

Advancing statistical approaches to sociophonetic variation: A case study of Ecuadorian Spanish

Christina García (*Ohio State University*)

Given that sociophonetics brings together two distinct fields, it must straddle the fence between what is considered sound methodology in both sociolinguistics and phonetics (c.f. Foulkes et al. 2012). Following the advances in phonetics, scholars working on sociophonetic variation are now predominantly using continuous acoustic measures as the dependent variable. This poses a challenge for the traditional statistical tools used in sociolinguistics. Thus, we must question what the most appropriate analysis is for this kind of variation in order to properly examine the interaction of linguistic and social variables. I will show that inflated beta and proportional odds modeling are well suited for sociophonetic variation by looking at the case study of intervocalic /s/ voicing in Highland Ecuadorian Spanish (HES).

The variable realization of /s/ as [z] (for example [losamiyos] ~ [lozamiyos] ‘the friends’) has been attested in a few modern dialects of Spanish, including HES (Robinson 1979, among others). For this study, thirty-two natives of Loja, Ecuador, half males and half females, were recorded in a sociolinguistic interview and a reading task. The latter included 64 tokens of intervocalic /s/ in word-initial, word-medial, and word-final position. From the sociolinguistic interviews, 20 tokens of each position were extracted for each speaker, and percent voicing of /s/ was measured for over 3,000 tokens. Following Campos-Astorkiza (2011), I consider voicing as a ternary category: unvoiced (less than 20%), partially voiced (20-90%), and fully voiced (100%). Each token was also coded for local speech rate, stress, preceding/following vowel, and position. For continuous analyses, inflated beta linear models were fit to the data with percent voicing as the dependent variable. For discrete analyses, proportional odds models were fit to the data with the dependent variable as the three voicing categories detailed above. Speaker was included as a random effect.

Overall, the results show that intervocalic /s/ voicing in Lojano Spanish is a gradient process conditioned by speech rate, position, stress, and speaker gender. Voicing increases in faster speech; in word final and initial contexts as opposed to medial; when /s/ is between unstressed syllables; and in males’ speech. Unlike other statistical tests used in sociolinguistics, inflated beta regression does not assume that the data is normally distributed. This is important since many phonetic variables, such as percent voicing, do not exhibit a normal distribution. The categorical analysis using proportional odds improves upon traditional analyses in that it allows for more than two categories, and it assumes the categories are ordered, that is, that *voiced* is a higher category than *partially voiced*, which is in turn a higher category than *unvoiced*. These tests not only represent more accurately the structure of sociophonetic data, but also their simultaneous use provides complementary information that allows for a more comprehensive analysis. Following other scholars working in sociophonetics (c.f. File-Muriel 2012), I argue that we must push the boundaries of conventional sociolinguistic analysis in order to better account for the types of data that are now commonplace.

References

Campos-Astorkiza, Rebeka. To appear. Sibilant voicing assimilation in Peninsular Spanish as gestural blending. In *Romance Linguistics 2011*, ed. Marie-Hélène Côté, Éric Mathieu and Shana Poplack. John Benjamins.

File-Muriel, Richard J. 2012. A laboratory approach to s-lenition in the Spanish of Barranquilla, Colombia. In Richard J. File-Muriel & Rafael Orozco (eds.), *Colombian Varieties of Spanish*, Madrid: Iberoamericana Vervuert. 127-140.

Foulkes, P., Scobbie, J.M. & Watt, D.J.L. 2012. Sociophonetics. In Hardcastle, W., Laver, J. & Gibbon, F. (eds.) *Handbook of Phonetic Sciences* (2nd ed.) Oxford: Blackwell. 703-754.

Robinson, Kimball L. 1979. On the voicing of intervocalic S in the Ecuadorian Highlands. *Romance Philology*, 33:1, 137-143.