A comparative assessment of embodied and computational topic extraction

James Carney
Department of Psychology, Lancaster University, Lancaster, Lancashire, United Kingdom

Abstract: Word embedding algorithms like word2vec (Mikolov et al., 2013) have enabled advances in topic modelling by training shallow neural networks on the co-occurrence of words in corpuses of sentences. However, it is not clear how this process reflects human cognition. This poster will compare the results of document classification using the word2vec skipgram model and the 20k sensorimotor word norms collected by the presenter and colleagues (Lynott & Connell 2013; Carney et al., in prep.) (These latter norms establish how concepts are processed by way of perceptual and motor schemes, and thus offer a useful proxy for human conceptual classification.) The results of the comparison will generate insights into the different ways in which higher-order concepts are inferred, and allow systematic biases in concept formation to be identified. It will also allow for machine learning processes to be finessed so as to more accurately reflect human-level modes of cognition.