The role of contextual frequency in morphosyntactic variation: The case of Spanish subject pronouns

Danielle Dionne (ddionne@bu.edu)

Boston University

Abstract

This study incorporates insights from Usage Based Grammar (UBG) into research on morphosyntactic variation. Specifically, this study investigates lexical frequency as well as a series of enhanced frequency metrics in a Variationist study of Spanish Subject Pronoun Production (SPPs). Results support Erker & Guy (2012), indicating that lexical frequency impacts pronoun production indirectly by activating or amplifying the effects of other linguistic constraints. However, results suggest that how often the verb appears in various Switch Reference/Priming contexts impacts the overall pronoun production of the verb more than simple verb frequency. This early finding aligns with the UBG prediction that the most-frequent *context* in which a verb appears will dominate the overall pronominal tendencies of the verb.

1. Background & RQs

Inspired by insights from Usage-Based Grammar (UBG) frameworks, recent variationist studies on Spanish subject pronoun production have investigated lexical frequency [11, 5, 2, 1, 10]. The prediction is that verb frequency will impact the mental representation of each verb and therefore affect pronominal tendencies.

However, verb frequency alone largely undertheorizes UBG [7, 3, 8, 4], which postulates mental grammars as webs rich memory. Moreover, results from these studies have been contradictory, leaving the role of frequency still in question.

Research Questions:

Q1. Are the frequency effects in Erker & Guy (2012) replicable?

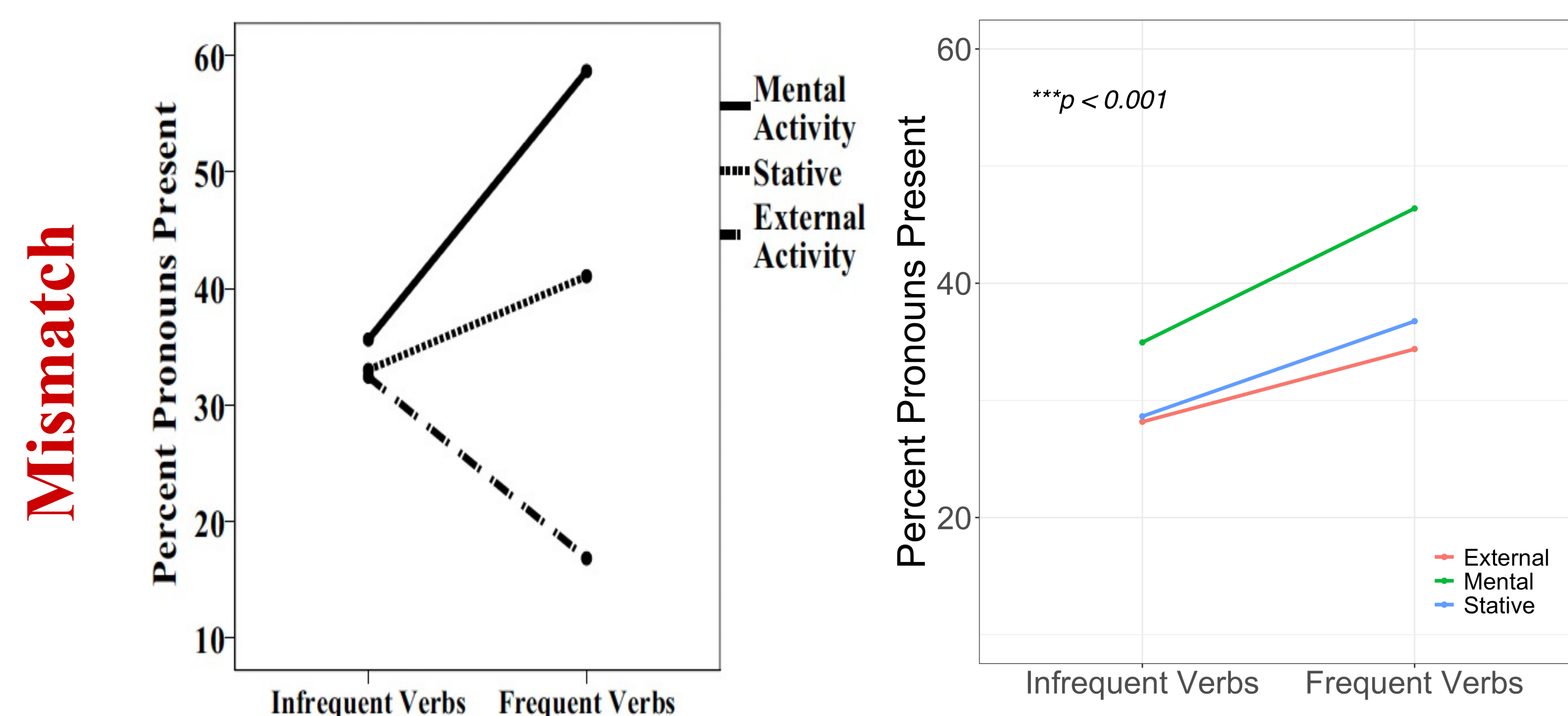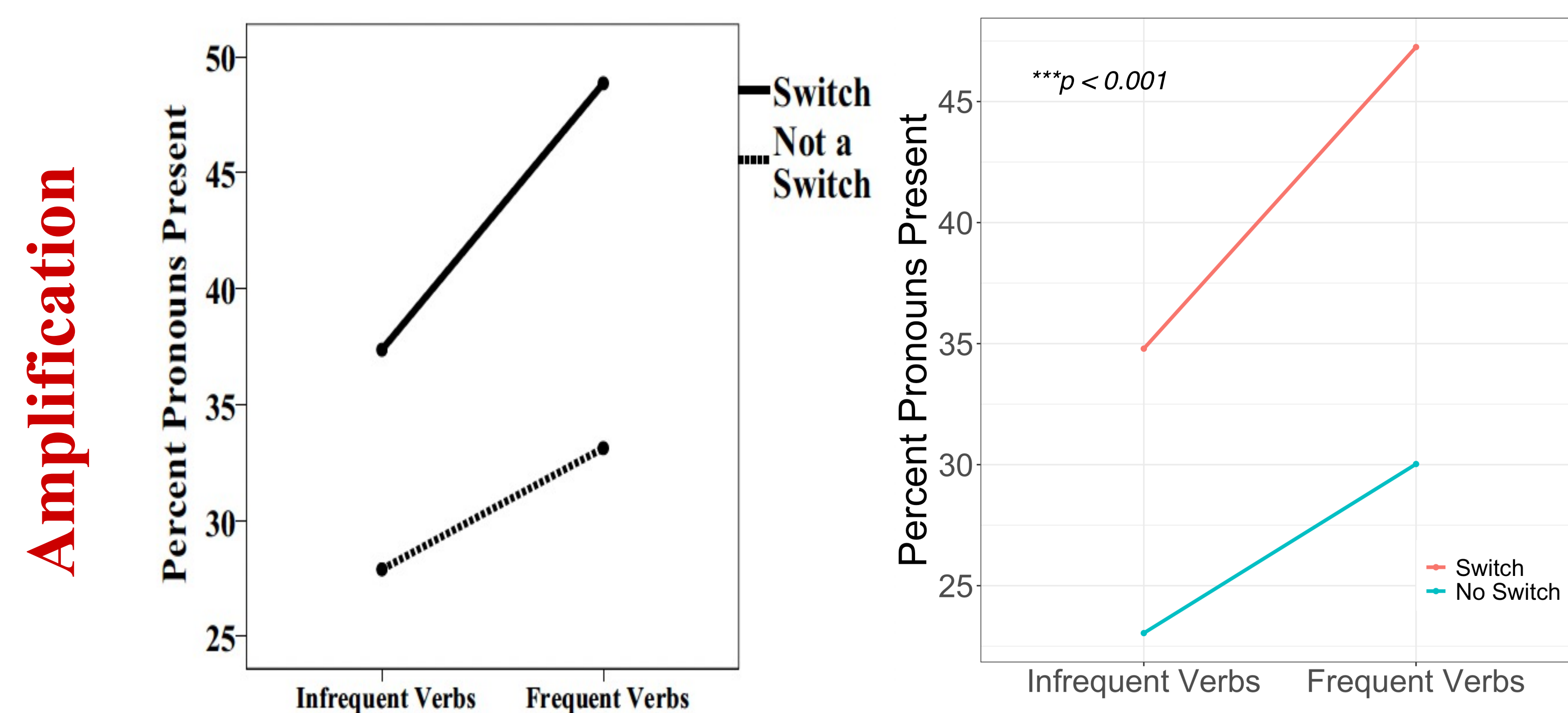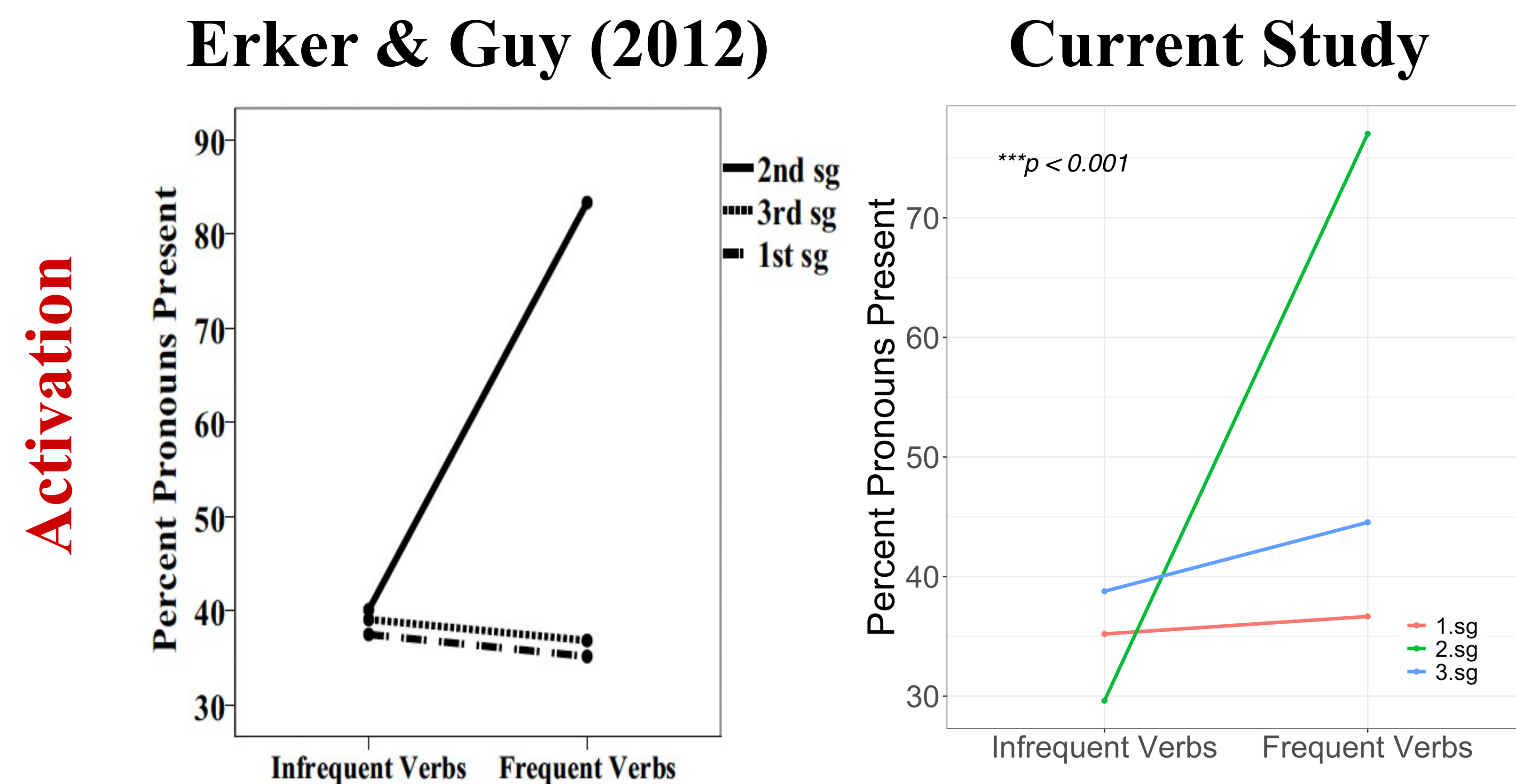Q2. Are context-specific frequency metrics more explanatory?

2. Methodology

Structure: (1) Investigate lexical frequency à la [5] in a different, larger dataset; (2) investigate six **contextual frequency metrics**:
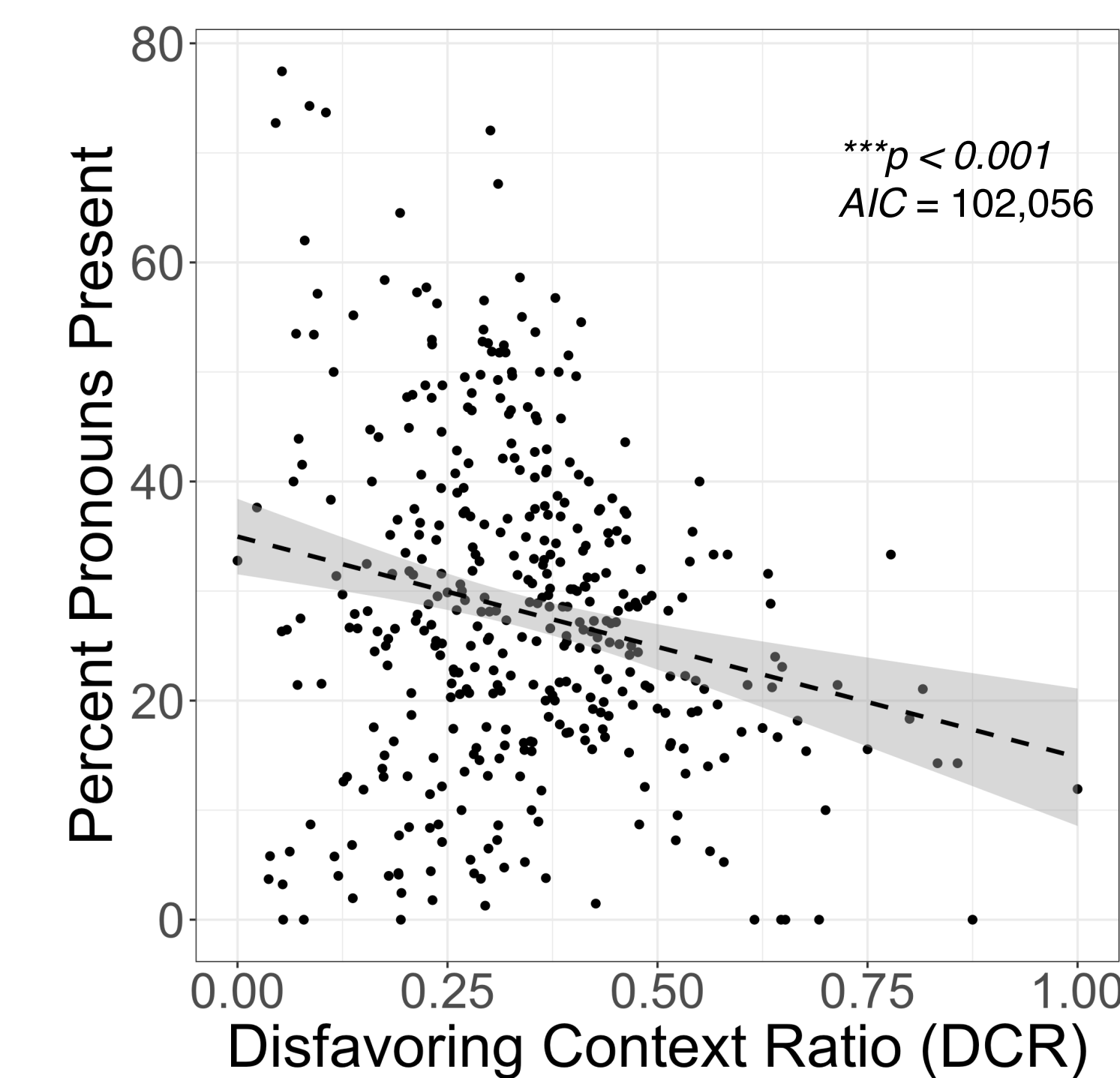
- The verb frequency in four combined contexts relative to SWITCH REFERENCE and PRIMING (switch/priming, switch/no priming, no switch/priming, no switch/no priming)
- **Context Ratios**: (1) ratio of the verb appearing in a pronoun-favoring context (with a switch *and* with priming) relative to its total frequency; (2) ratio of a verb appearing in a pronoun-disfavoring context (without a switch *and* without priming) relative to its total frequency

Data: 88,001 tokens from 141 Participants in Spanish in Boston Corpus [9] and Spanish in New York Corpus [6]

3. Lexical Frequency

Erker & Guy (2012)

Current Study

Activation

Amplification

Mismatch

4. Enhanced Frequency

5. Summary

Lexical Frequency:

- *Evidence of lexical frequency effects*: Current results fall mostly in-line with [5] such that **discrete frequency** interacts with other predictors to activate an effect or amplify an effect (see 3).

Enhanced Frequency:

- *Evidence of enhanced frequency effects*: All models with **enhanced frequency metrics** outperformed models with **lexical frequency metrics**.
- **Disfavoring Context Ratio** was the strongest predictor of pronoun use relative to all other frequency metrics.

These preliminary findings provide support for the UBG notion that speakers are aware of and retain information on the conditioning contexts of lexical items. Future research should explore enhanced frequency metrics across a wider range of speech communities and variables.

References

[1] Bayley, R., Greer, K. A., Holland, C. L. (2017). Lexical frequency and morphosyntactic variation: Evidence from U.S. Spanish. *Spanish in Context*, 14(3).

[2] Benevento, N. M. Dietrich, A. J. (2015). I think, therefore digo yo: Variable position of the 1sg subject pronoun in New Mexican Spanish-English code-switching. *International Journal of Bilingualism*, 19 (4).

[3] Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language*, 82.

[4] Bybee, J. (2010). *Language, usage, and cognition*. Cambridge University Press.

[5] Erker, D. Guy, G. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*.

[6] Erker, D. (2022). How social salience can illuminate the outcomes of linguistic contact: Data from Spanish in Boston. In Guy, G.R. and Beaman, K., eds., *The coherence of linguistic communities: Orderly Heterogeneity and Social Meaning*, Routledge Studies in Sociolinguistics.

[7] Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

[8] Goldberg, A. (2006). *Constructions at Work*. Oxford University Press.

[9] Otheguy, R., Zentella, A. C., Livert, D. (2007). Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language*.

[10] Torres Cacoullos, R. Travis, C. (2018). *Bilingualism in the Community: Code-switching and grammars in contact*. Cambridge University Press.

[11] Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language variation and change*, 19(2).