

Do Spaniards speak faster than Mexicans? Studying Spanish rhythm in natural speech

Fabian Santiago¹ & Paolo Mairano²

¹Université de Paris 8 & Sorbonne Nouvelle, ²University of Warwick

Background. Previous studies examining dialectal variation across the Spanish-speaking world have focused on the segmental level (vocalic and consonantal variations) and prosodic patterns (mainly intonational differences). Many studies have analysed differences in temporal patterns across different varieties of English (Jacewicz et al, 2010), French (Schwab & Avanzi 2015), Italian (Russo & Barry 2004) and many other languages, but only few studies have compared Spanish varieties. These have pointed out that articulation rate is affected by extra-linguistic factors such as age and professional profile of speakers, but not by different dialects (cf. the study of Schwab, 2015, comparing Costa Rican and Barcelonan Spanish).

Research Goals. This study aims to fill this gap by comparing temporal patterns across two different varieties of Spanish: Castilian Spanish (Madrid) and Mexican Spanish (Mexico City). Our first goal is to investigate variation in articulation rate across these varieties, also considering the effect of other factors such as gender and speech style. Our second goal is to examine whether vowel reduction phenomena (namely the compression of the vocalic space) are related to differences in articulation rate across these two varieties. We hypothesize that a faster articulation rate should result in a more compressed vocalic space.

Methodology. A total of 22 monolingual speakers (12 Mexicans and 10 Spanish, aged 21-41, SD = 5, gender-balanced groups) were recorded in a soundproof room, for a total of 5.75 hrs. of speech. All speakers were university students or staff at the time of recording. Participants performed two tasks: a spontaneous speech task including a semi-directed interview and the description of a painting, and a reading task. The corpus was orthographically transcribed and then automatically segmented into phones, syllables and words with Easyalign. A careful manual segmentation check was performed. Additionally, overlaps, disfluencies, hesitations, filled pauses and noticeable lengthened words were marked in the data and subsequently excluded from the analysis.

The corpus amounts to 45k words and 7,977 IPUs (Inter-Pausal Units, see Fig. 1). For the present analysis, we excluded all IPUs containing less than 3 syllables and all IPUs in which one or more word(s) contained disfluencies, hesitations, filled pauses, etc. Therefore, the remaining 4,618 IPUs (209 IPUs per participant on average) consisted of speech chunks of continuous phonation without noticeably lengthened words.

Metrics. We used two metrics. The first one is the articulation rate (AR) as measured in number of syllables per second (excluding pauses). The second one measured the vocalic space. We automatically extracted F1 and F2 values at the mid-point of all the vocalic segments with a *Praat* script (band lower than 5kHz for males, and lower than 6kHz for females, with a 25ms Gaussian window). Formant values were converted into Barks for normalization. We measured the vocalic space of each speaker in terms of 'degree of dispersion' following Gendrot & Adda-Decker (2007): (i) we identified the gravity centre of the acoustic space by averaging F1 and F2 values of all vowels produced by each speaker; (ii) we computed the Euclidean distances from the gravity centre to each vowel's position within the vowel chart, then averaged them for the 5 peripheral Spanish vowels (separately for stressed and unstressed vowels); (iii) we summed the resulting averaged Euclidean distances for the 5 peripheral Spanish vowels (separately for stressed and unstressed vowels).

Results. Data were analysed with linear mixed effects models (statistical significance was set at the 0.01 level). Our data show that, in a global perspective, AR is affected by dialectal variation: Castilian speakers articulate more syllables per second than Mexicans (Fig. 2a). Speech style also affects the AR: speakers articulate faster in spontaneous speech than in the reading task, independently of their origin (Fig. 2a).

Additionally, the analysis shows that female speakers have a tendency to speak faster than male speakers in the Mexican group, whereas gender does not have an effect among our Castilian group (Fig. 2b). However, Castilian female speakers read faster than Mexican female speakers, whereas male speakers of the two dialectal variants produce similar ARs (Fig. 2c). In spontaneous speech, female speakers articulate faster than males within the Castilian group, whereas the opposite pattern is observed in Mexican speakers (Fig. 2d).

Finally, we analysed the compression of the vocalic space depending on two conditions: dialectal variation and stress status. In Fig. 3, we plotted F1 and F2 values in Barks (averaged across 72k vowels), and the gravity centre (indicated by the asterisk). We can observe a degree of vowel space compression in the

unstressed condition independently of the dialect. However, the vocalic space seems to be more reduced in the Mexican group than in the Castilian one.

Discussion and conclusion. Our study shows differences in AR across Spanish varieties: Castilian speakers articulate faster than Mexicans. Our findings also show that speech style affects both groups in a similar way: AR increases in reading but decreases in spontaneous speech. As for gender, this factor affects the AR but not in the same way across the two varieties analysed: on average, our male participants speak faster than females, except for Castilian speakers in the spontaneous task (where female participants seem to articulate faster).

We observed noticeable vowel space compression among Mexican speakers, but less so among Castilian speakers, despite the fact that the latter articulate faster. This is contrary to our initial hypothesis. It can perhaps tentatively be explained by the fact that Mexicans tend to speak slower and mark more clearly the distinction between stressed vs. unstressed vowel; instead, the higher speech rate among Castilian speakers only allows them to produce a smaller distinction.

References

- Gendrot, C. & Adda-Decker, M. (2007). Impact of duration and vowel inventory size on formant values of oral vowels: an automated formant analysis from eight languages. *International Conference of Phonetic Sciences 2007*, 1417-1420.
- Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America*, 128, 839-350.
- Russo, M. & Barry W.J. (2004). Interaction between Segmental Structure & Rhythm. A look at Italian Dialects and Regional Standard Italian. *Folia Linguistica*, 38, 277-296.
- Schwab, S. & Avanzi, M. (2015). Regional variation and articulation rate in French. *Journal of Phonetics*, 48, 96-105.
- Schwab, S. (2015). Las variables temporales en el español de Costa Rica y de España: un estudio comparativo. *Filología y Lingüística*, 41 (1), 127-139.

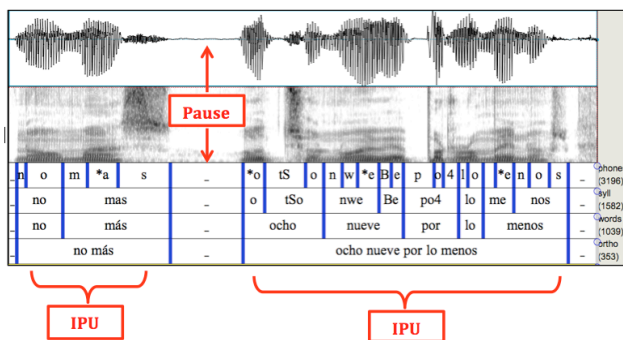


Figure 1

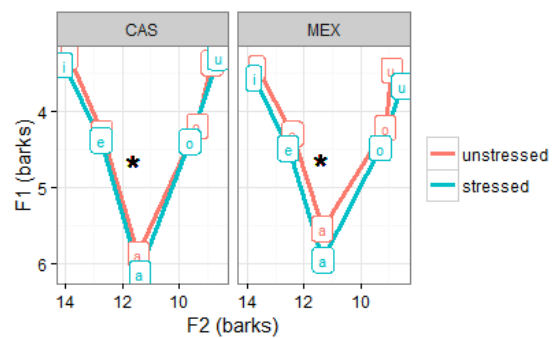


Figure 3

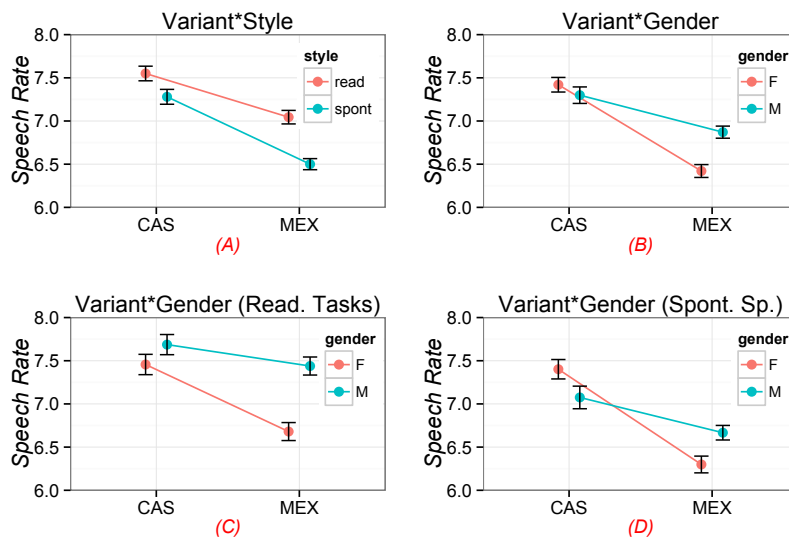


Figure 2