A MODEL OF CULTURAL CO-EVOLUTION OF LANGUAGE AND PERSPECTIVE-TAKING

MARIEKE WOENSDREGT[*1], SIMON KIRBY[1], CHRIS R. CUMMINS[1], and KENNY SMITH[1]

[*]Corresponding Author: mwoensdr@ed.ac.uk
[1]Centre for Language Evolution, University of Edinburgh, UK

Language relies on *theory of mind*, as language users have to entertain and recognise communicative intentions (e.g. Moore, 2016; Scott-Philips, 2014). Theory of mind abilities in turn profit from language, as language provides a means for expressing mental states explicitly (e.g. Bar-On, 2013), and for transmitting one's understanding of minds to others (e.g. younger members of the population) (Heyes & Frith, 2014). Given this interdependence, it has been hypothesised that language and theory of mind have co-evolved (e.g. Malle, 2002).

We present an agent-based model to formalise this hypothesis, which builds on a model described in Woensdregt, Kirby, Cummins, and Smith (2016). This model focuses on the ability to take others' perspectives, which is part of (the development of) theory of mind (e.g. Apperly, 2011), and has a clear and concrete role to play in everyday referential signalling (e.g. Clark & Wilkes-Gibbs, 1986). In this model, the communicative behaviour of an agent is determined not just by their language but also by their perspective on the world. Each agent has a perspective which, in interaction with a given context, determines the salience of potential topics to them. Because all objects in the world are considered potential topics, learners have no other way of inferring the intended referent of an utterance than by gaining knowledge of the speaker's perspective and lexicon (i.e. cross-situational learning is not possible). However, neither a speaker's perspective nor their lexicon are directly observable. Learners thus have to infer both of these simultaneously, using Bayesian inference. Simulation results show that learners can solve this task by bootstrapping one from the other, but only if the speaker uses a language that is at least somewhat informative (Woensdregt et al., 2016).

Given this model of how an agent's perspective in combination with a context gives rise to a mental state, we can compare two different types of agents. *Literal agents* choose an utterance purely based on whether or not the signal is associated with the intended referent in their lexicon. A listener then interprets the utterance by inverting this speaker model (using the lexicon and perspective hypothesis they selected for this speaker after a learning phase). Listeners of this type thus use their perspective-taking ability for interpreting utterances, but literal speakers do

not do so for production. *Pragmatic agents*, in contrast, reason about each other's minds in both directions. A pragmatic speaker optimises their utterances based on how a perspective-taking listener would interpret them (using the *rational speech act* model, Goodman & Frank, 2016). Moreover, pragmatic speakers can in some cases capitalise on the context (and the assumption that the listener knows their perspective) in order to mitigate the adverse effect of ambiguity in their lexicon.

We embedded these models of literal and pragmatic agents in an iterated learning model (where languages are transmitted over generations), in order to explore under which circumstances populations can build a fully informative language from scratch. With literal agents, we found that without any additional learning bias or selection pressure, populations do not establish any linguistic conventions. If we add selection on communicative success, however, populations do evolve a fully informative language. Interestingly, selection on correctly inferring perspectives also results in partially-informative lexicons emerging, which is sufficient for inferring others' perspectives. Note that the only thing that is transmitted over generations is the language; this is thus a model of cultural evolution, where better perspective-inference is reached only by virtue of the population establishing meaningful linguistic conventions. Selecting for the ability to infer perspectives introduces a pressure which leads to the cultural evolution of languages which permit such inference, i.e. the more informative ones. These results illustrate the potential for cultural co-evolution between language and perspective-taking, given the assumption that these two skills are interdependent in their development.

With pragmatic agents, we find the same result when no selection pressure is present. However, the results under selection look different, such that both pressures (selection for communication and selection for inferring perspectives) lead to similar levels of informativeness, and hence similar (high) levels of success at both communicating and inferring perspectives. This is because pragmatic agents can reach high levels of communicative success even with suboptimal languages, while both communication and inferring perspectives still rely on the language being somewhat informative.

To recap, the model described here assumes that language learning and perspective-inference develop interdependently. This is the same for both literal and pragmatic agents, but pragmatic agents do an extra step of reasoning about each others' minds when they communicate. This yields a cultural evolution model in which theory of mind is useful for communication, while language in turn helps to learn about others' minds. We show that when this is the case, a pressure to evolve one of these two skills can give rise to the other for free. In the case of literal agents, the level of communicative success that is reached under a selection pressure for the ability to infer perspectives is limited. However, in the case of pragmatic agents, both types of selection lead to equally high success in communication and perspective-inference, because these agents can communicate successfully even with languages that contain ambiguity.

References

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Psychology Press.

Bar-On, D. (2013). Origins of Meaning: Must We 'Go Gricean'? *Mind & Language*, *28*(3), 342–375.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190), 1243091–1 – 6.

Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Amsterdam: John Benjamins.

Moore, R. (2016). Gricean Communication and Cognitive Development. *The Philosophical Quarterly*, *67*(267), 303–326.

Scott-Philips, T. (2014). *Speaking our minds.* Palgrave Macmillan.

Woensdregt, M. S., Kirby, S., Cummins, C., & Smith, K. (2016). Modelling the co-development of word learning and perspective-taking. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 1241–1246.