COMPACT REPRESENTATIONS OF WORD LOCATION INDEPENDENCE IN CONNECTIONIST MODELS

Frédéric Dandurand & Jonathan Grainger

Laboratoire de Psychologie Cognitive, CNRS & Université de Provence, 3 Place Victor Hugo, 13331 Marseille Cedex 3, France

Research questions

1. What kinds of representations emerge in Cascade- Correlation neural networks when learning word location independence task?
2. Can we find evidence for open bigram representations?

Open bigrams

Coding that involves pairs of letters that allow for non-contiguous letters, but is order-preserving, ex: WITH --> WI, WT, WH, IT, IH and TH

Methods

Task: Location independent encoder task

Example: WITH in 3 locations

Input	Output
WITH##	WITH
#WITH#	WITH
##WITH	WITH

Corpus: 1179 words of four letters

Neural network coding: each letter slot is encoded using 26 binary values indicating the presence and absence of a given letter (Blanks: [0 0 0... 0])

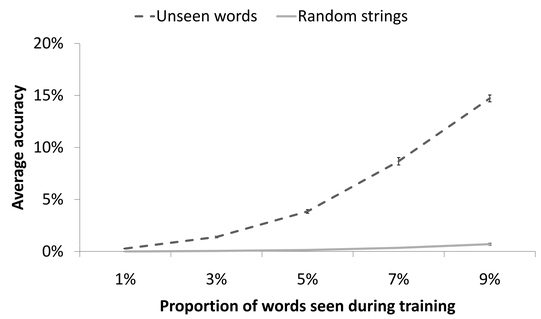
Example:

A : [1 0 0 ... 0],

Z : [0 0 … 1].

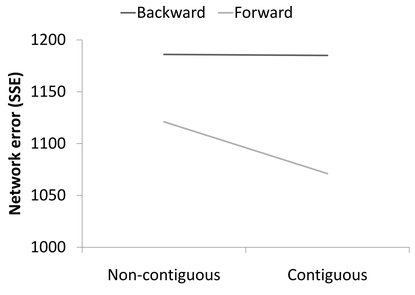
Results

a. Generalization to unseen words



b. Evidence for open bigrams

Train on ABCD, EFGH

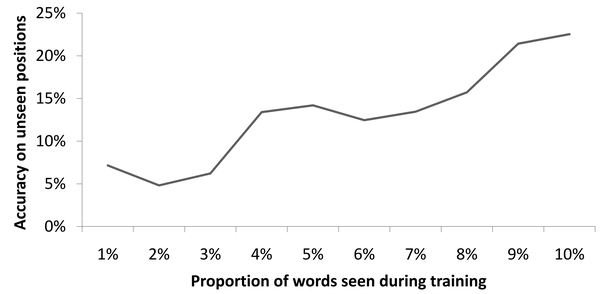


Test bigrams	Forward	Backward
Letter contiguous	ABEF; EFAB; ABFG; FGAB; ABGH; GHAB; BCEF; EFBC; BCFG; FGBC; BCGH; GHBC; CDEF; EFCD; CDFG; FGCD; CDGH; GHCD;	BAFE; FEBA; BAGF; GFBA; BAHG; HGBA; CBFE; FECB; CBGF; GFCB; CBHG; HGCB; DCFE; FEDC; DCGF; GFDC; DCHG; HGDC;
Letter non-Contiguous	ACEG; EGAC; ACEH; EHAC; ACFH; FHAC; ADEG; EGAD; ADEH; EHAD; ADFH; FHAD; BDEG; EGBD; BDEH; EHBD; BDFH; FHBD;	CAGE; GECA; CAHE; HECA; CAHF; HFCA; DAGE; GEDA; DAHE; HEDA; DAHF; HFDA; DBGE; GEDB; DBHE; HEDB; DBHF; HFDB;

Actual (real) bigrams are forward so we expected less error on forward than backward bigrams

c. Generalization to unseen positions

Example: train: ##WITH, WITH## test: #WITH#



Discussion

1. Cascor learned word structure regularities and location invariance simultaneously
2. Error patterns compatible with open bigram representations

References

[1] S. Dehaene, L. Cohen, M. Sigman, & F. Vinckier, The neural code for written words: a proposal, *Trends in Cognitive Sciences, 9* (7), 2005, 335-341.

[2] S. E. Fahlman, & C. Lebiere, *The cascade-correlation learning architecture* (In Advances in neural information processing systems 2, D. S. Touretzky (ed.), Los Altos, CA: Morgan Kaufmann, 1990, 524-532).

[3] J. Grainger, & W. J. B. Van Heuven, *Modeling letter position coding in printed word perception* (In P. Bonin (Ed.), The Mental lexicon. New York : Nova Science Publishers, 2003, 1-24).

[4] S. E. Fahlman, *Faster-learning variations on back-propagation: An empirical study* (In T. J. Sejnowski, G. E. Hinton & D. S. Touretzky (Eds.), the Proceedings of the 1988 Connectionist Models Summer School. San Mateo, CA: Morgan Kaufmann, 1988).

[5] R. Shillcock, & P. Monaghan, The computational exploration of visual word recognition in a split model, *Neural Computation, 13*, 2001, 1171-1198.

[6] J. McClelland, & D. E. Rumelhart, *Explorations in parallel distributed processing* (Cambridge, MA: The MIT Press, 1988).

Neurocomp ’08, Deuxième conférence française de Neurosciences Computationnelles, October 8-11, 2008, Marseille, France