

## Operationalizing reference in Spanish subject pronoun variation

Gabriela G. Alfaraz (*Michigan State University*)

After variationist studies discussed coreference as an internal variable constraining the variation of Spanish subject pronouns (Cameron, 1992; Silva-Corvalán, 1982), it became a core variable that is consistently found to have an effect on the variation. Switch reference, as it is known, is a binary factor that reflects the local relationship between a referent and an antecedent, expressed as same or switch. The switch reference variable is currently a staple in the growing body of variationist work on Spanish subject pronoun variation (Carvalho, Orozco, & Shin, forthcoming; Otheguy & Zentella, 2012). We have not, however, studied it closely in quantitative work, even though doing so can improve our understanding of the strength of internal constraints, particularly if we study claims that switch reference is weakening in some monolingual varieties (Bullock & Toribio, 2009) and that it can be lost in language contact situations (Flores-Ferrán, 2007).

Bayley and Pease-Álvarez (1997) proposed an interpretation of reference as discourse connectedness, which takes into account the larger discourse (Paredes Silva, 1993). Their comparison of switch reference and discourse connectedness showed that the latter was more effective in explaining the variation. Although it has been used in a few studies, discourse connectedness tends to be misunderstood and reduced to the binary switch-same reference (see Otheguy & Zentella, 2012). More recently, Travis and Torres-Cacoullous (2012), noted that variable use of the first person singular *yo* was influenced more by Intervening Human Subjects than by switch reference.

This paper presents an approach to operationalizing reference that aligns variationist interpretations with studies on reference in discourse. Abbott (2010: 251) noted that approaches used to explain the selection of NPs in discourse include Prince (1981, 1992), Ariel (1988, 1990), and Gundel, Hedberg, and Zacharski (1993, 2001). This research adopted the model in Gundel et al. (1993, 2003), which takes into account a referent's cognitive status, based on the addressee. Cognitive status is defined in terms of Givenness and represented on a scale in which each status entails all others below it. The statuses on the Givenness Hierarchy are: In focus > Activated > Familiar > Uniquely identifiable > Referential > Type identifiable.

To examine the viability of using this referential hierarchy, Rbrul was used. The data studied was from sociolinguistic interviews conducted in Santo Domingo, Dominican Republic. The sample included 3002 tokens from 20 individuals, equal numbers of men and women between the ages 18 and 60+ from three SES groups. The internal predictors included switch reference, givenness, person/number, verb type, priming, and polarity. The social predictors studied were age, gender, and socioeconomic status. Individual was included as a random effect.

Results showed that the significant predictors were person/number ( $p < .000$ ), givenness ( $p < .000$ ), priming ( $p < .000$ ), polarity ( $p < .000$ ), and verb type ( $p < .000$ ). The speaker random effect ( $SD = .25$ ) was also significant. Switch reference was not significant ( $p = .071$ ). For the givenness variable, it was found that in focus referents disfavored overt pronouns and that activated referents and the statuses above it on the scale favored them. It is argued that in conjunction with switch reference, other approaches are needed to understand the role of reference in subject pronoun variation.