Quantifying Iberian Spirantization: Acoustics and articulation

José Ignacio Hualde*
Miquel Simonet^
Ryan Shosted*
Marianna Nadeu*

*University of Illinois at Urbana-Champaign, ^University of Arizona

LSRL40, Seattle, March 26-28, 2010
Why Spanish spirantization?


- Recent phonetic work has challenged the correctness of traditional descriptions.
Traditional description

Spanish /bdg/ realized as:

- stops in utterance-initial position and after a homorganic nasal or lateral,
  - barro
  - damo
  - gata

- “fricatives” (= approximants) elsewhere.
  - loba
  - oda
  - lago

(Navarro Tomás 1977, Harris 1969:37)
Spectrographic analysis reveals substantial variation in the degree of constriction of /bdg/ in lenition contexts (Quilis 1981:224).
2 allophones in complementary distribution?

Recent experimental work suggests that the realization of /bdg/ involves a continuum of degrees of constriction.


Our study:

Goals:

Find out the effect of the preceding segment on the degree of constriction of /bdg/.

Is traditional description (2 allophones in complementary distribution) really wrong?
Our study

- Acoustics: estimation of degree of constriction from sound wave
- Articulation: EPG
Participants: 20 native speakers of Peninsular Spanish (10 females).

Task: Participants had to guess the identity of a famous person by asking at most 20 questions.

Recordings: Participants were wearing head-mounted microphones and high-quality recordings were obtained.

Data were originally collected for a different purpose. For more information, see Simonet (2008).
3 acoustic measurements

- 3 acoustic measurements
  - a) Energy difference in CV
  - b) Spectral tilt
  - c) Maximum Velocity (after Kingston 2008)

- Articulatory measurement: EPG
Focus on /d/:
- More balanced numbers of tokens in all relevant environments.
- Only segment for which EPG data can be obtained.
Acoustic analysis

/d/ = 241 tokens

after /n/ = 19
after /s/ = 30
after /i, u/ = 48
after /e, o/ = 63
after /a/ = 70
Data analysis

Manual segmentation of the portion of the sound wave corresponding to /d/ and the following segment (using PRAAT).

Measurements were obtained automatically with a script written for this purpose.
Energy difference in CV (EnDiff)
Examples

Spectrogram of más de ‘more than’
Examples

Spectrogram of *universidad* ‘university’
Examples

Spectrogram of *mundo* ‘world’ (expected stop)
Energy difference


Our improvement: Intensity was calculated after applying a Hann pass band filter between 500 and 10,000 Hz to the signal in order to exclude possible effects of (high and) low frequency noise. This filtering procedure excludes the effects of the voice bar (F0).
Spectral tilt

Difference in energy in the portion of the signal corresponding to the consonant between a low frequency band (50-500 Hz) and a high energy band (500-5000 Hz).
Maximum Rising Velocity (MaxVel)

Modified from Kingston (2008)

The intensity curve is extracted with the same settings as in our Intensity Difference measurement. After that, we calculate the first difference of the intensity curve in 0.001 s steps. Finally we extract the maximum value of the first difference that is found between the intensity minimum corresponding to the consonant and the maximum corresponding to the following vowel.

The maximum of the first difference indicates the maximum rising velocity between the midpoint of the consonant and the midpoint of the following vowel.
MaxVel (cont.)

This measurement reduces possible effects of variation in the intensity of the vowel and focuses on the abruptness of the transition between consonant and vowel.

More lenited consonants have a less abrupt transition irrespective of the time from the intensity minimum to the intensity maximum.
Hawkins (2010), JPhon, Fig 3: Acoustic correlates of [+cons] and [-cons], as defined by K.N. Stevens. There is an abrupt change in the [+cons] case, [b], and very little change in the [-cons] case, [w].
- Three mixed-effects models with speaker and lexical item as random factors, preceding segment (high, low, mid, nasal, sibilant) as fixed factor and one of the three acoustic measurements (IntDiff, MaxVel, Spec-Tilt) as response were fitted to the data.

- For these regression models the preceding level high (i.e. high vowels /i, u/) was used as the intercept and thus all other levels were compared to this one.
Results, acoustics: Density plots
Intensity difference /dV/ by preceding segment
### Statistics Intensity Difference

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>13.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-7.82</td>
<td>-5.09</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mid</td>
<td>-4.41</td>
<td>-2.76</td>
<td>= 0.006</td>
</tr>
<tr>
<td>Nasal</td>
<td>3.54</td>
<td>1.63</td>
<td>= 0.1</td>
</tr>
<tr>
<td>Sibilant</td>
<td>3.32</td>
<td>1.7</td>
<td>= 0.9</td>
</tr>
</tbody>
</table>
MaxVel of /dV/ by preceding context

![Box plot showing MaxVel of /dV/ by preceding context in Spanish.](image)
## Statistics Maximum Velocity

<table>
<thead>
<tr>
<th></th>
<th>( \beta )</th>
<th>( t )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>432.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-158.4</td>
<td>-2.77</td>
<td>0.005</td>
</tr>
<tr>
<td>Mid</td>
<td>-114.7</td>
<td>-1.93</td>
<td>0.05</td>
</tr>
<tr>
<td>Nasal</td>
<td>349</td>
<td>4.37</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sibilant</td>
<td>266.4</td>
<td>3.71</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
SpecTilt of /d/ by preceding context
## Statistics Spectral Tilt

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-10.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>8.15</td>
<td>7.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mid</td>
<td>2.86</td>
<td>2.53</td>
<td>= 0.01</td>
</tr>
<tr>
<td>Nasal</td>
<td>-2.31</td>
<td>-1.52</td>
<td>= 0.1</td>
</tr>
<tr>
<td>Sibilant</td>
<td>-2.48</td>
<td>1.81</td>
<td>= 0.07</td>
</tr>
</tbody>
</table>
Partial conclusions

1. Distribution according to all 3 acoustic measurements (density plots) is more compatible with an allophonic continuum than with two allophones in complementary distribution (unimodal distribution)

2. Influence of preceding segment on degree of constriction of /d/:
   a- > e-, o- > i-, u- > s-, n-
   (most clearly seen with MaxVel)
In order to find out whether /d/ tokens preceded by nasals and sibilants differed from each other, we refitted the three mixed-effects models with a change in the order of levels of the preceding segment factor.

In particular, the three models were refitted with sibilant as the intercept, instead of high vowel. None of the three models returned significant differences between /d/ tokens preceded by nasals and those preceded by sibilants:
/sd/ vs /nd/

IntDiff (\(\beta = 0.22, t = 0.09, \text{ ns}\)),
SpecTilt (\(\beta = 0.17, t = 0.1, \text{ ns}\)),
MaxVel (\(\beta = 82.64, t = 0.9, \text{ ns}\)).
All other results were consistent with those of the original three mixed-effects models.
/sd/ vs /nd/

In sum:
/d/ tokens preceded by nasals did not differ in their constriction degree from those preceded by sibilants.

The scale remains:

- a- > e-, o- > i-, u- > s-, n-
- open closed
Articulatory study: Electropalatography
CAVEAT

We need to confirm with more speakers!

(Author’s results only valid if they are later confirmed with data from other speakers)
/anda/
/alda/
/arda/
EPG. moment of maximum constriction in the production of /d/ in each sequence.
Results for speaker JIH: Presence of complete occlusion

Complete occlusion* obtains in:
0/4 [Vd] tokens (e.g. cada)
4/4 [nd] tokens (e.g. anda)
4/4 [ld] tokens (e.g. caldo)
3/4 [s(#)d] tokens (más da, desde)
0/4 [rd] tokens (e.g. cardo)

* Contact occurs across an entire horizontal row of the palate
Results for speaker JIH: Duration of occlusion (including occlusion of /n/ and /l/)

For those types where occlusion obtains, the entire duration* of occlusion is (avg)

130 ms [nd]
140 ms [ld]
17 ms [s(#)d]
Duration of [d]-portion of clusters

For [nd] and [ld] occlusions, closer inspection of the waveforms indicates that:
[d] in [nd] is on average 33 ms
[d] in [ld] is on average 52.5 ms
Diminution in voiced amplitude was taken as the primary correlate of [d] in both clusters. An additional physiological channel (nasal airflow or pressure) would be helpful for separating the two segments in [nd]. Separating [l] and [d] may be less tractable since lateral occlusion may be linguodental, thus not observable using EPG.
Conclusions from EPG study

- /d/ always [ð] in /ad/, /rd/
- /d/ always [d] in /nd/, /ld/
Conclusions from EPG study

- In /sd/
  - Traditional description: [ð]
  - Our acoustic study: no difference between /nd/ and /sd/

- EPG: sometimes occluded, but not always. Shorter and less extensive occlusion when found.
Vowel height effects
/pudo/
Acoustics vs. EPG

Acoustics: degree of constriction.

EPG: extent and duration of contact.
(interpretable as +/- full occlusion)
General conclusions

- /sd/
- /id/ vs. /ad/
Thank you very much!
References


References (cont.)


Thank you!