Acoustic and auditory analyses of Xhosa clicks and pulmonics

Keith Johnson
University of Alabama, Birmingham

Introduction

Usually, the phonological natural classes predicted by acoustic properties are the same as those predicted by articulatory properties.[1] For example, labial sounds may be characterized either by the low frequency emphasis in their acoustic spectra or by the presence of a constriction at the lips. Similarly, the presence of turbulent noise in the acoustic wave form of a sound defines roughly the same class of sounds as those which have narrow articulatory constrictions (fricatives). In these and most other cases, the predicted natural classes are the same whether one chooses to focus on sounds or articulations. Thus, for most features phonetic implementation may be defined in either acoustic or articulatory terms.

Clicks are interesting because, although the natural classes among them predicted by acoustic and articulatory properties are not different, the acoustic properties of clicks predict different cross-classifications of click and nonclick sounds than do their articulatory properties. Therefore, we expect to find that either the acoustic or the articulatory properties of clicks correctly predict the cross-classifications with pulmonics which are needed for the statement of linguistic phenomena.

This paper reports a quantitative analysis of acoustic and auditory properties of clicks and pulmonics in Xhosa. The study focused on the cross-classification of clicks and nonclicks and extended earlier work on the acoustics of clicks (Traill, 1992a, 1992b, Sands, 1991, Ladefoged & Traill, to appear) in two ways. First, the paper reports the use of two quantitative procedures (cluster analysis and moments analysis) to discover the similarities among sounds. Theoretical and perceptual studies (Stevens & Blumstein, 1978; Traill, 1992b) have suggested certain acoustic characterizations, but the most commonly used methods for discovering acoustic properties are still quite subjective. Cluster analysis provides a rigorous basis for statements of acoustic/auditory similarity, and moments analysis provides an objective basis for the description of spectral properties. Second, the study tested the hypothesis that auditory spectra provide a better basis for the comparison of speech sounds than do acoustic spectra. This hypothesis is based on the assumption that if we can simulate the properties of the human peripheral auditory system, we can get a representation that is closer to the listener's experience of speech than is the acoustic spectrum.

Classification of clicks

(1a) shows an articulatory and acoustic classification of coronal clicks based on extensive phonetic data. Note that the feature [lateral] is redundant for the lateral click and the somewhat antiquated feature [delayed release] is used to distinguish the affricated and unaffricated clicks. An alternative representation in terms of aperture sequences (Steriade, 1992) is possible if we assume, following McDonough (1993), that there is a lateral aperture, and that frication on lateral aperture is redundant for voiceless lateral sounds. The revised classifications are shown in (1b).

---

[1] I am assuming that phonetic similarity among sounds "predicts" phonological natural classes. Since the sounds in a natural class typically share some phonetic property, it seems reasonable to hypothesize that if we find that a phonetic property is shared by certain sounds these sounds may function as a phonological natural class.

(1a) Articulatory and acoustic classifications of clicks. Adapted from descriptions given by Traill & Ladefoged (1984, to appear), Traill (1985), and Sands (1991).

Articulatory properties	| Dental	|| Lateral	! Alveolar	ǂ Palatal	Acoustic properties
laminal	+	-	-	+	acute
delayed release	+	+	-	-	noisy
lateral		+			

(1b) Alternative representation with aperture sequences.

Articulatory properties	| Dental	|| Lateral	! Alveolar	ǂ Palatal	Acoustic properties
laminal	+	-	-	+	acute
release aperture	A_f	A_{lat}	A_{max}	A_{max}	noisy release

From the articulatory properties in (1) we can predict some click/nonclick cross-classifications. For example, dental and palatal clicks should pattern with laminal pulmonic consonants such as dentals as opposed to apical consonants such as alveolars. Of course, since clicks are coronal they should pattern with other coronal consonants. However, it is important to note that clicks have a velar or uvular closure in addition to, and simultaneous with the anterior closures shown in (1), so clicks are specified for two articulators ([coronal] and [dorsal]) and may participate in processes involving either. The acoustic properties shown in (1) predict the same natural classes among clicks as the articulatory properties. However, cross-classification of clicks and non-clicks predicted by the acoustic properties are different from those predicted from articulation. For example, the lateral and alveolar clicks are acoustically [+grave] and thus should pattern with other [+grave] sounds.

Acoustic versus articulatory features

Traill (1992a) argues that natural classes of clicks and pulmonic consonants predicted by acoustic properties correspond to the classes at work in synchronic phonological processes and diachronic sound changes, while classes defined by articulatory properties do not. He discusses two sets of data.

In the first, a synchronic process in !Xóõ, /a/ becomes [i] when preceded by [|] or [ǂ] and followed by [i], as the examples in (2) illustrate (note that there is a partial raising and centralizing of /a/ to [ɐ] after [⊙], [!] and [||]). Traill (1992a) also notes that 'there are a handful of cases' that show a change from /a/ to [i] "following the non-click consonants /t th s/". An earlier account of the process, outlined in Traill (1985), indicated that the class of sounds which conditions /a/-raising is "dental" and also includes /l/ and /n/.

(2) /a/-raising in !Xóõ (from Traill, 1992a).

/|ā̤i/ → [|ī̤i]	'aardwolf' (Proteles cristatus)
/ǂái/ → [ǂíi]	'steenbok' (Raphicerus campestris)
/⊙á'i/ → [⊙ɐ́'i]	'abomasum' (part of ruminant's digestive system)
/!ái/ → [!ɐ́i]	'sp. of tree' (Zisyphus mucronata Willd.)
/||à̤i/ → [||ɐ̤̀i]	'old (Class.1)

These data illustrate that [|] and [ǂ] pattern together in !Xóõ and must be grouped in a natural class contrasting with [⊙], [!] and [||], a pattern which can be predicted by either the articulatory or the acoustic properties of clicks ([± laminal] or [± grave]). Traill (1992a) argued that /a/-raising in !Xóõ is evidence that clicks and pulmonics should be cross-classified according to their acoustic properties rather than their articulatory properties because [|] and [ǂ] like dental pulmonics and /i/ are [-grave]. However, the descriptions given in Traill (1985) make it clear that the dental sounds in !Xóõ, like dentals in other languages, are laminal. Therefore, the articulatory description given in (1) makes the same predictions for /a/-raising in !Xóõ as the acoustic description, and Traill's (1992a) argument does not hold.

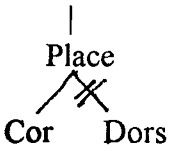
The second set of data discussed by Traill (1992a) provides stronger support for his hypothesis. These data (3) illustrate a diachronic process in Khoe dialects in which [!]-series clicks are replaced by pulmonic velars and [ǂ]-series clicks are replaced by pulmonic palatals.

(3) Diachronic click replacement in Khoe dialects (from Traill, 1992; Traill, 1986).

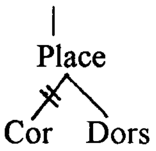
|Gwi	Ts’ixa	gloss
[!are]	[kare]	'cut into strips'
[!ŋaro]	[ŋgaro]	'chameleon'
[!ganee]	[ganni]	'chin'
[!hae]	[khae]	'pierce'
[ǂii]	[cii]	'call'
[ǂŋu]	[ŋɟuu]	'black'
[ǂgoa]	[ɟua]	'ash'
[ǂhuni]	[chuni]	'elbow'

Informally the sound change in Ts’ixa involves a change of nonaffricated clicks to pulmonics, and since clicks are specified for both [dorsal] and [coronal] we may formally express the change as a delinking process in which one of the place nodes in the click delinks. However, this formalization requires two types of delinking as shown in (4). The dorsal node delinks in the palatal click [ǂ] resulting in a non-click palatal, while the coronal node delinks in the alveolar click [!] resulting in a non-click velar. Without some inelegant stipulations this account cannot explain why coronal delinking applies to the alveolar click while dorsal delinking applies to the palatal click. Since the data in (3) indicate that at some point in the history of Ts'ixa there were rules like (4) in the synchronic grammar, these data suggest that phonological cross-classifications between clicks and pulmonic consonants can not be predicted from the articulatory properties of the sounds. However, since both the alveolar click [!] and the velar stop [k] are classified acoustically as [+grave], click replacement in Ts'ixa is consistent with Traill's hypothesis that the phonological cross-classification of click and nonclick consonants is based on acoustics, not articulation.

(4) ǂ > c



! > k



A rough sketch of an analysis based on the acoustic properties of clicks and pulmonics is shown in (5). The distinction between air stream mechanisms is represented as a difference in the root node and the process is simply that clicks become pulmonics. In this account, cross-classifications between clicks and pulmonics are based on what Jakobson, Fant & Halle (1963, henceforth JFH) called resonance features.

(5) ǂ > c
click → pulmonic
| |
acute acute

! > k
click → pulmonic
| |
grave grave

What is missing from this argument is convincing evidence that click and nonclick consonants in a click language pattern together acoustically. It has been noted previously (Traill, 1992a, 1992b; Traill & Ladefoged, to appear; Sands, 1991) that the spectrum of the alveolar click [!] has a concentration of energy in lower frequencies (compared to the spectra of other clicks), just as the spectrum of [k] has more low frequency energy than is typically found in the spectrum of [t]. These observations suggest that [k] and [!] share a phonetic property, but this important prediction has not been tested in any previous study of the acoustics of clicks. The study reported here was designed to fill this gap.

Methods

The speaker was Ncediwe Mdunyelwa. She is a native speaker of Xhosa from Capetown, South Africa who was a visiting scholar at UCLA during the 1991-1992 academic year. The recording was made in February, 1992 at the UCLA phonetics laboratory by Sujin Yi as an illustration of a project for an introductory phonetics course. A word list illustrating most of the distinctive sounds of Xhosa was recorded twice. Both recordings of a subset of this list (shown in 6) was analyzed in the present study.[2]

(6) List of Xhosa words used in the study.

sound	orthography	transcription	gloss
p^h	phala	[p^hála]	'go fast'
t^h	thala	[t^hála]	'ledge of the rock'
k^h	khala	[k^hála]	'cry out'
ɸ	fukama	[ɸuq'áma]	'lie on'
s	salisa	[salísa]	'make remain'
x	ruzula	[xuzúla]	'pull away'
ǁ	xela	[kǁɛ́la]	'tell, say'
ǀ	cuba	[kǀúba]	'tobacco'
!	qaba	[k!ába]	'paint'

---

[2] These words present a stiffer than usual challenge for a classification scheme because of varying coarticulation from the different contextual vowels.

The recorded utterances were digitized (20kHz, 12 bits) and in each word a 25.6 ms (512 samples) section of the acoustic wave form was identified for further analysis. In stops and clicks the section of wave form was centered around the release burst (as in Stevens & Blumstein, 1978). In fricatives the window was centered around the peak amplitude of the frication. FFT and LPC spectra were calculated from these wave form segments (the LPC had a filter order of 22).

Following Stevens & Blumstein (1978), the LPC spectra were taken to represent the acoustic spectrum. The auditory spectra were constructed from the FFTs using an auditory model. Three important characteristics of the human auditory system were modeled. First, the (100) bandpass filters used in the model were equally spaced along a nonlinear auditory frequency scale (the Bark scale, Zwicker, 1961; Schroeder, Atal & Hall, 1979). Second, the bandwidths of the filters increased as the center frequencies increased (Patterson, 1976). Because of these aspects of the filter bank, lower frequencies were emphasized relative to higher frequencies, and small differences at high frequencies were wiped out because the high frequency filters had large bandwidths. The third aspect of the human auditory system captured by the model is the relative sensitivity of hearing at different frequencies. This was modeled by applying an equal loudness contour (Fletcher & Munson, 1933) to the filter outputs (see JFH, p.27). This increased the amplitudes of frequency components between 1000 and about 3000 Hz, relative to other parts of the spectrum.

After the acoustic and auditory spectra had been calculated, the DC offset in each spectrum was removed by subtracting the average amplitude of the spectrum from each point. (The LPC spectra were down-sampled from 512 points to 100 points, so that the acoustic and auditory cluster analyses were based on the same number of points.) With the DC offset removed the focus of comparison was on the shape of the spectrum ignoring overall amplitude differences. Then the average spectrum (of the two productions) was calculated for each of the 9 consonants. Ward's method of cluster analysis (Everitt, 1980; SAS, 1982) was used to group the spectra into hierarchical similarity spaces. Spectral moments (Forrest, et al., 1988) of the average acoustic and auditory magnitude spectra were calculated without removing the DC offset.

Results

The acoustic spectra for stops (top left panel of Figure 1) agree with the theoretical predictions outlined by Stevens & Blumstein (1978). The spectrum of [t] (drawn with the thick solid line) has more high frequency energy than those of [p] or [k], and the spectrum of [k] (drawn with a thin solid line) shows a concentration of energy between 1 and 2 kHz. The same basic aspects can be seen in the acoustic spectra of the fricatives (middle left panel in Figure 1). However, [x] (thin solid line) has a peak near 4 kHz which is unexpected. In the acoustic spectra of the click bursts (bottom left panel of Figure 1), the alveolar click [!] shows a low frequency peak similar to the peak seen in [k]. The dental [|] and lateral [||] clicks have similar looking spectra, and although the dental click has more high frequency energy than the lateral click, it does not have the same high frequency emphasis seen in [t] and [s].

One problem with these spectra is knowing what to focus on. For instance, the acoustic spectrum of [p] shows a small peak at about 1 kHz which may or may not be important for the identification of [p]. The lowest peak in the spectrum of [ɸ] is at 1.8 kHz which may or may not correspond to the lowest peak of [p]. The alveolar click [!] has a low frequency peak similar to that of [k], but [x] and [!] have prominent peaks at about 4 kHz. Is the similarity of [!], [k] and

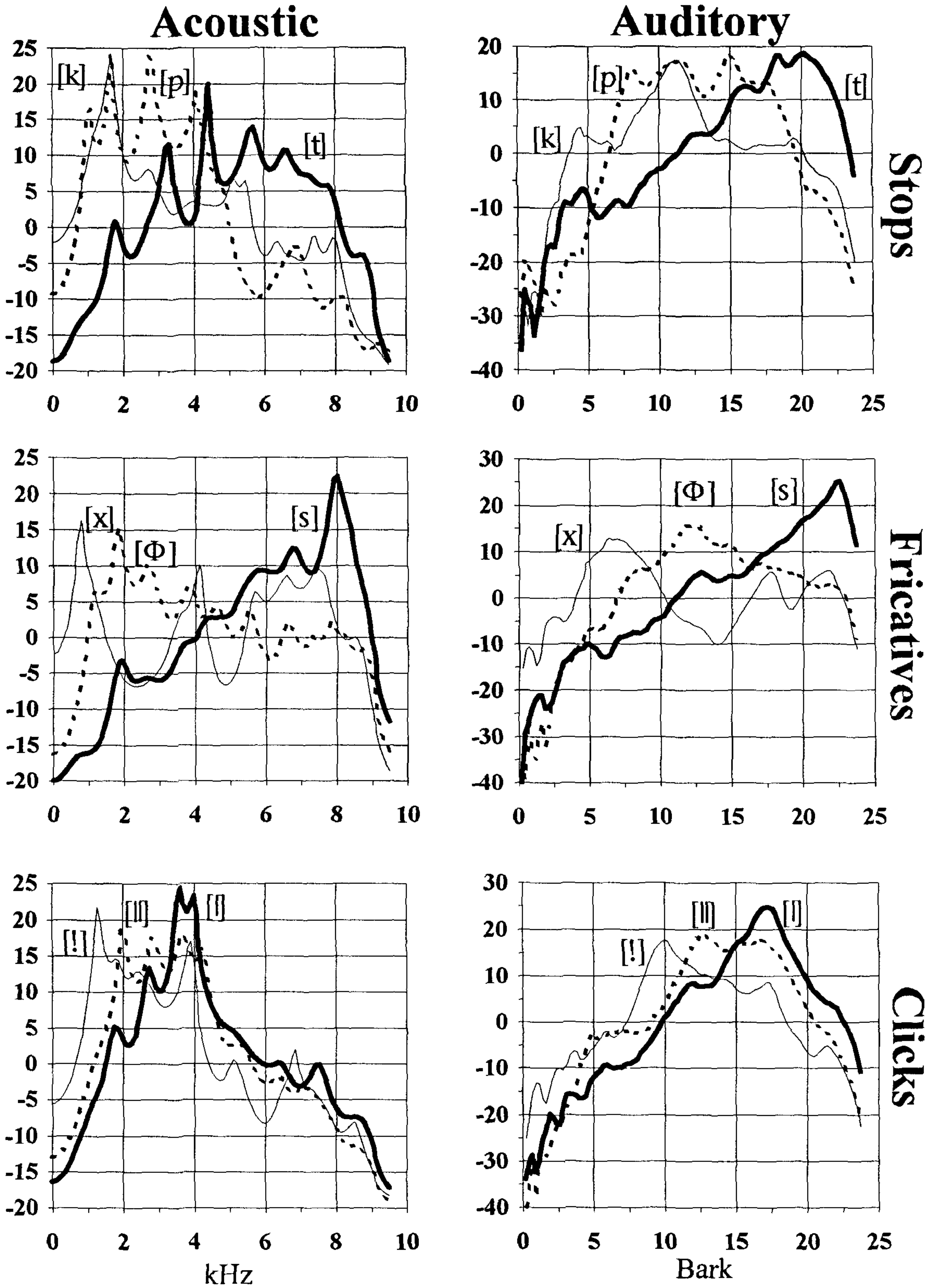


Figure 1. Acoustic and Auditory spectra of Xhosa stops, fricatives and clicks.

[x] below 2 kHz enough to offset their differences in the higher frequencies? Without *a priori* assumptions it is not clear how to assess the differences and similarities between these acoustic spectra.

The auditory spectra are shown in the right-hand panels of Figures 1. By compressing the high frequencies and expanding the low frequencies, some similarities among the pulmonic consonants are enhanced. Dentals have high frequency emphasis, labials have mid frequency emphasis, and velars show at least two broad regions of energy with one lower than the lowest peak in the labials. Additionally, the clicks show some similarities with the pulmonic consonants. The dental click [|] looks more like a pulmonic dental in the auditory spectra than it does in the acoustic spectra. Also the alveolar click [!] has a lower frequency peak than either of the other two clicks. So, some aspects of the spectra are visually enhanced in the auditory display, and the enhancement is non-arbitrary because it is the product of a psychoacoustic model of auditory processing. However, our visual inspection of these displays is still subjective. Just as with acoustic spectra, it is necessary to decide what aspects of the auditory spectra to focus on.

Figure 2 shows the results of the cluster analyses. The bottom panel shows that the auditory spectra of stops and fricatives were grouped by place of articulation. This analysis also produced an interesting picture of click/non-click cross-classifications. The alveolar click [!] was grouped with the velars as Traill (1992a) predicted. Note, however, that there was no evidence for [grave] in the JFH sense because labials did not pattern with velars. The analysis of acoustic spectra (shown in the top panel of Figure 2) is problematic. The natural classes predicted in the

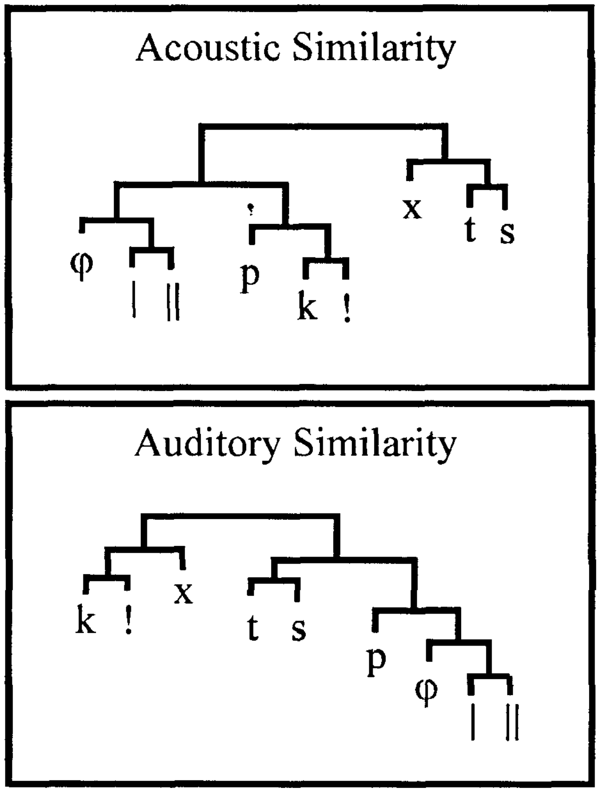


Figure 2. Results of the cluster analyses. Top panel: cluster analysis of LPC spectra. Bottom panel: cluster analysis of auditory spectra.

acoustic analysis are phonologically bizarre. The acoustic spectra of [p] and [k] were grouped together as the JFH definition of [grave] would predict, but the analysis did not result in coherent classification system because place of articulation among the pulmonic sounds was not preserved. For instance, the analysis incorrectly predicts that [x], [t], and [s] form a natural class.

Spectral Properties.

It is important to note that the cluster analysis of auditory spectra grouped the alveolar click [!] with the velar pulmonics. This supports Traill's argument that the historical association of alveolar clicks with velar pulmonic consonants in the Khoe dialects involves sound similarity. However, cluster analysis has an important possible limitation when applied to speech spectra. It focuses on spectral details rather than general (and perhaps somewhat abstract) patterns. So, spectral differences which may be due to a spectral frequency shift may affect cluster analysis as much as a difference in overall spectral shape. JFH pioneered some approaches to extracting spectral properties which reflect the general shape of the spectrum arguing that such general patterns are more linguistically relevant than absolute detailed spectral matching.

Euclidean distance matrices based on the acoustic and auditory spectra (shown in 7) support the JFH viewpoint. These matrices suggest that the similarity relations among the spectra may have been underspecified in the cluster analysis.[3] The matrix in (7a) shows that the acoustic spectrum of [|] was about as similar to the spectrum of [t] as to [p], and most similar to [ϕ], while the acoustic spectrum of [||] was very similar to both [p] and [ϕ]. The dental [|] and lateral [||] clicks had very similar spectra (especially above 4 kHz). On the other hand, although both [|] and [||] group with the pulmonic labials in the cluster analysis of auditory spectra, (7b) suggests that [|] is more similar to the dental sounds in the auditory analysis than in the acoustic analysis. The dental click [|] was about as similar to [t] (RMS=66dB) as to [||] (RMS=65dB) in spite of the fact that [|] was produced in the environment of a round vowel while [t] was not.

Yet cluster analysis missed the similarity between [|] and the pulmonic dentals that is apparent in (7b). In order to understand this it is important to realize that initially cluster analysis uses distances as in (7), but each time a cluster is identified the elements are merged and the average spectrum is used to represent the cluster. For example, the first step in an analysis of the auditory data would be to merge [t] and [s] since the distance between their spectra is smaller than any other distance in (7b). Then a new distance matrix is constructed with the average of the [t] and [s] spectra representing the new cluster. Note in (7b) that although [|] is more similar to [t] than it is to [||], it is not more similar to the average of [t] and [s] than it is to [||]. Therefore, cluster analysis puts [|] in with [||]. It seems inevitable that the distances in (7) reflect not only important phonetic differences, but also random variance which listeners disregard. So a more robust phonetic analysis may require that we pull out abstract spectral properties rather than rely on the unanalyzed spectrum. One method of extracting spectral properties which has been useful in previous research is to calculate the statistical moments of the spectrum.

The idea of using spectral moments for the quantification of spectral shape features in speech analysis appears to have originated with JFH (1963) and has been pursued more recently by Forrest et al. (1988). In moments analysis, some abstract properties of an acoustic or auditory spectrum are described by calculating the statistical "moments" of the spectrum. The first moment

---

[3] It should be noted that the cluster analyses used distance estimates derivable from those shown in (7).

(7) Root Mean Square (RMS) distances (in dB) between acoustic spectra and auditory spectra.

(a) Acoustic Distances

	t	k	ɸ	s	x	|	!	||
p	143.27	70.83	86.62	191.25	135.48	87.42	50.70	60.27
t		118.62	82.24	**78.51**	99.85	**86.80**	131.55	**105.80**
k			65.38	160.04	100.67	97.42	50.58	77.61
ɸ				114.83	94.41	68.51	66.74	65.19
s					115.24	**136.25**	171.55	**159.23**
x						115.24	109.42	124.86
|							85.68	**45.56**
!								63.21

(b) Auditory Distances

(b)	t	k	ɸ	s	x	|	!	||
p	146.56	99.85	79.46	171.77	155.95	108.90	72.68	75.33
t		128.56	93.36	**55.05**	138.52	**66.05**	137.12	**105.80**
k			75.98	150.38	101.25	126.44	55.59	95.48
ɸ				109.05	133.10	81.31	80.01	60.16
s					150.56	**100.85**	159.20	**136.83**
x						158.41	115.08	155.01
|							117.73	**65.45**
!								89.69

is the spectrum's weighted **mean.**[4] If the spectrum has more high frequency energy than low, the spectral mean will be high. If low frequency energy predominates, the spectral mean will be low. JFH suggested that the spectral mean is one possible correlate of the feature [grave] but they note that if we define gravity in terms of the spectral mean all vowels would be [+grave] while almost all consonants would be [-grave] and any value that the relative values of the spectral mean has in classifying sounds is lost because of these large absolute differences. The second moment is the spectral **variance**. JFH suggested the second moment as a possible correlate of the [compact]/[diffuse] distinction, but later research (Forrest et al., 1988) did not support this proposal. In the present study also the second moment did not produce a coherent division of the sounds for either the acoustic or the auditory spectra. The third moment of the spectrum is its **skew**. If the energy in the spectrum falls equally on both sides of the mean, the spectrum has a skew of 0. If there is more energy above the mean than below it the spectrum has a positive skew, and conversely, spectra with more energy below the mean than above it have negative skew. This measure is relevant for the definition of [grave] and was suggested by JFH as an alternative to the spectral mean because it is not sensitive to the absolute value of the mean. The fourth moment, **kurtosis**, is a measure of the peakedness of the spectrum which Forrest et al. suggested as a correlate of the compact/diffuse dimension instead of using variance as suggested by JFH.

---

[4] Jakobson, Fant & Halle (1963) used the term "center of area". Others call this the spectral "center of gravity".

Forrest et al. (1988) found that the acoustic spectrum of [t] had a higher mean than did those of [p] and [k]. That result is replicated in Xhosa (Figure 3, top panel). They also found that the acoustic spectrum of [t] had lower skew than [p] or [k], and this result was also replicated. However, the spectrum of [k] did not have greater kurtosis than the other stops as would have been predicted from the JFH description and Forrest et al.'s findings.

As in the cluster analysis of the acoustic spectra, the results of the moments analysis do not give a very coherent picture of Xhosa clicks and pulmonics. The velar fricative [x] is closer to [t] than to [k], and the labial fricative [ɸ] is closer to [k] than to [p]. Forrest et al. used spectral moments to classify fricatives (but not to cross-classify fricatives with stops) but reported only the classification results not the moments data which served as input to the analysis, so we don't know if the lack of coherence seen in the present analysis differs from their findings or not.

The results of the moments analysis of the auditory spectra also conform to Forrest et al.'s findings.[5] The auditory spectrum of [t] had a higher mean than did those of [p] and [k]. Also [k] and [p] were more positively skewed than [t] as they tended to be in Forrest et al.'s study. Interestingly, the spectral mean, skew and kurtosis of these Xhosa consonants were correlated with each other. The correlation between mean and skew was strongest ($r^2=0.938$), while the correlation between mean and kurtosis was only slightly weaker ($r^2=0.763$). This suggests that there is only one effective dimension among these measurements, and that the spectral mean, skew, and kurtosis are simply different ways of measuring the same thing for these particular sounds. It isn't clear whether this dimension is best expressed as the spectral mean (an absolute value) or skew (a relative value), or whether the same dimension will capture the same sorts of distinctions among other sounds as well.

As was the case in the cluster analysis, moments analysis of auditory spectra produced a coherent picture of the Xhosa pulmonics (stops and fricatives produced at the same place of articulation are close together in the bottom panel of Figure 3) and an interesting pattern of click/nonclick cross-classifications. However, unlike the cluster analysis, spectral moments analysis groups the dental click [|] with the dental pulmonic sounds, while (as in the cluster analysis) the lateral click [||] is closer to the labial pulmonics, and (as suggested by Traill, 1992a and the cluster analysis) the alveolar click [!] is quite similar to [k]. One other difference between the cluster analysis of the auditory spectra and this moments analysis is that the moments analysis resulted in a clear division between sounds that have a strong high frequency component ([t s |]) and other sounds, corresponding to the JFH description of the acute/grave distinction.

This coherent division of the sounds is also interesting because the consonants were produced in different vowel environments. Sands (1991) has shown that Xhosa click spectra are influenced by labial coarticulation, so the fact that the dental click [|] patterned with the dental pulmonics in the moments analysis is impressive because the click was produced before [u] while [t] and [s] were both produced before [a]. Similarly, the fricatives [ɸ] and [x] patterned with their homorganic stops in the auditory moments analysis despite coarticulatory rounding from the following [u].

---

[5] Forrest et al. (1988) used a Bark scale to perform an "auditory" analysis. It should be noted that they did not use an auditory model as I have here, but rather simply changed the frequency scale of the acoustic spectrum.

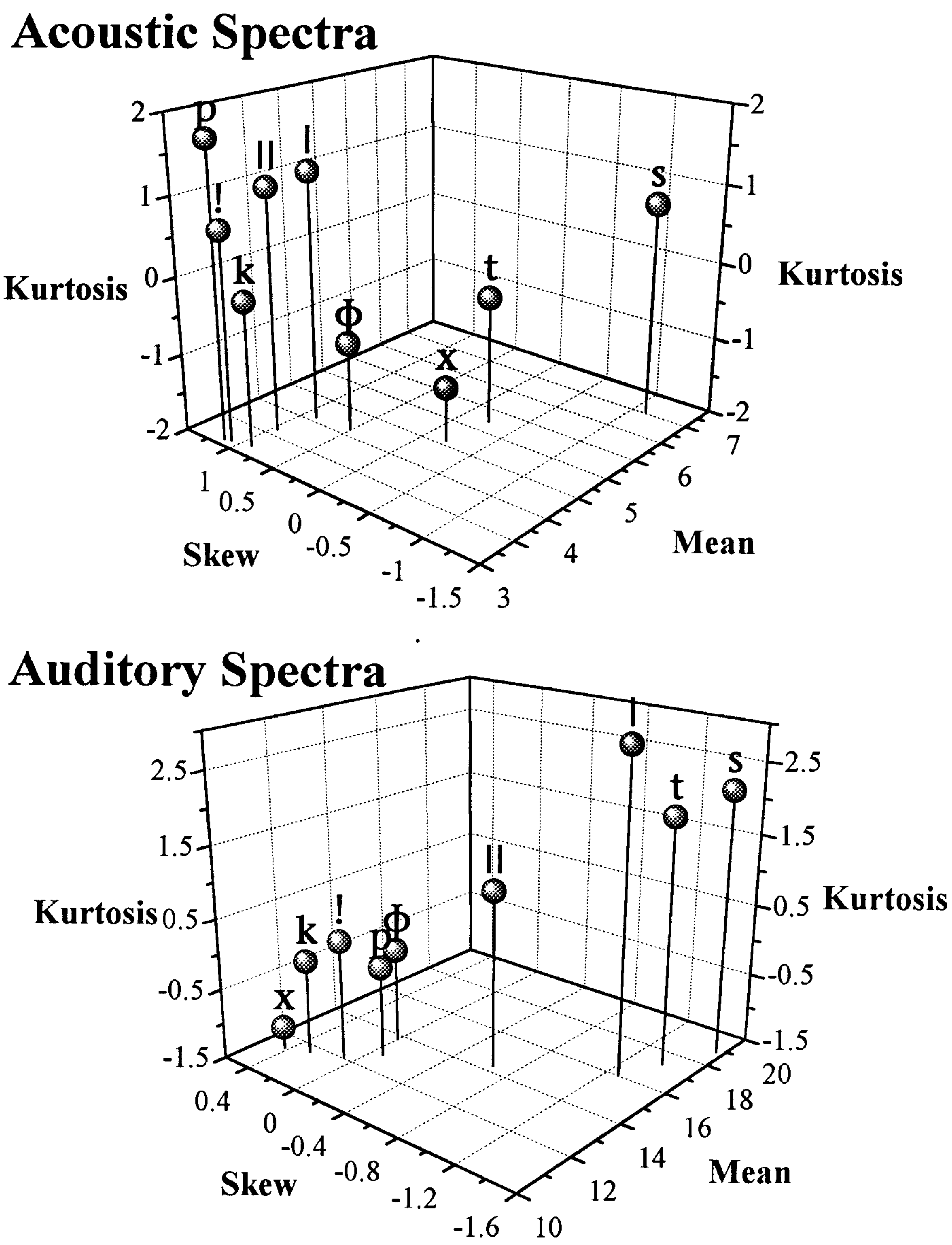


Figure 3. Results of the spectral moments analysis of the average acoustic spectra (top) and average auditory spectra (bottom).

Conclusions

The main finding of this study was that the alveolar click [!] and the velar stop [k] as produced by a native speaker of Xhosa were objectively similar to each other both acoustically and auditorily. This finding supports Traill's (1992a) contention that the proper statement of click replacement in Ts'ixa requires acoustically defined features.

An additional important finding was that analyses of auditory spectra produced different results which differed from the results of identical analyses of acoustic spectra of the same sounds. There are two reasons to prefer auditory spectra over acoustic spectra in studies that aim to predict linguistic natural classes. First, auditory spectra are preferable *a priori* because they represent a level of representation which is closer to the listener's experience of speech sounds. Second, in practical terms, this study found that auditory spectra produced phonologically coherent results in both cluster analysis and in moments analysis despite the varying vowel contexts in which the consonants had been produced, while acoustic spectra of the same sounds did not.

Finally, this study also suggested that in the auditory spectra of voiceless consonants spectral mean, skew, and kurtosis are highly correlated and seem to reflect one dimension along which these sounds differed. It was suggested that this dimension may have been most closely related to either spectral mean or skew, and that only further research can determine the potential importance of this finding.

Acknowledgments

Thanks to Ian Maddieson for his efforts in organizing the 1992 West Coast Phonetics Forum at Berkeley, CA and for asking me to make a presentation, and to Peter Ladefoged for his insightful comments on an earlier version of the paper. A truncated version was also presented to the 67th annual meeting of the Linguistic Society of America. I appreciate the comments and suggestions I received at WCPF and LSA, and also from my friends and colleagues at UCLA and UAB. My address is: Dept. of Biocommunication, UAB, Birmingham, AL 35294-0019

References

Everitt, B.S. (1980) *Cluster Analysis.* 2nd Ed., London: Heineman Educational Books Ltd.

Fletcher, H. & Munson, W.A. (1933) Loudness, its definition, measurement and calculation. *J. Acoust. Soc. Am.* **5**, 82-108.

Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R.N. (1988) Statistical analysis of word-initial voiceless obstruents: Preliminary data. *J. Acoust. Soc. Am.* **84**, 115-123.

Jakobson, R., Fant, G. & Halle, M. (1963) *Preliminaries to Speech Analysis.* Cambridge, MA: MIT Press.

Ladefoged, P. & Traill, A. (1984) Linguistic phonetic description of clicks. *Lg.* **60**: 1-20.

Ladefoged, P. & Traill, A. (to appear) Clicks and their accompaniments.

McDonough, J. (1993) On the phonological representation of the feature 'lateral'. Paper presented at the 67th Annual Meeting of the Linguistic Society of America.

Patterson, R.D. (1976) Auditory filter shapes derived with noise stimuli. *J. Acoust. Soc. Am.* **59**, 640-654.

Sands, B. (1991) Evidence for click features: acoustic characteristics of Xhosa clicks. *UCLA Working Papers in Phonetics.* **80**, 6-37.

SAS (1982) *SAS User's Guide: Statistics 1982 Edition.* Cary, NC: SAS Institute Inc.

Schroeder, M.R., Atal, B.S. & Hall, J.L. (1979) Objective measure of certain speech signal degradations based on masking properties of human auditory perception. In B. Lindblom & S. Öhman (eds). *Frontiers of speech communication research*, London: Academic Press.

Steriade, D. (1992) Closure, release and nasal contours. To appear in M.Huffman and R.Krakow (eds.) *Nasality: Phonological and phonetic properties*, Academic Press.

Stevens, K.N. & Blumstein, S.E. (1978) Invariant cues for place of articulation in stop consonants. *J. Acoust. Soc. Am.* **64**, 1358-1368.

Traill, A. (1985) *Phonetic and Phonological Studies of !Xóõ Bushman*. (Quellen zur Khoisan-Forschung, 1.) Helmut Buske, Hamburg.

Traill, A. (1986) Click replacement in Khoe. In *Contemporary Studies on Khoisan in Honor of Oswin Köhler*. Rainer Vossen & Klaus Keuthmann (Eds.) (Quellen zur Khoisan Forshung, 5) Helmut Buske, Hamburg.

Traill, A. (1992a) Place of articulation features for clicks: Anomalies for universals. MS. University of Witwatersrand, Johannesburg, Department of Linguistics.

Traill, A. (1992b) The perception of clicks in !Xóõ. MS. University of Witwatersrand, Johannesburg, Department of Linguistics.

Zwicker, E. (1961) Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *J. Acoust. Soc. Am.* **33**, 248.