Formant transition as a cue to place of articulation in Brazilian Portuguese coronal fricatives

Two kinds of cues are shown to be used in the distinction between coronal voiceless fricatives: the spectral shape differences in the frication noise (4-8kHz for [s], 2-4kHz for [S]) and the spectral changes in formant transitions between the noise and the adjacent vowel (Harris 1958; Heinz and Stevens 1961; Hughes and Halle 1956; Dorman, Raphael and Isenberg 1980 for English; Guerlekian 1981 for Spanish). In Portuguese, frication noise has been described to have center frequencies around 5kHz for [s] and 3kHz for [S] in the European variety (Lacerda 1982; Jesus, 1999), and in Brazilian Portuguese (BP), Haupt (2008) and Santos (1987) described values around those of English: for [s] 4,5-7,4kHz, for [S] 2-4,6kHz. Lacerda (1982) tested those frequencies for consonant recognition, but to the best of our knowledge formant frequencies were not investigated yet. We decided to test whether the formant transitions said to be useful in English are also a cue to place of articulation for coronal fricatives in Brazilian Portuguese.

Methods:

Subjects – 12 female subjects with age varying from 14-28 years with no history of hearing problems.

Stimuli – Four vowels were synthesized using HLSyn (Sensimetrics, Inc.): two tokens for [a] and two for [u]. One token of either vowel was manipulated in formant transitions: one from [s], the other from [S] (Tab.1). Formant values for the steady-state part were taken from Escudero, Boersma, Rauber & Bion (2009), and the transitions, from Nittrouer and Miller (1997). A raw frication noise (160ms) with no filtering was synthesized using the Klatt cascade model implemented in Praat (Boersma & Weenink 2011). Noise was subsequently single-pole filtered in different formant frequencies with a bandwidth of 230Hz. Noise filter frequencies were taken from a normally distributed, randomly generated 100 number sequence with mean = 4830 and sd = 50. With this procedure, each frication noise was not exactly the same, so that subjects would not get tired responding many times to one and the same stimulus, but at the same time the main effect would normally cluster around halfway between the center frequency of a [s] and that of a [S]. The 100 different frication noises were then concatenated with the four vowel tokens (150ms) to produce 400 synthetic syllables of 310ms duration.

Procedure – The 200 tokens for [(s)a] and [(S)a] were presented in a block, the 200 tokens for [(s)u] and [(S)u] in another block within the same session. Subjects were allowed to have a break between blocks. A rating task was used (Macmillan & Creelman 2005). Subjects were required to listen to a syllable and respond whether they heard [(s)a] or [(S)a] (1st. block) and [(s)u] or [(S)u] (2nd. block), and also to rate their confidence in their answers using a 3-pt scale. PercEval (LPL, CNRS/Univ. Provence) was used for sound presentation and response entry.

Results:

Hit and false alarm rates of all response alternatives to [(s)a]-[(S)a] and [(s)u]-[(S)u] were computed. Slope and A_z (Tab.2) were estimated with ROC-kit (Metz, Herman & Shen, 1998). So, the standard (here very strong) assumption of unit slope is unnecessary. As A_z values range from .5 (no sensitivity) to 1 (complete sensitivity), for [a] there is good evidence in our data that formant transitions are an important and sufficient cue to place of articulation in coronal fricatives. Also, formant transitions described for English, once adapted to the formant frequency values reported for BP, are also useful for BP listeners to distinguish between [sa]-[Sa].

For [u], however, things were more confuse. In the 1st run, Az under .5 in 3 out of 4 subjects mean here that subjects made more false alarms than hits. We expected that the results would not be like those for the [a] tokens, since [u] is acoustically less clear. We then re-synthesized the [u] tokens with a longer steady-state part (270ms) and re-run the experiment. Subject PGF was re-tested a month after the 1st run. Then, 8 subjects were tested in the 2 blocks. The longer [u] tokens resulted in better classification performance, but with results around .5 most subjects were barely sensitive to a difference between [(s)u] and [(S)u]. All subjects in the 2nd run were then pooled in a single set of results: [(s)a] and [(S)a] seemed to be almost 60% as different as [(s)u] and [(S)u]. Our results do not allow to decide whether the difference in formant transition between [(s)u] and [(S)u] is insufficient as a cue or if the actual formant values we used for the transitions are inadequate for BP.

Table 1 – Initial and Final Transition Formant values (Hz)

	F1		F2		F3	
	Initial	Final	Initial	Final	Initial	Final
[(s)a]	483	683	1449	1329	2624	2324
[(S)a]	483	683	1769	1329	2624	2324
[(s)u]	310	310	1391	761	2609	2309
[(S)u]	310	310	1711	761	2609	2309

Table 2 – Results

Subject	Sa-CHa				Su-CHu			
1st run	n(S/s)	slope	Az	sd	n(S/s)	slope	Az	sd
ATR	100/99	2,227	0,976	0,011	100/100	1,392	0,58	0,041
MTR	97/92	2,888	0,811	0,035	100/98	2,512	0,422	0,043
PGF	100/99	0,679	0,965	0,013	100/99	0,92	0,387	0,043
RPA	98/99	0,299	0,839	0,043	99/99	1,341	0,414	0,053
2nd run								
PGF					98/97	0,909	0,662	0,041
ACA	100/99	0,274	0,977	0,016	100/100	0,744	0,665	0,046
CM	90/62	0,685	0,867	0,034	69/66	1,11	0,529	0,054
DOS	100/100	0,869	0,967	0,011	100/100	0,72	0,559	0,044
DSF	100/100	2,564	0,984	0,012	100/100	1,168	0,544	0,044
FAP	100/99	2,406	0,964	0,018	99/100	1,04	0,943	0,017
LM	100/100	1,131	0,644	0,041	100/100	1,151	0,589	0,041
SHA	100/99	1,46	0,95	0,016	100/97	1,614	0,555	0,048
TPA	100/100	1,04	0,846	0,029	99/100	1,471	0,571	0,044
pooled	1090/1061	1,032	0,891	0,007	1064/1058	0,99	0,598	0,013

References

1. Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.40, retrieved 13 September 2011 from http://www.praat.org/
2. Dorman, M. F, L. J Raphael, e D. Isenberg. “Acoustic cues for a fricative-affricate contrast in word-final position”. *Journal of Phonetics* 8 (1980): 397-405.
3. Gurlekian, J. A. “Recognition of the Spanish fricatives/s/and/f”. *The Journal of the Acoustical Society of America* 70 (1981): 1624.
4. Harris, K. S. “Cues for the identification of the fricatives of American English”. *The Journal of the Acoustical Society of America* 26 (1954): 952.
5. Haupt, C. “As Fricativas [s],[z],[S] e [Z] do Português Brasileiro”. *Letras & Letras, Uberlândia* 24, nº. 1 (2008): 59-71.

6. Heinz, J. M, e K. N Stevens. “On the properties of voiceless fricative constants.” *Journal of the Acoustical Society of America* (1961).
7. Hughes, G. W, e M. Halle. “Spectral properties of fricative consonants”. *The Journal of the Acoustical Society of America* 28 (1956): 303.
8. Jesus, L. M.T., *Acoustic phonetics of European Portuguese fricative consonants*. Technical Report, Department of Eletronics and Computer Sciences. University of Southampton, 2000.
9. Lacerda, F. P. “Acoustic perceptual study of the Portuguese voiceless fricatives.” *Journal of Phonetics* (1982).
10. Macmillan, N. A, e C. D Creelman. *Detection theory: A user’s guide*. Lawrence Erlbaum, 2005.
11. Metz, C.E., Herman, B.A., Shen, J.H. Maximum-likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 1998;17(9):1033-1053.