ANIMACY AND THE EVOLUTION OF GENDER AGREEMENT SYSTEMS: A STUDY OF NORTH-WESTERN BANTU

Francesca Di Garbo*[1] and Annemarie Verkerk [2]

*FRANCESCA DI GARBO: FRANCESCA.DIGARBO@HELSINKI.FI

[1]Department of Languages, University of Helsinki, Helsinki, Finland
[2]Department of Language Science and Technology, Saarland University, Saarbrücken, Germany

Recent experimental studies within the iterated learning paradigm (Vihman et al., 2018) show that animacy distinctions in noun class systems emerge in response to a learning bias, and, in fact, increase the learnability of such systems. This experimental evidence is paired by a wealth of empirical data from historical-comparative linguistics (Igartua, 2018; Seifart, 2018). In this paper, we investigate the processes of language evolution through which animacy-based semantic restructuring unfolds in grammatical gender systems (i.e., agreement-based noun class systems) and the non-linguistic factors that may favor this type of restructuring. Our focus is on the evolution of patterns of gender agreement, and we present the results of two studies based on a sample of 255 North-Western Bantu (henceforth NWB) languages.

Study 1 investigates the distribution of patterns of animacy-based agreement in the languages of the sample, and discusses how animacy-based agreement is related to the emergence of highly eroded systems of gender marking. We present a bottom-up typology of gender systems in NWB, which accounts for the distribution of animacy-based gender marking as opposed to lexically-specified gender (see Figure 1). We argue that the different types of gender systems we distinguish form a continuum: highly eroded gender systems are explained as a result of the evolutionary dynamics by which animacy-based semantic agreement rises and spreads in more conservative languages.

In Study 2, we investigate whether animacy-based agreement in NWB can be explained as a function of population history. In particular, we test whether highly eroded and/or semantically more transparent gender systems are more common (1) in NWB languages of wider communication, under the assumption that community size is a proxy for degree of language contact, and that animacy-based agreement is favored in virtue of its higher learnability, and/or (2) in areas of NWB (i.e. the Central African rainforest) where language shift from indigenous non-Bantu languages is suspected to have occurred, under the assumption that animacy-based restructuring is a substrate effect.

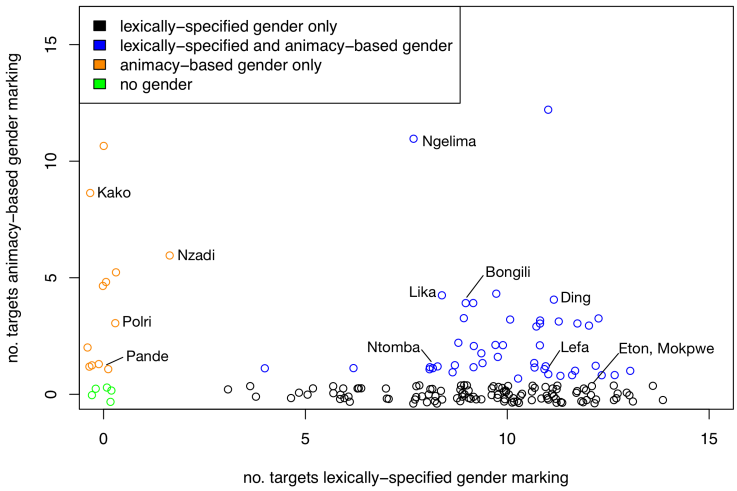

Figure 1. The number of agreement targets that are marked for lexically-specified gender (x-axis) and for animacy-based gender (y axis). Points have been jittered so they do not overlap. Colors reflect four-way typology introduced in Study 1. Exemplar languages are labelled.

We build generalized linear mixed effects models (GLMMs) using the package `brms` (Bürkner, 2017) in R (R Core Team, 2019). We model both a binary typology (contrasting languages with only lexically-specified gender vs. the other three types distinguished in Figure 1) as well as the counts of agreement targets displayed in Figure 1. We find that the number of L1 speakers as well as two geographical measures – (1) whether a language has an ancestor in the Central African rainforest (yes or no), and (2) longitude – are relevant predictors of the type of gender system (see Table 1). These analyses match evidence from a qualitative investigation of the sociolinguistics of 17 NWB languages with highly eroded gender systems, whose social history is clearly shaped by contact with non-Bantu neighbors or by pidginization/creolization. To our knowledge, this is the first quantitative cross-linguistic study that confirms the oft-repeated claim (Trudgill, 1999) that situations of intense language contact or language shift favor the restructuring and erosion of grammatical gender, and that animacy distinctions are a likely blueprint for restructuring.

Table 1. 95% Confidence Intervals of GLMMs

	no. L1 speakers	*ancestor rainforest*	*longitude*
binary typology	0.05-1.81	0.24-4.63	-0.27-1.00
target counts	-0.15- -0.01	-0.12-2.20	0.19-0.74

References

Bürkner, P-C. (2017). "`brms`: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software,* 80 (1), 1–28. https://doi.org/10.18637/jss.v080.i01.

Igartua, I. & Santazilia, E. (2018). How animacy and natural gender constrain morphological complexity: Evidence from diachrony. *Open Linguistics,* 4(1), 438–452.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Seifart, F. (2018). The semantic reduction of the noun universe and the diachrony of nominal classification. In W. McGregor & S. Wichmann (Eds.), *The diachrony of classification systems* (pp 9–32). Amsterdam: John Benjamins.

Trudgill, P. (1999). Language contact and the function of linguistic gender. Pozna´n *Studies in Contemporary Linguistics*, 35, 133–152.

Vihman, V-A, Nelson D. & Kirby, S. (2018). Animacy distinctions arise from iterated learning. *Open Linguistics,* 4(1), 552–565.