Sound change and rhythm in Altiplano Mexican Spanish  
Gillian Marchini and Michael Ramsammy  
University of Edinburgh

Vowel compression in Spanish is much debated. This paper presents evidence from Altiplano Mexican Spanish (AMS) that phonologically, prosodically and phonetically constrained coda-driven vowel compression is variety-specific in Spanish.

Vowel compression is the phonological process in which vowels are shortened relative to syllable structure, i.e., the greater number of segments within a syllable constituent, the more extreme the shortening. Thus, for example, the nucleus of an English CCVCC syllable like dreads exhibits more extreme compression than that of a CVC syllable, i.e., dead. Maddieson’s (1985) Closed Syllable Shortening Principle encapsulates this, claiming that compression is coda-driven in all languages, i.e., vowels in closed syllables are shorter than those in open syllables.

Language-specific research however questions the universality of compression effects. Findings from Katz (2012) show that, in English, durational differences between vowels in open and closed syllables are negligible and that coda complexity has no effect on degree of compression. Moreover, cross-dialectal comparisons reported in Aldrich & Simonet (2019) reveal that onset complexity causes greater compression than coda complexity in Spanish. In both studies, onset-induced compression is consistent across all speakers whilst coda-induced compression is speaker-specific. Particular to the present study is the following quote: "we are aware of no published finding suggesting that, in Spanish, coda presence (or complexity) drives compensatory vowel shortening" (Aldrich & Simonet, 2019:268)

Although this may be true for the Aldrich & Simonet’s aggregate findings – based on cross-dialectal data – dialect-specific analysis in this paper reveals that coda-induced compression exists in Spanish. Accordingly, we observe pairs such as those in (1) in which the closing of an open syllable (here through inflectional operations) triggers compression of the nuclear vowel.

(1) Vowel compression in AMS

<table>
<thead>
<tr>
<th>ˈɡɾa.ɲ’dɛ</th>
<th>ˈɡɾa.ɲ’dɛs</th>
<th>ˌes.ˈta.βa</th>
<th>ˌes.ˈta.βa.n</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘big’ (SING)</td>
<td>‘big’ (PLURAL)</td>
<td>‘he/she/it was’ (3rd PERSON SING)</td>
<td>‘they were’ (3rd PERSON PLURAL)</td>
</tr>
</tbody>
</table>

Our data comes from experimental methods designed to test for compression effects in AMS, a variety of Spanish local to Mexico’s central valley and highlands. Continuously read speech was first elicited from six female participants and segmented using the Montreal Forced Aligner. Once manually checked and coded, the durational and formant-frequency measurements for stressed and unstressed, low and mid vowels (/e, o, a/) were compared. Results reveal that coda-driven compression shortens, centralises and reduces the intensity of atomic vowels in this variety. In particular, statistical analysis noted that these effects are:

1. Phonologically constrained: compression occurs in word-final, post-tonic closed syllables. Although complex onsets partially shorten the vowel, these are negligible when compared to coda effects.
2. Phonetically constrained: compression effects are most extreme when the coda is /s/.
3. Prosodically constrained: compression is most extreme in domain-final positions of large prosodic units, i.e., utterance phrases, rather than smaller ones, i.e., syllables.

The results therefore suggest that coda-driven compression occurs in AMS and that, potentially, it is a variety-specific phenomenon. The existence of the compression effects summarised above bears further upon a theoretical point relating to phonological timing in Spanish: namely, that with regard to compression effects, dialect-specific phonetic-
phonological interactions may allow certain varieties of Spanish, such as AMS, to behave in a way typically associated with stress-timed languages, i.e., English. This is significant given that previous analyses of phonological timing categorised Spanish as a prototypical, syllable-timed language (Dauer 1983; Pike 1945).

We further link these findings to wider debates concerning phonological rhythm, specifically the role of timing and prominence. Firstly, the variation in both quantitative and qualitative properties between stressed and unstressed vowels supports the theory that phonological rhythm does not solely arise from timing, but rather from the marking of prominence. Although timing may contribute to the marking prominence, other acoustic properties, e.g., intensity, formant-frequency, may interact to indicate stress. As such, the terms stress-based and less stress-based seem more appropriate for a cross-dialectal characterisation of Spanish than the fixed, dichotomous categorisation of stress- and syllable-timed languages (Dauer 1983). Secondly, our results contribute further evidence in support of the claim that there is not one universal metric for rhythm. On the contrary, rhythm is a complex phenomenon that may depend, to various degrees, on phonological and prosodic constraints of the language and variety, as well as dialect-specific phonetic interactions with phonology (Arvaniti, 2009, 2012; Turk & Shattuck-Hufnagel 2013). Thus, the acoustic realisation and uniformity of rhythm (prominence) are conditioned by unique phonologies of specific language varieties (Turk & Shattuck-Hufnagel 2013; Smith & Rathcke 2020).

In view of our results from AMS, we argue that further work is needed into the apparent fluidity in how languages, and specific varieties within them, express prominence (Arvaniti, 2009, 2012; Smith & Rathcke 2020, Turk & Shattuck-Hufnagel 2013). We see particular advantages in further research that would seek to understand variation in prominence in Spanish along a dialectal continuum, potentially crossing into the phonetic territory of what have previously been thought to be rhythmically distinct languages, e.g., Portuguese (Frota & Vigário, 2001).

References


Non-binary language forms in Spanish: consciously using it facilitates processing during comprehension?
Noelia A. Stetie (CONICET / UBA) nstetie@conicet.gov.ar
Gabriela M. Zunino (CONICET / UBA) gmzunino@conicet.gov.ar

There is empirical evidence in different languages on the computation of gender morphological marks during psycholinguistic processing and the incidence of this process in the construction of mental representations on gender stereotypes (Bradley 2020, Prentice 1994, Prewitt-Freilino et al. 2012). A central phenomenon of this line of studies is the extent to which the so-called generic masculine functions as a generic and works to represent groups of people with a non-uniform gender. Empirical studies in German, Italian, French and Polish (Cacciari et al. 2007; Gygax et al. 2008, Maciuszek et al. 2019, Misersky et al. 2019) show that the bias generated by the generic masculine is identifiable and seems to have a sustained effect on cognition. However, there are not many analyses of the underlying psycholinguistic process in Spanish, especially studies that consider the two non-binary morphological variants usually used: -x and -e. Our aim is to analyze if: 1. the three morphological variants (-o, -e, -x) generate a generic representation; 2. there are biases on the preference of speakers regarding non-binary morphological variants; 3. there are any processing costs associated with non-binary forms; 4. the stereotipicality of the role nouns, specially associated with the frequency of use of non-binary forms, has an effect on processing costs.

In order to evaluate the processing of non-binary forms in Spanish, we programmed and performed a sentence comprehension task in PCIbex (Zehr et al. 2018). We presented sentences with simple noun phrases (see 1, 2 and 3). After reading the sentences, participants had to answer a multiple choice question about the possible reference of that nominal phrase. They had to indicate if it referred to a group of women, men or a mixed group. Sentence reading times, response types and times were recorded. The task had a 3x3 factorial design: morphology (-o, -x, -e) and stereotipicality of role names (low, medium and high). The items were presented in 3 counterbalanced lists: 18 items and 30 fillers.

1. Low: Los niños siempre empiezan el jardín con un período de adaptación inicial.
   Children always start kindergarten with an initial adaptation period.
2. Medium: Los cocineros siempre organizan los ingredientes sobre la mesada para cocinar.
   The chefs always organize the ingredients on the cooking table.
3. High: Los plomeros con matrícula pueden hacer trabajos en edificios y consorcios.
   Licensed plumbers can do work in buildings and consortiums.

We evaluated 538 participants (age: $M=34.24$; $SD=11.63$; min=18; max=98): 386 women, 131 men and 21 non-cisgender. Participants were asked if they used any form of non-binary language and 136 stated that they did not, 111 that they used it occasionally, and 291 that they used it frequently.

We used Linear Mixed Effects Models for the analysis. Regarding the sentence reading times (Graphic 1), morphological variants and stereotipicality of the role name didn’t show a statistically significant difference. Instead, the times did vary depending on the frequency of use of non-binary forms ($\beta_0=6574.4, t=22.714, p=<2e-16; \beta_1_{\text{little}}=-169.0, t=-0.416, p=0.67756; \beta_1_{\text{frequent}}=-929.2, t=-2.794, p=0.00539$). Particularly, the difference was between participants who declared using on a regular basis non-binary forms and those responding they didn’t use them at all ($p=0.00539$), the first ones took shorter reading times.

Considering the response times for the mixed groups responses (Graphic 2), we found a main effect of morphology, stereotipicality and frequency of use of non-binary forms. The morphology factor generated a main effect ($\beta_0=6945.61, t=40.981, p=2e-16; \beta_1_{\text{x}}=-1360.16$,
In the opposite direction to what we have hypothesized: phrases with generic masculine morphology required longer response times than those with either of the two non-binary morphological variants. Besides, as expected, for high stereotipicality role nouns participants took longer response times ($\beta_1 = -1306.18$, $t = -10.833$, $p = 3.63e-15$). We also found a main effect for the use of non-binary forms ($\beta_2_{medium} = -19.08$, $t = -0.170$, $p = 0.86619$; $\beta_2_{high} = 321.09$, $t = 2.804$, $p = 0.00755$). We also found a main effect for the use of non-binary forms ($\beta_3_{little} = -293.93$, $t = -1.460$, $p = 0.03693$; $\beta_3_{frequent} = 344.80$, $t = -2.092$, $p = 0.03693$): participants that frequently use non-binary forms took shorter reading times that those who use them occasionally and never.

When analyzing the type of responses, we found that, in the face of non-binary forms, all participants consistently chose a mixed reference. However, with the generic -o, there was an interaction with the stereotipicality factor: with role names that were not so strongly stereotyped, the answer of the mixed group was chosen (65%), but with those of high stereotipicality, the most chosen answer implied the exclusive reference to men (66%). These results taken together suggests that non-binary forms do not require a higher processing cost, and generate systematically generic representations, as opposed to the generic masculine which fails to generate a generic representation for all role nouns, and it depends on the stereotipicality of the role noun.

References
Misersky et al. (2019). Grammatical gender in German influences how role-nouns are interpreted: Evidence from ERPs. Discourse Processes, 56(8), 643-654.
Zehr et al. (2018). PennController for Internet Based Experiments (IBEX).
Sociolinguists have observed an ideological link between African American English (AAE) and the performance of certain ‘masculine’ personae. For example, AAE can serve as a resource for non-Black men to index physical power/toughness [1, 2], and for Black drag performers to exaggerate a stylistic contrast when flouting a hegemonically feminine ‘white woman’ persona [3]. These links, as well as the robust association between F0 and gender identity [4, 5], contribute to the prevailing folk linguistic notion that Black Americans have lower pitched voices than American speakers of other ethnicities. Several studies have sought to investigate whether such a difference exists between the speech of Black and white men, but have returned mixed results [6, 7]. Some studies claim that Black men employ a wider pitch range, while others claim they use a narrower range, and still others argue they have an overall lower F0. This lack of consensus may be partially due to differences in the measures, measurement sites, and linguistic content comprising the data in these studies. Moreover, there is a paucity of work exploring F0 differences between Black and white women, especially as a controlled comparison to men. In this study, we investigate Black and white speakers’ mean F0 and pitch range in identical read phrases. Results from linear mixed effects regression models show differences according to self-reported gender for both Black and white speakers, and according to self-reported ethnicity (Black vs. white) for female—but not male—speakers.

A total of 94 recordings of 47 Black (male = 33, female = 14) and 47 white (male = 32, female = 15) speakers (mean age: 21.4) reading the same passage (the number of phrases = 33; average length = 7.9 words/phrase) were drawn from the Memphis corpus [8]. Phrase boundaries were identified according to the text, regardless of idiosyncratic pausing/disfluencies, and were manually placed using Praat textgrids. Measurements for F0 mean, max, min, and standard deviation were automatically extracted from each phrase with the pitch floor at 75 Hz and ceiling at 300 Hz. Following established practice in the literature [9, 10], we computed a set of pitch span measures: +/- 2 standard deviations around the mean (SD4), as well as 80% range, 90% range and 100% range, computed from the semitone-transformed (ST) F0 max and min. Group averages for each measure are summarized in Table 1.

We fit separate linear mixed-effects models for mean F0 and pitch range, with Ethnicity (Black/white) and Gender (male/female) as fixed effects, and Speaker and Phrase as random effects. In order to generate all relevant comparisons we refit the same model configuration with different reference levels and contrast coding schemes. Results show that Black women have significantly lower mean F0 than white women ($\beta = -0.12$, $p = 0.01$). However, Black men do not significantly differ from white men in this respect ($\beta = -0.00$, $p = 0.92$), as shown in Figure 1. For both ethnic groups, men have significantly lower mean F0 than women (white men < white women, $\beta = -0.60$, $p < 0.001$; Black men < Black women, $\beta = -0.49$, $p < 0.001$). In terms of pitch range, as shown in Figure 2, Black women do not significantly differ from white women ($\beta = -0.57$, $p = 0.28$), and Black men do not significantly differ from white men ($\beta = 0.24$, $p = 0.50$). There is still a main effect of Gender as white men have significantly narrower range than their female counterparts ($\beta = -2.15$, $p < 0.001$), as do Black men ($\beta = -3.14$, $p < 0.01$). These patterns still hold when we compare ethnic and gender differences in 90% range, 80% range and SD4 span (Figure 3).

In short, for both ethnic groups, men use a lower mean F0 and narrower range (e.g., 100%, 90%, 80% range and SD4) than women. Meanwhile, even though Black women tend to have lower mean F0 than white women, they do not differ in these range measures. Conversely, no difference is found between Black and white men in both mean F0 and pitch range measures. Our findings therefore do
not support the basic stereotype of Black men having a lower F0, and underline the need to further explore these variables in women’s speech. Moving forward, we are conducting SSANOVA analyses on the F0 contour shapes of individual phrases, to investigate group differences in F0 implementation across the phrase.


![Figure 1: Mean F0 (Hz) for Black and white speakers](image1)

![Figure 2: 100% range (ST) for Black and white speakers](image2)

![Figure 3: SD4 (Hz) for Black and white speakers](image3)

| Table 1: Mean F0 and pitch range measures for Black and white speakers |
|------------------|------------------|------------------|------------------|------------------|------------------|
|                  | Mean F0 (Hz)    | 100% range (ST) | 90% range (ST)  | 80% range (ST)  | SD4 (Hz)         |
| Black women      | 183.95          | 6.94            | 6.25            | 5.55            | 68.58            |
| Black men        | 112.48          | 5.61            | 5.05            | 4.49            | 35.29            |
| white women      | 206.16          | 7.52            | 6.77            | 6.02            | 81.71            |
| white men        | 113.30          | 5.37            | 4.83            | 4.30            | 35.05            |
This paper reports on an experiment designed to measure how listeners’ perceptions of speaker age and ethnicity condition identification of lexical items with THOUGHT / LOT vowels in New York City English (NYCE). Several independent studies have recently reported evidence of THOUGHT-lowering and/or LOT/THOUGHT merging in NYCE led by younger non-whites. Spoken corpus data by Wong (2012) and Becker (2010) suggest rapid THOUGHT lowering, particularly in Asian and Latinx communities. Similarly, younger Asian and Latinx NYCE speakers favor merged LOT/THOUGHT responses in controlled homophony judgment tasks (Johnson 2010, Haddican et al. 2016). Moreover, matched-guise results by Becker (2014) suggest that raised THOUGHT is associated mainly with older whites. Unaddressed in this literature is whether listeners use perceived social information about the speaker—i.e. perceptions of age and ethnicity—in their phonemic categorization of low back vowels in comprehension of NYCE (Rubin 1992, Hay, Warren and Drager 2006). Here, we report results from a forced-choice lexical identification experiment intended to investigate this.

Subjects were 140 self-reported native speakers of New York City English from a range of self-reported ages and ethnicities, who participated via a web-based application in the spring of 2020. In each trial, participants were presented with (i) a picture of a speaker, and (ii) a single-word auditory stimulus produced by one of six different NYCE-native white women, all native LOT/THOUGHT distinguishers, whose LOT/THOUGHT items were clearly distinguished in acoustic space. The task was to identify a lexical item (using standard English orthography) associated with the auditory stimulus, i.e. [kɑt] as ‘caught’ or ‘cot’. Each auditory stimulus was accompanied by a picture of one of six women, which judges, in a prior norming stage, categorized into one of two age categories (older, $M=46.5$ vs. younger $M=23.4$) and one of three ethnicities (Latina, White, Asian). Faces and stimulus voices were rotated across six lists. Subjects heard a total of 72 different lexical items including six THOUGHT and six LOT words accompanying each of the six different faces (older/younger Latina, older/younger White, older/younger Asian).

Generalized linear mixed effects regression modeling revealed a main effect of stimulus picture ethnicity such that LOT stimuli were more likely to be categorized as CAUGHT when accompanied by a picture of an East Asian or Latina face than a white face (Figure 1). The analysis revealed no significant main effect or interaction for stimulus picture age, nor, importantly for subjects’ reported age and ethnicity.

These findings, in keeping with results on ethnicity effects in recent production and social perception studies, suggest that NYCE-native listeners from a range of age and ethnicities use social information about speakers in lexical categorization of LOT/THOUGHT items, and more generally support recent findings suggesting that social information is recruited in speech perception.
REFERENCES


Figure 1: **Mean accuracy by face ethnicity and vowel type.**
A higher proportion of merged responses for stimuli accompanied by East Asian and Latina faces vis-a-vis White faces.