

THE ROLE OF GRAMMATICAL RELATION PROBABILITIES IN THE SPREAD OF LANGUAGE CHANGE: AN EXAMPLE WITH SPANISH EXISTENTIAL *HABER* 'THERE (BE)'

Javier Rivas & Esther Brown
University of Colorado, Boulder

Spanish existential verb *haber* 'there (be)' typically occurs in constructions such as (1), in which its sole argument (*fiestas patronales*) is regarded as a direct object, since it allows for accusative clitic replacement, and it does not display verbal concord [verb (*hubo*, 3rd person singular), N(oun)P(hrase) (*fiestas patronales*, plural)]:

(1) Interview 1, 9

I: ... *Y este año no hubo fiestas patronales...*

T: ... And this year they didn't celebrate their patron saint's holiday...

However, in many varieties of Spanish, this construction co-exists with another one in which the verb *haber* agrees in person and number with its sole argument, as example (2) illustrates [verb (*hubieron*, 3rd person plural), NP (*fiestas*, plural)]. For these cases, it may be argued that the sole argument of *haber* is the subject of the construction:

(2) Interview 16, 80

M: *Hubieron fi-, hubieron fiestas en todos los pueblos menos en ése.*

'M: They celebrated their patron saint's holiday in all towns except in that one'

Previous studies (Montes de Oca 1994, Díaz-Campos 2003, Hernández Díaz 2006, D'Aquino Ruiz 2008) maintain the presentative *haber* construction is undergoing a change in progress whereby the construction in which *haber* agrees with its sole argument is likely to eventually replace the impersonal construction. *Haber* undergoes a process of analogical regularization which, by displaying agreement with its subject, levels *haber* with the other Spanish verbs. In this study we will examine the way in which this analogical extension spreads.

To achieve this goal, we extract all cases of existential *haber* that show variation (N=346) from a corpus of approximately 370,000 words of spoken Puerto Rican Spanish (Cortés-Torres 2005). In addition to the linguistic variables that have been considered in the previous literature (number of NP referent, tense, polarity, word order, presence/absence of a quantifier, definiteness of the NP, and human vs. non human referent of the NP), we also include a previously untested probabilistic measure, namely, subject grammatical relation probability for the NP; in other words, the likelihood that each NP will act as a subject of a sentence in overall usage. We submit our coded data to variable rule analysis using Varbrul (Rand & Sankoff 2001). Our results show that, in addition to a plural NP referent, regularization of *haber* directly reflects the degree to which the NP being presented is used with subject function in spoken Spanish. That is, NPs with a low rate of usage with subject function trigger *haber* regularization less than NPs with a high rate of usage with subject function. All the other variables are not significant.

Previous usage-based studies have shown that speakers' knowledge of probabilities of certain phones, words, and syntactic structures affects language production (Bush 2001, Jurafsky, Bell, Gregory & Raymond 2001, Gahl & Garnsey 2004, Bybee 2010). This current analysis extends these findings to examine grammatical relation probabilities for the NP. Results of this current study provide evidence for lexicalized NP grammatical relation-based probabilities. The implications of this work are therefore twofold. Specifically, this work contributes a new perspective and innovative methodology to the widely studied problem in Hispanic Linguistics of *haber* regularization in Spanish. Also, importantly, findings in this work contribute to the development of usage-based models of language variation and change by testing new types of probabilistic knowledge.

References

- Bush, Nathan. (2001). Frequency effects and word-boundary palatalization in English. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 255-280.
- Bybee, Joan. (2010). *Language, usage and cognition*. Cambridge: Cambridge Univ. Press.
- Cortés-Torres, Mayra E (2005). *La perífrasis estar + -ndo en el español puertorriqueño: variación dialectal o contacto lingüístico?* Doctoral dissertation, University of New Mexico.
- D'Aquino Ruiz, Giovanna. (2008). El cambio lingüístico de *haber* impersonal. *Núcleo* 25: 103-123.
- Díaz-Campos, Manuel. (2003). The pluralization of *haber* in Venezuelan Spanish: A sociolinguistic change in real time. *IU Working Papers in Linguistics* 03-05.
- Gahl, Susanne, & Garnsey, Susan M (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80: 748-75.
- Hernández Díaz, Axel. (2006). Posesión y existencia. La competencia de *haber*, *tener* en la posesión y *haber* existencial. In C. Company Company (Ed.). *Sintaxis histórica de la lengua española. Primera parte: la frase verbal*. Vol. 2. México: Univ. Nacional Autónoma de México: Fondo de Cultura Económica. 1053-1160.
- Jurafsky, Daniel, Bell, Alan, Gregory, Michelle, & William D Raymond. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 229-254.
- Montes de Oca-Sicilia, M. del Pilar. (1994). La concordancia con *haber* impersonal. *Anuario de letras* 32: 7-35.
- Rand, David & Sankoff, David. (2001). *GoldVarb: a variable rule application for Macintosh*.