ACOUSTIC PROPERTIES OF NON-SIBILANT FRICATIVES

Allard Jongman and Joan A. Sereno
Cornell University, Ithaca, N.Y., U.S.A.

ABSTRACT

Two recent classification metrics, spectral moments and locus equations, were employed in an attempt to distinguish English labiodental fricatives /f, v/ from dental /θ, ð/. Preliminary results suggest that these two classes of fricatives are distinct, both in terms of spectral moments (primarily skewness and kurtosis) and slope and intercept of locus equations.

1. INTRODUCTION

A fundamental issue in speech research concerns whether distinctions in terms of place of articulation are more successfully captured by local (static) or global (and/or dynamic) properties of the speech signal. Most studies of place of articulation have investigated stop consonants (e.g., [1-3]). In contrast, fricatives are less well-known, and it is uncertain whether the classification metrics proposed for distinguishing place of articulation in stop consonants can be successfully applied to fricatives.

Although fricatives have been the subject of much research, the cues which serve to classify English fricatives according to place of articulation are still not fully understood. Acoustic studies focusing on the frication noise show that properties of the spectrum, amplitude, and duration of the noise can all serve to distinguish the sibilant /s, z, ʃ, ʒ/ fricatives from the non-sibilant /f, v, θ, ð/ fricatives (e.g., [4-6]). Within the class of sibilant fricatives, spectral properties serve to distinguish /s, z/ from /ʃ, ʒ/. However, none of the noise properties seems adequate to distinguish /f, v/ from /θ, ð/. Most research (e.g., [7]) suggests that acoustic cues to this distinction might be located in the fricative-vowel transitions.

The present study focuses on two recent classification metrics that, with appropriate modifications, seem particularly promising to investigate the role of these transitions as cues to the /f, v/ - /θ, ð/ distinction: spectral moments and locus equations.

Spectral moments analysis involves a statistical approach, capturing both local (spectral peak) and global (spectral shape) aspects of obstruents. Specifically, FFTs are calculated at different locations in the speech signal, and each FFT is then treated as a random probability distribution from which the first four moments (mean, variance, skewness and kurtosis) are computed. Mean and variance reflect the average energy concentration and range, respectively; skewness refers to spectral tilt and kurtosis is an indicator of the peakedness of the distribution. Previous research [8] using spectral moments has primarily examined the information derived from the first 20 ms of obstruent-vowel sequences. This approach reliably distinguished /s/ from /ʃ/, but failed to distinguish the non-sibilants from each other.

The locus equations approach is also statistical in nature. Locus equations are derived based on the second formant (F2) at vowel onset and at vowel midpoint (e.g., [9]). Locus equations constitute a dynamic representation of speech sounds since they express a relation between F2 at different points in the speech signal. Previous results indicate that the F2 starting frequency of a vowel preceded by an obstruent provides unique information about the articulatory configuration used to generate the consonant. Although locus equations have recently been successful in the classification of place of articulation in voiced stop consonants, researchers have only just begun to apply this method to fricatives. Recent research using locus equations to analyze fricatives has reported contradictory results. Fowler [10] shows consistently different slopes for /v/ and /ð/, while Sussman [11] does not.

In their present form, neither of the two approaches just described has been entirely successful at uniquely distinguishing (English) fricatives in terms of place of articulation. Specifically, none of the metrics is capable of reliably distinguishing /f, v/ from /θ, ð/.

Nevertheless, with appropriate modifications, spectral moments and locus equations seem particularly promising in capturing the defining properties of fricatives.

The present study extends the spectral moments approach by using a larger window size (40 ms instead of 20 ms) and by additionally examining possible cues present later in the frication noise and transition region. In addition, locus equations will be computed in an attempt to shed light on previous contradictory results. Since locus equations specifically encode information about F2 at vowel onset and vowel midpoint, they may provide a very appropriate metric to investigate the role of transition information.

2. METHODS

Three native speakers of American English (2 males, 1 female) were recorded in the Cornell Phonetics Laboratory. Targets were of the form CVC, with the first consonant being /f, v, θ, ð/, the vowel being /i, e, æ, ɑ, o, u/, and the last consonant always being /p/. Three repetitions of each of these targets were produced in the carrier phrase "Say ____ again".

All recordings were sampled at 22 kHz with 16 bit quantization using Waves+ software running on a SparcStation LX. Two types of spectral measurements were made. For spectral moments, FFT spectra were computed using a 40-ms full Hamming window at each of four locations: onset, middle, and offset of the fricative noise, and centered over vowel onset. For each window location, the first four spectral moments (mean frequency, variance, skewness, and kurtosis) were calculated. These moments were calculated from both linear and Bark transformed spectra.

For locus equations, LPC spectra were computed, using a 23.3 ms full Hamming window at two locations: vowel onset, and centered over vowel midpoint. Spectral peaks were picked from the LPC spectral displays.

3. RESULTS

a. Spectral moments

All four spectral moments were derived for each stimulus at four separate locations: onset, midpoint, and offset of the fricative, as well as centered over vowel onset. For each moment and window location, the moment data for the labiodental fricatives /f, v/ were contrasted to the dental fricatives /θ, ð/. Each analysis was performed on both the linear and Bark data.

At fricative midpoint, labiodentals can be distinguished from dentals both in spectral mean and skewness, with labiodentals showing a higher spectral mean and a more negative skewness. Additional differences in kurtosis are found at fricative onset and fricative offset, with labiodentals having more diffuse peaks compared to dentals. For the window location centered over vowel onset, there are significant differences in skewness and kurtosis, although at this location, labiodental fricatives show greater positive skewness and less diffuse peaks compared to dental fricatives.

Interestingly, the analysis of the Bark moment data show few differences distinguishing fricative place of articulation, except at the middle of the fricative noise, where all four moments show distinct differences between the fricatives /f, v/ and /θ, ð/.

b. Locus equations

Table 1 shows slopes and intercepts of the locus equations for each fricative for each of the three speakers.

Table 1. Summary of locus equation slopes and y intercepts (Hz) for each speaker for /f/, /v/, /θ/, and /ð/. F and M indicate female and male speakers, respectively.

	/f/		/v/		/θ/		/ð/	
Speaker	slope	y interc	slope	y interc	slope	y interc	slope	y interc
F1	.742	383	.645	542	.417	1083	.321	1301
M1	.759	373	.597	585	.592	660	.565	797
M2	.790	224	.779	273	.545	778	.575	789

As can be seen from Table 1, the two places of articulation have distinct slopes and intercepts. Labiodentals have high slope values and low intercepts while dentals have lower slopes and higher intercepts. Paired two-tailed t-tests confirm that the difference in slope is significant [t(5) = 4.80, p = .0049], as is the difference in intercept [t(5) = -5.66, p =.0024].

Figure 1 shows locus equation scatterplots for labiodentals (top) and dentals (bottom) for all three speakers. The regression line equation is $y=.728x+379$, $r^2=.95$ for the labiodentals and $y=.499x+894$, $r^2=.81$ for the dentals.

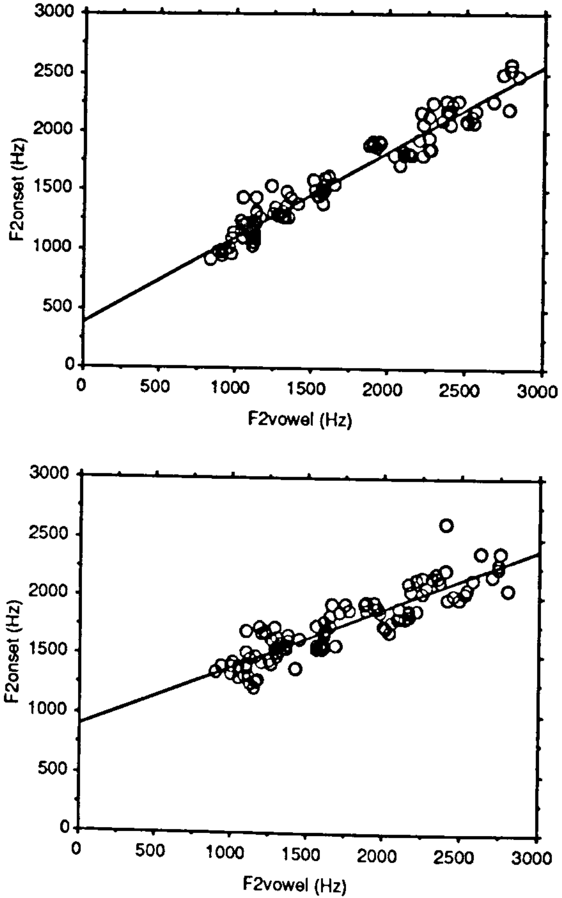

Figure 1. Locus equations for /f, v/ (top) and /θ, ð/ (bottom) for all six vowel contexts for all three speakers (two male, one female).

4. DISCUSSION

The present preliminary results show a number of interesting patterns. Spectral moments derived at the center of the frication noise and at the transition between noise and vowel seem most promising in distinguishing the non-sibilant fricatives in terms of place of articulation. The second moment, variance, does not seem to contain much distinctive information (cf. [8]). The use of the Bark scale does show some significant differences in place of articulation for the non-sibilant fricatives, but only when analyzing frication midpoint. Forrest et al. [8] found little effect of using non-linear transforms, perhaps because only onset information was examined. The present result suggests that location of the analysis window is crucial in determining place of articulation in non-sibilant fricatives.

Locus equations derived for /f, v/ and /θ, ð/ for three speakers are dissimilar. The two places of articulation seem to have distinct slopes and intercepts, with labiodentals having high slopes and low intercepts while dentals have low slopes and high intercepts. These present values are very similar to those reported for /v/ and /ð/ [10], both in terms of slope and intercept values. Although Fowler [10] questions the use of locus equations as cues to place of articulation across stops and fricatives, the use of locus equations as cues to place of articulation within each manner class is appealing since robust cues to the stop-fricative distinction are immediately available [12].

We feel the current results are encouraging. Both the moment analyses and the locus equations provide potential approaches for successfully distinguishing labiodental from dental fricatives in English. However, data for many more speakers are needed in order to conduct the necessary statistics (e.g., discriminant analyses for category discrimination both for the moment data and for slope and intercept values). A full dataset will be presented at the conference.

5. ACKNOWLEDGEMENT

We thank Scott Gargash for his technical expertise.

6. REFERENCES

[1] Liberman, A.M., Cooper, F.S., Shankweiler, D.P., and Studdert-Kennedy, M. (1967), Perception of the speech code, *Psychological Review*, 74, pp. 431-461.

[2] Stevens, K.N., and Blumstein, S.E. (1981), The search for invariant acoustic correlates of phonetic features, In P.D. Eimas and J.L. Miller (Eds.), *Perspectives on the Study of Speech*, New Jersey: Erlbaum, pp. 1-39.

[3] Jongman, A. and Miller, J.D. (1991), Method for the location of burst-onset spectra in the auditory-perceptual space: A study of place of articulation in voiceless stop consonants, *Journal of the Acoustical Society of America*, 89, pp. 867-873.

[4] Hughes, G.W., and Halle, M. (1956), Spectral properties of fricative consonants, *Journal of the Acoustical Society of America*, 28, pp. 303-310.

[5] Strevens, P. (1960), Spectra of fricative noise in human speech, *Language and Speech*, 3, pp. 32-49.

[6] Behrens, S.J., and Blumstein, S.E. (1988), Acoustic characteristics of English voiceless fricatives: A descriptive analysis, *Journal of Phonetics*, 16, pp. 295-298.

[7] Harris, K.S. (1958), Cues for the discrimination of American English fricatives in spoken syllables, *Language and Speech*, 1, pp. 1-7.

[8] Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R.N. (1988), Statistical analysis of word-initial voiceless obstruents: Preliminary data, *Journal of the Acoustical Society of America*, 84, pp. 115-124.

[9] Sussman, H.M., McCaffrey, H.A., and Matthews, S.A. (1991), An investigation of locus equations as a source of relational invariance for stop place categorization, *Journal of the Acoustical Society of America*, 90, pp. 1309-1325.

[10] Fowler, C.A. (1994), Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation, *Perception & Psychophysics*, 55, pp. 597-611.

[11] Sussman, H.M. (1994), The phonological reality of locus equations across manner class distinctions: Preliminary observations, *Phonetica*, 51, pp. 119-131.

[12] Jongman, A. (1989), Duration of fricative noise required for identification of English fricatives, *Journal of the Acoustical Society of America*, 85, pp. 1718-1725.