ARTICULATORY ADJUSTMENTS IN INITIAL VOICED STOPS IN SPANISH, FRENCH AND ENGLISH

Maria-Josep Solé

Universitat Autònoma de Barcelona, Spain
mariajosep.sole@uab.cat

ABSTRACT

This work reports cross-language differences in the voicing of initial voiced stops, and in the use of active maneuvers to achieve closure voicing, using multiparametric aerodynamic data. Oral pressure, oral and nasal flow, and acoustic data were obtained for utterance-initial /b d p t m/ for 10 speakers of Spanish, 6 speakers of French and 5 speakers of English. Voiced stops were classified as fully voiced or devoiced, and by shape of the oral pressure pulse (implosivized, other cavity enlarging maneuver) and/or occurrence of nasal flow (prenasalized) or oral flow (spirantized) during the stop phase in an attempt to relate aerodynamic data and actual glottal vibration to vocal-tract gestures and maneuvers to facilitate voicing. Maneuvers that favor the initiation (and sustaining) of voicing are then related to (i) language-specific differences in the use of glottal vibration during the constriction as a cue to voicing, (ii) place of articulation, and (iii) speaker dependent variation.

Keywords: voiced stops, aerodynamics, Spanish, French, English

1. INTRODUCTION

A number of studies have investigated and modeled the aerodynamic conditions required to maintain voicing during a stop closure [5, 6, 10], and have described the characteristic voicing patterns in stops e.g. [2]. Most aerodynamic studies have focused on postvocalic word-initial, intervocalic or word final stops, in part due to the difficulty to unambiguously identify the beginning of the stop closure post-pausally. Utterance-initial stops, however, involve somewhat different initial conditions for the respiratory system and the laryngeal system. For example, the subglottal air pressure must increase from atmospheric pressure, the vocal folds must approximate and be properly tensed, and vocal fold vibration has to be initiated rather than sustained (which requires a larger transglottal pressure differential, 3-4cmH_2O vis-à-vis 1-2cmH_2O [1]). Because oral and subglottal pressure rise in synchrony during utterance-initial stops, the pressure differential is insufficient to set the vocal folds into vibration and voicing is unlikely to occur without additional adjustments. Thus the laryngeal and aerodynamic conditions in utterance-initial stops are less conducive to voicing than in medial or final position, where voicing continues from the preceding vowel/sonorant.

This study reports aerodynamic and acoustic data of voicing in initial stops in Spanish and French, with typically prevoiced stops, and in American English, where devoiced and prevoiced stops occur. These data are then related to adjustments used by speakers to diminish oral pressure, and thus achieve sufficient transglottal pressure differential for the initiation of voicing. Specifically, this study presents a qualitative analysis of the data, whereas a separate study [9] presents a quantitative analysis of prenasalization in Spanish voiced stops.

Previous studies have provided evidence of a number of ways in which speakers may slow down the buildup of oral pressure during voiced stop closures, and thus allow phonation to initiate earlier and/or last longer. Three basic mechanisms can be used: (i) passive expansion of the oropharyngeal walls by reducing the level of muscle contraction, (ii) active enlargement of the oral cavity (lowering the larynx, elevating the soft palate, advancing the tongue root, depressing the tongue body); (iii) releasing airflow through an incomplete oral or velopharyngeal closure [8, 10]. The specific adjustments used vary within and across speakers [10] and possibly across languages.

In this study we will examine a multiparametric representation of utterance-initial voiced stop production, involving oral pressure, oral airflow, nasal airflow and acoustic data. Such data can provide information on oral cavity expansion gestures (as indicated by concave pressure trajectories, [5, 6]), oral leakage, and nasal leakage during the stop closure, and more generally provide insights into pressure lowering mechanisms used in prevoiced stops.

To investigate the use of such voicing-initiation adjustments, one needs to look at languages in which utterance initial stops are prevoiced, such as Spanish or French, and languages where voicing lead may not be present, such as English. We assume the position of the vocal folds in English is conducive to voicing in utterance-initial stops, as shown by laryngographic data [3], and aero-dynamic perturbation studies indicating that stops are voiced when the aerodynamic conditions are varied [7].

2. METHOD

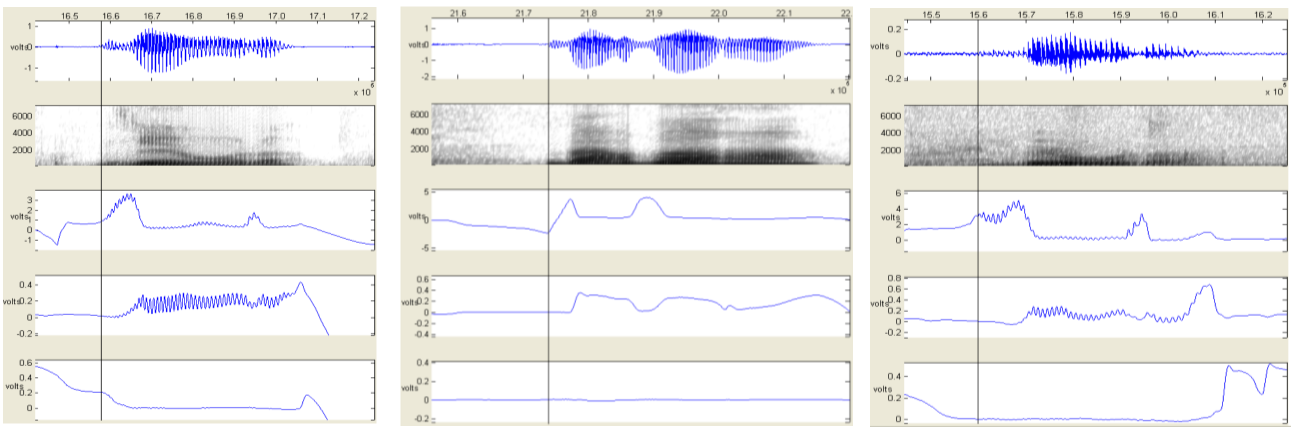
Simultaneous oral pressure, oral airflow, nasal airflow, and audio signal were obtained for ten Spanish speakers (three female (S1, S7, S9) and four male speakers (S3, S4, S5, S6) of continental Spanish; one female Mexican speaker (S2), and one female and one male Uruguayan speakers (S8, S10)); five continental French speakers, three females (F2, F4, F5) and two males (F1, F3); and six English speakers, one female (E6) and four male speakers (E1, E2, E3, E5) from California, and one female speaker from Australia (E4). They were instructed to produce the following sentences as if they were a dialogue between A and B, in order to obtain two isolated utterances and that the segment of interest (the /b/ in Barbara, the /d/ in Deborah, etc) was utterance initial.

- Spanish: A: ¿Cómo se llama ella?
  B: **B**árbara [**D**ébora, **P**aula, **T**ábata, **M**arta]
- French: A: Comment s'appelle-t-elle?
  B: **B**arbara [**D**ebora, **P**amela, **Th**érèse, **M**arta]
- English A: What's she called?
  B: **B**arbara [**D**eborah, **P**amela,**T**abitha,**M**artha]

The subjects repeated the dialogue with each test word, providing 10 to 13 repetitions of utterance-initial voiced stops /b d/, voiceless stops /p t/, and nasal /m/. Velar stops were not included because the catheter recording the oral pressure interfered with the articulation of back stops. The vowel following the stop was always a nonhigh vowel and the test words were kept as similar as possible in the three languages. In English and Spanish the test words had stress on the first syllable, whereas in French the stress was on the last syllable. There were thus 550 tokens for Spanish (average of 11 repetitions × 5 test words × 10 speakers), 330 for English and 275 for French.

The subjects' productions were recorded using National Instruments PCI-6013 data acquisition hardware and the Matlab Data Acquisition Toolbox (20kHz sample rate per channel and 16 bits/sample). Oral pressure was obtained by a catheter introduced through the center of the lips, ending just behind the lips for labials and behind the alveolar ridge for alveolars, and connected to a pressure transducer (Biopac TSD160C). Oral airflow and nasal airflow were collected with a split Rothenberg mask and Fleisch pneumotachographs with Biopac TSD160A pressure transducers. The airflow and oral pressure signals were digitally filtered with a 400Hz lowpass filter, a cutoff that allowed for the simultaneous visualization of relatively slow-changing aspects of the aerodynamic signals and the effects of voicing. The audio signal was recorded with an AKG C520 microphone and M-Audio AudioBuddy microphone preamp.

Figure 1: Waveform, 0-7 kHz spectrogram, oral pressure, oral flow, and nasal flow. Left: Example of a prenasalized and spirantized token of Spanish /d/ in *Débora* (S6). Middle: Implosivized token of French /b/ in *Barbara* (S4). Right: Concave token of Spanish /d/ in *Débora* (S3), classified as 'other'. The vertical line indicates onset of vocal fold vibration.

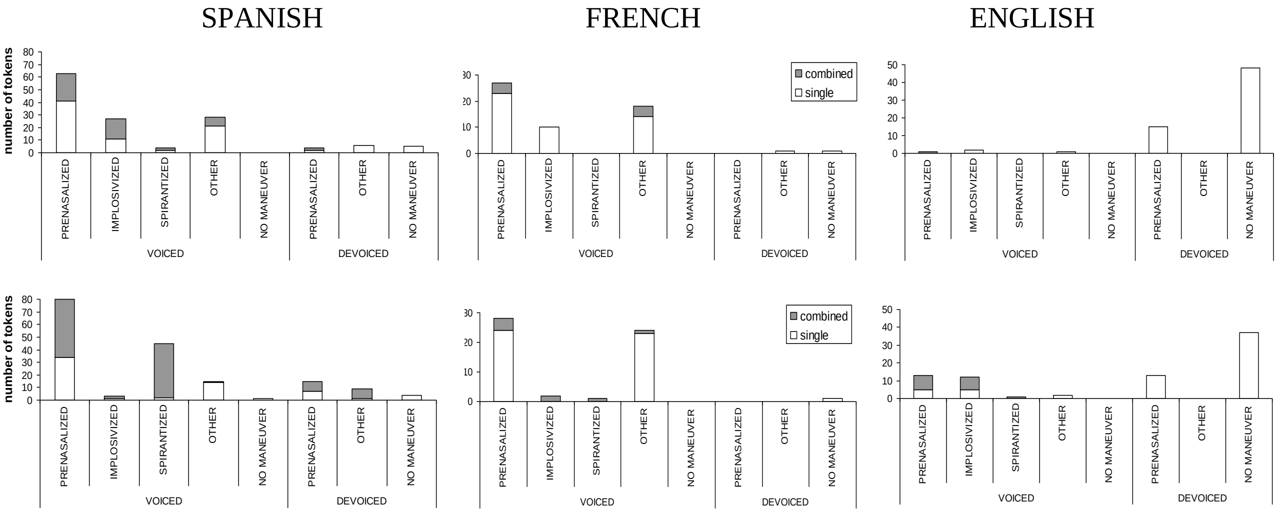


The data were analyzed qualitatively and quantitatively. First, each voiced token was visually classified as prevoiced (if voicing lead was present) or devoiced. Then tokens were classified by shape of the pressure pulse and/or by the occurrence of nasal or oral flow during the stop closure. If nasal airflow was present during (typically the first part of) the stop constriction, either continuing from the rest position or as a momentary velum opening gesture, the token was classified as *'prenasalized'*. Tokens showing oral airflow during the stop constriction were classified as '*spirantized'* (Fig. 1 left). Tokens were also classified by shape of the oral pressure, in an attempt to indicate whether it showed (i) negative pressure before pressure buildup (*'implosivization'*) (Fig. 1 middle), (ii) a concave pressure rise which will be assumed to be associated with a cavity enlargement maneuver [6] (*'other'*) (Fig. 1 right) —the difference between the two most probably reflects a timing difference in the volume expansion gesture — or (iii) a gradual (linear) pressure rise (*'no maneuver'*), characteristic of passive expansion of the oropharyngeal walls. Tokens were added to multiple categories where required. For example, if a token was prenasalized and spirantized, it was classified in both categories and was also counted as a 'combined' maneuver (indicated by the stacked dark bars in Fig. 2). This classification primarily attempted to relate aerodynamic results and glottal vibration to vocal-tract gestures and maneuvers to facilitate voicing. For stops showing prevoicing these five categories were used, but for devoiced stops only the categories *'prenasalized', 'other'* (including the few spirantized, implosivized and 'other' tokens) and *'no maneuver'* were used. Because a certain degree of visual judgment was being deployed in classifying the tokens, even with clear patterns, all the tokens were reclassified by the investigator in order to ascertain how consistently the different adjustments were being identified. Correlations between the original and the reclassified data showed high reliability (r values 0.90 or higher and all p values < 0.001). The within-subject coefficient of variation was 8.2%.

3. RESULTS

The results are plotted in Fig. 2, which shows the distribution of voiced and devoiced stops by type of maneuver in the three languages. The use and non-use of specific voicing maneuvers seems to be conditioned by language, that is, whether the language uses prevoicing as a cue to the voicing contrast (Spanish, French) or not (English). Thus Fig. 2 shows that in Spanish and French the great majority of tokens have *voicing lead* – 85.6% and 97% respectively – and involve some type of maneuver: prenasalization in 62.4% of the cases in Spanish and 53.9% in French; implosivization in 13.1% and 11.7% of the cases, predominantly in labial stops (top Figs.); spirantization in 21.4% of the cases in Spanish (only one case in French), predominantly in apical stops (bottom Figs.), and 'other' likely cavity enlarging adjustments in 18.7% and 41.2% of the cases in Spanish and French respectively (labial and apical stops pooled; single and combined adjustments pooled). In English, on the other hand, the great majority of tokens (83.2%) are *devoiced*. The aerodynamic

Figure 2: Distribution of [b] (top) and [d] (bottom) realization per voicing (voiced/devoiced) and type of maneuver (prenasalization, implosivization, etc.) for Spanish (left), French (middle), and English tokens (right). White bars indicate single maneuvers, gray bars combined maneuvers. See text.



data reflects fewer cases of voice facilitating adjustments than in the Romance languages and, when such adjustments are present, a higher rate of failure to initiate voicing (e.g., 84.8% of the tokens with prenasalization –as a single maneuver– failed to voice in English *vs* 10.7% in Spanish and 0% in French). Failure to voice initial stops when prenasalization is present may be due to differences in the timing (e.g., early velum closure) or magnitude of the nasal gesture, or to a relatively later glottal adduction in English vis-à-vis the Romance languages.

The production type furthermore depends not only on the language but also on the place of articulation of the consonant. As expected, apicals tend to be less conducive to voicing than labials – most likely due to lesser area of compliant tissue [8]. This is shown by voiced apicals involving more cases of prenasalization than labials in the three languages and, in Spanish, by (i) a higher number of devoiced apical vis-à-vis labial realizations in spite of voicing maneuvers (e.g., prenasalization) and (ii) more instances of combined maneuvers. In English, however, the data show more instances of voicing maneuvers and actual voicing in /d/s than in /b/s, though all the voiced /d/s are contributed by one speaker. Another place-related difference is that labials show more cases of implosivization than apicals in Spanish and French, parallel to typological patterns [4], while apicals show a preference for spirantization in Spanish. For those Spanish speakers that consistently spirantize (4 out of 10), oral flow typically occurs during the latter portion of the stop closure most likely to sustain voicing (characteristically accompanied by prenasalization during the initial part of the stop). For these speakers, lack of spirantization tends to extinguish voicing in the latter part of the stop. In Spanish and French, prenasalization shows a higher rate of occurrence than the other maneuvers, but if the categories 'implosivized' and 'other' are pooled as cavity enlarging adjustments, then the occurrence of prenasalization and cavity expansion is comparable.

Another dimension of variability is within and across speaker variability. For example, 78.3% of the prevoiced tokens in English are contributed by 2 of the speakers. In spite of the variability, all speakers (except two English speakers, E2 from California and E4 from Australia) show cases of prenasalization of voiced stops. In 3 out of 5 French speakers, 8 out of 10 Spanish speakers, and 4 out of 6 English speakers, prenasalization is the most common adjustment, usually as common as all the other maneuvers combined. Similarly, all speakers of Spanish, French and English (except E4, the Australian speaker) show cases of active volume expansion during voiced stops.

4. CONCLUSIONS

The data reflects different articulatory adjustments, used singly or in combination, directed to lower the oral pressure and achieve the pressure difference for voicing initiation. In French and Spanish, where voicing is typically present and maintained throughout the stop closure, maneuvers such as cavity enlargement and nasal or oral leakage have a greater rate of occurrence than in English. The type of adjustment is often speaker- or even phoneme-specific (e.g., spirantization in apicals).

5. ACKNOWLEDGEMENTS

This study was supported by grants FFI2010-19206 from the Ministry of Science and Innovation, Spain and 2009SGR003 of the Generalitat de Catalunya.

6. REFERENCES

[1] Baer, T. 1975. *Investigation of Phonation Using Excised Larynxes*. Unpublished PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA.

[2] Docherty, G. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris.

[3] Flege, J.E. 1982. Laryngeal timing and phonation onset in utterance-initial English stops. *J. Phonetics* 10, 177-192.

[4] Greenberg, J.H. 1970. Some generalizations concerning glottalic consonant, especially implosives. *International Journal of American Linguistics* 36(2), 123-145.

[5] Koenig, L.L., Lucero, J.C. 2008. Stop consonant voicing and intraoral pressure contours in women and children. *JASA* 123, 1077-1088.

[6] Müller, E.M., Brown, W.S. 1980. Variations in the supraglottal pressure waveform and their articulatory interpretation. In Lass, N. (ed.), *Speech and Language: Advances in Basic Research and Practice*. Vol. 4. New York: Academic, 317-389.

[7] Ohala, J., Sprouse R. 2003. Effects on speech of introducing aerodynamic perturbations. *Proc.15th ICPhS* Barcelona, 2913-2916.

[8] Rothenberg, M. 1968. *The Breath-Stream Dynamics of Simple-Released-Plosive Production*. Basel: Karger.

[9] Solé, M.J., Sprouse, R., 2011. Voice-initiating gestures in Spanish: Prenasalization. *Proc. 17th ICPhS* Hong Kong.

[10] Westbury, J.R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *JASA* 73,1322-1336.