	white
be^{3}	to cut	**ʃa**^{1}	road
pe^{3}	to shiver		
p^{h}e^{4}	to uproot	**ga**^{2}	to winnow
		ka^{2}	to spoon out
de^{4}	to hem	k^{h}a^{2}	to prohibit
te^{2}	to catch		
t^{h}e^{3}	to squeeze/wring	pfə^{4}	to carry
		za^{3}	to place pot on stove
ʒa^{1}	big	**tsa**^{1}	little

m̥^{h}e^{4}	to blow
me^{1}	mouth
n̥^{h}e^{4}	to blow the nose
ne^{1}	to push
ɲ̊^{h}e^{4}	to plaster onto a wall
ɲe^{3}	thousand
ŋu^{2}	to see
va^{3}	to illuminate
se^{4}	to deceive
ze^{4}	to sleep
ha^{3}	to breathe
ɹ̥ə^{3}	to plan
ɹə^{3}	to aim at
ɹ̥a^{3}	to afford
ɹa^{3}	village
j̥a^{4}	to operate/dig out
ja^{1}	to shade
ʍi^{4}	to roam
we^{2}	we
l̥^{h}a^{3}	to rummage
la^{1}	for
pɹə^{2}	hail stone
p^{h}ɹə^{3}	to read
kɹu^{1}	to flow
k^{h}ɹu^{2}	to bury
g^{w}e^{3}	to be thin
ke^{3}k^{w}e^{2}	to tidy
k^{hw}e^{4}	cloth
gvə^{2}	to bear fruit
kvə^{3}	to climb/step up
k^{h}fə^{4}	bitter
pi^{2}	mushroom
pe^{3}	to shiver
ta^{3}	to chew
po^{4}	one
pu^{3}	to bloom
pfə^{4}	to carry
mi^{1}	fire
me^{2}	always
ma^{3}	price
mo^{3}	to moo
mu^{4}	sweet
mə^{4}	to refuse
mie^{1}	body hair
g^{w}e^{1}	to bud
g^{w}e^{2}	to occupy
g^{w}e^{3}	to be thin
g^{w}e^{4}	physique
su^{1}	to wash face
su^{2}	in place of
su^{3}	to block (as of view)
su^{4}	deep
ke^{3}ɹi^{1}	to twist
ke^{3}ɹi^{2}	to marry
ke^{3}ɹi^{3}	to be ill at ease
ke^{3}ɹi^{4}	to mix
ke^{3}ba^{1}	snare
ke^{3}ba^{2}	time
ke^{3}ba^{3}	to place on top of each other
ke^{3}ba^{4}	to play or mess about in mud
pe^{3}zu^{3}	urine
pe^{3}zu^{2}	to urinate
ti^{3}ɹə^{3}	rain
ti^{1}ɹə^{3}	to rain
ɹə^{3}za^{3}	injury
ɹə^{3}za^{2}	to be injured
te^{3}ɹha^{3}	an item of traditional dress
te^{3}ɹha^{4}	to put on the te^{3}rha^{3}
vo^{3}kɹ(ə)^{1}	female pig
vo^{3}kɹ(ə)^{2}	male pig

Navajo vowels and phonetic universal tendencies

Joyce McDonough, Peter Ladefoged and Helen George

1. Introduction

Accounts of Navajo have transcribed the Navajo vowels as in (1), a 4 vowel system. The vowels are high front /i/, mid front /e/, low central /ɑ/ and mid back /o/.

(1)

i			
	e		o
		ɑ	

There is a long / short distinction and a full set of nasal vowels, so that in all there are 16 contrasts, not including tone, as shown in (2).

(2) oral:

short				long			
i				iː			
	e		o		eː		oː
		ɑ				ɑː	

nasal:

short				long			
į				įː			
	ę		ǫ		ęː		ǫː
		ą				ąː	

The transcription in (2) is not in accordance with the recommendations of the IPA (International Phonetic Association, 1989) in that nasality is marked with a subscript [̨]
in deference to the Athabaskan spelling conventions. We consider this departure from IPA practice to be necessary in order to make our results more available to our Navajo colleagues, who regularly use the system adopted here.

We will consider two points concerning the Navajo vowel system. First, the possibility of additional distinctions in vowel quality that are not represented in this transcription system, in particular the amplification of the long / short contrast by a peripheral / interior distinction. Secondly, we will investigate whether the lack of a high back vowel makes this system unbalanced, and thus a less favored system, of particular interest in the formulation of explanatory models of vowel distribution. We will be considering the phonetic details of Navajo vowels from two aspects: the number and kind of vowel qualities and the dispersion and variance of the vowels in the vowel space.

On the first point, many researchers have noticed that the long / short distinction has been amplified by a Tense / lax or peripheral / interior distinction, though there has been no general agreement about the extent and nature of the contrast. The primary source on the Navajo language from the 40's to the present has been Young and Morgan's extraordinary series of grammars and dictionaries. Young and Morgan (1980, 1986) suggest that there are 5 vowel qualities, with a tense / lax distinction (which they did not transcribe) only for the front high vowel. Interestingly, their earlier work (1972) notes an additional tense / lax difference in vowel quality for the back vowel. Sapir and Hoijer (1967) note these tense / lax distinctions in the front vowels. Maddieson (1984), based on his observations, suggests that there are 6 vowel qualities, with a tense / lax distinction for both the front high and the mid back vowels, as shown in (3):

(3) short long

i			iː		
					oː
ɛ		ɔ	ɛː		
	ɑ			ɑː	

On the second point, the overall distribution of vowels in the vowel space, the primacy of [i, u, a] as the peripheral vowels has been established by databases such as UPSID, (Maddieson, 1984). Navajo is an example of a disfavored vowel system, in that there is a gap in the Navajo system, the high back vowel being missing. The vowel system then is important for evaluating explanatory models of vowel distribution such as Lindblom's dispersion theory. In a discussion of the UPSID database, Maddieson (1984) calls systems with high back gaps 'skewed' systems, that is systems that do not fully exploit the differentiation in vowel quality that is possible within a triangular vowel space. Despite this gap, he maintains that a principle of sufficient contrast, consistent with dispersion theories, is operable in languages such as Navajo. Sufficient contrast predicts that when 'skewed' systems like Navajo do occur, evidence of compensation for the gap will be found. In order to meet the demands of adequate dispersion, the mid back vowel will move up to fill the gap, so that it is higher than its front counterpart, as shown in (4):

(4)

iː		
		oː
ɛː		
	ɑː	

2. Data and methods

A set of recordings was made in a classroom at Navajo Community College, Shiprock, New Mexico, in April 1992. The speakers were educated native speakers of Navajo, who live on the reservation and use Navajo daily. They were recorded while repeating a word list that illustrated the phonemic contrasts. All the speakers were consulted on the accuracy and acceptability of the word list before making the recording. The words were repeated both in isolation and in a frame sentence.

From these recordings we extracted a database of 119 words. Navajo has a notoriously complex verbal morphology and a rich system of inflectional prefixes (Young and Morgan 1986, 1992, Kari 1976, 1990, McDonough 1990, 1992). Phonemic contrasts are best observed in the stems. Simple nouns consist of a stem, which is usually monosyllabic, and a prefix such as /bi/ 'his/hers'. About 80% of our words were of this form. The remainder had other prefixes, or disyllabic stems, or, in a few cases, were verbal forms. All observations were made on the vowels in the stems. Tokens from the prefix 'bi' were entirely avoided because of the severe reduction of the inflectional prefix that often occurs in fluent speech. The complexity of the language made it difficult to find minimal pairs illustrating the vowels. All the vowels were in a CV context, but other aspects of the context, including tone, were not strictly controlled.

The present report is based on an analysis of 7 speakers, 5 female and 2 male, producing the citation forms of the words. 500 oral vowel tokens were measured using a CSL (Computer Speech Lab, Kay Elemetrics). The utterances were sampled at 10 kHz, and FFT and LPC spectra were computed in the middle of each vowel. For the LPC spectra, 10 and 12 coefficients were used for the female and male speakers respectively. Measurements were taken of the vowel formants and duration. In a preliminary assessment of the data, F1 and F2 means and standard deviations were calculated for each vowel. Tokens more than two standard deviations from the mean for that vowel were considered to be errors, either in measurement or in pronunciation, and were removed.

3. Vowel duration

As can be seen from the data in Table 1, Navajo speakers make very clear distinctions between long and short vowels, at least when producing citation forms. Short vowels were less than half the length of long vowels. One might suspect, therefore, that length distinctions would not be enhanced by quality distinctions. However, as we will show below, this is not the case.

Table 1. Vowel duration means and standard deviations in milliseconds.

	women	men
long	266 (90)	264 (80)
short	114 (50)	114 (50)

4. Vowel qualities.

A familiar way of observing vowel quality is by means of a formant chart. We plotted the data from the male and female speakers on separate charts, showing the first formant, F1, against (F2'-F1) the difference between the first formant and F2', a form of the second formant that takes into account the influence of the third formant, F3. Figure 1 shows the vowels of the female speakers. In order to reduce the number of symbols, the length marks have been omitted on this graph. Even so, because of the large number of vowels involved, this figure is somewhat confusing. Accordingly, Figure 2 shows just the means and an indication of the standard deviations. The ellipses enclose all data points for a given vowel that are within two standard deviations of the mean of that vowel. The axes of the ellipses are along the first two principal components of variation. The values of the means and standard deviations are listed in Table 2.

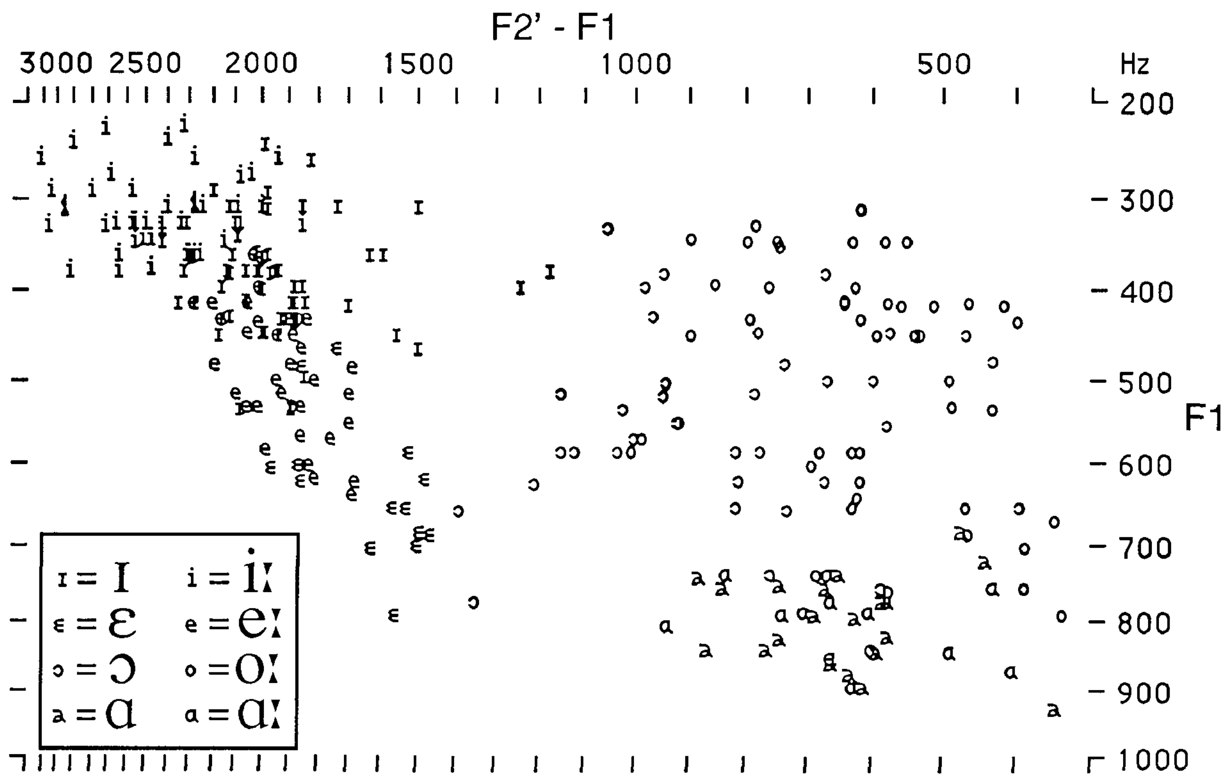


Figure 1. A formant chart of the vowel tokens produced by the 5 female Navajo speakers.

For these speakers, three of the four long / short pairs clearly involve additional differences between peripheral and interior vowels. These results confirm the tense / lax enhancement of length for all but the low vowel. As we noted above, the length differences are considerable. If this were not the case, one might even consider the quality difference as being primary for some of the vowels, with the length difference being an enhancement that could be predicted from the quality.

Table 2. Mean formant frequencies and standard deviation of the vowels of the female speakers of Navajo.

	long	short	long	short	long	short	long	short
	iː	ɪ	eː	ɛ	oː	ɔ	ɑː	ɑ
F1	315 (40)	391 (60)	498 (72)	619 (98)	488 (132)	558 (109)	802 (63)	808 (50)
F2	2528 (279)	2069 (256)	2200 (124)	2017 (149)	943 (143)	1176 (243)	1279 (122)	1299 (99)
F3	3228 (279)	2926 (207)	2829 (213)	2880 (131)	2702 (477)	2843 (374)	2660 (392)	2314 (534)
# of tokens	48	60	31	15	43	39	20	13

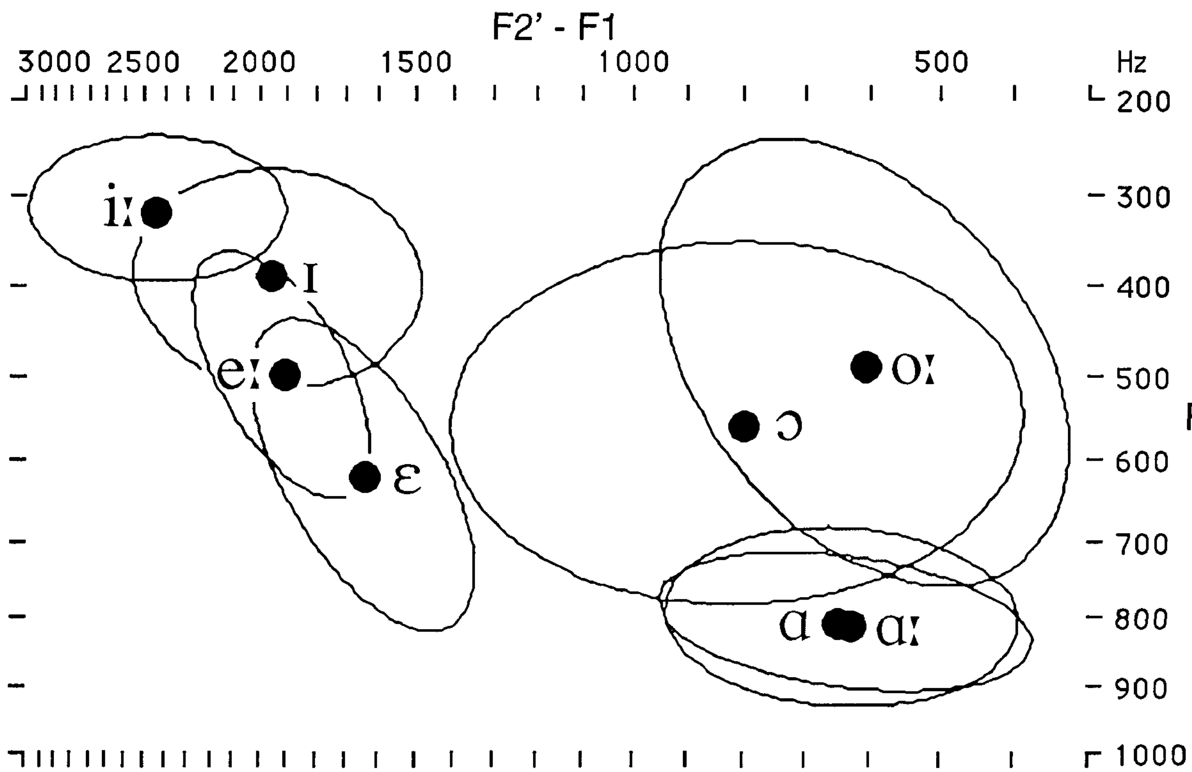


Figure 2. Mean locations of the vowels of the 5 female speakers of Navajo. The ellipses enclose areas corresponding to two standard deviations along the first two principal components of variation for each vowel.

The same plotting and statistics were done for the data on the vowels of Navajo men, but with less clear results (Figures 3 and 4). As we presently have only two male speakers, and we are dealing with a smaller numbers of tokens, we do not consider these data to be very reliable. It is not clear that they show the same general tendencies as the women's. There are still observable distinctions between the members of the peripheral / interior pairs, but all the front vowels are very close together.

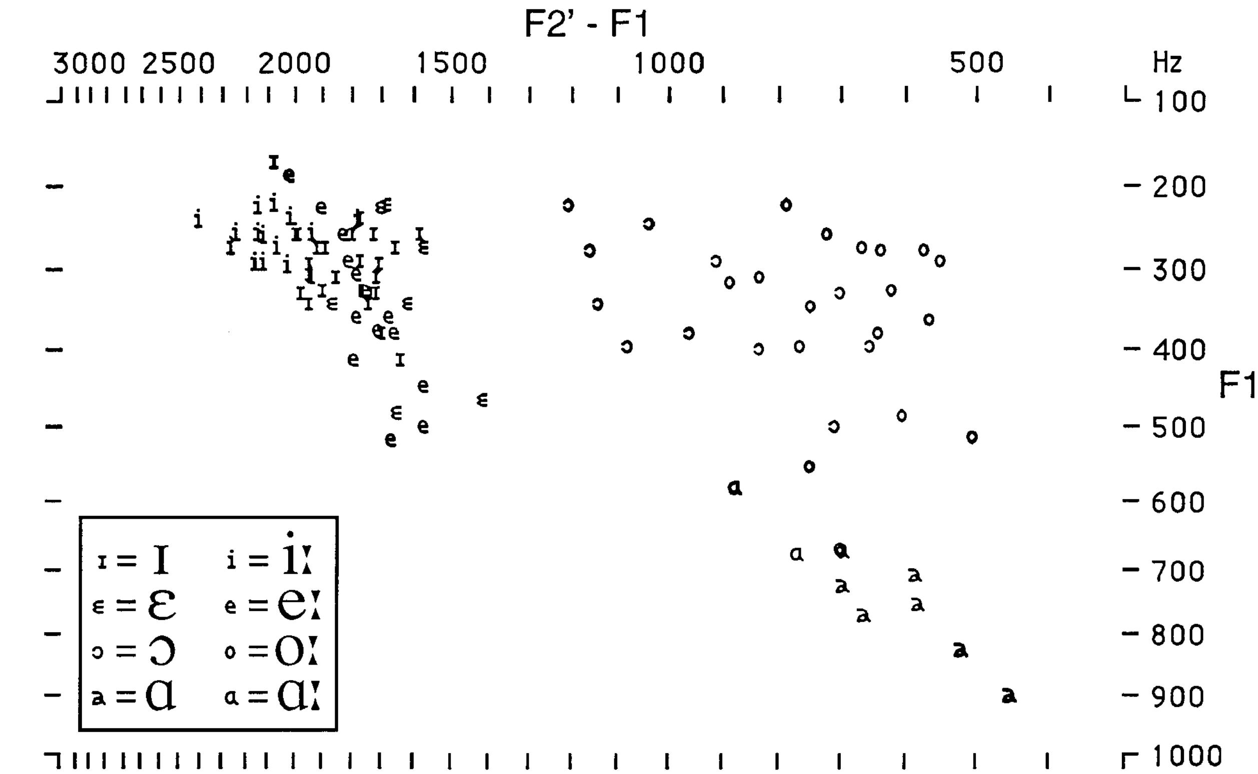


Figure 3. A formant chart of the vowel tokens produced by the 2 male Navajo speakers.

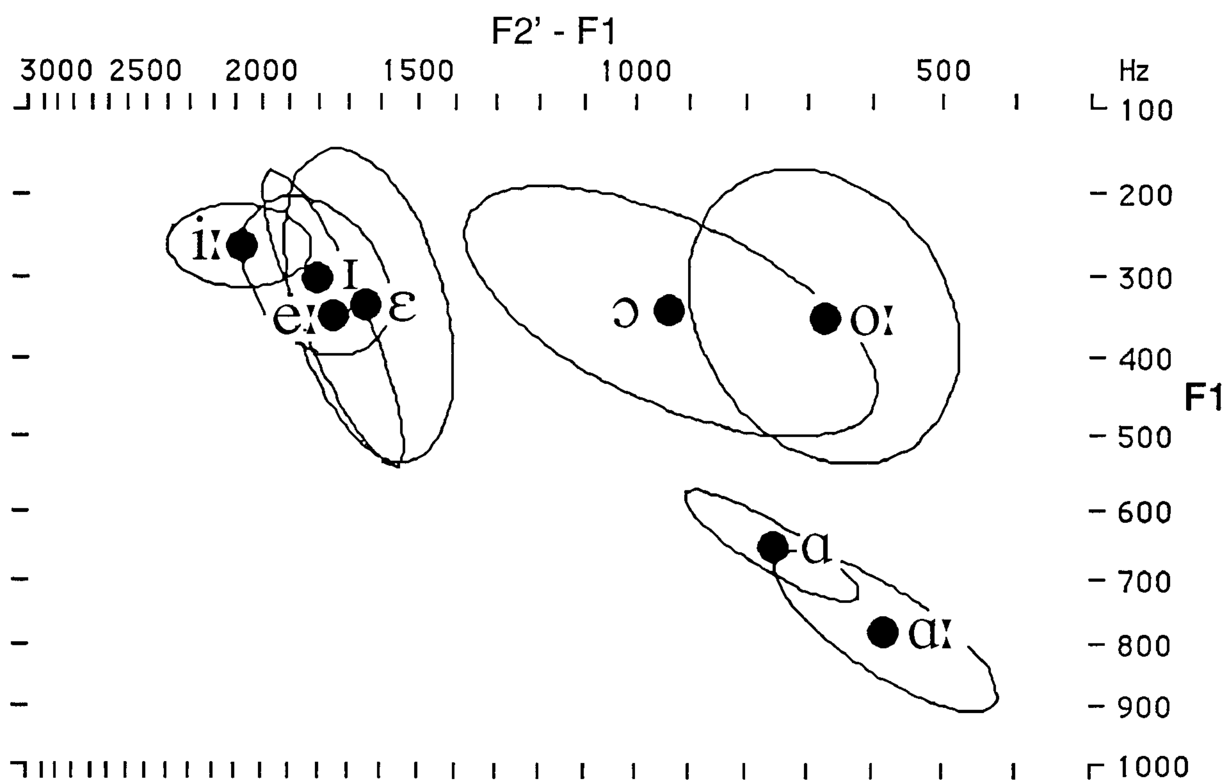


Figure 4. Mean locations of the vowels of the 2 male speakers of Navajo. The ellipses enclose areas corresponding to two standard deviations along the first two principal components of variation for each vowel.

In Navajo the length distinction in our most reliable data has become associated with a property of the segments, resulting in distinct vowel qualities for most of the long and short vowels. There are then 7 vowel qualities in Navajo. We have not yet made an acoustic analysis of the nasalized vowels of Navajo, but our auditory impression is that these vowels have similar qualities to the corresponding oral vowels. Considering the oral / nasal contrasts, but excluding tone, we may say that there are 16 vowels in Navajo, which, in a narrow IPA transcription, are as in (5). We should also note that the short mid back vowel is somewhat centralized, so that in an even narrower transcription it would be possible to transcribe it as [ɵ] instead of [ɔ].

(5) Navajo vowel contrasts

oral:							
		short				long	
	ɪ				i		
		ɛ		ɔ		eː	oː
			ɑ			ɑː	
nasal:							
		short				long	
	ɪ̃				ĩː		
		ɛ̃		ɔ̃		ẽː	õː
			ɑ̃			ɑ̃ː	

5. Dispersion and variance

The results also confirmed the spatial asymmetry of the vowels. As we can see from Figure 1, there is no high back vowel or vowel pair. It has been suggested (Maddieson 1984) that in systems like this, in accord with a theory of dispersion, compensation for lack of a high back vowel might occur in the form of the raising of the mid back vowel above the level of the mid front vowel. As we saw in Figure 2, the mid back short vowel is indeed slightly higher than the mid front short vowel, but there is virtually no difference in the height of the mid long vowels.

This skewed asymmetric system is therefore not exhibiting the full compensatory effects predicted to occur to mid back vowels. However we do find what we might consider a compensatory effect: there is considerably more variance in both of the mid back vowels than in any of the other vowels, as may be seen in Figure 2. Each of the ellipses for the back vowels encloses almost three times the area of the ellipses for any of the other vowels.

The question is — can this greatly increased variance in the back vowels be seen as the result of the expansion of these vowels into the high back area, a kind of compensatory effect? The answer we believe is no. Our belief is based on two studies: Goldstein's (1984) articulatory simulations and Keating and Huffman's (1984) study.

5.1 Skewed systems

The Navajo pattern , in which there is great variance in the back vowels, accords to a great extent with Keating and Huffman's (1983) findings in their work on Japanese vowels. Japanese is another example of a 'skewed' system, in that the high back vowel is unrounded, making F2 higher and the top right corner of a standard formant chart comparatively empty.

Keating and Huffman found that the variation in the vowels near empty spaces may be greater than the variation of other vowels,. Nevertheless, the variation did not eliminate the skewing; the high back space remained comparatively empty. This is also true of the Navajo vowel space. The high back area is avoided and the gap remains; the system maintains its skewing.

These are interesting facts in the light of the Goldstein (1984) study, which examined the relationship between random articulatory variation and the resulting acoustic pattern. Entertaining the hypothesis that random variation engenders vowel shifts, Goldstein used an articulatory synthesizer to subject vowels to random articulatory perturbations. He found that front and back vowels exhibit different acoustic patterns of variance. Front vowels are affected mainly on the F1 axis, resulting in overlapping vowel heights. But variation in the back vowels produces values of F1 and F2 that are positively correlated and do not result in overlapping vowel heights, as is the case with the front vowels. Goldstein's (1984) simulations suggest that the kind of variation we are seeing in Navajo in the back vowel is a result of simple random articulatory variation, and thus not a result of any compensatory effects.

Furthermore, an interesting hypothesis is suggested by the Goldstein simulations in conjunction with the Keating–Huffman study and this Navajo data. Given the distinction between front and back vowels in their patterns of acoustic variance under random articulatory variation — front vowel variations can result in overlapping vowel heights, whereas back ones do not — we might expect skewed systems like the Japanese and Navajo systems with high back vowel gaps to resist compensatory effects,. In other words, they would be more stable, because articulatory variation does not produce acoustic overlap in this area, and hence there are no confusions in vowel quality that would lead to changes in the system.

6. Summary

In summary in our study of Navajo vowels we have found: (1) the existence of a tense / lax distinction enhancing the length distinction for all but the low vowel; (2) the vowel space is skewed, with the asymmetry being preserved despite the variance of the tokens. The variation does not extend significantly into the high back vowel area, contrary to what we might expect under a theory of dispersion.

*This work was supported by grants from the NSF and NIH

Bibliography

Goldstein, L. (1984). Vowel shifts and articulatory-acoustic relations. In V. d. Broeke & Cohen (Eds.), *Proceedings of the 10th International Congress of Phonetic Sciences*. (pp. 267-269). Dordrecht: Foris.

Kari, J. 1990. "Affix positions and zones in the Athabaskan verb complex: Athna and Navajo " *International Journal of American Linguistics* V55.4.

Kari, J. 1976. *Navajo Verb Prefix Phonology*. Garland. New York

Keating, P. and M. Huffman. 1984. "Vowel variation in Japanese", *Phonetica 41*: 919-207.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (Eds.), *Experimental Phonology* (pp. 13-44). Orlando: Academic Press.

Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

McDonough, J. 1993. "The bi-partite structure of the Navajo verb", ms., UCLA.

Sapir, E. & H. Hoijer, 1967. *The Phonology and Morphology of the Navajo Language*. University of California Publications in Linguistics, 50.

Stevens, K. 1989. "On the quantal nature of speech", *Journal of Phonetics* 17.3-46.

Young, R & Morgan, W. 1986. *The Navajo Language*. Albuquerque: University of New Mexico.

Navajo stops

Joyce McDonough and Peter Ladefoged

Navajo consonants can be divided into two groups: (1) oral stops, with both an articulatory closure and a velic closure; the plosives, ejectives and affricates, fall into this group; (2) the remaining consonants, consisting of the nasals, fricatives and approximants. The stops are distinct from the non-stops in that many of them exhibit a three way contrast, aspirated, unaspirated and glottalized (ejective). In addition, they do not participate in the alternations in the phonology such as the infamous Athabaskan ‘d-effects’ process (Sapir 1932, Stanley 1969, Howren 1971). Nasals are included in the non-stop group and not with the stops because they share with the fricatives and sonorants the pattern of exhibiting a ‘d-effects’ alternation. This paper will be concerned with a phonetic description of the stops discussing the non-stops only in so far as they share common phonetic characteristics with the stops such as place of articulation.

So as to make this paper as useful as possible to both Navajo scholars and other linguists, we will first list the Navajo consonants in the conventional orthography, and then provide an account of the consonant system in an IPA transcription. Table 1 shows the Navajo consonants as given by Young and Morgan (1980, 1986). This is the orthography preferred by the Navajo nation.

Table 1. Navajo consonants in Young and Morgan’s orthography. The consonants to be discussed in this paper are shown in the top half of the table.

p	t, d, t’		k, g, k’	kw, gw	‘
	ts, dz, ts’	ch, j, ch’			
	tl, dl, tl’				
	s, z	sh, zh	x,gh	h	
	ł, l				
m	n				
w		y			

In this orthography the symbol ‘h’ represents both a voiceless velar fricative and a glottal fricative. The palato-alveolar fricatives, the voiced velar fricative and the affricates are represented with digraphs. Similar descriptions can be found in Sapir (1932), Sapir and Hoijer (1967), Hoijer (1963). Differences among all these descriptions are primarily in the place of articulation of velars, and in the characterization of the glides. Young and Morgan characterize the velar stops as palato-velars or ‘back palatals’ (1986:xii). The descriptions of the glides vary, depending on different phonological considerations that will not be considered here.

The complete set of Navajo consonants is given in Tables 2 and 3 in an IPA transcription. Table 2 shows the Navajo stop consonants, and Table 3 shows the remaining Navajo consonants.

Table 2. Navajo stops in an IPA transcription.

	labial	alveolar	palato-alveolar	velar	labialized velar	glottal
plain	p	t^{h} t t'		k^{h} k k'	k^{hw} k^{w}	ʔ
laterally released		t^{ɬ} t^{l} t^{ɬ}'				
affricated		ts^{h} ts ts'	tʃ^{h} tʃ tʃ'			

Table 3. Navajo consonants other than stops in an IPA transcription.

	labial	alveolar	palato-alveolar	palatal	velar	labio-velar	glottal
fricative		s z	ʃ ʒ	ç	x ɣ	ɣ^{w}	h
nasal	m	n					
approxi-mants	w			j	ɰ		

Stop Contrasts

The stops of Navajo contrast in place of articulation, as can be seen in Table 1. The place of articulation distinction has less functional load than might appear from the table. The labial consonants /p/ and /m/ are uncommon, appearing in only a handful of morphs. Outside of these rare labials, there are only two primary places of articulation in Navajo: coronal and dorsal (velar). Each of these may be further subdivided into two possibilities, the coronal stops being either anterior or non-anterior, and the velar stops being either labialized or not. This type of place of articulation distinction is characteristic of Athabaskan languages.

At each primary place of articulation there are contrasts in the manner of release. There are three possibilities: a regular stop release, as in [t] or [k], a fricated release, as in [ts] and [tʃ], and a lateral release, as in [tɬ]. Both the latter possibilities will be considered to be affricates. All the affricates are coronal in place of articulation. The centrally released affricates contrast in anteriority, being either alveolar or palato-alveolar. The lateral affricates are alveolar.

Voicing is not significant among Navajo stops, but there are nevertheless three types of contrast that depend on laryngeal action. The first two types depend on differences in the voice onset time (VOT); stops can be aspirated, or unaspirated. The aspirated stops have very long periods of voicelessness after the release burst, and are characterized by noisiness. Observations of heaviness or frication in the aspirated period of the consonants have been made by most of the linguists working on Navajo cited above. The other laryngeal distinction is often termed 'glottalization'. The glottalized plosives are ejectives, that is they differ from the non-glottalized ones in the airstream mechanism employed in their production, glottalic vs. pulmonic. Ejective sounds are produced by making both an articulatory and a velic closure, while at the same time forming a glottal stop. While all these closures are maintained, the closed glottis is raised, compressing the air in the oral cavity. The articulatory closure is released shortly before the release of the glottal stop. This is a particular type of consonant and differs from the sequence glottal stop + stop, in which the glottal stop is not simultaneous with the release of the oral stop.

The glottalized consonants in Navajo are all ejectives, except for the glottalized reflexes of verb stem initial nasals and glides (Hoijer 1967), which involve the superimposition of a glottal stop or creaky voice on a voiced sound. This pattern is true of the Athabaskan family as a whole, and in contrast with the pattern in the neighboring families in the Northwest, Salishian and Wakashian, which also have glottalized fricatives and sonorants.

The Navajo stops are divided into two groups: affricates and non-affricates. Table 4 shows the stop contrasts dependent on laryngeal action for both groups. The bilabial /p/ and the glottal stop /ʔ/ are omitted from this table, as is the labialized velar stop. These segments do not participate in the three way contrasts shown in Table 4. As may be seen, only the alveolar stops have the complete set of three release possibilities (shown in rows 1,3 and 5), and three laryngeal possibilities (shown in the three columns), giving a total of nine contrasts.

Table 4. Contrasts among Navajo stops that are dependent on laryngeal action.

Unaffricated stops	aspirated	unaspirated	glottalized
1 coronal	t^{h}	t	t'
2 velar	k^{h}	k	k'
Affricated stops			
3 alveolar	ts^{h}	ts	ts'
4 palato-alveolar	tʃ^{h}	tʃ	tʃ'
5 lateral	tɬ^{h}	tɬ	tɬ'

Young and Morgan (1980, 1986) differ from the analysis given here in that they group the aspirated coronal stop [t^{h}] with the affricates rather than with the stops. We will discuss the phonetic characteristics of the aspirated coronal stop below.

Phonetic analysis procedures

In order to examine the phonetic nature of the contrasts, a set of recordings was made in a classroom at Navajo Community College, Shiprock, New Mexico, in April 1992. The speakers were educated native speakers of Navajo, who live on the reservation and use Navajo daily. They were recorded while repeating a word list that illustrated the phonemic contrasts. All the speakers were consulted on the accuracy and acceptability of the word list before making the recording. The words were repeated both in isolation and in a frame sentence.

From these recordings a database of 119 words was extracted . Navajo has a notoriously complex verbal morphology and a rich system of inflectional prefixes (Young and Morgan 1986, 1992, Kari 1976, 1990, McDonough 1990, 1992). Phonemic contrasts are best observed in the stems. Simple nouns consist of a stem, which is usually monosyllabic, and a prefix such as /bi/ 'his/hers'. About 80% of the words were of this form. The remainder had other prefixes, or disyllabic stems, or, in a few cases, were verbal forms.

The present report is based on an analysis of 7 speakers, 5 female and 2 male, producing the citation forms of the words. Tokens were measured using a Computer Speech Lab (Kay Elemetrics CSL). The utterances were sampled at 10 kHz, and displayed so that duration could be measured on wide band spectrograms, as illustrated in the figures. For the greater part, the stops were measured intervocalically, in a VCV context. The offset of the vowel, the release of the articulation, and the onset of the formant structure for the following vowel are all well determined, so that measuring the stop closure duration and the Voice Onset Time was straightforward. We measured VOT as the interval between the release of the articulatory stop closure and the onset of the voicing of the following vowel, as shown in Figure 1. This measure was used for both plain and affricated stops, and thus includes any voiceless central or lateral fricative portion of the stop release.

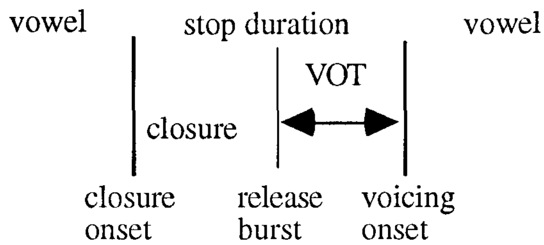


Figure 1. The divisions used in measurements of closure and VOT. The stop duration is taken to be the sum of these two measurements.

Duration measurements

The VOT durations for the unaffricated stops are given in Table 5. The unaspirated bilabial and coronal stops have VOT's that are basically the size of the release burst.; the VOT of the velar unaspirated stop is a little longer. But, despite the orthographic representation of these stops as 'b,d,g', it is clear that they are voiceless unaspirated plosives, and not voiced stops. Not surprisingly, the aspirated stops have long VOT's that are statistically very distinct from the unaspirated stops (t = 10.62, p .0001). The glottalized stops are also distinct from the unaspirated stops (t = 6.94, p .0001). There is a smaller, but still significant difference, (t = 3.36, p .002) between the VOT's of the aspirated and the ejective stops, the ejective VOT's being slightly shorter.

Table 5. Mean VOT of Navajo unaffricated stops in ms. The standard deviations are given in parentheses.

Place of articulation	unaspirated	aspirated	glottalized
bilabial	12 (5)	— — —	— — —
coronal	6 (2)	130 (29)	108 (31)
velar	45 (9)	154 (43)	94 (21)

The major point to note concerning the aspirated and glottalized stops is the magnitude of the VOT measures: they are longer than comparable consonants in other languages. The duration of the aspiration is more than double that reported by Klatt (1975) for English monosyllables. The VOT for the velar ejective is also more than double that reported by Lindau (1984) for the comparable stop in Hausa, and the coronal ejective is even longer. As will become evident from all the duration measurements to be reported here, Navajo stops (and, we believe, other consonants, although we so far have measurements only for the stops) are on the whole much longer than similar sounds in other languages.

Navajo follows the universal phonetic tendency for velar stops to have a longer VOT than stops made at more anterior positions. The unaspirated velars are longer than the unaspirated coronals by about 40 ms, a highly significant difference (t = 9.87, p .0001). The difference in VOT between the aspirated velars and coronals is in the same direction, but is not significant. The unaspirated bilabials are only 6 ms longer than the unaspirated coronals. This difference is also not significant.

Table 6. Mean VOT of Navajo affricated stops in ms. The standard deviations are given in parentheses.

Place of articulation	unaspirated	aspirated	glottalized
alveolar	91 (20)	149 (12)	142 (41)
palato-alveolar	68 (19)	153 (17)	144 (24)
lateral	42 (20)	149 (29)	157 (40)

VOT measurements for the affricated stops are shown in Table 6. Again, the main point to observe in this table is that many of these releases are very long. There are also a number of particular points of interest. In the alveolar and palato-alveolar affricates by far the majority of the fricative release is voiceless. As we noted above, these measurements are from the stop release to the onset of voicing, so these measurements include the central or lateral fricative. An alternative definition of VOT might have entailed our measuring from the release of the fricative to the onset of voicing, but we found it difficult to make reliable determinations of the moment when the fricative articulation could be said to be released. The fricative energy usually faded away as the vowel began. As in the unaffricated stops, the aspirated and glottalized stops are statistically very distinct from the unaspirated stops (for the combined group $t = 7.24$, $p < .0001$). But in the case of these stops there is no difference between the VOT's of the aspirated and the ejective stops.

The laterally released stops are somewhat different. In the unaspirated lateral plosive the voicing starts during the lateral, making these sounds statistically very different ($t = 3.62$, $p < .002$) from the other unaspirated stops in Table 6. As we will discuss below, it is arguable that these sounds are not affricates, and hence belong in Table 4, in which the unaspirated stops have comparable VOT's. The aspirated and glottalized lateral stops are, however, clearly affricates. Their VOT's are not distinguishable from those of the other aspirated and glottalized sounds in Table 6.

Table 7. Mean closure duration of unaffricated and affricated stops in ms. The standard deviations are shown in parentheses.

Closure -stops	unaspirated	aspirated	glottalized
coronal	162 (22)	118 (11)	149 (41)
velar	102 (22)	97 (15)	76 (29)
bilabial	154 (18)	------------	-------------
Closure -affricates			
alveolar	98 (15)	62 (5)	121 (14)
palato-alveolar	101 (21)	84 (16)	98 (17)
lateral	147 (42)	112 (15)	119 (33)

Table 7 shows the closure durations of both the unaffricated and the affricated stops. The unaspirated stops have longer closures (mean 132 ms, S.D. 36.5) than the aspirated stops (mean 96 ms, S.D. 23.0). This is a highly significant difference ($t = 4.867$, $p = .0001$). The decrease in closure duration when aspiration is added does not result in the total length of the consonant remaining the same, regardless of whether it is voiced or voiceless; but there is a significant move in that direction. The glottalized stops are not significantly different in closure duration from either of the other two types of stops. In addition, we should note that many of these closure durations are very long. Young and Morgan (1986:xv) have remarked that the consonants of Navajo are doubled between vowels; but it is the case that consonant gemination is not contrastive, i.e., the stops act like single segments. Furthermore consonants in general are long, not just intervocalic ones.

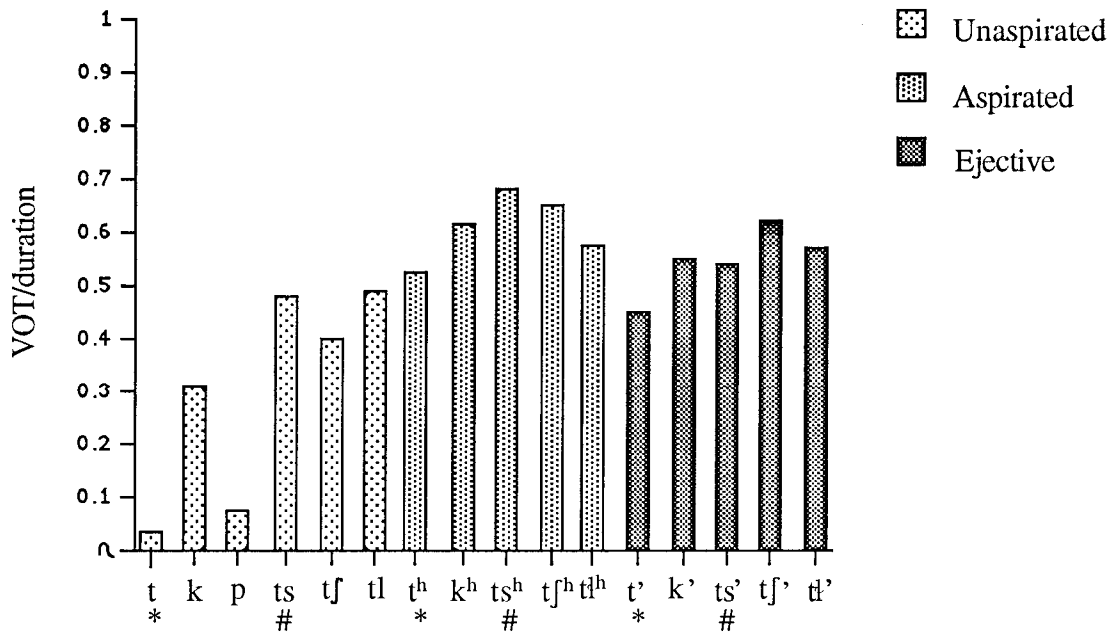


Figure 2: Bar graph of the ratio between the VOT and the entire consonant duration (closure + release, until the onset of the following vowel) of the Navajo stops

Figure 2 illustrates another way of considering consonantal length in Navajo by showing the ratio between the VOT and the length of the stop as a whole, defined as in Figure 1. For all the stops except the unaspirated stops [p, t, k], the VOT makes up around half of the duration of the consonant. Note also the difference in the bar graph between the unaffricated coronals [t, t^h,t'], marked by *, and the affricated coronals [ts, ts^h, ts'], marked by #. There is a sharp difference in VOT between the unaspirated coronal stop [t] and the aspirated and ejective coronal stops [t^h, t']. But in the affricated series, the difference between [ts] and [ts^h, ts'] is smaller. This difference between the two series may be influenced by the fact that the affricate measurement includes the fricated release. We have noted that we cannot reliably measure the aspiration, that is to say the voiceless interval in which there is only glottal friction, between the release of the consonant, be it stop or fricative, and the onset of the vowel. If we could we might be able to show that there is a difference in aspiration between the affricated and the unaffricated coronals: if there is a fricative release, then there is less aspiration.

Considering all the duration measurements, it appears that the aspirated and ejective stops pattern together on the one hand and the unaspirated stops on the other. The aspirated and ejective stops have a long VOT of similar duration, while the unaspirated ones are very short. The VOT for all but the unaspirated stops makes up around half of the intervocalic consonant durations in Navajo.

Spectral characteristics

In this section we will present illustrations of our findings concerning those acoustic aspects of Navajo stops that are apparent in spectrographic records. We have not yet quantified many of these findings, and do not know how reliable they are. The following account should be taken as a qualitative preliminary description rather than a statistically valid account of the spectral properties of Navajo stops.

In the spectrogram of the voiceless unaspirated [t] shown on the left of Figure 3, there is a burst at the stop release, with virtually no subsequent fricative component. This is as we would expect, given the VOT measures reported above. The principal energy in the burst is in the higher part of the spectrum around 7-8,000 Hz. In the unaspirated velar stop [k], illustrated in the second word in this figure, there is a longer VOT and fricative energy is present until the onset of voicing in the vowel. The coming together of the second and third formants as the stop closure is formed, and the release burst with energy around 3,000 Hz are typical of a velar stop before a high front vowel. The length of both the intervocalic and the final consonants are also apparent in these spectrograms; all these consonants are over twice the length of any of the vowels.

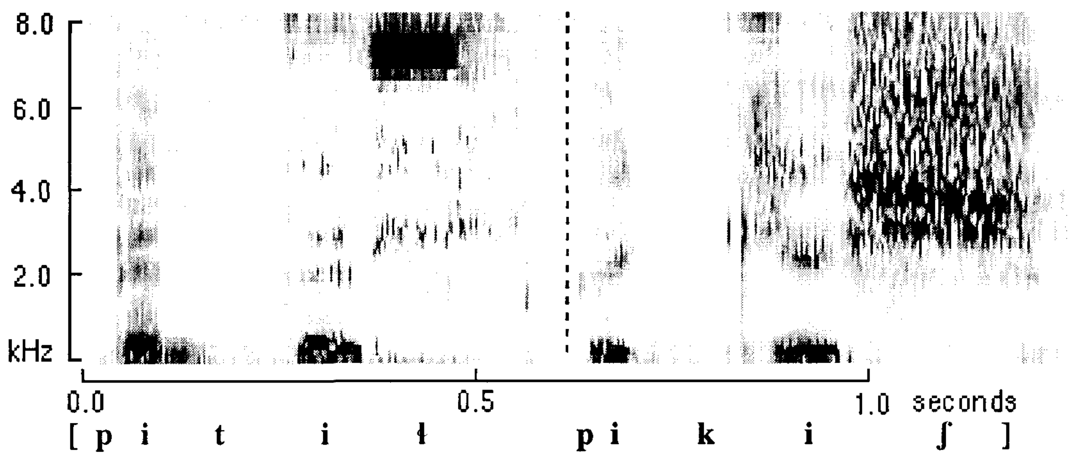


Figure 3. The unaspirated coronal [t] and velar stop [k] in [pitil] 'bidil' (his blood) and [pikiʃ] 'bigish' (his digging stick).

The spectrograms of the unaspirated coronal affricates [ts] and [tʃ] in Figure 4 can be clearly divided into two parts, the stop closure and the fricative release. These are affricated but not aspirated stops. There is no period of voiceless aspiration after the friction; the onset of voicing and vowel formants follow immediately after the fricative. This pattern of close association between the onset of voicing and the onset of the vowel formants is characteristic of all the Navajo unaspirated consonants, except for the lateral affricates which we will discuss below.

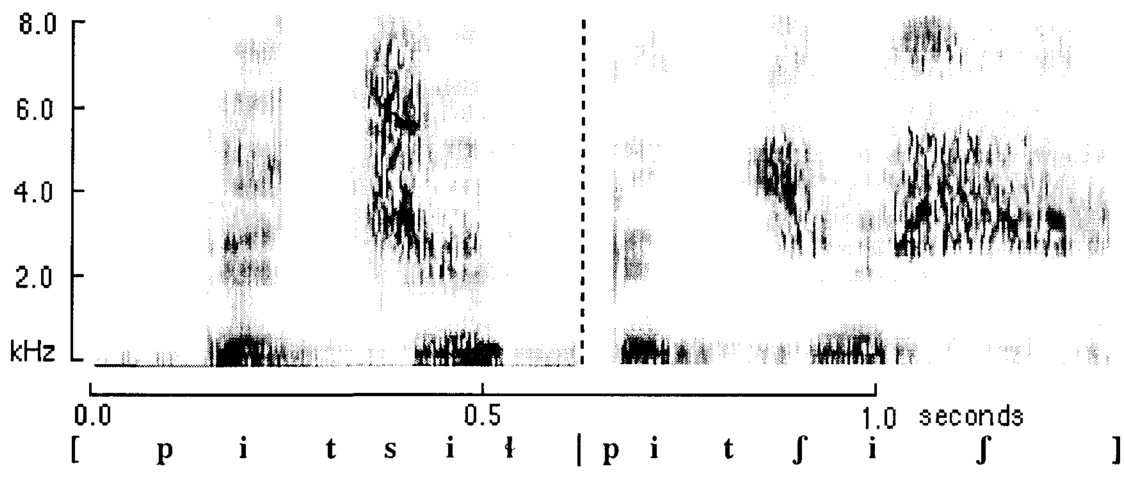


Figure 4. The unaspirated affricates[ts] and [tʃ] in [pitsɨɬ] 'bidzɨɬ' (his mountain) and [pitʃiʃ] 'bijish' (some day).

The palato-alveolar [tʃ] differs from the alveolar [ts] in both the location and the intensity of the energy in the fricative release. The [ts] has more energy in the higher frequencies and greater intensity than the [tʃ] in keeping with the typical difference between palato-alveolar and alveolar fricatives.

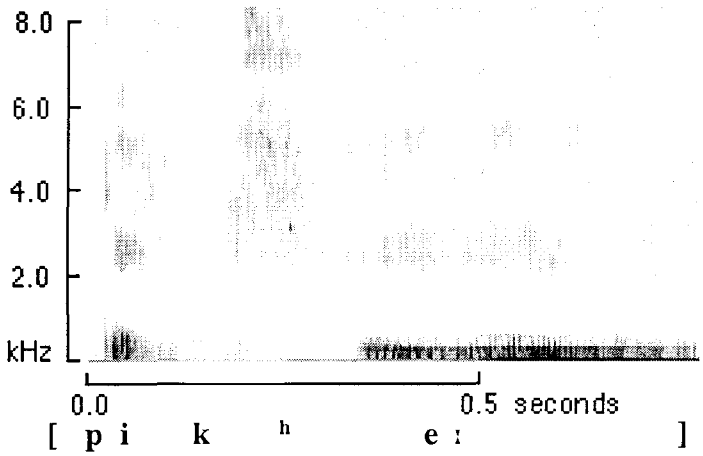


Figure 5. The aspirated velar stop [k^h] in [bik^heː] 'bikee' (his foot, toe).

There is also a great deal of fricative energy in the aspirated stops. As illustrated in Figure 5, the aspirated velar stop [k^h] is realized as a velar affricate which might well have been symbolized [kx]. The fricative intensity increases somewhat towards the middle of the voiceless interval and then diminishes slightly, but there is no interval in which there is no fricative noise and only voiceless aspiration. The fricative energy is centered in the same region as that in the burst at the stop release.

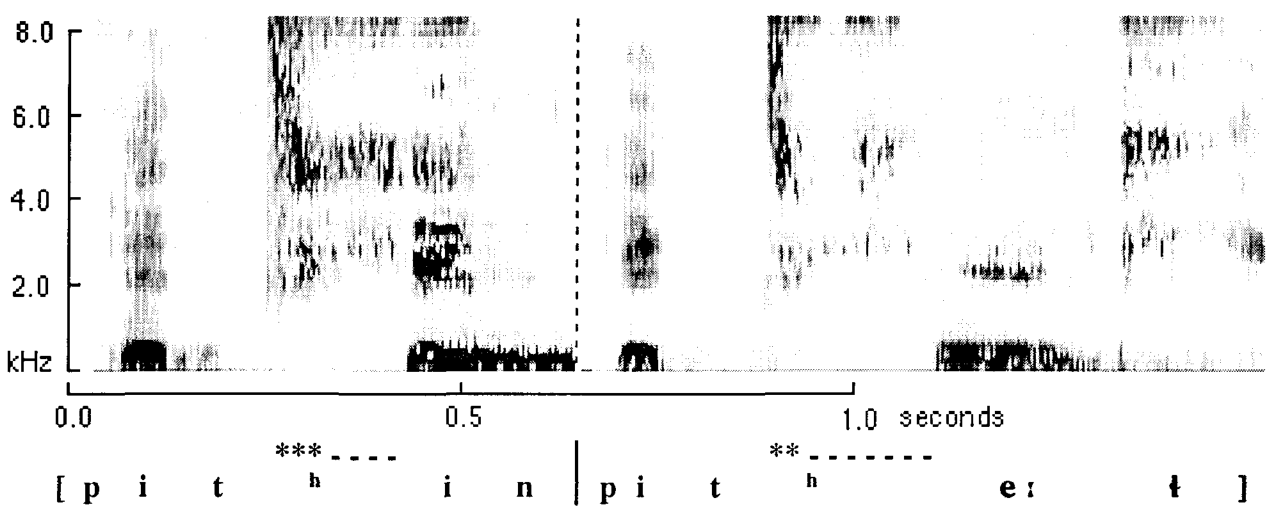


Figure 6. The aspirated coronal stop [t^h] before front vowels in [pit^hin] 'bitin' (his ice) and [pit^heːɬ] 'biteeł' (his cat tail).

The aspiration in the coronal stop [t^h] is more unusual. It typically consists of two parts, as illustrated in the spectrograms in Figures 6 and 7. Immediately after the release of the stop closure there is a burst with energy in the higher part of the spectrum, as appropriate for an alveolar stop. This section is marked by asterisks in Figures 6 and 7. After this there is a much longer section, marked with dashed lines in the figures, in which the greatest energy is in a lower spectral region. This is a velar fricative comparable to the one that occurs in the release of the aspirated velar stop [k^{h}], as in Figure 5. The exact quality of this fricative depends slightly on the quality of the following vowel. In Figure 6, in which the aspirated coronal stops are before front vowels, there is a band of energy at around 3,000 Hz, but also slightly more energy between 4,000 and 6,000 Hz, indicating a fronted velar fricative. In Figure 7, which illustrates the same stop before back vowels, the spectral peaks are at a somewhat lower frequency, indicating more purely velar sounds.

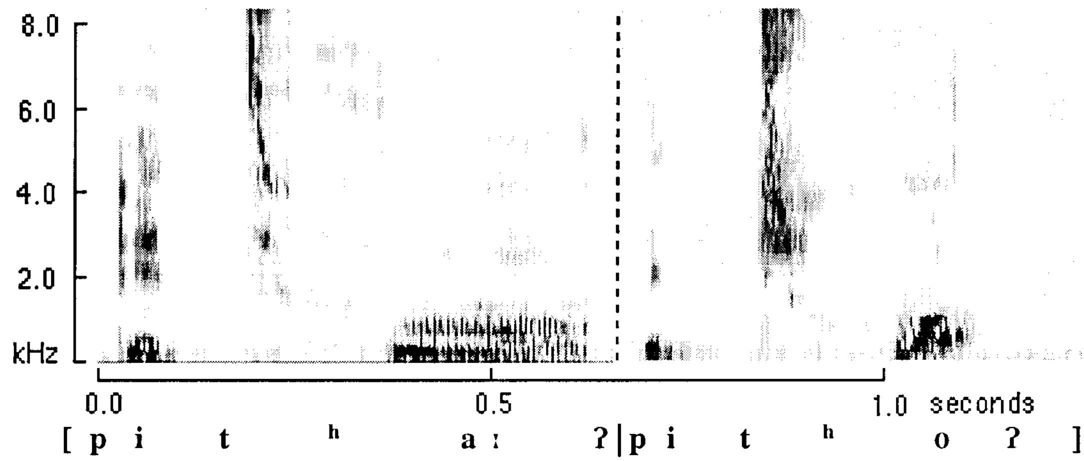


Figure 7. The aspirated coronal stop [t^{h}] before back vowels in [pit^{h}aːʔ] 'bitaa' ' (his father) and [pit^{h}oʔ] 'bito' ' (his water).

The notion that the period of aspiration is fricated, or 'strong' is in keeping with the observations of Hale (1972) and Young and Morgan (1980, 1986). In their phonemic inventory, Young and Morgan (1986:xv) include the aspirated coronal stop with the affricates, i.e. the stops with fricated releases, and not with the plain stops. Although it is easy to see why this has been done, we will regard [t^{h}] as belonging in the same series with the aspirated velar stops, partly because this increases the symmetry in the phonological descriptions of the consonants (although nothing depends on this from the point of view of phonological rules), and partly because we consider affricates to be sequences of stops + homorganic fricatives. It is an arbitrary decision.

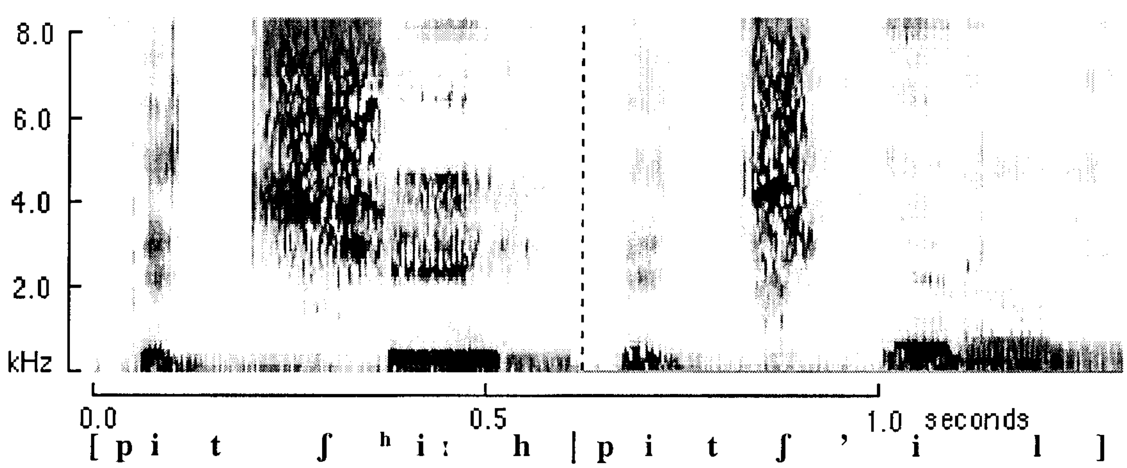


Figure 8. The aspirated coronal affricate [tʃ^{h}] in [pitʃ^{h}iːh] 'bichiih' (his red ochre) and the ejective coronal affricate [tʃ'] in [pitʃ'il] 'bich'il' (his plant)

The aspirated affricates [ts^h,tʃ^h] are interesting in that they, too, do not have much in the way of a period of voiceless aspiration; there is only a very brief gap of 10-20 ms after the fricative and before the onset of the vowel, as may be seen in [tʃ^h] in [pitʃ^{h}iːh] in Figure 8. This pattern contrasts with the unaspirated affricates which do not have any such gap; as we saw in Figure 4, the voicing for the vowel starts at the very end of the fricative. In addition to the small gap, another difference between the aspirated and unaspirated affricates is that the fricative portion is considerably longer. What we are calling a difference between aspirated and unaspirated affricates is realized principally in this extra length of the fricative portion. Both these differences, the small gap between the fricative and the vowel, and the extra length of the fricative, are consistent across subjects and across tokens in our data.

Figure 8 also illustrates the ejective coronal affricate [tʃ'] in [ptʃ'il]. The ejective affricates are different from the plain stops because of the mechanics of their release. Ejectives have two closures, oral and glottal. The closures are made approximately simultaneously, but the release of the oral closure occurs well before the glottal release. In non-ejective affricates, the frication has a pulmonic source. By contrast, for ejective affricates, this frication does not have a pulmonic source because the glottis is closed. Any frication must be the result of the movement of air trapped above the glottis. After the release of the stop closure, the air is pushed out through the fricative constriction by raising the glottis. As Navajo ejectives have a rather long period between the oral release and the release of the glottal closure, given the limitations of volume it is not surprising that only part of this period is fully fricated.

As we saw in Table 6, there is no significant difference in VOT between the aspirated stops and the ejectives. This point is illustrated by the words in Figure 8; the VOT is almost exactly the same in the two tokens. In general, the difference between [ts^h,tʃ^h] and [ts',tʃ'] lies in part in the extent to which the VOT interval is filled with fricative noise, and in part in the glottal onset to the vowel.

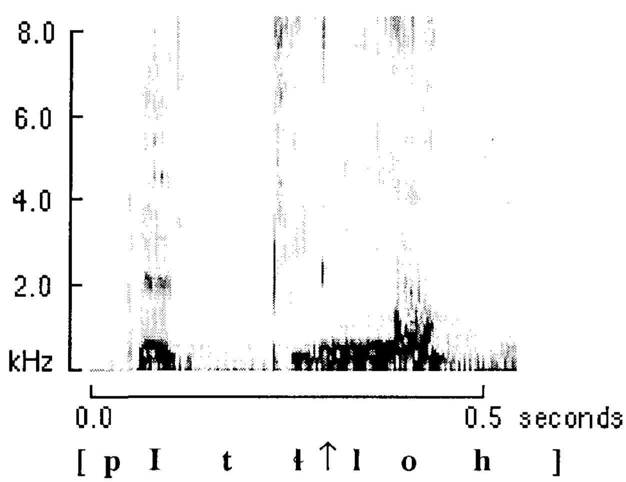


Figure 9. The voiceless unaspirated lateral affricate in [pɪtɬoh] 'bidloh' (his laughter). The transcription under the spectrogram distinguishes between the voiceless [ɬ] and the voiced [l] parts of the lateral release, with the onset of voicing being marked by an upward arrow.

The lateral affricates pattern with stops and affricates in their three way laryngeal contrasts, aspirated, unaspirated and glottalized. However they have some interesting properties that distinguish them from the other stops. The illustration of the voiceless unaspirated lateral affricate in Figure 9 makes it apparent that the voicing starts about one third the way through the lateral. As we noted in the discussion of Table 6, the VOT in these sounds is much shorter than in the other sounds we have been considering as affricates, and is actually less than in the unaspirated velar plosive. At the time we raised the question of whether these sounds should be considered as affricates. From spectrograms such as that exemplified in Figure 9, it appears that the first part of the sound has fricative energy, but the last part, in which there is voicing, does not, so that it would be quite reasonable to regard this sound as [dl].

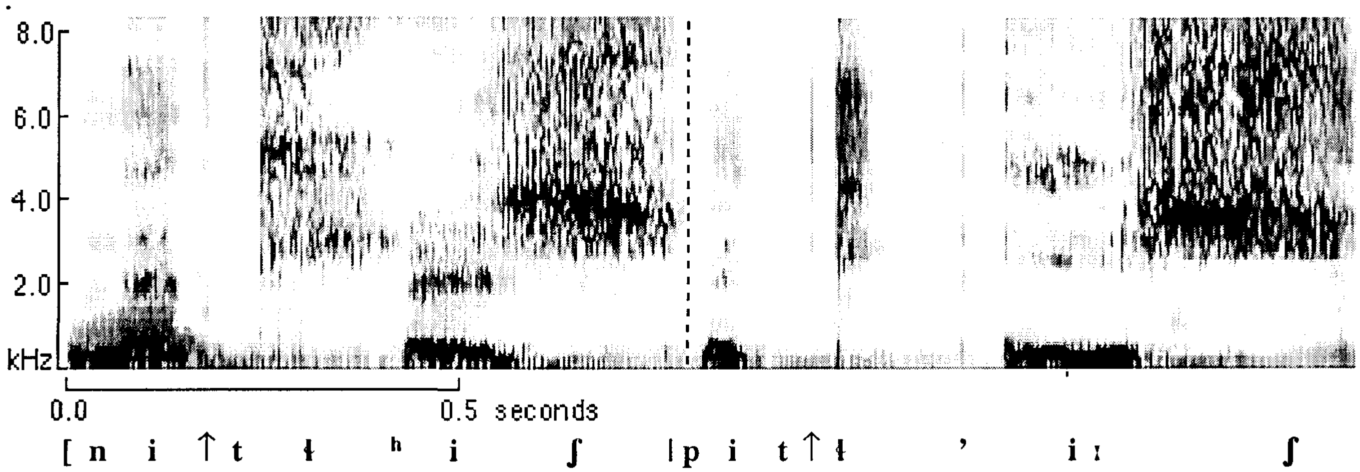


Figure 10. The aspirated lateral affricate [tɬʰ] in [nitɬiʃ] 'nitlish' (it has arrived) and the ejective coronal affricate [tɬ'] in [pitɬ'iːɬ] 'bitl'iish' (his snake). (See text for discussion of the arrows.)

The spectrograms in Figure 10 make it clear that the other laterally released stops are affricates. There is a long fricative portion in the aspirated lateral stop [tɬʰ], with only a very small gap for the aspiration. Both the increase in length of the fricative and the small gap for the aspiration are similar to those in the other aspirated affricates, such as the palato-alveolar aspirated fricative [tʃʰ] illustrated in Figure 8.

The ejective lateral affricates have a noticeable but short fricative period, as the source of frication is necessarily supraglottal. The shortness of this period in the lateral ejectives is consistent with the general pattern of frication in ejectives in Navajo in which the duration of the frication is governed by the amount of air trapped above the glottis.

One feature of the lateral ejectives did not occur in the other ejectives. The glottal closure was often released and then re-made, producing a spike on the spectrogram, as immediately above the ['] symbol in Figure 10. There was then a short interval in which there was a glottal stop which was finally released as the voicing in the vowel began. For some speakers this interval was filled with two or three pulses of a creaky voice phonation. This kind of glottal action never occurred in the other ejectives, in all of which, after the glottal closure was released, there was an immediate sharp onset of the vowel formants as voicing began.

(Objectivity in reporting our results requires us to note that there is a mystery in some of these laterally released stops. Quite consistently, for two of our speakers, one of whom produced the tokens illustrated in Figure 10, there is a spike on the spectrograms at the positions shown by the arrows in the figure. This is apparently during the alveolar stop closure. We must await further aerodynamic investigations before we can be sure of our interpretation of this acoustic event.)

Navajo has an interesting intensifier, known as the augmentive, which is written as secondary velar frication on a stop or affricate by Young and Morgan (1980). Examples of its use can be found in the words [litso] 'litso' (yellow) and in the augmentive form, an intense yellow, [litsˣo] 'litsxo' (orange). Spectrograms of these two words are shown in Figure 11. The major difference between these two tokens is in the addition of the velar articulation after the alveolar fricative [s].

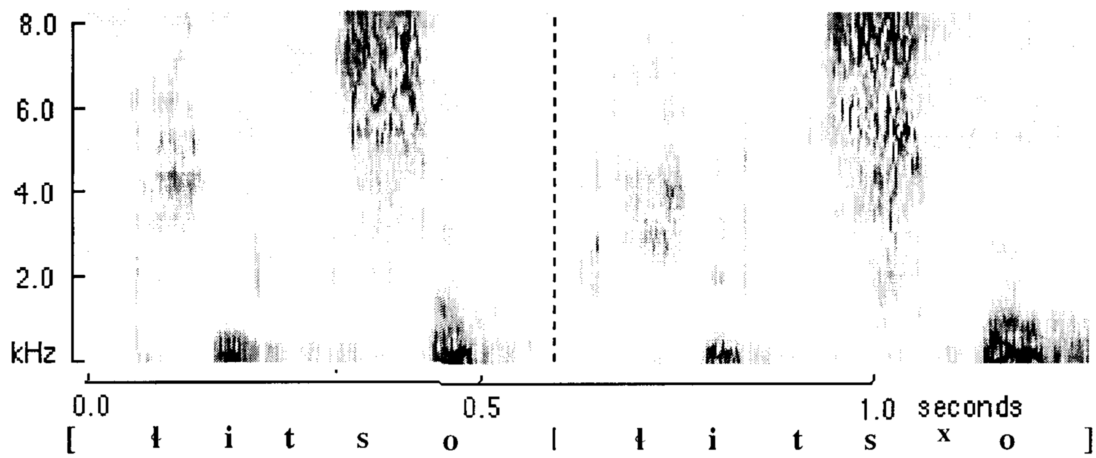


Figure 11. The coronal affricate [ts] in [litso] 'litso' (yellow) and in the augmentive form in [lits^{x}o] 'litsxo' (orange).

While it is clear that the augmentive involves the imposition of velar articulation on the segment, there was some variation in the actual realization of this articulation. In some speakers a distinct velar stop could be heard during the frication of the affricate. In the second word in the figure spike bursts can be seen in the period near the end of frication. This speaker had a perceivable velar stop imposed on the segment, a kind of [tsk^{x}]. The augmentive always involves the imposition of a velar articulation on a segment, but there is some speaker variation as to whether the velar is a fricative or stop.

Young and Morgan report two labialized velar consonants 'gw' and 'kw', which in IPA transcription would be an unaspirated [k^{w}] and an aspirated [k^{ʍ}] (see Figure 1). They suggest that the labialized aspirated velar stop is "comparable to the 'qu' of English 'quick'..." (1986:xiv #6). Our data suggests that there are some differences. We recorded two items containing the aspirated velar stop, both before front vowels, [k^{ʍ}iː],'kwii', (here) and [k^{ʍ}eː] 'kwee' (there). As may be seen in the spectrogram of the first of these words in Figure 12, there is a double burst of the kind that is typical of velar stops. This is followed by a comparatively steady state high back rounded approximant. After this approximant, there is a sharp rise to F3, very similar to that found in the English 'w'. All our speakers had this steady state voiceless portion during the release of the stop. The comparable English sound has no steady state portion. The labialized aspirated velar stop has a VOT of 145 ms (S.D. 39) in initial position (as it is in all our recordings).

As noted, the labialized velar does not have a broad distribution. Related to this fact is the quasi-phonemic existence of a labialized glide 'w' (Young and Morgan 1986:xv, #33). Recall that the Navajo vowel system is skewed; there is no high back vowel, though one is posited in the proto-language (Leer 1981). If there is a historically similar source for both sounds in the de-syllabification of [u] (Maddieson 1984, Ohala and Lorenz 1977), the lack of this vowel in the present system may account for the marginal status of these segments.

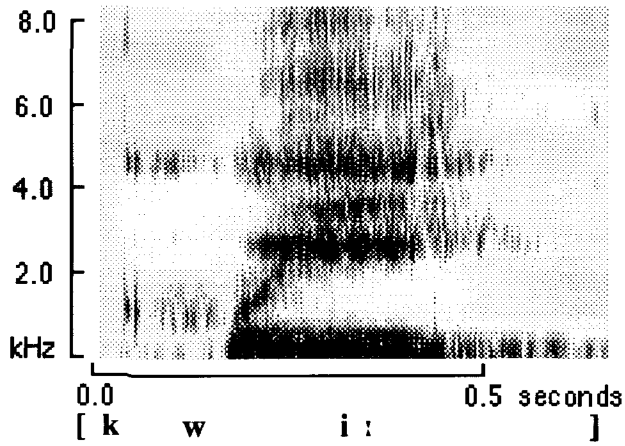


Figure 12. The labialized velar fricative in [kwiː] 'kwii' (here).

Conclusion

The stops in Navajo show no alternations in the phonology. The phonological processes that effect alternations among consonants in Navajo for the most part involve voicing and addition of closure (d-effects) to segments. As the stops already have closure they can participate in this process only vacuously. There are also sets of glide/fricative alternations (McDonough, forthcoming) that are not relevant to the stops.

Phonetically, the stops are characterized primarily by two things: their very long durations and the 'heavy' aspiration and frication, that is, noise in the low frequencies. The source of this noise is either the glottis or a constriction in the vocal tract at some place other than the primary place of articulation. For all but the plain unaspirated stops, the VOT makes up around half the duration of the consonants. The result of a long VOT in ejectives means that there is a long period between the release of the oral gesture and the release of the glottis, in marked contrast to ejectives found in other language families where the glottal release follows more directly after the oral release. This also means that the frication in ejectives is completed well before the release of the glottis. Thus the events of the VOT are means of both marking the contrasts among stops and of being a distinctive acoustic property of Navajo sound system in general.

The spectrograms that have been presented in Figure 4 - 12 provide a good overview of the major characteristics of Navajo stops. On looking through all these illustrations, the overwhelming impression is of the extraordinary length of the consonants, particularly when they are compared with the lengths of the vowels, which are no longer than they would be in citation forms of bisyllabic English words. The consonantal length is a particular phonetic fact about Navajo which must be recognized in any complete account of the language.

The system of contrasts in Navajo represented by Table 1 is very characteristic of the Athabaskan family as a whole: a phonologically inactive series of stops and affricates contrasting with the more active fricatives in laryngeal specifications, and a limited set of place specifications among the consonants. The stops contrast in the release mechanism employed, plain, fricated and lateral. Some of the features of the stops, such as the velar fricative accompanying the plain aspirated stops, have been recorded for other languages in the family (Hoijer 1943, Hoijer and Jeel 1962) and are apparently a quite systematic and persistent part of the grammar.

Acknowledgments

Special thanks to the staff and faculty of the Navajo Community College for their cooperation and interest. Particular thanks to Clay Slate, the director of the college, and Martha Austin-Garrison. Thanks also to Ian Maddieson, Richard Wright, Helen George, and others in the Phonetics lab for their valuable discussion of earlier versions of this work. Besides funding by the NSF, this work supported additionally by the NIDCD under grant number TC-32-00029.

Bibliography

Abramson, A. (1977). Laryngeal timing in consonant distinctions. *Phonetica, 34*, 295-303.

Hale, K. (1972) *Navajo Linguistics*. Ms. MIT.

Ladefoged, P. (1993). *A Course in Phonetics* (3rd. ed.). New York: Harcourt Brace Jovanovich.

Leer, J. (1979). *Proto-Athabaskan Verb Stem Variation*. Fairbanks: Alaska Native Language Center.

Lindau, M. (1984). Phonetic differences in glottalic consonants. *J.ournal of Phonetics, 12*, 147-155.

Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

McDonough, J., P. Ladefoged and H. George. (1993). "Navajo vowels and phonetic universal tendencies" (this volume).

McDonough, J. (1990). *The Phonology and Morphology of Navajo Verbs*. Ph.D, University of Massachusetts, Amherst.

Ohala, J., & Lorentz, J. (1977). 'The story of [w]: An exercise in the phonetic explanation for sound patterns'. *Proceedings of the Berkeley Linguistics Society, 3*, 577-599.

Sapir, E. & H. Hoijer, 1967. *The Phonology and Morphology of the Navajo Language*. University of California Publications in Linguistics, 60.

Young, R & Morgan, W. 1992. *An Analytic Lexicon of Navajo*. University of New Mexico, Albuquerque.

Young, R & Morgan, W. 1980, 1986. *The Navajo Language*. University of New Mexico, Albuquerque.

Tone and tonogenesis in Navajo

Ken deJong and Joyce McDonough

I. Introduction.

The present paper outlines some issues in the description of Athabaskan tone, and reports the results of a preliminary phonetic study of tone in Navajo. To our knowledge, the most recent instrumental study of Athabaskan tone is P.I. Goddard's kymographic investigation of Kato, Chipewyan, and Hupa, performed in the first decades of this century (Goddard, 1907, 1911, 1912). We will begin by outlining some proposals concerning the historical origin of tone in Athabaskan languages. In addition, we will outline some aspects of Navajo morphology and phonology which shed light on these proposals. Finally, we report on the results of a preliminary phonetic study of f0 patterns on a corpus of Navajo words. Specific questions addressed in this study are whether Navajo has a pitch accent system like that in Japanese, as well as the phonetic effects of various consonant-types on the f0 pattern.

2. Athabaskan tonogenesis.

Several scholars have argued (Leer 1981, Krauss 1980, Kingston 1990) that the tone in Athabaskan languages is the historical result of consonantal effects on vowels. As such, the developments in Athabaskan would parallel developments in other languages of the world (Hyman, 1978). Leer (1979) and Krauss (1981) have argued that two historic processes that involve changes on the vowel parallel one another, nasalization and 'constriction[1]'. Phonemically nasalized vowels have arisen by the absorption of consonantal information into stem vowels. That is, historic vowel-nasal sequences have become nasalized vowels, probably through coarticulatory nasalization of the vowel and subsequent deletion of the oral articulation of the nasal consonants. Similarly, they argue, tonal contrasts have arisen from coarticulatory effects of glottalized consonants on previous stem vowels, and subsequent deletion of the glottalized consonant. The coarticulatory effects take the form of perturbations in the fundamental frequency on the vowel, presently apparent as tonal contrasts.

This analysis of tonogenesis in Athabaskan has several problems. Not all Athabaskan languages are tonal, and those that are differ in at least one striking way: in tonal polarity. That is, some languages have high and some have low default tone. Thus, following this historical account, putative glottal perturbations on fundamental frequency (f0) have given rise to distinctively high pitched vowels in some languages, but distinctively low pitched vowels in others. Kingston (1990) has provided an account of this polarity. Another problem is the relation between the diachronic account and the present state of affairs; we have found that consonantal effects on tone in Navajo are not consistent with diachronic accounts. A third problem is the variety of tonal systems among Athabaskan languages. There is a great deal more variation among the tonal patterns in Athabaskan than the stipulation of Athabaskan tone as high-marked or low-marked implies.

---

[1] "'Constriction' ...involves the transformation of a sequence of a vowel plus glottal stop or ....a vowel followed by a glottalized consonant into a vowel which was suprasegmentally modified by glottalization." (Leer 1979:26)

Navajo is a low tone default language. In Leer's (1981) diachronic analysis, high tone in Navajo, the marked tone, arose in non-obstruent closed stems. All other stems, those with 'reduced' vowels and with sonorant closed and constricted vowels, bear low tone. Since high tones have only arisen in a restricted environment, Krauss (1982) claimed this as the reason for a statistical predominance of low toned stems in Navajo. Stanley (1969), in a similar, synchronic account, argued that stem high tones may come from absorption of abstract nasal and glottal increments into the stem. In addition, prefix vowel high tones are derived via nasality. In a discussion of Sacree, which like Navajo has a low tone as the default tone, Sapir (1925) proposes two kinds of high tones, those in stems and those in prefixes. Tones in stems are specified lexically, while those in prefixes are derived. To understand the distribution of tonal distinctions, it is necessary to understand a bit about the morphology of Navajo.

3. Navajo morphology

There are two morphological categories: stems and affixes. Stems are canonically monosyllabic and the carry tonal contrasts. Affixes are prefixal. Thus stems fall on the right edge of their domain, as shown in (1).

(1) [(af) | stem]

Navajo syllable structure is CVV only, with the exception that coda consonants can appear in stems. Stems are monosyllabic. Prefixes are consonantal; their vowels are epenthetic. Prefixes receive default tone or in some cases tone by rule (as described below).

Nominal morphology is relatively simple as compared to that of verbal complexes. Athabaskan does not have a large inventory of nouns, many nouns are deverbal. Tone in nouns is very simple, the vowels either bear a H or L tone. Contour tones exist only for some derived long vowels. Nouns are minimally monosyllabic and can be monomoraic, though they are lengthened when prefixed. Many nouns require possessor prefixes. The following are examples of unprefixed (a.) and prefixed (b.) noun stems:

(2) a. kǫ́ 'fire'

b. bidiɫ '(his) blood'
bi + diɫ

Navajo has a very rich verbal morphology. For the structure of verbs, we assume the bipartite constituent model laid out in McDonough (1990, 1993). In this model there are two constituents, verb (v) and infl (I), and each constituent has a stem and set of prefixes, the misnamed 'classifier' (cl) prefixes and the aspect (ASP) prefixes respectively. In addition, there is a set of agreement markers (AGR), like the ones on nouns, that are prefixed to the infl-verb compound. The following is a diagram of the model:

(3) clitics # $[_{Wd}$ AGR + $[_{Wd}$['Asp' | STEM $]_{I}$ + ['cl' | STEM $]_{V}$]]

The stems are the obligatory parts of the verb, which accounts for a disyllabic minimality of verbs. Verbs have at least two syllables.

(4) minimal verb:

a. [$_{Wd}$ [STEM]$_{I}$ | [STEM]$_{V}$]

b. yishcha 'I cry'
[$_{Wd}$ [ish]$_{I}$ | [cha]$_{V}$]

The stems (and clitics) in Navajo carry the tonal contrasts. Both verbal stems carry paradigmatic tone. Tone on the prefix vowels is predictable by either the spread of high tone or default specification (Kari 1976). Deverbal nouns have the structure of verbs, with its telltale disyllabicity and internal consonant cluster. Deverbal nouns often appear with a clitic on the end [igii] or [i].

4. Tone rules

Kari (1976) has a fairly complicated discussion of tone in Navajo. However, many of his tone rules are artifacts of his morphological analysis of verbal complexes. In the positional class 'template' morphology which Kari (1976) used, all morphemes in the verbal complex except the verb stem are considered prefixes. Because he considers them to be prefixal, he must account for tonal variations in the second (i) stem by positing phonological rules. In a bipartite view, these tonal variations are simply the result of underlying tonal specifications on the second stem; assuming a bipartite structure eliminates many of the tone rules from Kari's analysis.

An example of one class of remaining tone rules is given in (5). In this example, the iterative clitic, *ná-*, bears an underlying high tone. This tone is realized on the surface on both the clitic vowel and the next vowel after the clitic and is subject to interference from high tone specifications on the stem.

(5) High tone spread rule, the iterative prefix /ná/:

	hanishchaad	vs	hanáníshcha' (iterative)
	ha # [nish] [ł chaad] [stem] [af | stem]		ha**ná** # [nish] [ł chaad] [stem] [af | stem]
high tone spread			H |‾ ‾ ‾ ha**na** # [nish] [ł chaad]

There are two conditions on this rule: long vowels characteristically block tone spread, and the gamma perfective conjugation does not allow high tone throughout the paradigm (Kari 1976:131). This condition will block an occurrence of the high-tone spread rule. An example of the long vowel of the 1dual imperfective *iid* blocking the spread rule is given in (6). With one exception, /v́n-/, the verbal prefixes do not carry underlying tone. This high tone spreading applies to tones introduced by clitics, or in the rare case the prefix. Tones on stems do not spread.

(6) néiilzééh
ná # [iid] [l zééh]
[stem] [af l stem]

H
|
ná # [iid] [l zééh]

high tone spread (blocked)

Most descriptions of Navajo assume it is a tone language, with highs and lows specified on each syllable. However, tonal contrasts exist almost exclusively in stems, due possibly to the genesis of tone and the nature of the Navajo morphological system . Stems are monosyllabic, and affixes do not carry the tonal contrasts. Considering this distributional fact, and the fact that Carrier, another Athabaskan language has been analyzed as a pitch accent language, it is possible that Navajo is an example of a pitch accent language. Since there is some question as to the tonal typology of Navajo, we will analyze a corpus of Navajo data to shed light on this question. However, before doing so, it is worth considering such typological labels as pitch accent and tone in more detail.

5. Pitch accent versus tone.

A traditional distinction has been drawn between languages in their use of pitch specifications in the lexicon between pitch accent and tone languages. In such works as McCawley (1978), pitch accent languages are languages in which lexical items are distinguished by the location of a tonal specification. The archetypal language in this category is Japanese, in which words differ tonally only in terms of which syllable has a high-low pattern associated with it. Such syllables are termed accented syllables. As has been shown experimentally by Pierrehumbert and Beckman (1988), other syllables have a fundamental frequency pattern determined by interpolation between f0 values determined by tones associated with the accented syllables. That is, unaccented syllables are underspecified for tonal values.

Tone languages, by contrast, have pitch specifications more densely distributed throughout each lexical item. Specifically, each lexical item will have one tonal specification per tone bearing unit, where tone bearing units may consist of morae, vowels, or syllables. An archetypal example of a tone language would be Yoruba, where each syllable in a word may contrast between high, low, or mid specifications. Each of these specifications is expressed phonetically.

A reconsideration of this tonal typology points out several, perhaps orthogonal, ways in which language tone systems may differ. The first has to do with the density of tonal specifications in the lexicon. Tone languages, as already noted, may differ in whether tones reside on all morae, or just on morae containing vowels, or just on each syllable. Pitch accent systems fall further out on this continuum — allowing only one tone per morpheme, or word.

A second way in which languages may differ is as to whether there are syntagmatic contrasts in tone placement within the lexicon. Japanese, as noted above, makes distinctions between lexical items by the location of a tone pattern. Other languages with equally sparsely specified tones, such as Kyung-Sang Korean, do not make syntagmatic contrasts in tone. In this dialect of Korean (Chung, 1991), tonal contrasts are only specified on the initial syllable of a word. Similarly with other dialects of Korean, tones appear at predictable locations within the prosodic word — and thus, the presence of tones allows for the determination of the extent of a prosodic domain (Jun, 1990).

A third way, orthogonal to the first two ways, that languages may differ is in the number of paradigmatic distinctions in tone. Yoruba, as mentioned above, has three possible tones on any tone-bearing unit (not counting contours). Kyang-Sung Korean, along with many traditionally identified tone languages, has two tonal possibilities. Japanese, in the analysis of Pierrehumbert and Beckman (1988), has no paradigmatic distinctions; there is only one tonal possibility for each accented syllable. Note that the number of paradigmatic distinctions is at least somewhat orthogonal to whether a language is traditionally described as a tone or a pitch accent language. For example, Swedish is a traditional pitch accent language (in that it makes syntagmatic distinctions in tone placement) which has two possible tonal patterns per accented syllable (Bruce, 1975, Gårding, 1973). That is, in addition to the specification of pitch accent location, the tonal value of the pitch accent must be specified as well.

Finally, languages may differ in whether they use tone as part of their system of marking prominences. Traditional pitch accent and stress accent languages (such as Japanese and English, respectively; Beckman, 1986), have pitch specifications as markers of one or another level of prominence. This use of pitch specifications seems to be somewhat independent from the three distinctions noted above, even the first distinction in density of tonal specifications. Oklahoma Cherokee seems to be a case of a language in which there is a dense distribution of tonal specifications (like a traditional tone language), as well as a tonal specification associated with prominence which overrides the tonal specification on the prominent syllable (Wright, in preparation).

Given that tonal systems can vary in many different ways, one question raised by the present study is that of how to characterize Navajo. Given the relatively sparse underlying specifications of tonal contrasts in Navajo, we wonder whether Navajo might in fact be a pitch accent language, with transcribed high tone syllables being accented, and transcribed low tone syllables being unaccented and possibly phonetically unspecified for tone. Some Athabaskan languages, such as Carrier, seem to be pretty clear examples of pitch accent languages.

6. Carrier 'pitch accent'

The tonal patterns found in Central Carrier (Northern Athabaskan) indicate that tone is not paradigmatically contrastive in this language. While it has been described as a tone language (Pike 1986, Story 1984), the tonal system operates like the Japanese system: the underlying contrast is between accented and unaccented words (McDonough 1989). In accented words the placement of the tonal melody, HL, is not predictable; it can appear on any of syllable in the word including the final syllable. The accented syllable initiates a drop, that is the high associates to the accented syllable and the L appears on the following syllable. Syllables following the accented syllable all are transcribed with low tones, presumably up to the end of some prosodic domain. McDonough (1989) does not discuss what the domain is, except to point out that the domain of spread is greater than a morpheme or word. High tones subsequent to the accented syllable do not occur within this domain -- just as in the Japanese accentual phrase.

In unaccented words, the tone melody is realized on the final syllable and differs from the accented morphemes' melody. The default melody applied to unaccented words is simply H. Thus accented words and unaccented words will contrast in the pattern they impose on the morphemes or words that follow them. Unaccented words have no L to spread. Following are two monosyllabic words, accented and unaccented, respectively. The accented word [yes] 'wolf' spreads its L to the syllables of the following word [niɫ'én] 'he sees':

(7)	yés nìɫ'èn	he sees the wolf
	xóh niɫ'én	he sees the goose

The morpheme [ʔiɫó] is accented and will lower the tone of any following morpheme. In the following, the [yés] itself is affected by the HL of the preceding accent:

(8)	/ʔiɫó /	+	/yés/	----->	[ʔiɫó yès]
	one		wolf		

In Pike's (1984) account of Central Carrier, there is reason to assume that the syllables preceding the accented syllable are not specified for tonal values.

7. The Navajo sound system

In order to understand the investigation of Navajo tonal specifications reported below, it is necessary to review the Navajo sound system. Especially important are the Navajo consonants. First, since they are implicated in theories of Athabaskan tonogenesis, it would be useful to take note of the synchronic effects of the different consonants on the f0 pattern. Second, in trying to ascertain the tonal specifications, one must factor out the influences of other factors on the f0 pattern. Most troublesome of these other factors, especially in an obstruent-rich language, such as Navajo, are glottal specifications on consonants appearing in the neighborhood of the measured f0 values.

Table 1 gives an inventory of the Navajo consonants in Young and Morgan's (1986) orthography. (See Mcdonough and Ladefoged (this volume) for IPA transcriptions corresponding to each orthographic symbol.) Many of the consonants appear with phonetic alternations. Fricatives often appear intervocalically as approximants. Voicing is a property of continuants, stops are not voiced. Stops have a three way laryngeal contrast: unaspirated, aspirated and glottalized (ejectives). The consonants in Navajo have two characteristic properties; they are very long in duration, almost twice the length of English consonants, and for all but the unaspirated plain stops, the period from the release of the stop closure to the onset of voicing and vowel formants is often as long as the duration of the acoustic implosion. Fricatives are often phonetically voiced in the same environment as that in which they appear as approximants. Only stem onsets show the full set of consonantal contrasts: coda consonants are a reduced set. The prefix consonants with few exceptions are also taken from a reduced set.

The Navajo vowel system is asymmetric; it has four vowels, two front, a low central vowel and a mid back vowel [i e a o]. The system lacks a high back vowel. In addition, there is an oral and nasal, and a long and short distinction in each quality. The length distinction is enhanced by vowel quality differences for the front vowels, and depends on whether the speakers are men or women, for the low and back vowels.

Table 1. Navajo Consonants
(after Young and Morgan 1986)

p	t, d, t'	k, g, k'	kw, gw		'
	ts, dz, ts'	ch, j, ch'			
	tl, dl, tl'				
	s, z	sh, zh	x, gh		h
	ł, l				
m	n				
w		y			

8. The corpus.

A set of recordings was made at Navajo Community College, Shiprock, New Mexico, in April, 1992. The speakers were educated native speakers of Navajo, who live on the reservation and use Navajo daily. The present report is based on an analysis of 7 speakers, 5 female and 2 male. They were recorded while repeating a word list that was designed to illustrated the phonemic contrasts of Navajo. All the speakers were consulted on the accuracy and acceptability of the word list before making the recording. The words were repeated both in isolation and in the frame sentence, *ashiké* ____ *da:ní* '(The boy ___ he said').

As phonemic contrasts are best observed in the stems, the corpus was largely composed of simple nouns. These consist of a stem, which is usually monosyllabic, and a prefix such as *bi-* 'his/hers'. About 80% of the words were of this form. The remainder had other prefixes, or disyllabic stems, or, in a few cases, were verbal forms.

9. Tonal Distribution.

The first way in which Navajo appears to be a pitch accent language is that tonal markers seem to be somewhat restricted as to where they can occur, and as to the relative proportion of tonal shapes. This oddity of tonal distribution is readily apparent in the present corpus. Given the free distribution of the tones, high and low, across all syllables, as the traditional account would have it, one would expect x^n tonal patterns, where x = the number of paradigmatic contrasts, 2 in this case, and n = the number of tone-bearing units in the item. We assume for the present that the tone-bearing units in this case are syllables, since bimoraic syllables always are transcribed with level tones. With a few exceptions consisting of complex contracted forms, none of the long vowels are transcribed as having contours. In addition, we would expect a more or less even distribution of items across the tonal categories.

Grouping the target words in the corpus by transcribed tonal pattern, however, reveals a rather skewed distribution of items having each tonal pattern. The number of tokens of each tone pattern are given in Table 2. Of the single syllable items, two thirds have low tones. Considering disyllables, the tendency toward more low tones is even more pronounced; nearly three fourths of the tokens have final low tone, and most of those exhibit a low - low pattern. Some of the preponderance of initial low tone is due to the choice of the low-toned *bi-* as a classifier for most of the nouns in the corpus. Taken in all, there is a preponderance of low tones in the present corpus, and also a tendency

Table 2 Number of tokens of each tonal pattern.

One syllable

High	Low
8	16

Two syllables

	All items		Morphologically simple items	
	First syllable			
	High	Low	High	Low
High	7	14	4	7
Second syllable				
Low	5	53	2	44

Three syllables

	All items	Morphologically simple items
LLL	3	1
LLH	3	1
LHL	2	1
HLL	0	0
LHH	1	1
HLH	0	0
HHL	1	0
HHH	0	0

for high tone to appear only on the last syllable of each token. When we eliminate the items in the corpus which are transparently verbal complexes of some sort, the pattern becomes even stronger. The number of remaining items appears to the right in Table 2. Of the two syllable items, only one in seven remaining items bears final high. The proportion of items with initial high is even smaller.

This skewing of transcribed tone types might be indicative of Navajo being a pitch accent language in which one tone (the high tone) occurs on specified, accented syllables. Low, in this case, might not be specified to occur at all, but rather, transcribed low tone syllables might simply be unspecified for tone at the phonetic level.

10. Instrumental analyses.

A second aspect of the f0 pattern found in this corpus points toward a pitch accent analysis as well. An inspection of the corpus of data shows that many syllables transcribed as bearing high tone actually have a rising f0 contour. One example of a low-high contour is shown in Figure 1. Many pitch accent languages, such as Japanese, and Swedish (and possibly Carrier as well) have contours which occur on accented syllables. Thus, if Navajo were a pitch accent language of the same kind, the highs might actually be contour low-high patterns. The occurrence of a contour tone would also fit well into a consonantal theory of the origin of the high tone in Navajo. Consonantal effects of f0 appear superficially as tonal contours; a loss of the initiating consonant which leaves the tonal effects behind would produce a contour tone. For example, final glottal stops seem to have given rise to the rising contours of Vietnamese (Haudricourt, 1954; Hombert, 1978). Although many of the syllables in the present corpus appear with rising f0 patterns, some do not, but rather, occur with a more or less flat, high f0 pattern. The low-high patterns may arise as a result of consonant-related f0 depression at the onset of the syllable. Or conversely, the flat patterns may arise from consonantal raising effects overlaid on a low-high pattern — effacing the low tone specification.

To more accurately assess the effects of consonants on the f0 pattern, as well as to get a general picture of the phonetics of the transcribed tonal specifications, f0 measurements were performed on a majority of the corpus. All high-toned syllables in the corpus were first measured. Then each high-toned syllable was matched with a separate low toned syllable with the same vowel, the same vowel length and the same or a similar previous consonant. Although a low toned syllable with an exactly matching preceding consonant could not be found for each high toned syllable, the consonants were perfectly matched in glottal specifications. Four f0 measurements were taken from each token, as illustrated in Figure 1. The f0 in the target vowel was measured by noting the frequency of the 5th harmonic at a time in which the f0 was at a steady-state, or nearly steady-state, portion. Typically, this occurred in the later half of the vowel. The second measurement was the f0 at the beginning of the vowel, taken at the onset of voicing after voiceless consonants, or at the time of release following voiced sounds. A third measurement was taken at an inflection point between the two other measurements. This measurement will not be discussed further here. Finally, the f0 at a zero velocity point in the last syllable of the frame sentence previous to the target item, *kɛ*, was taken as a reference to help factor out speaker differences and intra-speaker pitch range differences across the tokens. The frequency of the 5th harmonic was divided by 5 to find the estimated value of the fundamental, and then subtracted from the estimate of the fundamental during this reference syllable. Thus all frequencies reported below are given in Hertz above or below the reference high tone value.

Given the hypothesis that Navajo is a pitch accent language, one would expect the following effects to be apparent. First, if the transcribed low-toned syllables are phonetically underspecified for tone, one would expect the f0 on low-toned syllables to be a function of various factors, such as the glottal specification on the neighboring consonant, and the tone values of neighboring syllables. The effect of these factors should be much less apparent on specified high syllables. Second, if the transcribed high tone syllables are actually low-high contours, one would except the transcribed high to have a lowering effect on the f0 at the onset of the syllable. In comparison, the transcribed low-toned syllables should either have no effect, if they are phonetically underspecified, or they should have the same or similar effect to the putative low-high contour.

11. Factors in fundamental frequency values.

The measurements of f0 during the vowel and at the onset of voicing (measurements 1 and 2, in Figure 1) were submitted to a series of ANOVA's with various aspects of the items as independent variables. The various factors, listed in Table 3, were either transcribed aspects of the tone pattern of the utterance, or segmental aspects of the vowel and consonants around the vowel which are known to have effects on f0 patterns. In addition, vowel quantity was examined, since the length of the vowel might have some effect on the strength of coarticulatory influences on the f0 in the vowel, simply by reason of its effect on the temporal proximity of neighboring consonants and vowels. For the purposes of the present analyses, [h] is included with aspirated consonants, and [ʔ] is included with ejectives in the class of glottalic consonants.

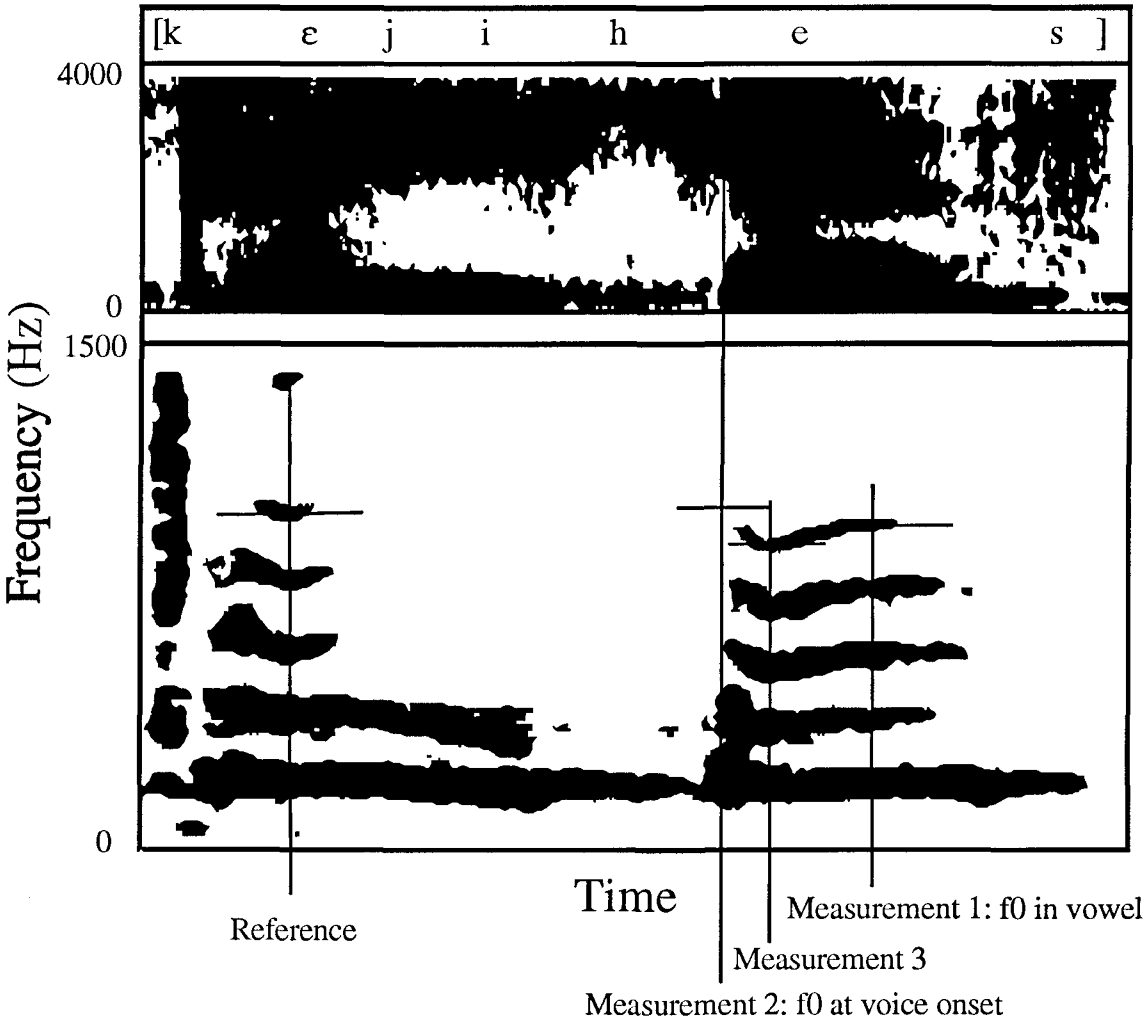


Figure 1 A 50 Hz narrow-band (below) and a 300 Hz broad-band (above) spectrogram of the example Navajo word, *jihé̃é̃s* ("it itches"). Shown here are measurements of the frequency of the 5th harmonic.

To simplify the discussion, we will first consider the factors which will be eliminated from further discussion. The first of these is the preceding consonant's supra-laryngeal specification. This factor could not simply be used in the present corpus, since there is a relationship between the laryngeal and supra-laryngeal grouping of consonants — for example, there are no phonemically voiced stops in Navajo; also there are no voiceless sonorants. This relationship makes it difficult to interpret effects for this factor in a way which is independent of effects of laryngeal specifications on the consonant. Supra-laryngeal effects will not be discussed further in this paper.

Another aspect of the present data which deserves further analysis, but will not be discussed in detail here is the extent to which the present patterns are consistent across speakers. An ANOVA of the f0 in the vowel (measurement 1 in Figure 1), with the tone on the vowel and speaker as factors shows that there is some effect of speaker on the f0 in the target syllable. In this analysis, there is an extremely strong main effect of tone

Table 3 Factors used in Analysis of f0 Values.

Tonal Specifications:	Tone on target syllable Tone on preceding syllable Tone on following syllable	High or Low High or Low High or Low
Aspects of the Vowel:	Vowel height Vowel quantity	High, Mid, or Low Long or Short
Aspects of Consonants:	Following glottal stop	Yes or No
	Preceding consonant's laryngeal specification	Aspirated, Voiceless, Voiced, or Glottalic
	Preceding consonant's supra-laryngeal specification	Stop, Fricative, Sonorant, or Nothing

($F(1, 468) = 870.84$, $p < 0.0001$). In addition, there's a weaker main effect of speaker ($F(5, 468) = 22.38$, $p < 0.0001$) and a significant interaction of speaker and tone ($F(5, 468) = 35.17$, $p < 0.0001$). Speaker differences might be reduced or eliminated by using a different method of normalization.

Vowel height also has a significant effect on the f0 in the vowel. A two-way analysis of variance with tone and vowel height shows significant effects of tone ($F(1, 474) = 584.37$, $p < 0.0001$) and vowel height ($F(2, 474) = 8.54$, $p < 0.0005$). Higher vowels have a higher f0. There was no significant interaction between tone and vowel height ($F(2, 474) = 2.55$, $p > 0.05$). Thus, the method of estimating f0 used here is sensitive enough to detect the subtle, though well-attested (Black, 1949; House and Fairbanks, 1953; Hombert, 1978) difference in f0 due to vowel height. Since the vowel height effect does not interact with the tonal specification's effect on f0, we will not consider it in any more detail here.

One further factor tested was that of the presence of a following glottalic consonant. It turns out that this factor had no significant effect on the f0 in the vowel. In a two-way analysis of variance with tone, there was a strong main effect of tone ($F(1, 476) = 607.68$, $p < 0.0001$), but the presence of a glottal stop had no significant main effect ($F(1, 476) = 0.06$, $p > 0.05$), nor interaction with tone ($F(1,476) = 0.26$, $p > 0.05$).

12. Tone pattern and consonantal interactions.

We will now consider factors more directly involved in the tonal analysis of Navajo, the three tone factors, and the preceding consonant glottal specification. First, we will consider their effect on the f0 at the onset of voicing (measurement 2 in Figure 1). The results of a three-way analysis of variance with the tone on the target syllable (henceforward "target tone"), preceding tone, and glottal specifications on the consonant are summarized in the upper left portion of Table 4. Both the transcribed target tone and the transcribed tone on the previous syllable had strong main effects on the value of f0 at the onset of voicing. These effects were independent of one another (in that their interaction was not significant). The magnitude of the two effects is about the same —

Table 4. Summary of ANOVA's.

independent variables		dependent variable				
		f0 at onset of voicing			f0 in vowel	
		d.f.	F	p	F	
preceding tone, target tone, consonant type	preceding tone:	1	106.80	< 0.0001	36.32	< 0.0001
	target tone:	1	83.19	< 0.0001	490.13	< 0.0001
	consonant type:	3	11.15	< 0.0001	1.31	0.27
	prec. tone by targ. tone:	1	0.16	0.69	0.00	0.99
	cons. type by prec. tone:	3	8.09	< 0.0001	9.07	< 0.0001
	cons. type by tone:	3	4.42	0.004	4.29	0.005
	three-way interaction:	3	0.89	0.45	2.66	0.047
	residual:	464				
preceding tone, target tone, following tone	preceding tone:	1			6.60	0.011
	target tone:	1			170.63	< 0.0001
	following tone:	1			0.26	0.87
	prec. tone by targ. tone:	1			1.81	0.18
	foll. tone by targ. tone:	1			0.50	0.48
	prec. tone by foll. tone:	1			0.04	0.84
	three way interaction:	1			0.81	0.37
	residual:	472				

a difference of 18.3 Hz between high and low target tone environments, and a difference of 18.4 Hz between preceding high and low tone environments.

As is expected, there was also a significant main effect of consonant type on the f0 at the onset of voicing. Aspirated and voiceless consonants have a mean raising effect of about 15 Hz, as compared to the voiced consonants. Since many of the voiced consonants were sonorants, we expect that voiced consonants will not be creating large f0 perturbations; thus, we can analyze this difference as indicating a raising effect of aspiration and voicelessness. Such effects of aspiration, at least, are well attested (see Hombert, 1978, for a review). Glottal consonants have a somewhat smaller mean raising effect. These effects, however, are not independent of the background tonal configuration, as is evident in significant interactions between consonant type and the preceding tone, and consonant type and target tone.

Figure 2 plots the mean f0 values following each consonant type divided by preceding tone and target tone. The two interaction effects have somewhat different causes. The amount of effect of the preceding tone (right panel of Figure 2) is reduced to

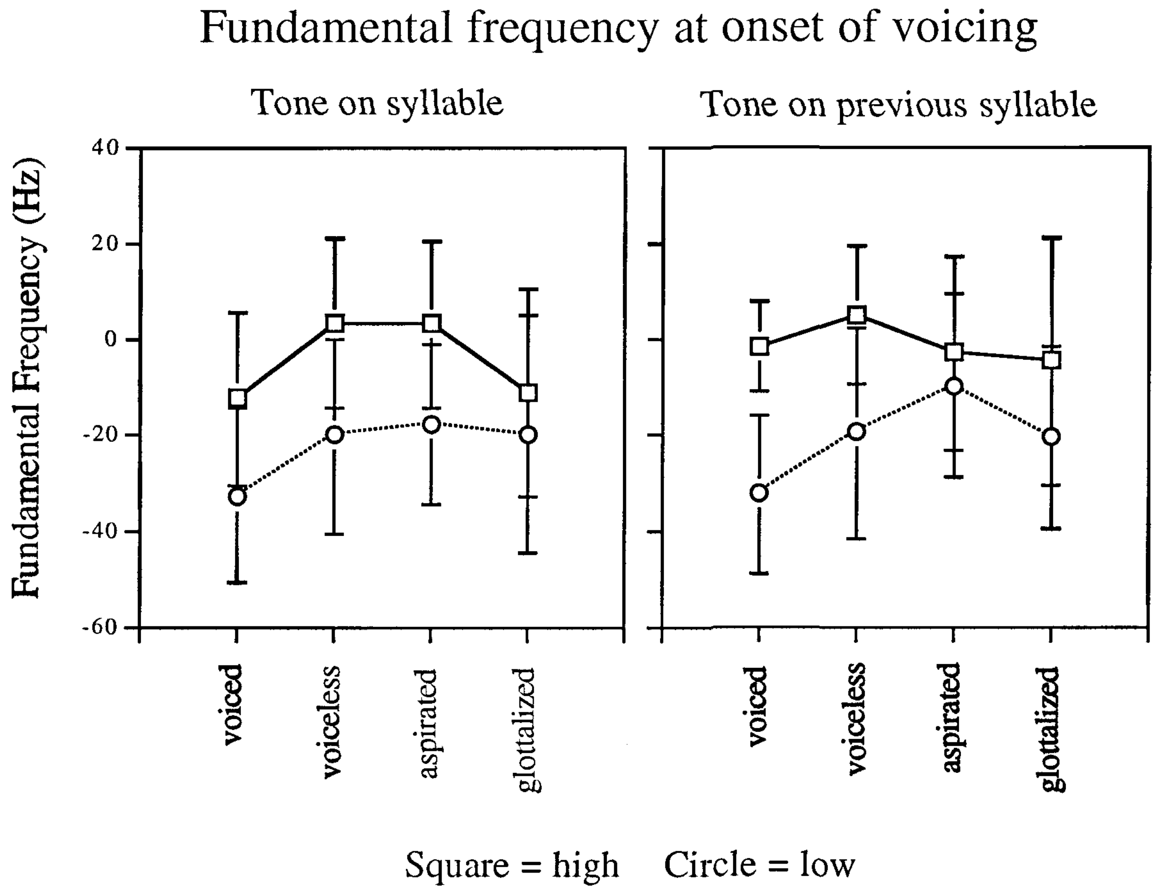


Figure 2. Mean f0 at the onset of voicing, broken down by consonant type and preceding tone or tone on target syllable. Error bars indicate one standard deviation.

less than half following aspirated consonants. Especially aspirated consonants are blocking perservatory coarticulation effects. The effects of the target tone (left panel of Figure 2), are being reduced in the environment of glottalic consonants. This may be due to glottal consonants only having a raising effect on low-toned syllables. Mean values with glottal consonants and low toned vowels are similar to those with aspirated and voiceless consonants; however, mean values with glottal consonants and high toned vowels are more like those with voiced consonants.

Now, we will go on to consider the results of the same three-way analysis of f0 values in the vowel (measurement 1 in Figure 1), also summarized in (the upper right part of) Table 4. The general pattern of results is the same as that for f0 at the onset of voicing, with two differences. First, the magnitude of the effects of tonal specifications are somewhat different; the tone on the vowel has an even stronger effect of greater magnitude (33.4 Hz). The effect of preceding tone is smaller (10.4 Hz).

The other major difference here is that there is no significant main effect of consonant type on the f0 in the vowel. This suggests that the tone specifications of the utterance have effects over a much longer duration than f0 perturbations due to

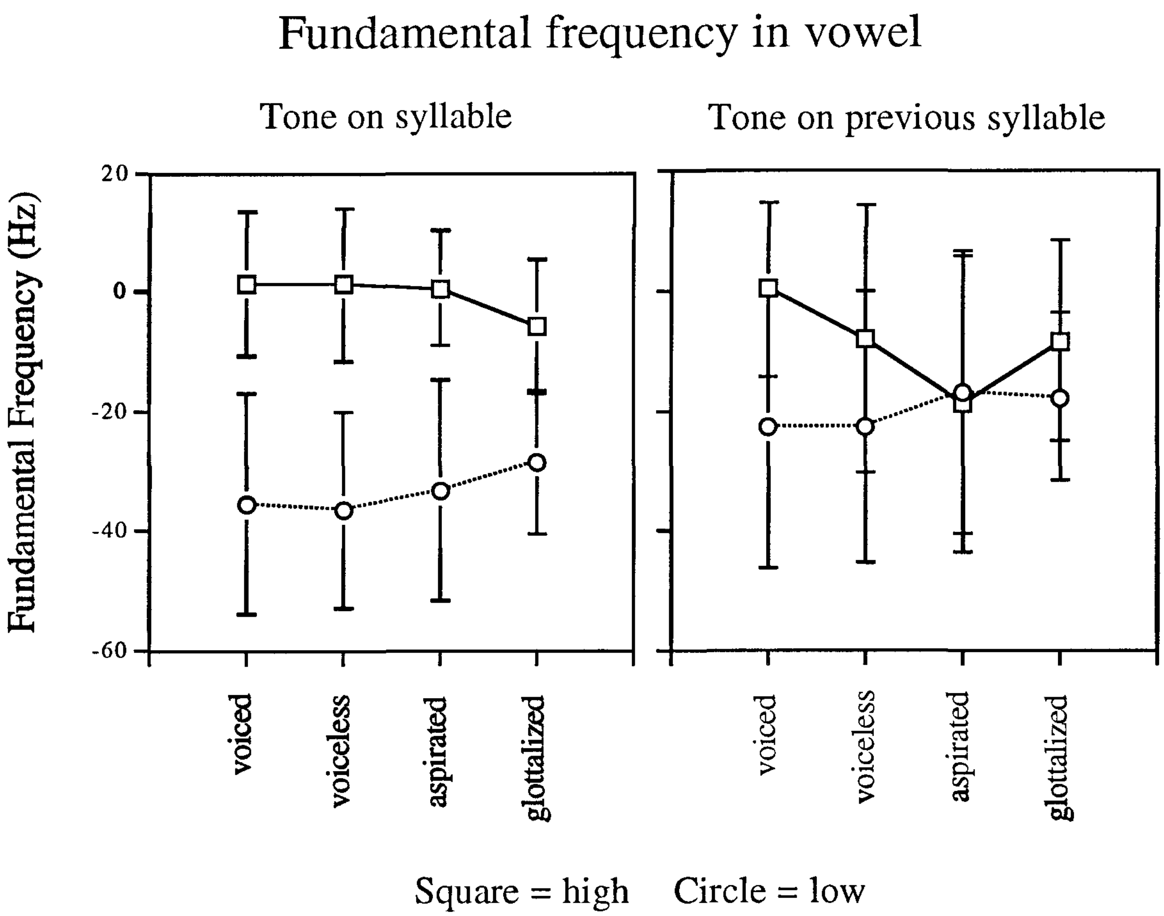


Figure 3. Mean f0 at the f0 steady state in vowel, broken down by consonant type and preceding tone or tone on target syllable. Error bars indicate one standard deviation.

consonants. This asymmetry is somewhat like that found for consonant and vowel effects on formant movements. Vowel effects on formant values extend over longer periods of time, and vowel coarticulatory effects are found between vowels (Öhman, 1967); consonants, by contrast, have a much more temporally localized effect on formant movement.

This is not to say, however, that consonant type has no effect on the f0 in the vowel; its effect is found in an interaction with preceding tone and an interaction with the tone on the syllable. Again, as with the effect on the f0 at voice onset, the two interactions arise in somewhat different ways. The interaction with the preceding tone is a result of some of the consonant types reducing the coarticulatory effects of the preceding tones. This type of effect -- a blocking of tonal coarticulation by voiceless stops seems to have been phonologized in Ngizim (Schuh, 1971, 1978) as a conditioning on the spread of low tones. Aspirated consonants appear to be opaque to tonal coarticulatory influences. A parallel situation in the analysis of formant movement is that of secondary articulations blocking vowel-to-vowel coarticulatory effects (Keating, 1985; Hussein, 1990). It is alternatively possible that the difference in the opacity of consonant types to tonal coarticulation is an indirect effect of the typical duration of time taken up by the different types. Glottal and aspirated consonants are typically longer than the unaspirated consonants (McDonough and Ladefoged, this volume). This increase in the temporal distance between the preceding tone and the target tone may also account for part of the reduction of coarticulatory effects found here.

The interaction involving the target tone (Figure 3, to the left) arises in large part by a reduction of the f0 difference between high and low tones after glottal consonants. There is also a small raising effect of aspirated consonants detectable on low toned syllables which possibly contributes to this interaction. Thus, glottalic consonants seem to be having a slight neutralizing effect on the tone of the syllable.

One final analysis of note here points out an extreme temporal asymmetry in the tonal effects. Table 4 (bottom) summarizes the results of an ANOVA of f0 in the vowel with preceding tone, following tone, and the tone on the target syllable as factors. As with other analyses, preceding tone and tone on the target syllable both have independent main effects. Following tone has no significant main effect. None of the interactions reach significance at the 0.01 level. Thus, though there are strong perseveratory coarticulatory effects of tone, this study finds no evidence of anticipatory effects. In general, it appears that tones seem to be most fully realized in the later portions of their syllables. The steady state measurements analyzed here typically occurred in the later half of the syllable. This misalignment of the tones seems to account for the presence of such clear rising tonal contours on transcribed high tone syllables. The low part of the contour may perhaps be a reflection of a previous low realized near the onset of the high toned syllable. Thus the asymmetry of the coarticulatory effects may be due simply to a late alignment of the tones in Navajo. Indeed, late tone realization may provide an alternate explanation for the tone-spreading rule mentioned above in section 4. A more focused durational study should shed more light on these issues.

13 Conclusion.

Finally, we will briefly consider the relationship of the results of this study to the suggestion that Navajo is a pitch accent language. The analyses reported here yield no evidence that high toned syllables really bear complex tones. The f0 values near the beginning of the syllable are neatly predicted by a simple coding of the following and preceding tones as high or low. There were no unexpected effects of a putative untranscribed low tone. Also of note, low tones seem to have essentially the same status as high tones; thus this study has uncovered no evidence that low tones are phonetically unspecified. Cells marked off as low toned exhibit the same order of variation as their high toned counterparts. More to the point, there are no interactions between tone specifications, which would have suggested that high tones are having an inordinately large coarticulatory effect on low tones. Rather, the background f0 pattern seems to be determined straightforwardly as a sequence of highs and lows. It is possible that there may be some deeper level of grammatical description at which Navajo tones are specified in the same manner as those of Japanese, but that the tones are filled in on the surface. McCawley (1978) argues for this type of analysis for Ganda. Further study into this matter is called for.

This study has pointed out three suggestive oddities of the Navajo tone system. First, there is a strong preponderance of low tones. Second, tones seem to occur inordinately late with respect to the syllables they reside on. Third, glottal consonants have a reducing effect on the following tones. These three oddities provide some circumstantial evidence that the present tone system may have arisen from presence of consonants which have diachronically disappeared. The limited number of high tones, then, would reflect the limited distribution of the missing consonant or consonant series.

The late realization of tones, similarly, could be an artifact of the tone-producing consonants residing after the stem syllable. Finally, the effect of glottal consonants on the size of the tone difference suggests that the consonants and the tonal system are not entirely independent even at the present time.

14. Acknowledgement.

Special thanks to the staff and faculty of the Navajo Community College for their cooperation and interest. Particular thanks to Clay Slate, the director of the college, and Martha Austin-Garrison. Thanks also to Ian Maddieson, Richard Wright, Helen George, and others in the Phonetics lab for their valuable discussion of earlier versions of this work. Besides funding by the NSF, this work supported additionally by the NIDCD under grant number TC-32-00029.

15. References.

Beckman, M.E. (1986). *Stress and Non-Stress Accent.* Dordrecht: Foris.

Black, J.W. (1949). "Natural frequency, duration, and intensity of vowels in readings." *Journal of Speech and Hearing Disorders*, **14**: 216 - 221.

Bruce, G. (1975). "Stockholm accents in focus." *Lund University Phonetics Laboratory Working Papers,* **10**: 37 - 52.

Chung, Y.-H. (1991). *Lexical Tone in the Northern dialect of Kyung-Sang Korean.* Ph.D. dissertation, Columbus, Ohio: Ohio State University.

Fromkin, V.A. (1978). *Tone: a Linguistic Survey.* New York: Academic Press.

Gårding, E. (1973). *The Scandinavian Word Accents.* Available as *Lund University Phonetics Laboratory Working Papers*, **8**.

Goddard, P.I. (1907). "The phonology of the Hupa language." *University of California Publications, American Archaeology and Ethnology*, **5** (3).

Goddard, P.I. (1907). "Kato texts." *University of California Publications, American Archaeology and Ethnology*, **5** (3).

Goddard, P.I. (1911). "Athapaskan (Hupa)." In F. Boaz (ed.), *Handbook of American Indian Languages.* Washington, D.C.: Government Printing Office.

Goddard, P.I. (1912). "Elements of the Kato language." *University of California Publications, American Archaeology and Ethnology*, **11**(1).

Haudricourt, A.G. (1954). "De l'origine des tons en vietnamien." *Journal Asiatique*, **242**: 69 - 82.

Hombert, J.-M. (1978). "Consonant types, vowel quality, and tone." In Fromkin (1978), pp. 77 - 103.

House, A.S., and G. Fairbanks (1953). "The influence of consonant environment upon the secondary acoustical characteristics of vowels." *Journal of the Acoustical Society of America*, **25**: 105 - 113.

Hussein, L. (1990). "VCV Coarticulation in Arabic." *Ohio State University Working Papers in Linguistics*, **38:** 88 - 104.

Hyman, L. 1978. "Historical tonology." In Fromkin (1978), pp. 257 - 268.

Jun, S.-A. (1990). "The accentual pattern and prosody of the Chonnom dialect of Korean." *Ohio State University Working Papers in Linguistics*, **38**: 121 - 140.

Krauss, M., & Leer, J. (1974) *Proto-Athabaskan-Eyak *y and the Na-Dene Sonorants* . Fairbanks: Alaska Native Language Center.

Krauss, M., & Leer, J. (1981). *Athabaskan, Eyak and Tlingit Sonorants* Available as *Alaskan Native Language Center Papers* , **3.**

Leer, J. (1979). *Proto-Athabaskan Verb Stem Variation, Part One: Phonology . Alaskan Native Language Center Research Papers*. Fairbanks: Alaska Native Language Center, University of Alaska.

Leer, J. (1982). *Navajo and Comparative Athabaskan Stem List..* Ms. Ala Fairbanks.

Kari, J.M. (1976). *Navajo Verb Prefix Morphology*. New York: Garland.

Keating, P.A. (1985). "CV phonology, experimental phonetics, and coarticulation. *UCLA Working Papers in Phonetics*, **62**: 1 - 13.

McCawley, J. D. (1978). "What is a tone language?" In Fromkin (1978), pp. 113 - 132.

McCawley, J. (1970). "Some tonal systems that cone close to being pitch accent systems but don't quite make it." In *Proceedings of the Chicago Linguistic Society,* 6 . Chicago: University of Chicago, Department of Linguistics.

McDonough, J. M. (1989). "Tone and Accent in Carrier." *University of Massachusetts Occasional Papers, 14.*

McDonough, J. (1990) *Topics in the Morphology and Phonology of Navajo Verbs*. Ph.D dissertation. Amherst, Mass.: University of Massachusetts.

McDonough, J.M (1993). "The bipartite structure of the Navajo verb." Ms., UCLA

McDonough, J.M., and P. Ladefoged (1993) "Navajo stops." This volume.

Öhman, S.E.G. (1966). "Coarticulation in VCV utterances: Spectrographic measurement." *Journal of the Acoustical Society of America*, **39**: 151 - 168.

Pike, E. (1986). "Tone Contrasts in Central Carrier." *International Journal of American Linguistics, 52***(4).**

Sapir, E. (1949). "Glottalized continuants in Navajo, Nootka, and Kwakiutl". In D. Mandlebaum (Eds.), *Selected Writings of Edward Sapir* (pp. 225-251). Berkeley, Calif.: University of California Press.

Sapir, E. (1925). "Pitch accent in Sacree, an Athabaskan language". *Journal de la Societe des Americanistes de Paris,* **17**:185-205.

Schuh, R.G. (1971). "Verb forms and verb aspects in Ngizim." *Journal of African Linguistics*, **10**: 47 - 60.

Schuh, R.G. (1978). "Tone rules." In Fromkin (1978), pp. 221 - 256.

Story, G. (1984). *Babine and Carrier Phonology: A Historically Oriented Approach..* Arlington, Texas: SIL.

Wright, R. (in preperation). "Tone in Oklahoma Cherokee."

Young, R., and W. Morgan (1987). *The Navajo Language.* Albuquerque, N.M.: University of New Mexico.

Young, and Morgan (1992). *Analytical Lexicon of Navajo.* Albuquerque, N.Mex.: University of New Mexico Press.