

Grammatical gender marking in a corpus of L2 Spanish: A longitudinal analysis of feminine modifiers

Aarnes Gudmestad,^a Amanda Edmonds,^a and Thomas Metzger^a

^aVirginia Polytechnic Institute and State University and ^bUniversité Paul-Valéry Montpellier 3

In the current study, we apply variationist methods to a corpus-based investigation of grammatical gender marking in second-language (L2) Spanish. In Spanish all nouns have masculine or feminine gender, and modifiers (adjectives and determiners) must agree in gender with the noun they modify. A rich literature exists on grammatical gender marking in Spanish among adult L2 learners and bilinguals (cf. Alarcón, 2014). Within this body of work, the object of study has been accuracy, meaning that researchers seek to account for the errors that learners make with gender marking. An error refers to a mismatch between the gender of a noun and its modifier (e.g., *una* [feminine] *problema* [masculine] ‘a problem’). One consistent finding across studies is that learners tend to overgeneralize masculine forms of modifiers, which leads to higher rates of accuracy with masculine nouns and suggests that the use of feminine modifiers develops more slowly in learner language than masculine modifiers (e.g., Finnemann, 1992; Montrul, Foote, & Perpiñán, 2008; Schlig, 2003; White, Valenzuela, Kozłowska-Macgregor & Leung, 2004). Although error analyses of learner language are common in the field of second language acquisition (SLA; cf. Ellis, 2008), one subfield – variationist SLA (cf. Geeslin with Long, 2014) – has advocated for examinations of the forms that learners use and the linguistic and extra-linguistic factors that condition this use, rather than whether learner production is accurate or not. In this vein, we aim to build on previous research on grammatical gender marking in L2 Spanish by shifting the focus from the errors that learners make to an explanation of the forms that learners use and how they vary. Given the evidence that masculine and feminine modifiers develop at different rates in learner language, our analysis centers on the gender of the modifier that learners use.

We examined data from LANGSNAP (<http://langsnap.soton.ac.uk>), a longitudinal learner corpus, with the goal of identifying the factors that impact use and that characterize development of feminine modifiers over the course of a 21-month period, which included an academic year abroad in a target-language environment (Spain or Mexico). Our dataset consists of 21 adult English-speaking, university-level learners of Spanish. We coded data that were collected at three time periods: before going abroad (pre-stay), one year later (in-stay: after nine months in the target-language context), and eight months after returning to the United Kingdom (post-stay). At each data-collection period the participants completed a written, argumentative essay and two oral tasks – one interview and one narration. We coded every possibility for gender marking when a noun was modified by a determiner or adjective ($K = 16,357$). The current dataset includes modifiers that show overt gender marking and whose noun is in the immediate clause or one of the preceding 10 clauses ($k = 11,833$). The dependent variable was modifier gender (masculine or feminine). Each token was coded for 11 linguistic and extra-linguistic factors that were independent, fixed effects in the analysis. These factors were identified in previous research as potentially influencing the development of grammatical gender marking (noun gender, noun ending, noun class, noun number, modifier type, log-frequency(language), frequency(individual), syllable distance between the noun and modifier, task, data-collection time, and initial proficiency score). We also examined possible interactions between time and other fixed effects

and we included participant as a random effect in analysis. Before fitting a regression model, we used chi-square methods to determine whether there was a strong correlation between fixed effects and we assessed the potential importance of the remaining fixed effects using bootstrapping. This latter step enabled us to further reduce the set of fixed effects we examined and helped us to avoid overfitting the model. We then fit a generalized linear mixed model using SAS version 9.4.

The results revealed that several factors impacted learners' variable use of feminine modifiers: noun gender, modifier type, noun ending, syllable distance, initial proficiency, and task. There were also significant interactions between time and noun ending and time and noun gender, which suggest that noun ending and noun gender help to explain changes in the variable use of feminine modifiers over time in L2 Spanish. We conclude with reflections on what a variationist analysis contributes to the understanding of the acquisition of grammatical gender marking.

References

- Alarcón, I. (2014). Grammatical gender in second language Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 202-218). Malden, MA: Wiley-Blackwell.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed). Oxford: Oxford University Press.
- Finnemann, M. D. (1992). Learning agreement in the noun phrase: The strategies of three first-year Spanish students. *International review of Applied Linguistics in Teaching*, 30(2), 121-136.
- Geeslin, K. L., with Long, A. L. (2014). *Sociolinguistics and second language acquisition*. New York: Routledge.
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 503-553.
- Schlig, C. (2003). Analysis of agreement errors made by third-year students. *Hispania*, 86(2), 312-319.
- White, L., Valenzuela, E., Kozłowska-Macgregor, M., & Leung, Y-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25, 105-133.