P2-4

Computational models of integration in Fricative-Vowel syllables

S. Fernández and S. Feijóo
Departmento de Física Aplicada, Universidad de Santiago de Compostela, Spain
fasanti@usc.es

This study explores different models of integrating the information relative to fricative perception contained in fricative-vowel (FV) syllables. A set of syllables formed by the fricatives /f, θ, s, ʃ, x/ with vowels /a, e, i, o, u/, was acoustically and perceptually analysed. 21 listeners identified the fricative present in: 1) The isolated fricative noise; 2) The fricative noise + 100 ms of the following vocalic portion. Identification of fricatives was significantly better in condition 2 than in condition 1 (86% vs. 63%). Two hypothesis were considered: a) The two segments (fricative and vocalic portion) are evaluated separately and later combined in a single percept; b) Both cues are evaluated jointly. A Quadratic Discriminant function (QDA) was used to take into account contextual effects with different acoustic characterizations (LPC cepstral, FFT cepstral, FFT mel ceptral, mel filters and linear filters). QDA allows us to estimate both separately and jointly the amount of fricative information contained in the F and V segments as the *a posteriori* probability of pertenence to each of the 5 fricative classes. If $P_i(F)$ represents the amount of information relative to the fricative i in the fricative noise; $P_i(V)$ represents the i fricative information of the vocalic portion; P(i) represents the a posteriori probability of pertenence of the FV segment to class i, then:

a) If both the F and V segments are evaluated separately, we have two choice functions:

1) OR function: $P(i)=P_i(F)+P_i(V)-P_i(F)P_i(V)$. P(i) would represent the probability of a particular fricative i being determined by either the F or V segments. This corresponds to the perceptual hypothesis predicting that only one of the segments determines the identity of the fricative [1].

2) AND function: $P(i)=P_i(F)P_i(V)$. P(i) would represent the probability of fricative i being determined by both segments. This is in accordance with hypothesis predicting that a decision is made taking into account both cues, coarticulatory information playing an important role [2, 3].

b) If both cues are processed as a unitary representation, the whole FV segment should constitute the basis for determining P(i). This can be done by training the QDA with the whole FV segment.

Both the classification scores and the correlation between P(i) and the results obtained by listeners were computed for each of the three possibilities, and used as indexes of the perceptual validity of the aforementioned hypothesis. The results show the inadequacy of models that consider the separate evaluation of cues to achieve a final percept. Nevertheless the computational models were unable to extract the same benefits obtained by listeners when the vocalic part is included in the stimuli. An explanation of the perceptual results might be based on some unknown interaction between both segments at the auditory level [4].

[1] Whalen, D.H. (1999) Three lines of evidence for direct links between production and perception in speech. In *Proceedings of the 1999 International Congress of Phonetic Sciences*, San Francisco, 2, 1257-1260.
[2] Whalen, D.H. (1981) Effects of vocalic formant transitions and vowel quality on the English /s/-/š/ boundary. *Journal of the Acoustical Society of America* 69, 275-282.
[3] Johnson, K. (1991) Differential effects of speaker and vowel variability on fricative perception. *Language and Speech* 34, 265-279.
[4] Kurowski, K., and Blumstein, S.E. (1984) Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *Journal of the Acoustical Society of America* 76, 391-404.