

Preserving the sounds of disappearing languages

Peter Ladefoged
Linguistics Department, UCLA

1. INTRODUCTION

Language is the heart of a culture. It holds people together and establishes their identity as group. But nothing is more ephemeral than the sounds of a language. The sounds will live only as long as the language is spoken. When the sounds are those of elderly speakers whose children belong to another world, then soon those sounds will be gone forever. All that can remain are whatever records we have been able to archive.

There are about 7,000 languages in the world today, but there will probably be only 3,000 or so in 100 years time. Most languages are spoken by comparatively few speakers. Over half the 7,000 languages are spoken by less than 10,000 speakers, and more than a quarter of them by less than 1,000 speakers. These are too small numbers to ensure survival. The big languages that will swallow them up are really big – over 48% of the world's population are first language speakers of one of the world's top 10 most spoken languages: Standard Chinese, English, Spanish, Bengali, Hindi, Portuguese, Russian, Arabic, Japanese, and German.

Languages die because children no longer speak their mother's tongue. Very often this is because their mothers don't speak it to them. Schools and employers reward people who speak one of the major languages. TV, radio and films also encourage the use of major languages. When a language goes, much of the local culture goes with it, often causing great suffering among those affected. In addition, for those of us who are concerned with the sounds of languages, the disappearance of a language is a loss of a resource for the scientific study of human speech communication. Our notions of what constitutes a possible human language depend on studies of actual human languages. Many linguistic concepts are shaped by ideas of what is normal or unmarked in a language. This paper is based on a project (conducted with my colleague Ian Maddieson) that aims to provide descriptions of the salient phonetic structures of some of endangered languages before it is too late.

Making a record of the phonetic structures of an endangered language is, in principle, no different from making a record of the phonetic structures of any language. From a linguistic point of view, we want to record the facts about each language so that we can compare them with other languages. This implies that we have some universal phonetic framework that will accommodate the description of the sounds of any language. Less obviously, it also implies that we are working within some phonological framework that can describe all the differences in sound systems that we are likely to observe. We cannot describe the sounds of a *language*, without first considering the patterns of sounds, the phonological contrasts that are necessary for it to be a language. As phoneticians we are interested in the physical manifestations of these patterns of sounds.

2. MATERIAL TO BE RECORDED

All languages have consonants and vowels, and we can start describing the salient phonetic structures of a language by describing these entities. Consonants and vowels occur only as parts of words, the smallest bits of a language that can occur on their own. Generally the first task in describing the phonetic structures of a language is to draw up a list of words exemplifying the minimal contrasts. Ideally we want to record the language as it occurs in normal, everyday, spoken utterances. But this notion conflicts with the desire to describe all the contrasting sounds in a manner such that they can be compared with those of other languages. If we want to describe, for example, the vowels of English, it is impractical to just record a lot of free speech and hope that it will contain utterances in which somebody says each vowel between the same consonants, and with the same degree of stress and the same intonation. We cannot wait around in hopes that our speakers will all say *bead, bid, bade, bed, bad, bod, bud, bud, bird, bood*, etc. We must ask them to say the words that we want to record. It may not be the most natural form of speech, but a word list provides the only way in which one can record the complete set of sounds of a language as produced by a number of speakers.

In the project with which I have been associated, our normal procedure is to try to make a word list illustrating each vowel in at least two different stressed syllables, one beginning with a coronal stop (a sound like **t** or **d**), and one with some other consonant. If it is appropriate, we also record examples of unstressed vowels. Thus our word list for Banawa, an Arawakan language with four vowels, spoken in the Amazonian rain forest, included the words in Table 1. Stress in Banawa occurs on the penultimate syllable.

Table 1. Words illustrating the vowels of Banawa in stressed syllables after **t** and **b**, and in unstressed syllables after **b** and **f**.

Stressed		Unstressed	
tifa	drink water	tafi	eating
tefe	food (m.)	tafe	food (f.)
tafa	to eat	tafa	to eat
tufa	to block in	tafu	to eat
bita	mosquito	ibi	each other
befa	other	ibe	a strip
bata	to pick	iba	to put/place
bufa	put on water	ibufa	to dump into water

The structure of the language often does not permit the compilation of such simple sets of words. Languages with complex morphology, or a large number of consonants and vowels that contrast in tone, length and nasalization, may not have many minimal pairs. As a result, it may not be easy to compile a word list that includes all the contrasts in comparable forms. In Apache, an Athabaskan language with only four vowels, the morphological complexities and the elaborate consonantal system forced us to use a word list that was not as elegant as that for Banawa. Each of the four vowels can contrast in nasalization, length and tone as well as quality. In order to observe the interactions of all these features, nearly all of which can co-occur, we needed to record the 24 contrasting items shown in Table 2. As is evident, these words do not form neatly contrasting minimal sets, as they do in Banawa.

Table 2. The Apache vowels (in the traditional orthography) each in the root syllable (in the first column), in which they occur after a glottal stop ' and before a velar consonant g. Short nasalized vowels do not occur.

SHORT, LOW TONE		
'ig	ch'ig	blanket
'eg	k'egowa	home, upper home
'og	bésh dist'og	arrowhead
'ag	yidits'ag	he hears/obeys him
SHORT, HIGH TONE		
'íg	bá'its'íg	he is lucky
'éḡ	gode'égo	finally
'óḡ	ik'án nást'óḡ	dough
'ág	bitsit'áḡee	top of his head
LONG, LOW TONE		
'iig	nágonech'iigi	locust
'eeg	dálet'eego	same way
'oog	tú hayit'oog	he pumps water
'aag	hadich'aag	it oozes
LONG, HIGH TONE		
'íig	na'it'íigí	gopher
'e[e]	díyat'égo	awful
'óog	náyit'óog	he soaks it
'áag	hagodn'áḡee	corner (an elevated area, mound)
LONG, LOW TONE, NASALIZED		
'iig	ch'a'áníst'iig	he is hiding
'eeg	ch'íyingeçz	he is unlocking it
'oog	kee náts'oog	tennis shoes
'aag	goz'aḡee	place
LONG, HIGH TONE, NASALIZED		
'íig	hát'íig o	he wants (volunteer)
'éḡ	shihēhik'éḡ	one article fell on me
'óḡ	its'óḡ	it is bent
'áag	dahaḡee o'i'áḡo	some evening

The form of the word list required for illustrating consonants depends on the structure of the language. In a language with only CV or CVV syllables and no final consonants, we try to record each consonant preceding at least two different vowels in a stressed or low tone syllable. In Pirahã, a Mura language of this type spoken in the Amazonian rain forest, we were able to extend the word list so that we recorded each of the eight consonants before each of the three vowels. But in Toda, a Dravidian language spoken in the Nilgiri Hills in India, it was more important to record the consonants in final position, as this language contrasts three places of articulation word initially, but six places in final position.

Devising an appropriate set of material to record is the key element in investigating the phonetic structures of any language. Our normal procedure before going into the field is to examine the phonology of the language as thoroughly as possible, considering dictionaries, grammars, and whatever is available, and often consulting extensively with other linguists who have been working on the language. Then, with a good knowledge of the probable phonetic structures and a preliminary word list in hand, we go into the field. We spend a considerable amount of time checking this word list with one or two language consultants, often revising it extensively as we hear the language being spoken. We profit extensively from the comments of native speakers on our own attempts to produce the sounds. We find that we cannot over prepare before going into the field, and we cannot over check a list of words in the field before making a recording.

3. SPEAKERS

When we are satisfied that we have an appropriate set of material we record several speakers saying these words. A language is the property of a group of speakers, and it cannot be validly illustrated without recording several individuals. Any one speaker may have some trick of the voice that is not typical of the language as a whole. Successful comparisons of one language (or one dialect) with another require measurements of the phonetic characteristics of a group of people, so that the mean and standard deviation of various properties can be determined. We like to record at least six women and six men, although we sometimes have to settle for less. It is important to record both sexes as, quite apart from the physiological differences in the vocal tracts, the men may have a different dialect from the women. In Pirahã, the language previously mentioned in which there are only eight consonants, the women have a palatal fricative $\ç$ for one of those consonants, where the men have **h**.

Speakers need to be carefully chosen from several points of view. Firstly good dialect control is necessary. It is usually possible to ask speakers whether they regard themselves as typical speakers of the local dialect, and also to ask them whether they can hear any different local characteristics in the speech of the other individuals being recorded. If possible, we try to avoid speakers that are hard of hearing. Profoundly deaf speakers almost always have different speech characteristics, and even the slightly deaf may have been affected. Shy and easily embarrassed speakers are also difficult to record. Middle aged or younger speakers with strong voices and all their wits and teeth make the best language consultants for phonetic purposes, but they are often not available in endangered languages. We have to take what we can get, which sometimes includes speakers who have to be disregarded in the final account of the language.

4. AUDIO RECORDING

The principal data in any study of the phonetic structures of a language are the audio recordings. At the moment the best way of making such recordings is on a DAT (Digital Audio Tape) machine. A short while ago analog tape or cassette recorders were considered most suitable; and in a year or two recordings will probably be made directly onto a computer and then transferred onto a more permanent storage medium. DAT and computer recordings have the unfortunate property that they can become completely unusable because of a single small fault in the recording medium that would cause only a local noise in an analog magnetic recording. It is therefore vital to make more than one backup copy of any DAT or computer recording. Fortunately copies of digital recordings can be made without any loss of quality. It is probably advisable to make analog copies of digital recordings as well. The field is changing so rapidly that it is difficult to reproduce digital records made more than a dozen years ago. But magnetic recordings made in 1959 were still usable for producing spectrograms in 1995 for a survey of sounds of the world's languages (Ladefoged and Maddieson 1996); and the old wax or tin cylinders made at the turn of the century may be noisy but they are still a usable source of material for a contemporary dictionary (Bergsland 1994).

The type of microphone used for making the recording is just as important as the type of recording machine. Many recordings of endangered languages have to be made in noisy surroundings, in villages in which there are children crying, cocks crowing, and all the sounds of community life. The best way of dealing with this situation is to use a close-talking noise-canceling microphone, held by a head mount a few centimeters to the side of a speaker's lips.

Audio recordings are the prime source of data for the detailed description of the phonetic structures of every language. Of course much information is gained from direct listening and observation of speakers. But far more is learned from recordings that can be played over and over again, in small sections, and, if appropriate, at a slow rate without loss of quality. We make a great deal of use of computer systems that permit audio recordings to be examined in detail while still in the field as well as on returning home, where acoustic analyses of these recordings may be carried out.

Other recording techniques

Making a recording for later acoustic analysis is not the only way of getting data on the sounds of a language. If you want to study how the sounds are made, there are some additional techniques that can be used. The most obvious is a video camera for recording lip movements. Figure 1 shows the contrast between a bilabial (using the two lips) and a labiodental (only the lower lip and the upper teeth) in Logba, a language spoken in Ghana.

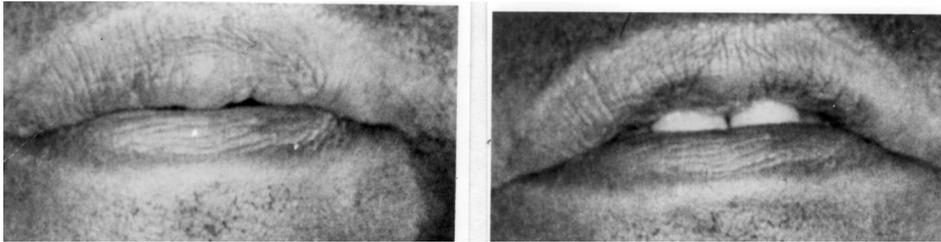


Figure 1. The lip positions in the consonants in two Logba words, $\dot{u}\beta\acute{a}$ "measles" and $\dot{u}v\acute{a}$ "side".

In addition it is possible to make records of the articulations involving the tongue by using palatography. In the form of the technique that we use a (tasteless) mixture of olive oil and charcoal powder is painted onto the roof of a speaker's mouth. The speaker then says a word containing the sound in question, and some of the mixture is transferred from the roof of the mouth to the tongue, revealing which parts made contact with the roof of the mouth during the pronunciation of the whole word. The contact pattern on the tongue is then photographed so as to make a permanent record. Subsequently the roof of the mouth is cleaned, more of the mixture is painted onto the tongue, the word is repeated and the contact pattern on the roof of the mouth is photographed using a mirror inserted at 45° to the plane of the upper teeth. Photographs of a speaker of Toda, a Dravidian language spoken by about 400 people in the Nilgiri Hills in India, are shown in Figure 2. This language has four different sibilants (s-like sounds).

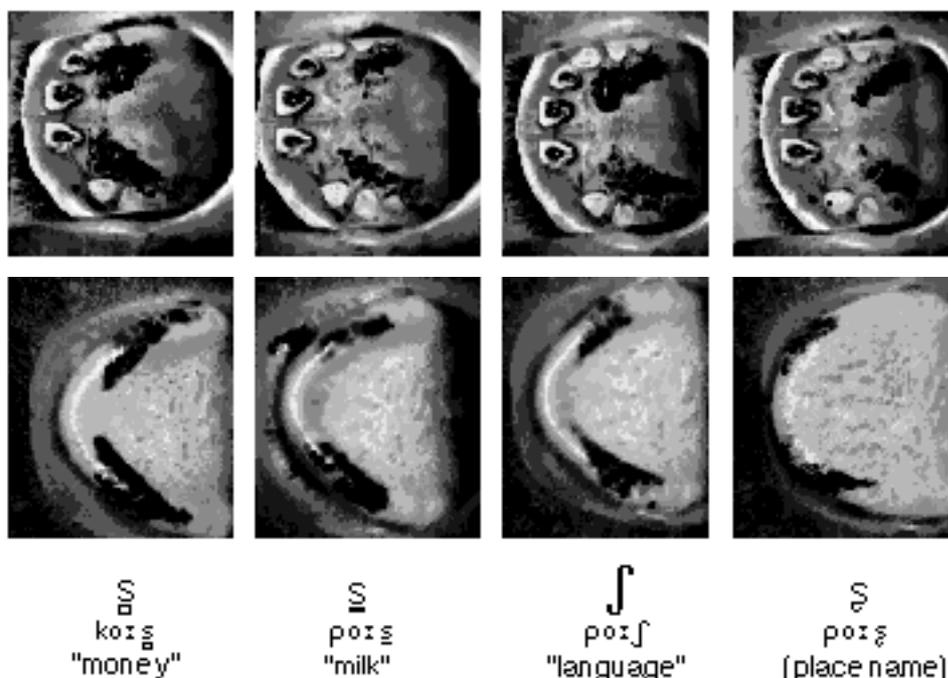


Figure 2. Palatograms and linguograms of four Toda words exemplifying the four voiceless fricatives in the language

Palatography is a good way of defining the precise place of articulation of a sound, but it is not the only way of providing archival information about how a sound was made. In the project being discussed we have obtained a great deal of information from aerodynamic records. Figure 3 illustrates the difference between a voiceless nasal η in Burmese and in Angami, a related language spoken in northeastern India. The top line shows the audio waveform the second line the airflow from the mouth, and the third line the airflow from the nose. In the middle of the Burmese word, the airflow stops coming out of the mouth at the same time as it starts coming out of the nose. The same is true but to a lesser extent in Angami. In Burmese the vocal fold starts vibrating at the time indicated by the second vertical line, while air is still coming out of the nose, and airflow from the mouth does not begin till later (at the time indicated by the third vertical line). But in Angami airflow from the mouth starts abruptly, at the time of the second vertical line, while air is still coming out of the nose, and the vocal folds do not start vibrating till later. These are similar but quite distinct sounds, as can be recorded by means of the aerodynamic records.

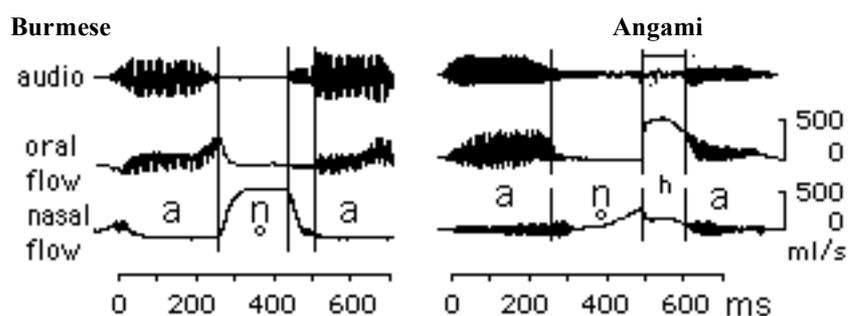


Figure 3. Aerodynamic records showing the difference between a voiceless nasal η in Burmese and in Angami.

4. THE TASK OF RECORDING THE SOUNDS OF ENDANGERED LANGUAGES

If we ask how many different sounds there are in the 7,000 present day languages, the answer depends on what we mean by a different sound. Phoneticians would say that two sounds are definitely different if they can change the meaning of a word in a language — they belong, in the technical jargon of linguistics, to different phonemes. When we try to say whether a sound in one language is the same as a sound in another language, the situation becomes more complicated. Generally speaking, when there are no contrasts to rely on, we can only guess that a particular sound is sufficiently different from all other sounds to be capable of changing the meaning of a word. In the 10 most widely spoken languages there are about 100 distinct consonants. But there are probably about 600 consonants in the languages of the world, many of them in small languages that are endangered,

Counting the number of vowels in the world's languages is even more difficult. Some languages have as few as three vowels. Many of the well-known languages have a large number of vowels. The most common number of vowels is 5. The vowels of about 20% of all languages can be symbolized by **a, e, i, o, u**. This includes such widely different languages as Spanish, Hawaiian, Swahili, and Japanese. Some European languages have a much larger number of vowels. There are Germanic dialects that will not be spoken a few years hence that have 24 vowels, 12 distinct qualities, each of which can be short or long (Gussenhoven, C. & Aarts, 1999).

In addition to varying in quality and length, vowels can differ in tone (pitch). Chinese, Cantonese are well known examples. Most languages use pitch in this way; there are more tone languages than languages like English and Italian that do not have significant tones. Some languages also have vowels that differ in the quality of the voice (vocal fold action). They may distinguish a regular voice, a creaky voice and a breathy voice. If we consider all the possible combinations of vowel qualities along with differences in tones and voice quality, there are probably about 200 different vowels.

Many sounds are fairly common but do not occur in the most widely spoken languages. Ejectives — sounds in which the closed glottis rises and pushes out air — occur in about 20% of all languages, implosives in which the glottis descends, occur in about 10% of all languages. Endangered languages may contain even more unusual phenomena, e.g. Oro Win, spoken in Brazil has a kind of [t] combined with a trill of the lips. This kind of trill has not been found in any other language, except for the related language Wari'.

In 100 years there will be far fewer languages, and we'll have less opportunity to observe what the human mind can use to communicate. Many of the sounds of the world's languages are fading away.

BIBLIOGRAPHY

Bergsland, K. (1994) *Aleut Dictionary*. Fairbanks, Alaska: Alaska Native Language Center, University of Alaska.

Gussenhoven, C. & Aarts, F. The dialect of Maastricht *Journal of the International Association*, 29, 1999, pp.55-56.

Ladefoged, P. & Maddieson, I. *The sounds of the world's languages*. Blackwells, 1996.