

Takac, M., Knott, A., Benuskova, L.: The Production of Idiomatic Language in a Simple Recurrent Network, *The 4th Computational Cognitive Neuroscience Conference (CCNC-09)*, Boston, MA, 18-19 Nov 2009 (poster).

Abstract:

## **The production of idiomatic language in a simple recurrent network**

Takac, M., Knott, A., Benuskova, L., Univ. of Otago, New Zealand

The question of how to model idiomatic language is an important one in current linguistic theory. Idioms are fixed multi-word expressions whose meaning is defined holistically, rather than through the meanings of individual words. Idioms can be continuous phrases in a sentence ('kick the bucket') or discontinuous ('take NP to task'). Idioms are hard to model within Chomskyan generative grammar, and have been used as evidence for alternative construction-based models of grammar (see e.g. Jackendoff, 2002). However, the neural mechanisms involved in the representation and processing of idioms have not yet been explored in detail. In this paper we present a neural network model of sentence generation, which is able to produce both continuous and discontinuous idioms within regular compositional sentences. The model is a simple recurrent network (Elman, 1990), extended to include a semantic episode representation as an extra input. It is trained on pairs of episode representations and word sequences. The network learns to rely on a mixture of semantic and syntactic representations to generate compositional (non-idiomatic) language, and on its own context representations to generate idiomatic language. We explore the network's representation of idioms by selectively lesioning its semantic and contextual inputs, and by examining the performance of a control network with no semantic inputs. We conclude that simple recurrent network architectures can learn to recognise idioms, and to switch between idiomatic and non-idiomatic language during a sentence generation task.