INVARIANT ACOUSTIC CORRELATES FOR PLACE OF ARTICULATION IN CATALAN VOICELESS STOPS

JOAQUIM LLISTERRI

Laboratori de Fonètica
Universitat Autònoma de Barcelona
Bellaterra, Barcelona, Spain

MARTIN WEST

Dept.ofAppliedAcoustics
University of Salford
Salford M5 4WT, UK.

ABSTRACT

The results of an acoustic analysis of Catalan voiceless stops are presented and discussed in terms of the invariant correlates for place of articulation within this class of sounds. Both gross spectral shape- a combination between frequency peaks in the burst and starting frequency of the formant transitions - and temporal parameters - specially VOT - can be considered as invariant acoustic correlates that distinguish between labial, dental and velar place of articulation in Catalan.

1. INTRODUCTION

The theory of acoustic invariance proposed by Stevens and Blumstein claims that a particular phonetic dimension must show invariant properties in the speech signal across all languages [1]. Invariant acoustic correlates for place of articulation have been found so far for English, French and Malayalam stop consonants [2] but there is still a need to carry out research on other languages. Catalan is a Romance language with three non aspirated unvoiced stops contrasting bilabial, dental and velar place of articulation ; its vowel system contains seven vowels in stressed position : [i], [e], [ɛ], [a], [ɔ], [o], [u] and a schwa [ə] appearing only in unstressed contexts. The purpose of this paper is to give the results of an acoustic analysis of Catalan stops, reviewing both frequential and temporal parameters that could provide invariant correlates for place of articulation in this language.

2. ACOUSTIC ANALYSIS

2.1. SUBJECTS AND UTTERANCES

Three male and two femalenative speakers of Catalan from Barcelona were asked to read a list of 24 carrier sentences containing bisyllabic words with initial voiceless stop-vowel groups in stressed position, except for the schwa that was unstressed. The eight Catalan vowels were combined with [p], [t] and [k], giving a total of 120 utterances studied.

The recordings were made in anechoic conditions using a Revox A77 tape recorder and a Sennheiser MD 441N cardioid microphone, placed at constant distance from the mouth.

2.2. METHOD

The recordings were low-pass filtered at 5kHz, sampled at 10 kHz and stored on a PDP-11 computer. 14 coefficient LPC spectra were calculated using a 12.8 ms Hamming window positioned at the consonantal release and automatically moved along the signal in 2.5 ms steps until the steady-state of the vowel was found. Formant frequencies were extracted using an automatic peak-picking program. Temporal information was obtained from manual measurements of digitised oscillograms, and checked for accuracy in the waveform display of a Brüel & Kjær 2033 narrow band analyser. Narrow band spectra were also used to check the accuracy of the steady-state formant frequencies for vowels.

2.3. RESULTS

2.3.1. Release Burst

The release of Catalan voiceless stops is accompanied by a short burst of acoustic energy. Its duration values

averaged across all speakers and vowel contexts are given in Table 1:

	Min	Max	Mean	SD
[p]	1.1	11.02	3.71	2.17
[t]	1.88	19.22	6.71	3.67
[k]	4.3	46.4	14.12	8.85

Table 1: Duration values for burst in ms.

It can be observed that burst duration is greater from labial to velar place of articulation; this same trend is observed in the data reported also for Catalan by Martí [3] ([p]: 8.6; [t]: 13.6; [k]: 20.5ms); differences in the absolute values might be explained by the fact that Martí made his measurements on spectrograms. However, this has not been observed for Castilian Spanish : [p]: 15.38; [t]: 15.62; [k]:21.86ms (Poch [4]), a language which shows the same contrasting places of articulation for stops as Catalan.

Frequency values were obtained for the first two prominent peaks (K1 and K2) in the LP spectrum with the window center positioned at the burst onset; they are shown in Table 2:

	K1	K2
[p]	1341.91 (321.32)	2060.64 (413.71)
[t]	1787.59 (250.68)	3027.74 (508.62)
[k]	1868.59 (689.86)	2949.49 (674.43)

Table 2: Burst frequency values and standard deviation (in parentheses) in Hz averaged for the two male speakers across all vowel contexts.

2.3.2. Voice Onset Time

Voice Onset Time values have been measured for all the speakers with the following results:

	Min	Max	Mean	SD
[p]	3.68	19.69	11.75	4.09
[t]	7.7.4	44.54	17.35	7.43
[k]	12.27	59.54	28.37	9.76

Table 3: VOT values in ms averaged across all vowel contexts.

It can be seen from the table above that Catalan voiceless stops show a short time interval between the consonantal release and the starting of voicing. The well known correlation between VOT and place of articulation is maintained. These results agree with those previously given by Julià [5] ([p]: 3; [t]: 16; [k]: 35ms) and Martí [3] ([p]: 10.2; [t]: 16.1; [k]: 26.1ms). They follow the same pattern as those reported for Spanish [6] ([p]: 17.18 ms; [t] : 19.75 ms; [k]: 30.01 ms), Italian [7] ([p] :12 ms; [t]: 17ms; [k] : 30 ms), French [8] ([p]:28.5 - 27.6; [t] 31.4-35.4; [k]: 53-51.7) or Modern Greek [9] ([p]: 9; [t] : 16; [k]: 30) although in some cases absolute values might differ.

2.3.3. Formant Transitions

Formant transitions will be described by means of two parameters:starting frequency and duration from the onset to the steady-state of the vowel. Both have been calculated by examining the successive LP spectra starting at the onset of voicing. The average starting frequencies for two male speakers are given across all vowel contexts in Table 4:

	F1	F2	F3
[p]	338.91	1019.02	2066.08
	(79.73)	(390.51)	(400.15)
[t]	351.11	1613.5	2640.2
	(90.04)	(313.96)	(236.47)
[k]	358.34	1728.99	2326.74
	(102.47)	(596.88)	(288.91)

Table 4: Averaged starting formant frequencies in Hz (standard deviation in parentheses) for two male speakers across vowel contexts.

Both these results and Martí [3] show very similar values for F1 in [p], [t] and [k], lower values for F2 and F3 in labial contexts, and higher values for F3 in contact with the dental stop.

Transition durations are presented in Table 5. The extent of the transition has been taken from the onset of the formant to the steady-state of the vowel. The dental consonant tends to show longer F2 and F3 transitions than the labial, while the velar appears to have longer transitions than the other two stops. A similar trend is observed in Martí [3].

	F1	F2	F3
[p]	14.4 (8.2)	20.8 (4.9)	23.2 (13.6)
[t]	18.8 (8.8)	25.1 (12.1)	19.6 (9.2)
[k]	25.6 (9)	24.8 (6.5)	23.6 (8.9)

Table 5: Mean duration of consonant-vowel transitions in ms (standard deviation in parentheses) averaged across all vowel contexts for two male speakers.

3. DISCUSSION

Acoustic data presented so far will be discussed in terms of its potential role as invariant correlate for place of articulation in Catalan stops.

3.1. FREQUENCY CORRELATES

Stevens and Blumstein [10] suggested that an approximation to the shape of the onset spectrum could be obtained by plotting the frequency of the burst against the difference between F3 and F2 frequency values at the onset of transitions. This is shown in Fig 2:

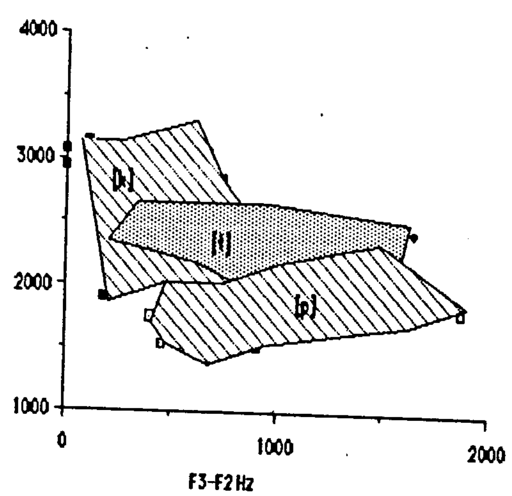

Fig 2: Burst frequency (mean value of the two spectral peaks measured in LP spectra) vs F3-F2 onset formant frequencies.

The degree of overlapping is similar to the one obtained when plotting only the first prominent peak (K1) or only the second one (K2).

It can be seen that this is not a completely satisfactory classification in terms of place of articulation, particularly because it fails to make a clear difference between dental and velar consonants. A more careful examination of the parameters involved shows that the difference between K1 for [p] and [t] is statistically significant (t= 3) (1) but it is not so for [t] and [k]. Similar results are obtained for K2: [p] and [t] show significant differences (t= 4.1), but [t] and [k] are not significantly different.

As for the onset of formant transitions, F2 can distinguish between [p] and [t] (t= 4.7), but again the onset of the F2 transition alone fails to separate [t] from [k]. The results for the onset of F3 transitions show a somewhat different patters: [p] is distinguished from [t] (t= 4.9) and [t] differs from [k] (t= 3.3). Significant differences are also found between F3-F2 onset frequencies for [t] and [k] (t= 2.9) but they do not appear for [p] and [t].

It has not been possible to measure our spectra in terms of the metrics developped by Blumstein and Stevens [11]. An informal observation of data presented in Table 2 and Fig. 2 suggests the predominance of low frequency spectral peaks for labial consonants and a high peak at the starting frequency of the second formant for dental stops. The high standard deviation found in K1 for velars (a range of 2046 Hz for [k] vs. 628 Hz for [t]) seems to suggest the confluence of K1 and K2 in a high-medium range of frequencies. This is in agreement with Blumstein and Stevens templates, the more problematic being the dental consonant.

Examining changes in energy distribution between the burst and the onset of transitions it appears that difference between the frequency of spectral peaks in the burst spectrum and at the onset of F2 and F3 transitions are more significant for [t] (t= 10.9) than for [p] (t= 6.5) or [k] (t= 9.2), suggesting a stronger change in spectral shape for dental place of articulation in front of labial or velar stops [2].

3.2. TEMPORAL CORRELATES

Burst duration has been found to provide a mean of differentiating between the three classes of stops studied: differences are significant for [p]-[t] (t= 4.4) (2) and for [t]-[k] (t= 4.8), the duration of the burst increasing from labial to velar place of articulation.

The same results hold for VOT, with significant differences between [p]-[t] (t= 4.1) and even more for [t]-[k] (t= 5.6). According to these results, both can provide temporal acoustic cues for place of articulation within the class of voiceless stops.

Burst duration and VOT show a strong correlation (Spearman's rho = 0.6 for 40 paired observations) although no correlation has been found between VOT and mean vowel duration in the Catalan stop-vowel groups studied.

The duration of the transitions does not seem to provide a cue for place of articulation, since no significant differences were found between the three classes of stops.

Finally, a strong correlation (Spearman's rho = 0.9 for 40 paired observations) was found between total stop-vowel and vowel duration, although this is independent of place of articualtion.

4. CONCLUSIONS

In summary, Catalan voiceless stops may be characterized by a very brief release burst and a short VOT, both increasing in length from labial to velar and by abrupt transitions into the following vowel. Frequency parameters for the burst or for the starting point of F2 and F3 transitions alone are not able to discriminate among the three places of articulation, but a gross characterization of the spectral shape at the consonantal release and at the onset of transitions combining the parameters measured seem to provide invariant cues for this phonetic dimension. In our data the differenciation is clearer between [p] and [t] than between [t] and [k].

Both burst duration and VOT can discriminate between place of articulation, these two parameters being strongly correlated. Differences in VOT between [t] and [k] are stronger than between [p] and [t]. This seem to suggest an interaction between temporal and frequency acoustic cues that would be used in conjunction for the discrimination of place of articulation independently of other factors studied as vocalic context or speaker variation. This fact agrees with the concept of dynamic invariance proposed by Blumstein [2].

(1) All t values are given for p≤ 0.05 and 30 degrees of freedom. (2) All t values are given for p≤ 0.05 and 77 degrees of freedom.

REFERENCES

[1] Blumstein, S.E. (1986) "On Acoustic Invariance in Speech" in J.S. Perkell & D.H. Klatt (Eds) *Invariance and Variability in Speech Processes.* Hillsdale: LEA. pp.178-193.

[2] Lahiri, A. et al. (1984) "A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study", *JASA*, 76,2: 391-404.

[3] Martí, J. (1986) *Estudi acústic del català i síntesi automàtica per ordinador.* Unpublished PhD. Universitat de València.

[4] Poch, D. (1984) *Las oclusivas sordas del español.* Unpublished PhD. Universitat Autònoma de Barcelona.

[5] Julià, J. (1981) "Estudi contrastiu dels oclusius de l'anglès i del català. Un experiment acústic", *Estudi General*, 1,2: 75-85.

[6] Poch, D. (1985) "Caractérisation acoustique des occlusives de l'espagnol: le problème du VOT", *Revue de Phonétique Appliquée*, 77: 477-490.

[7] Vagges, K. et al. (1978) "Some acoustic characteristics of Italian consonants", *Journal of Italian Linguistics*, 3,1: 69-86.

[8] Durand, P. (1985) *Variabilité acoustique et invariance en français. Consonnes occlusives et voyelles.* Paris: Editions du CNRS.

[9] Fourakis, M. (1986) "A timing model for word-initial CV syllables in Modern Greek", *JASA*, 79,6: 1982-1986.

[12] Stevens, K.N.- Blumstein, S.E. (1978) "Invariant cues for place of articulation in stop consonants", *JASA*, 64,5: 1358-1369.

[13] Blumstein ,S.E.- Stevens, K.N. (1979) "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants", *JASA*, 66,4: 1001-1017.