Lenition of /d/ in spontaneous Spanish and Catalan

Miquel Simonet¹, José I. Hualde², Marianna Nadeu²

¹Dept. of Spanish & Portuguese, University of Arizona, Tucson, Arizona, USA
²Dept. of Spanish, Italian & Portuguese, University of Illinois at Urbana-Champaign, Illinois, USA

simonet@u.arizona.edu, jihualde@illinois.edu, mnadeur2@illinois.edu

Abstract
The present study explores the acoustics of /d/ in two corpora of Spanish and Catalan spontaneous speech. Three acoustic metrics were developed as indexes of articulatory weakening. The findings suggest that variations in the implementation of /d/ result from gradient modulations in constriction degree on a unimodal statistical-acoustic distribution. The preceding segment is a strong predictor of the weakening of Catalan and Spanish /d/.

Index Terms: Spanish, Catalan, lenition, phonetics, phonology.

1. Introduction
The systematic lenition or “spirantization” of voiced plosives is a phonological areal feature of most of the Iberian Peninsula. In the languages of this area, voiced plosives are systematically realized as voiced approximants, called “fricatives” in earlier work, in some contexts. In the Iberian Peninsula, this phenomenon has been reported for Spanish [e.g., 1, 2], Catalan [e.g., 3, 4], Galician [5], Basque [6], and as an optional phenomenon in Northern-Central Portuguese [7].

Spirantization in these languages has been described as the complementary distribution between two categorically distinct allophones, stop and spirant. For instance, standard Spanish /b d / are said to have two types of allophones in complementary distribution: stop allophones in utterance-initial position and after a homorganic nasal (or lateral in the case of /d/, but not /b/ or /g/) and spirant allophones elsewhere [1, 2]. However, it has been pointed out that there is substantial variation in the constriction of these sounds in different word tokens [8, 9, 10]. This suggests that, even if realizations can be classified as presenting or lacking complete occlusion, the phenomenon may be better understood as a continuum of constriction degrees.

In line with this observation, it has been reported that, in addition to the preceding context (i.e., the standard phonological description), speaking rate, lexical stress, morpheme identity and the specific timbre of the surrounding vowels condition degree of constriction in at least some of these sounds, [e.g., 10]. Yet, while some degree of gradience has indeed been attested, this could be an artifact of a continuous distribution that might turn out to be bimodal if examined in detail. A bimodal gradient distribution could call for the application of a phonological rule acting at some higher level of abstraction; that is, two categorically distinct allophones. Is spirantization unimodally gradient in Iberian Romance?

A question that arises is how to obtain information on the articulatory parameter of degree of consonantal constriction from the non-linear, indirect information that sound waves provide. A first approach is to classify spectrographic tokens into categories (and subcategories) on the basis of visual inspection [5]. The problem with this approach is that it forces researchers to divide a continuum into categories whose boundaries can only be fuzzy. A second approach is to use RMS (Root Mean Square) energy measurements as indicative of degree of constriction. The energy output or intensity of the segment can be considered an acoustic correlate of degree of constriction: open, vowel-like realizations of /b d g/ are expected to have much greater acoustic intensity than very closed stop realizations.

Acoustic intensity, however, can only be measured in relative terms, since it greatly depends on linguistically irrelevant factors such as generally softer or louder elocution, distance to the microphone, etc. Several normalization methods have been developed that have used different window sizes as references to express acoustic intensity in relative terms. For instance, references that have been used include the whole word in which the consonant appears [11], the VCV sequence (i.e., the average of the intensity maxima of the two surrounding vowels) [12] and the peak intensity of the following vowel [e.g., 10].

An entirely different approach has been to explore velocity curves derived from acoustic intensity contours [13]. In this procedure, minima (highest falling velocity, i.e. from previous segment to target consonant) and maxima (highest rising velocity, i.e. from target consonant to following segment) values are extracted from velocity curves. The actual minima and maxima of the intensity curves are ignored; the velocity values are key. This procedure is based on the premise that more constricted consonants will have more extreme velocity values due to more abrupt segment-to-segment transitions [13]. In laboratory speech a good correlation has been found between these acoustic measurements and articulation [14], although not in contexts such as /nl/ and /ld/, where the acoustic output may not reflect the presence of complete occlusion [15].

The present study examines the effect of the preceding segment on the degree of constriction of /d/ in two comparable corpora of conversational speech, one for Spanish and one for Catalan, using three different acoustic measurements. Majorcan Catalan shows less extensive lenition of /b/ than Iberian Spanish [16]. For /d/, the difference appears to be smaller, but in the same direction [17, p. 322]. Our findings suggest that Iberian (Catalan and Spanish) spirantization is a non-categorical weakening process conditioned (in part) by the degree of constriction of the preceding segment. This gradualness extends to the intervocalic context. Even though this is the prototypical context for spirantization, we find that higher preceding vowels condition more constricted allophones than lower vowels.
2. Method

2.1. Recordings

A total of 40 speakers were recruited for participation in an interactive task with one of the authors. The data were collected on the island of Majorca, Spain, a Catalan-Spanish bilingual speech community. Twenty (10 females) Catalan-dominant bilinguals were recorded in Catalan, and 20 (10 females) Spanish-dominant bilinguals were recorded in Spanish. Language dominance was assessed by means of a language background questionnaire and a series of accent rating tasks.

The participants were asked to engage in an interactive speech game. Speech was recorded through a head-mounted microphone into a solid-state recorder (44.1 kHz, 16-bit). For each participant, we have approximately 15 minutes of speech.

2.2. Acoustic analyses

Three different acoustic metrics were used as correlates of degree of constriction. One measurement, intensity difference (ID), is the difference in acoustic intensity between the consonant minimum and the following vowel’s maximum: the more open the constriction, the smaller the difference with respect to the following vowel. A pass Hann band filter (500-10,000 Hz) was applied to the signal prior to obtaining the intensity curve. This filtering procedure excludes the effects of //f/. This is hypothesized to maximize the difference in this metric between approximants and voiced stops. Therefore, if a slight difference in constriction degree between the two sounds exists, this method will increase the likelihood of detecting it.

A second metric, spectral tilt (ST), is obtained by calculating the difference in acoustic energy in the portion corresponding to /d/ (hand-segmented) between a low frequency band (50-500 Hz) and a high-energy band (500-5000 Hz); the more open the constriction, the greater the amount of energy is expected in the higher zone of the spectrogram, resulting in a smaller difference in energy between the lower and the higher frequency bands.

In a third metric we calculate the first difference of the intensity curve in steps of 1 ms. Then we extract the maximum value (MV) in the velocity curve between the intensity minimum corresponding to /d/ and the maximum corresponding to the following vowel. This measurement reduces possible effects of variation in the intensity of the following vowel and focuses on the abruptness of the transition between consonant and vowel: more weakened consonants have a less abrupt transition irrespective of the time from the intensity minimum to the intensity maximum and, therefore, a lower MV than more constricted consonants. This procedure is a simplified adaptation of the method described above [13].

We manually segmented the portion of the wave corresponding to /b/, /d/ or /g/ and the following segment, using Praat [18]. Segmentation was done in two parts: (1) the onset and offset of the consonant were located by hand, and (2) the intensity minimum (during the consonant) and maximum (during the post-consonantal vowel) were hand-marked through visual inspection of the Praat-generated intensity curve. Tokens were coded as to the preceding segment, their position in the word (initial, medial) and their lexical stress configuration (stressed, unstressed). Utterance-initial tokens were excluded.

We report results only for /d/, since for the other two consonants we judged that an insufficient number of tokens had been obtained in crucial environments for any meaningful statistical treatment. For the statistical analysis we applied mixed effects regression modeling, using individual speaker and lexical item as random factors. The comparisons that we make are slightly different for the two languages due to differences in the vowel inventories.

3. Results

3.1. Spanish

The dataset comprised a total of 241 Spanish /d/ tokens from 20 speakers. The data were visually explored in order to see whether unimodal or bimodal distributions were found. Kernel density plots were obtained for all three acoustic metrics, ID, ST and MV. Density plots display the estimated probability density function of a continuous random variable. These plots are similar to histograms but provide a continuous estimate of the distribution of a variable, rather than grouping observations into discrete bins. The kernel density plots clearly showed unimodal distributions for all three metrics (see Figure 1).

![Figure 1. Density plots for three acoustic indexes of constriction degree in 241 tokens of Spanish /d/ recorded by 20 speakers.](image-url)

For the statistical explorations we classified the tokens as a function of the preceding segment in the following groups: after /n/ (“nasal”, N = 19), after /s/ (“sibilant,” N = 30), after /u/ (“high,” N = 48), after /e o/ (“mid,” N = 63), and after /a/ (“low,” N = 70). There were only three tokens after /l/ and therefore these were excluded from our analysis. Lexical stress and word position were initially included in the statistical models. The three datasets, one per acoustic metric, were submitted to mixed-effects regression models with three fixed predictors (preceding sound, stress, word-position) and speaker and lexical item as random factors. However, stress and word-position yielded non-significant results. Consequently, the models were refitted with the first one as the only predictor. These are reported below.

All models returned significant effects of preceding segment. For all initial regression models the preceding level “high vowel”
(/i u/) was used as the intercept and all other levels were compared to this one. Regarding ID, the model revealed that both preceding “low” (t = -5) and “mid” (t = -2.7) vowel contexts differed from /d/ tokens in high-vowel contexts, while neither “nasal” nor “sibilant” contexts differed from preceding high vowels. In particular, low- and mid-vowel contexts resulted in more weakened /d/ tokens than “high” vowel contexts do, while “nasals” and “sibilants” do not precede /d/ tokens significantly different from those preceded by high vowels. Finally, regarding MV, the regression model returned significant effects for “low” vowels (t = -2.7), “nasals” (t = 4.3) and “sibilants” (t = 3.7) but not for “mid” vowels. In sum, preceding low vowels were found to trigger more weakened /d/ tokens than preceding high vowels while preceding nasals and sibilants triggered more constricted /d/ tokens than preceding high vowels.

However, the three regression models were not designed to explore potential differences between the /d/ tokens in nasal context and those in sibilant context. Thus, the three models were refitted with a change in the order of levels of the fixed predictor: “sibilant” was the new intercept. None of the three models returned significant differences between /d/ tokens in nasal context and those in sibilant context. All other results were consistent with those of the original three regression models.

3.2. Catalan

The Catalan dataset included a total of 264 /d/ tokens from 20 participants. Kernel density plots corroborated that the three acoustic metrics revealed unimodal distributions for Catalan /d/ as well (see Figure 2).

Figure 2. Density plots for three acoustic indexes of constriction degree in 26 tokens of Catalan /d/ recorded by 20 speakers.

The tokens were classified according to the preceding segment in the following groups: after /n/ (“nasal”, N = 31), after /s/ (“sibilant”, N = 53), after /i u/ (“high”, N = 37) and after schwa (N = 99). Note that Majorcan Catalan schwa is phonemic and allowed in stressed syllables unlike in other Catalan dialects. Other contexts were not included because not enough tokens were found to run meaningful statistical tests. Two other factors were initially considered: lexical stress and word position. Although both factors yielded significant effects for at least one of the three acoustic metrics, significant correlations were detected between these predictors and the first fixed factor, “preceding segment”. The decision was made to exclude both of these prosodic factors from follow-up tests, which are reported here. These models included “preceding factor” as the sole fixed predictor and speaker and lexical item as random intercepts.

The intercept in all three follow-up models was the high-vowel context. Regarding ID, it was found that /d/ tokens in all (other) three preceding segments differed from those in high-vowel context: “schwa” (t = -2.1), “nasal” (t = 2.6), “sibilant” (t = 5); that is, /d/ tokens are more weakened when preceded by schwa than by high vowels and more constricted when preceded by /n/ or /s/ than by high vowels. On the other hand, regarding ST, only the “sibilant” context (t = -3.8) differed from the “high” context, but not the “schwa” or “nasal” contexts. Finally, regarding MV, the two consonantal contexts (/n s/) were found to be different from the intercept (“sibilant” t = 6; “nasal” t = 5.6), but /d/ tokens preceded by schwa were not different from those preceded by high vowels.

The models were refitted with “sibilant” as the new intercept. Interestingly, the intercept was found to be slightly different from the context “nasal” in two of the three models: ID (t = 2) and ST (t = 2.1), but not MV (t = 0.3). It appears that, regarding MV, /d/ tokens preceded by consonants were different from those in vowel contexts but no differences were detected between the two vowel contexts nor between the two consonant contexts. The other two metrics revealed differences between /d/ tokens in nasal context and those in sibilant context, with the latter resulting in more constricted /d/ tokens than the former. Only one of the three measurements was able to reveal a difference between the /d/ tokens preceded by schwa and those preceded by high vowels, with the latter being more constricted than the former. The fact that only one measure found this difference may suggest that it should not be considered more than a mere tendency, unlike in Spanish, where differences in the height of the preceding vowel triggered robust differences in the degree of constriction of /d/. However, note that the reliable difference in Spanish was between high- and low-vowel contexts (mid-vowel contexts were not robustly different from high-vowel contexts, at least not for all three acoustic metrics) while in Catalan we could only explore possible differences between /d/ tokens in schwa- and high-vowel contexts.

3.3. Catalan vs. Spanish

Do Spanish and Catalan /d/ tokens differ generally in their constriction degree? Data were submitted to three independent mixed-effects models with ID, ST and MV as responses, respectively, and language (Catalan vs. Spanish) as fixed factor. Random factors were individual speaker and lexical item. The two measurements based on the intensity curve, MV and ID failed to yield significant effects of “language”: MV (Catalan (intercept) 〈β = 482.7, Spanish β = -70.9, t = -1.4), ID (Catalan (intercept) 〈β = 12.8, Spanish β = -2, t = -1.4). On the other hand, the measure based on the spectral envelope, ST, yielded a significant “language” effect: Catalan (intercept) 〈β = -11.8, Spanish 〈β = 3.9, t = 3.7, p < .001). Thus, even though the
average values show a tendency for Catalan /d/ to be slightly more constricted than Spanish /d/, the evidence that Catalan and Spanish /d/ differ in their overall constriction degree is not robust, at least when we take into account intensity-based measures. The attested ST difference suggests that there might be a difference in the overall spectral structure of Catalan and Spanish /d/, with Catalan /d/ tokens presenting lower (negative) values than Spanish, and thus potentially more constricted.

The three metrics seem to capture (some aspect) of the constriction/weakening differences between /d/ tokens in different phonetic environments, across two languages. However, they are only modestly correlated with each other, except for the two measurements based on the intensity curve, of course (MV * ID r = 0.85; MV * ST r = -0.5; ID * ST r = -0.51).

4. Discussion

The present paper has examined the effects of the preceding sound on the constriction or weakening degree of /d/ in two corpora of spontaneous, interactive speech, one per language. The corpora provided sufficient data to perform meaningful analyses only for the following contexts: postvocalic, after /s/ and after /t/. On the other hand, the post-vocalic tokens were sufficiently numerous to enter vowel height in the statistical analyses.

Traditionally, /d/ is said to have two allophones in complementary distribution (stop vs. approximant or fricative), depending exclusively on the nature of the preceding segment. The acoustic metrics examined here provide evidence for the existence of a continuum in degree of constriction so that, even within the intervocalic context, the relative height of the preceding vowel introduces differences in the degree of constriction of /d/. This effect was robust in the Spanish data (high-, mid-, vs. low-vowel contexts) and a trend in the Catalan data (high-vowel vs. schwa contexts). Our acoustic methods were not able to detect any differences in their degree of constriction between /d/ tokens in /s/ vs. /n/ contexts, at least for Spanish. In fact, the methods reported a higher constriction degree in /s/ tokens than in /n/ tokens in the case of Catalan. These findings may be due to a limitation of the acoustic metrics used here to index degree of constriction. Recent comparative (articulatory-to-acoustic mappings) research has shown that in the /n/ context acoustic methods may not accurately reflect the presence of oral constriction [15].

The acoustic methods designed to index constriction degree in Spanish and Catalan /d/ provide quantitative data distributed over unimodal densities. This is true for all cases. The evidence suggests, therefore, that, for the dialects examined here, differences in constriction do not result (only) from the application of a categorical rule that acts over discrete allophones.

Examinations of all three acoustic metrics result in reasonably similar patterns, with /d/ being least constricted after a low vowel and most constricted after a consonant. However, there are differences of detail as a function of the specific metric. Is Spanish /d/ different from Catalan /d/ in its constriction degree? The metric based on the spectral envelope of the consonant (ST) revealed a difference between Catalan and Spanish /d/ in their constriction degree. However, this difference was not detected in other acoustic metrics. Interestingly, a difference exists for /b/ between these two languages [16].

5. Conclusion

We have obtained a more accurate understanding of spirantization allophony (in Iberian, Majorcan) Spanish and (Majorcan) Catalan, adding to other work on this topic. Although the presence/absence of complete occlusion can be treated as a binary feature, our results suggest that there is a unimodal continuum in the degree of constriction of /d/.

The present investigation contributes to the study of the nature of speech reduction processes. Iberian spirantization (as reflected in Majorcan Catalan and Spanish) seems to be a gradient speech-reduction process rather than, only the result of applying a categorical, phonological rule. Beyond a binary choice between stop and approximant allophones, the specific degree of constriction of /d/ is conditioned by the nature of the preceding segment, so that, even among tokens in the postvocalic context, different degrees of constriction are found depending on the height of the preceding vowel.

6. References


