RECONSIDERING THE CODE MODEL OF COMMUNICATION FOR SIMULATIONS OF LANGUAGE EVOLUTION

MATTHIJS WESTERA

Graduate School of Natural Sciences, Utrecht University,
Leuvenlaan 4, 3584 CE Utrecht, The Netherlands
Matthijs.Westera@phil.uu.nl

In all human languages, the link between forms and meanings is highly *systematic*. Particular forms correspond to particular meanings, and particular compositions of forms correspond to particular compositions of meanings (*compositionality*). The conditions under which compositional languages may evolve have been extensively studied, often employing computational models (Briscoe, 2002). Such conditions are for instance the presence of a *learning bottleneck* (Kirby, 2002) and the presence of certain innate or acquired *learning biases* (Smith, 2003).

Smith (2003) developed a model to investigate which learning biases are required for a compositional language to evolve in a population of agents. In the model, each agent is modelled as an association network linking signals to meanings. The algorithms for signal production and signal interpretation strongly resemble encoding and decoding. For instance, interpreting a signal amounts to retrieving the meaning to which the signal is most strongly associated.

Smith (2003) concluded that two learning biases are necessary for the emergence of a highly compositional language. First, the agents in the population need a bias in favour of one-to-one mappings between signals and meanings. Second, the agents need a bias in favour of decomposing signals and meanings into smaller parts. In a population of agents lacking one or both of these biases, a compositional language cannot be maintained through a learning bottleneck.

This conclusion holds, at least, when the agents involved do not possess any inferential capabilities. The model in (Smith, 2003) is based on the *code model of communication*, which assumes that signal production and interpretation can be fully described as a matter of encoding and decoding (Shannon & Weaver, 1949). The central position of the code model in many computational models of language evolution has been food for discussion during previous editions of the Evolang conference. What will happen to the necessity of certain learning biases, as proposed by Smith (2003), if the assumptions underlying the code model of communication are dropped?

I will present a re-implementation of the model by Smith (2003) that addresses this question in two ways. First, following Langacker (1987) the model attempts to embed linguistic knowledge into the more general framework of conceptual knowledge. Utterance production and interpretation are regarded as the same cognitive process, rather than the antagonistic encoding and decoding. Additionally, following Hoefler (2009) the (cognitive) distinction between signals and meanings is entirely dropped.

Second, the model incorporates the *inferential account of communication* as formulated within Relevance Theory (Sperber & Wilson, 1995). The inferential account of communication describes utterance interpretation as an inferential process, rather than a mere decoding. Relevance Theory grounds inference in the fundamental principle that all cognitive processes tend to the maximisation of relevance, i.e. obtaining the highest effect, which is context-dependent, with the least effort. Context, finally, is modelled using the notions of *ignorable* and *inferable* information as introduced by Hoefler (2009).

Abandoning the code model as such has resulted in a synthesis of the model developed by Smith (2003) to simulate learning biases and the model developed by Hoefler (2009) to simulate the role of context. The results obtained with this synthesis confirm the findings of Smith (2003) and Hoefler (2009) individually. However, exploring the parameter space further, by combining different learning biases with different kinds of contexts, has so far revealed an interesting and complex interplay of language learning and language use.

I will explain how the learning bottleneck, learning biases and inference may together determine the evolution of systematic languages. Based on the results obtained, I will discuss the validity of the code model of communication for simulations of language evolution.

References

Briscoe, E. J. (2002). *Linguistic evolution through language acquisition: Formal and computational models.* Cambridge: Cambridge University Press.

Hoefler, S. H. (2009). *Modelling the role of pragmatic plasticity in the evolution of linguistic communication.* Edinburgh: Edinburgh University Press.

Kirby, S. (2002). Learning, bottlenecks and the evolution of recursive syntax. In E. J. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models.* Cambridge: Cambridge University Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford, California: Stanford University Press.

Shannon, C. E., & Weaver, W. (1949). *A mathermatical model of communication.* Urbana, IL: University of Illinois Press.

Smith, K. (2003). *The transmission of language: models of biological and cultural evolution.* Edinburgh: Edinburgh University Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition.* Oxford: Blackwell.