Haiga/Haya Variation in Two Mexican Spanish Varieties

In Spanish there exists an alternation in the present subjunctive forms of the verb *haber*. The standard conjugation with a palatal fricative /**j**/ in stem-final position, as in *haya*, alternates in spoken Spanish with forms that have /**g**/ in the same position, such as *haiga*. Penny (2004) points out that the use of these forms is found at a rural level in many varieties of Spanish.

Most of the research on the use of the velar /**g**/ in stem-final position has adopted a historical perspective, examining the insertion of /g/ as an analogical extension (Penny 2002, Lloyd 1987, Elson 1988, Fondow 2010). This study is the first to quantitatively explore the linguistic and extra-linguistic factors that govern the variation between the two forms.

The data used for this study were extracted from three corpora of Mexican Spanish - one from Monterrey, Mexico (Rodriguez Alfano, 2006), and two from Mexico City (Butragueño & Lastra, 2011 and Lope Blanch, 1976). All standard and non-standard occurrences of present subjunctive *haber* were extracted, yielding a total of 423 tokens from 169 speakers. The tokens were coded for the following independent linguistic variables: type of use, number, person, type of clause, phonological form and presence or absence of negation. Auxiliary uses of *haber* were further coded for person, animacy of the subject, overtness of the subject and the type of participle on the verb. Presentational uses of *haber* were further coded for the plurality of the entity being introduced. All the tokens were also coded for four social factors: city of residence, age, gender and education level (taken as a measurement of social class).

A mixed effects model using speaker as a random effect was developed using the lmer function (Bates, Maechler & Bolker, 2011) in R (R Development Core Team, 2007). The role of the fixed and random effects was first tested on the complete data set, and then on auxiliaries and presentationals separately. Finally, the tokens from the two geographic varieties included in the study were also tested separately.

The results of the mixed-effects model on the complete data set show that speaker and education level affect the choice between *haya* or *haiga* such that less educated speakers use the use of *haiga* significantly more than more educated speakers (p<.01). Results for the Monterrey speakers show that in addition to the effects of speaker and education, *haiga* is used significantly more in presentationals than in auxiliaries (p=.01). While further research would be necessary to test this, we propose that in the Spanish of Monterrey the association of each variant to a different type of construction points to a desire on the part of the speaker to differentiate two semantic properties of *haber*. The difference between presentational and auxiliary *haber* and the productivity of this verb might explain why the variation between velar and standard forms still takes place in modern Spanish, while most of the verbs that underwent analogical extension of /g/ completely lost one of the variants.

References

Bates, Douglas, Maechler, Martin & Bolker, Ben. 2011. *lme4: Linear mixed-effects models using s4 classes* (R package version 0.999375-39)

Butragueño, Pedro Martín & Lastra, Yolanda. 2011. *Corpus sociolingüístico de la ciudad de México*. México: El Colegio de México, manuscrito.

Elson, Mark J. 1988. "The Synchronic Status and the Evolution of the *g* in Spanish *vengo, salgo,* etc. Revisited." *Hispania*, Vol. 71, No. 2, pp. 392-400.

Fondow, Steven Richard. 2010. Spanish Velar-insertion and Analogy: A Usage-based Diachronic Analysis. PhD dissertation. The Ohio State University.

Lloyd, Paul M. 1987. *From Latin to Spanish*. Philadelphia: American Philosophical Society. Lope Blanch, Juan M. 1976. *El habla popular de México: materiales para su estudio*. México: Universidad Nacional Autónoma de México (UNAM).

Penny, Ralph. 2002. *A history of the Spanish language.* Cambridge: Cambridge University Press.

Penny, Ralph. 2004. *Variation and change in Spanish.* Cambridge: Cambridge University Press.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna: Author.

Rodriguez Alfano, Lidia. 2006. *El habla de Monterrey, proyecto sociolinguïstico*. PRESEA.