THE ROLE OF STRUCTURAL PRIMING, SEMANTICS AND POPULATION STRUCTURE IN WORD ORDER CONVENTIONALIZATION: A COMPUTATIONAL MODEL

Loïs Dona[*1] and Marieke Schouwstra[2]

[*]Corresponding Author: l.m.dona@students.uu.nl
[1]Information and computing sciences, Utrecht University, Utrecht, The Netherlands
[2]Humanities, University of Amsterdam, Amsterdam, The Netherlands

1. Population structure and word order regularization

There is an ongoing debate about the influence of population structure on the emergence of linguistic rules. Some stress the importance of *population size*, or factors often confounded with it (Raviv, Meyer, & Lev-Ari, 2019; Thompson, Raviv, & Kirby, 2020; Lupyan & Dale, 2010; Wray & Grace, 2007; Lou-Magnuson & Onnis, 2018), while others argue for the importance of *network connectivity* (the degree of diversity in agent encounters) (Segovia-Martín, Walker, Fay, & Tamariz, 2020; Richie, Yang, & Coppola, 2014). In a series of simulations, we investigate the influence of population structure on the conventionalization of basic word order. Of the languages in the world, many use a dominant order of subject, object and verb for conveying who did what to whom (Dryer, 2013). Previous research has suggested that in the early stages of language emergence, people use word orders variably, depending on event semantics: SOV order for extensional events (in which the direct object is specific and concrete; e.g. pirate throws ball), and SVO for intensional events (in which the direct object is more abstract, and possibly dependent on the verb; e.g. pirate thinks of ball) (Schouwstra & Swart, 2014). Structural priming is a potential mechanism driving increased regularity in word order (Christensen, Fusaroli, & Tylén, 2016; Schouwstra, Smith, & Kirby, 2020). A key question is how properties of the population influence regularization.

2. Local regularization: interaction in pairs of agents

Our first model investigates the relative contributions of semantics and structural priming over time. We compare three different simulations of dyadic interaction, one where the influence of structural priming increases over time and that of semantics decreases (*increasing-priming*), one where it is the other way around (*decreasing-priming*), and one where their influences are equal and constant (*constant-priming*). The increasing and decreasing influence of priming are modelled by an exponential function with time as the exponent. The influence of semantics is conversely modelled as $1 - priming$. The influence shifts were accomplished by treating the probability of choosing SOV or SVO word order as a linear combination of previously observed and produced word orders that were SOV or SVO (structural priming), modelled as 1 and 0 respectively, and event semantics. Regularization was measured by calculating the change in mutual information over time. A comparison of the simulations shows that only when the influence of priming increases relative to the influence of event semantics, word order becomes more regular. By contrast, when structural priming is constant over time, this does not lead to increased regularity.

3. Simulating population properties

To assess the effects of population properties on word order regularization, we simulated populations with different sizes and densities, copying the agent properties of the increasing priming simulation above. We measured network connectivity of these populations using the distribution of the number of observations made by every agent from every other agent after every thirty two interactions. We used the slope parameter of a curve fitted to mutual information data over time as a measure for regularization speed.

The results show a positive correlation between population density and connectivity, and a negative correlation between population size and connectivity (using Spearman's rank correlation Myers & Sirois, 2004). Both size and density correlate significantly with regularization speed (see figure 1), but density correlates significantly stronger (Fisher Z-transform.; $z = -16.74, p < 0.001$).

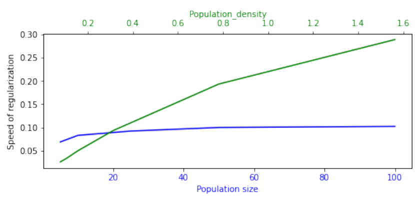

Figure 1. Population size and density plotted against regularization speed (k).

The relatively modest effect of population size on regularization speed indicates that population size alone is not an explanatory factor for the differences in convergence rates and linguistic simplicity between communities, but often correlates with these effects due to it being confounded with network connectivity.

These findings contribute to the unraveling of the forces at play during word order conventionalization on the cognitive and population level. Moreover, the computational structure of the model in this paper allows for easy incorporation of other potential influences (e.g. effects of language contact). Recent work has shown that viewing word order as continuous rather than discrete might be more realistic (Levshina et al., 2021), and that word order regularization is also reflected in learning (Motamedi, Wolters, Naegeli, Schouwstra, & Kirby, 2021). The current framework makes it possible to explore these new insights further.

References

Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, *146*, 67–80.

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Levshina, N., Namboodiripad, S., Allassonnière-Tang, M., Kramer, M. A., Talamo, L., Verkerk, A., Wilmoth, S., Rodriguez, G. G., Gupton, T., Kidd, E., et al.. (2021). Why we need a gradient approach to word order.

Lou-Magnuson, M., & Onnis, L. (2018). Social network limits language complexity. *Cognitive Science*, *42*(8), 2790–2817.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*(1), e8559.

Motamedi, Y., Wolters, L., Naegeli, D., Schouwstra, M., & Kirby, S. (2021). Regularisation, systematicity and naturalness in a silent gesture learning task. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43).

Myers, L., & Sirois, M. J. (2004). Spearman correlation coefficients, differences between. *Encyclopedia of statistical sciences*, *12*.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Larger communities create more systematic languages. *Proceedings of the Royal Society B*, *286*(1907), 20191262.

Richie, R., Yang, C., & Coppola, M. (2014). Modeling the emergence of lexicons in homesign systems. *Topics in cognitive science*, *6*(1), 183–195.

Schouwstra, M., Smith, K., & Kirby, S. (2020). The emergence of word order conventions: improvisation, interaction and transmission.

Schouwstra, M., & Swart, H. de. (2014). The semantic origins of word order. *Cognition*, *131*(3), 431–436.

Segovia-Martín, J., Walker, B., Fay, N., & Tamariz, M. (2020). Network connectivity dynamics, cognitive biases, and the evolution of cultural diversity in round-robin interactive micro-societies. *Cognitive science*, *44*(7), e12852.

Thompson, B., Raviv, L., & Kirby, S. (2020). Complexity can be maintained in small populations: a model of lexical variability in emerging sign languages.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.