

Shalom Lappin

University of Gothenburg, King's College London and Queen Mary University of London

Joint work with Jean-Philippe Bernardy, University of Gothenburg

Title: Learning Syntax with Deep Neural Networks.

We consider the extent to which different deep neural network (DNN) configurations can learn syntactic relations, by taking up Linzen et al.'s (2016) work on subject-verb agreement with LSTM RNNs. We test their methods on a much larger corpus than they used (a 24 million example part of the WaCky corpus, instead of their ~1.35 million example corpus, both drawn from Wikipedia). We experiment with several different DNN architectures (LSTM RNNs, GRUs, and CNNs), and alternative parameter settings for these systems (vocabulary size, training to test ratio, number of layers, memory size, and drop out rate). We also try out our own unsupervised DNN language model. Our results are broadly compatible with those that Linzen et al. report. However, we discovered some interesting, and in some cases, surprising features of DNNs and language models in their performance of the agreement learning task. In particular, we found that DNNs require large vocabularies to form substantive lexical embeddings in order to learn structural patterns. This finding has significant consequences for our understanding of the way in which DNNs represent syntactic information. We also achieved significantly better accuracy with our language model for unsupervised prediction of agreement than Linzen et al. report in their LM experiments.