ACOUSTICAL CUES FOR VOICED AND BREATHY FINAL STOPS IN GUJARATI LANGUAGE

Christine Langmeier

München, Germany

The problem of recognizing and discriminating stops is well known to those working in the field of speech recognition. To discriminate between voiced and breathy final stops of monosyllables several acoustical parameters of the preplosive vowel have been examined as predictor variables. Classification scores up to 90% in judged groups proved as highly tolerable performance.
The results confirm Fisher-Jørgensen [1] who claimed that the amplitude of the first harmonic (H1) is most selective to voiced versus breathy phonation.
On the other hand we have to consider the working relations of several acoustical parameters as there are the overall intensity, the first formant and the bandwidth of the first formant.
In one case making use of the difference measure between the first and second harmonic , a technique discussed by Ladefoged et al [2] generates better classification scores than computing on the measurements of H1.

1. INCTRODUCTION

Intensive research has been carried out to investigate the acoustical cues of different stop categories.
For initial stops a great impact on the following vowel has been admitted. For example Schiefer [3][4] found that in Hindi language fundamental frequency and overall intensity are cues to differentiate four stop categories in word-initial position: 1) voiceless aspirated 2) voiceless 3) voiced and 4) breathy stops.
We might doubt that comparable oppositions hold also for a word-final position because the contrasts could be in some way neutralized. Can we notice a regressive assimilation on the foregoing vowel?
A word-final distinction of voiced and voiceless stops has been discussed in the anglo-american literature as a phenomenon of vowel and stop closure duration, e.g. by Wardrip-Fruin [5].
A voiced final stop should lengthen the foregoing vowel and should be characterized by a relatively short closure duration (it is understood that duration measurements are always relative and give evidence only in controlled surroundings).
Lately there appeared also spectral arguments by Van Summers [6]. The first formant was found to stay somewhat lower in a voiced environment.
That there is indeed a physiological basis for a fourfold stop distinction in indo-arian languages even in word-final position has been shown by Yadav [7].
Employing a fiberoptic technique he found four grades of glottis opening when a speaker of Maithili produced the test words.
At the moment of stop release the glottis is 1) still closed for voiced final stops 2) slightly open for breathy stops after a roughly periodic phase during the closure 3) slightly open for voiceless stops after an open phase during the closure and 4) "widely" open for voiceless aspirated stops.
The proposed investigation examines several acoustical cues of two final stop categories: voiced and breathy final stops in mono-syllables of Gujarati language.

2. METHOD

The recordings of five native Gujarati speakers were made in Baroda by L. Schiefer and Bh. Modi.

Modi (female, 53 years, grown up in Bombay and district Surat, living for 29 years in Baroda)
Dalal (female, 47 years, grown up in Baroda)
Shah (male, 40 years, grown up in Ahmedabad, living for 14 years in Baroda)
Patel (male, 29 years, grown up in Baroda)
Vora (male, 26 years, grown up in Bombay and Baroda).

The following meaningful words were produced several times (it came up to a sample size of about 120 words for each person with 8 to 9 occurances of the same word):

rab	a̲b^h
sad	sa̲d^h
vaδ	ma̲δ^h
rag	va̲g^h
mσd	mσ̲d^h
rσδ	rσ̲δ^h
d^sσg	d^sσ̲g^h

porridge	sky
voice	accomplish
fence	musical form
raga	tiger
joy	honey
cry	stubborn
world	fragrance

You may read the vowel with the underscore as a breathy vowel; the sigma-letter as the vowel schwa; the delta-letter as a retroflex dental stop and "d^s" as a palatal affricate.

3. SEGMENTATION

The computer aided segmentation of the single word oscillograms was as follows: The initial consonant, the vowel, the closure and the burst of the final stop were cut by visual and auditive inspection of the digitalized acoustical signal.
There was no need to define a transition phase between the vowel and the stop closure part of the signal, the cut was set as soon as the amplitude was small enough to corresponde with the articulatory closure phase and the typical oscillating pattern of the vowel blurred.
The beginning of the burst then was defined as a small but abrupt vertical impuls of the signal. The final phase of the stop melted into the tape noise and therefore its duration couldn't be measured. The duration of the closure was measured proportional to the duration of the initial consonant and the following vowel.

4. ACOUSTICAL ANAYLSIS

The acoustical analysis was performed with the facilities of the Institute of Phonetics, München.
For all speakers a pitch synchrone discrete Fourier transformation of the vowel part of each word was calculated and for three speakers also a formant analysis by LPC-procedure.

Alltogether the following acoustical parameters computed for each word could enter as canonical variables into the discriminant analysis:

- overall intensity
- fundamental frequency
- intensity of H1
- intensity of H2
- intensity difference between H1 and H2
- relative closure duration
- first formant F1
- bandwidth of F1

5. DISCRIMINANT ANYLSIS

Working with the statistical package SPSS-PC+ a discriminant analysis was computed for every speaker.The jackknife-validation-procedure (one-leaving-out-method) was used to reduce the misclassification bias.
For three samples the assumed multivariate normality and equality of group covariance matrices didn't hold. Consequently the degrees of freedom had to be reduced.
One speaker (Modi) showed different characteristics for vowel /a/ and /σ/, therefore the sample had to be devided.
For another speaker (Shah) only one variable remained in the analysis.

6. DISCUSSION

A view of the discrimination score figures reveals that the first harmonic is the most selective variable for all speakers except one. In that case a transformation measure of H1 (the H1-H2-difference measure) shows best results. Indeed 75% of speaker Patel's correct discrimination scores are expressed by the Diff-variable.

All formant measurements showed up to be quite selective between the voiced and the breathy group.

In the cases of Modi and Patel the hight of the first formant is inversely related to breathiness, in the case of Vora the broader bandwidth within the breathy group makes the difference.

Looking at the canonical discriminant function coefficients of the Dalal-sample the overall intensity contributes to the discrimination function because it correlates negatively with the first harmonic in the breathy group but not in the voiced group. The coefficients can be seen in the figure showing the discrimination scores of the source groups.

The impression that the closure duration in breathy final stops could be somewhat longer than in purely voiced stops could't be verified for the whole samples. Indeed closure duration is strong enough to discriminate the phonation type in samples of retroflex dental articulations.

7. FINAL REMARK

In minimal pairs the mean maximum intensity differences of H1 between the two groups of phonationtype are around 3 db, the Diff-measure which corrects in a sense for overall intensity variation shows even lower means.

If we consider that Bickley [8] used 3-db-steps to test perception differences between clear and breathy vowels and found a 75% discrimination score at the 9-db-borderline, we may find that discrimination analysis (or let's say a computer) is much more sensitive than human perceptionists.

8. LITERATURE

[1] FISHER-JØRGENSEN, E. (1967). Phonetic analysis of breathy (murmered) vowels in Gujarati. Indian Linguistics 28, 71-139.

[2] LADEFOGED, P.; ANTONANZAS-BARROSO, N. (1985). Computer measures of breathy voice quality. UCLA-WP 61, 79-86.

[3] SCHIEFER, L. (1987). The role of intensity in breathy voiced stops: a close link between production and perception. Proc. 11th ICPhS Tallinn, Vol. 5, 362-365.

[4] SCHIEFER, L. (1987). F0-Perturbations in Hindi. Proc. 11th ICPhS Tallinn, Vol. 1, 150-153.

[5] WARDRIP-FRUIN, C. (1980). On the status of the temporal cues to phonetic categories: preceding vowel duration as a cue to voicing in final stops. PhD Diss. Stanford University

[6] VAN SUMMERS, W. (1988). F1 structure provides information for final-consonant voicing. JASA 84, 485-492.

[7] YADAV, R. (1984). Voicing and aspiration in Maithili: a fiberoptic and acoustic study. Indian Linguistics 45, 1-30.

[8] BICKLEY, C. (1980). Acoustic analysis and perception of breathy vowels. MIT Speech Communication Working Papers Vol. 1.