

## Cross-Linguistic Homogeneity and Heterogeneity in Pronominal Expression

Gregory R. Guy, Rafael Orozco & Danny Erker

Variationist sociolinguistics has long recognized that both features of linguistic contexts and speech communities' social characteristics condition variant choice. Recent work (Labov 2010; Tamminga et al. 2016) adds *cognitive constraints* as a third category in this typology. Much work in the field (Carvalho et al. 2015) assumes —implicitly or explicitly— that certain constraints have consistent effects across speech communities. The present paper probes that assumption using a comparative approach to constraints on the alternation between null and overt pronominal subjects in Spanish and Portuguese. We analyzed 44,605 tokens from sociolinguistic interviews with 208 speakers in 7 locales: Barranquilla, Boston, Lisbon, Medellín, New York, São Paulo, and Xalapa. We hypothesize that *cognitive* constraints, based on universal properties of the human mind, should have the most consistent effects across languages, while *linguistic* constraints will be affected by each language's structural idiosyncrasies, and *social* constraints will reflect particular speech communities' properties, practices and ideologies.

Our findings support this hypothesis. *Priming* —reflecting that property of cognition/neurobiology underlying human sensitivity to repetition— has a consistent cross-linguistic and cross-community effect: the realization as null or overt in one clause favors the same realization in successive clauses. *Reference chains* are similarly systematic: new referents favor overt pronominal subjects in all our datasets, a consequence, we argue, of the common cognitive task of tracking referents across discourse. Linguistic constraints on pronominal expression are more diverse: Spanish-speaking communities have consistent effects of *reflexivity* (reflexive verbs have lower pronoun rates), *person/number* (plural subjects have lower pronoun rates), and *tense/mood* (imperfectives favor overt pronouns while preterites disfavor them). These constraints diverge in the substantially different Portuguese pronominal system, where a new pronoun *a gente*, is replacing first-person plural *nós*, reflexives are rare, and 2<sup>nd</sup> sg *tu* is non-existent in São Paulo. Portugal and Brazil also differ in the treatment of second-person singular polite *você*: it favors overt forms in Brazil but disfavors them in Portugal. Notwithstanding, a functional analysis —a presumably cognitive requirement that more overt pronouns are needed when verbal inflections are less distinctive — is partially supported by the results from all speech communities.

Social distribution also varies considerably across speech communities. One prominent effect is *dialect*: Caribbean speakers use the most overt pronouns among the Spanish dialects, and Brazil has over twice the pronoun rate of Portugal (68% vs 32%). *Gender* differences although prominent in Lisbon, São Paulo, Barranquilla and New York —with women favoring overt subjects— are absent in the other communities. *Age grading* is found in Portuguese, and the Spanish monolingual corpora, where younger speakers favor lower pronoun rates. Our findings suggest that similar usage patterns will arise across languages and communities for a given linguistic variable insofar as they are linked to general aspects of human cognition. Crosslinguistic and cross-community heterogeneity, in contrast, is expected to emerge in relation to specific structural properties of individual language varieties. Similarly, the social dimensions routinely implicated in synchronic linguistic variation and in historical change are likely to be vectors of cross-community heterogeneity, reflecting local interactional roles, norms, and dynamics. (500 words)

### References

- Carvalho, Ana M., Rafael Orozco, & Naomi Lapidus Shin. 2015. *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Washington: Georgetown University Press.
- Labov, William. 2010. *Principles of linguistic change, volume 3: Cognitive and cultural factors* (Vol. 36). John Wiley & Sons.
- Tamminga, M., MacKenzie, Laurel, & Embick, David. 2016. The dynamics of variation in individuals. *Linguistic variation*, 16(2), 300–336. <https://doi.org/10.1075/lv.16.2.06tam>