LANGUAGE COMPLEXITY AS AN INTERACTION BETWEEN SOCIAL STRUCTURE, INNOVATION, AND SIMPLICITY

MATTHEW SPIKE

matthew.spike@anu.edu.au
Centre of Excellence for the Dynamics of Language, Australian National University

There are several claims regarding how social factors determine linguistic complexity: the world's most complex languages seem to be spoken by small, isolated populations, with dense social networks and a high proportion of child learners (see Trudgill, 2011), and simpler languages are spoken by larger populations (e.g. Lupyan & Dale, 2010), in language contact situations (e.g. Trudgill, 2011) and in particular emerge in the formation of creoles (e.g. McWhorter, 2001). Whatever the empirical status of these claims, the causal mechanisms offered up by these accounts are often conflicting. For example, is simplification driven mainly by speakers (e.g. Wray & Grace, 2007) or L2 learners (e.g. Trudgill, 2011)? Is complexification the normal directional trend in languages which are 'left alone' (Trudgill, 2011, p.235), or are complex features *selected for* by child learners (Lupyan & Dale, 2010)? Modelling work by Reali et al. (2014) suggests that simpler (i.e. more easily learned) conventions will predominate in larger populations, but how does this interact with differently structured populations?

I present results from a simple computational model which investigates the effect of all these factors, and argue that all these mechanisms can be understood in terms of their contribution towards i) maintaining complexity, and ii) the fixation of novel complexity in the language of a sub-population. This work supports some conclusions found in the established literature — for example, both speakers and L2 learners can drive simplification — but also provides a number of novel insights: for example the importance of *innovation* and its impact on different populations.

Individual agents are modelled as collections of tokens — a modified Moran model (Moran, 1962) or Pólya-Hoppe urn (Hoppe, 1984), where each token represents a specific *convention*, after Reali et al. (2014). However, where Reali et al. assume that complexity is an *inherent* property, capturing a difference between *easy* and *hard* conventions, the literature usually characterises complexity as *systemic* property, i.e. an increase in irregularity, redundancy, and the number of categories, along with a decrease in obvious systematicity (Trudgill, 2011). As such, complexity in this model is expressed quantitatively, i.e. the number of conven-

tions which are shared across whole and sub-populations. This corresponds with the idea that the more complex the language, the more information is required to acquire it.

The parameters of the model allow us to compare the effects of varying i) the overall population size; ii) the population structure (complete, random, small-world, and connected-caveman graphs); iii) an *innovation* bias, (the probability of an agent originating a novel convention); iv) a *simplicity* bias which abstracts over production, learning, and memory (a preference for using majority over minority variants); and v) the L1/L2 ratio, (modelled in two ways: a) L1 agents having more linguistic exposures, and b) different degrees of heterogeneity in the initial states of the population).

There are several main results: first, the overall population size is less important than the structure of that population, which determines the *effective* population size for any given agent. Simply put, if the agent is connected to many others via relatively few steps (i.e. a low average shortest path length), then its effective population size is large, and vice versa. Second, whereas Trudgill predicts that smaller populations should maintain high levels of complexity, results show drift effects and a subsequent loss of complexity taking over unless there is a strong *anti-simplicity* bias. Larger effective populations are *more* able to maintain complexity unless speakers or learners have a strong preference for simplicity. Third, the innovation of novel complexity plays an important dual role: in small effective populations, novel complexity fixates easily, in line with the strength of the bias. In larger populations, on the other hand, fixation of novel complexity becomes impossible and, surprisingly, high levels of complexity innovation have the paradoxical effect of leading to a *collapse* in complexity unless the simplicity bias is strong, as this overwhelms the population-wide system with noise. Finally, high-contact situations such as found in the emergence of creole languages only lead to simplification in large effective populations with a strong bias towards simplicity: in other conditions, such as when populations are small or the simplicity bias is weak, then we can see an *increase* in complexity.

In general, the simplicity bias acts against both the maintenance of existing complexity and the fixation of novel complexity, while the innovation bias promotes that fixation but works against maintenance. These both interact with the effective population size: larger populations promote the maintenance of existing complexity, but prevent the fixation of novel complexity. These results suggest that we should look to populations which speak highly complex languages for evidence of individual biases for maintaining complex forms and innovating new forms. Finally, we should consider the possibility that these biases are themselves culturally inherited, and so vary between social and linguistic groups.

References

Hoppe, F. M. (1984). Pólya-like urns and the Ewens' sampling formula. *Journal*

of Mathematical Biology(20), 91–94.

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

McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, *5*, 125–166.

Moran, P. A. P. (1962). *The statistical processes of evolutionary theory*. Clarendon Press.

Reali, F., Chater, N., & Christiansen, M. H. (2014). The paradox of linguistic complexity and community size. In S. . Roberts, H. . Lyn, & H. . Cornish (Eds.), *Evolang 10* (pp. 270–277). Vienna: World Scientific.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

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