DEEP LEARNING MODELS OF LANGUAGE PROCESSING AND THE EVOLUTION OF SYNTAX

WILLEM ZUIDEMA
Institute for Logic, Language and Computation
Amsterdam Brain & Cognition
University of Amsterdam
Amsterdam, the Netherlands
zuidema@uva.nl

Jackendoff's (2002) observation that 'Your theory of language evolution depends on your theory of language' is nowhere more true than in accounts of the origins of syntax and grammar. Researchers arguing that no specialised cognitive machinery is needed for syntactic processing, naturally locate the explanation for the emergence of syntactic language somewhere in the hominin lineage in non-linguistic domains such as changes in social structure, brain capacity or cooperativity. Researchers that favor connectionist or constructionist models of syntax and grammar with language-specific parameters & mechanisms, naturally argue for gradual, evolutionary adaptation of a prexisting brain structures. And finally, researchers that assume discrete, symbolic models of syntax, tend to reject gradualist accounts in favor of a 'saltationist' view, where the crucial machinery for learning and processing syntactic structure emerges in a single step (as a side-effect of known or unknown brain principles or as a macro-mutation). A key argument for the last position is that the vast amount of work on describing the syntax of different natural languages makes use of symbolic formalisms like minimalist grammar, HPSG, categorial grammar etc., and that it is difficult to imagine a process of any sort that moves gradually from a state without to a state with their key components. The most often repeated version of this argument concerns 'recursion' or 'merge', which is said to be an all or nothing phenomenon where you can't have just a little bit of it (Reuland, 2009).

In my talk I will review this argument in the light of recent developments in computational linguistics where a class of models has become popular that we can call 'vector grammars', that include the Long-Short Term Memory networks, Gated Recurrent Networks and Recursive Neural Networks (e.g., Goller & Küchler, 1996; Hochreiter & Schmidhuber, 1997; Socher, Manning, & Ng, 2010; Socher, Bauer, Manning, & Ng, 2013; Karpathy, Johnson, & Fei-Fei, 2015; Rohrmeier, Zuidema, Wiggins, & Scharff, 2015). These models have over the last two years swept to prominence in many areas of NLP and shown state of the art results on many benchmark tests including constituency parsing, dependency parsing, sentiment analysis and question classification. They have in common that they are much more expressive than the older generation of connectionist models (roughly speaking, vector grammars are just as expressive as grammar formalisms popular in theoretical linguistics), but replace the discrete categorical states from those formalisms with points in a continuously valued vector space and hence quite easily allow accounts of gradual emergence (as exploited when these models are trained on data using variants of backpropagation).

I will conclude that these new type of models, although mostly evaluated in a language engineering context, hold great promise for cognitive science and language evolution. For theories of the evolution of grammar and syntax they effectively undermine a key argument: that the only explanatory models are incompatible with gradualist evolution; rather, they show that syntactic category membership, recursion, agreement, long-distance dependencies, hierarchical structure etc. can all gradually emerge, both in learning and in evolution.

References

Goller, C., & Küchler, A. (1996). Learning task-dependent distributed representations by backpropagation through structure. In *International conference on neural networks* (pp. 347–352).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Jackendoff, R. (2002). *Foundations of language.* Oxford, UK: Oxford University Press.

Karpathy, A., Johnson, J., & Fei-Fei, L. (2015). *Visualizing and understanding recurrent networks* (Tech. Rep.). arXiv preprint arXiv:1506.02078.

Reuland, E. (2009). Imagination and recursion: issues in the emergence of language. In F. C. Sophie de Beaune & T. Wynn (Eds.), *Cognitive archaeology and human evolution.* Cambridge University Press.

Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *370*(1664), 20140097.

Socher, R., Bauer, J., Manning, C. D., & Ng, A. Y. (2013). Parsing with compositional vector grammars. In *Proceedings of the ACL conference.*

Socher, R., Manning, C. D., & Ng, A. Y. (2010). Learning continuous phrase representations and syntactic parsing with recursive neural networks. In *Proceedings of the NIPS-2010 Deep learning and unsupervised feature learning workshop.*