All over the place!
Whole-phrase frequency effects in a full-decomposition model

R. Harald Baayen & Peter Hendrix

Overview

Arnon & Snider (2010) report a whole-phrase frequency effect for 4-word phrases, such as "*all over the place*". They interpret this finding as evidence for phrasal representations in our mental lexicon.

We successfully simulated the whole-phrase frequency effect in a parameter-free full-decomposition model, which shows that whole-phrase frequency effects can emerge in a model that has no phrasal representations.

Introduction

Arnon & Snider (2010) show that phrasal decision latencies for high frequency phrases, such as "*all over the place*" are shorter than those for low frequency phrases, such as "*all over the city*". This effect could not be recuded to frequency effects of single words or substrings. They interpret this result as evidence that multi-word phrases are units of representation.

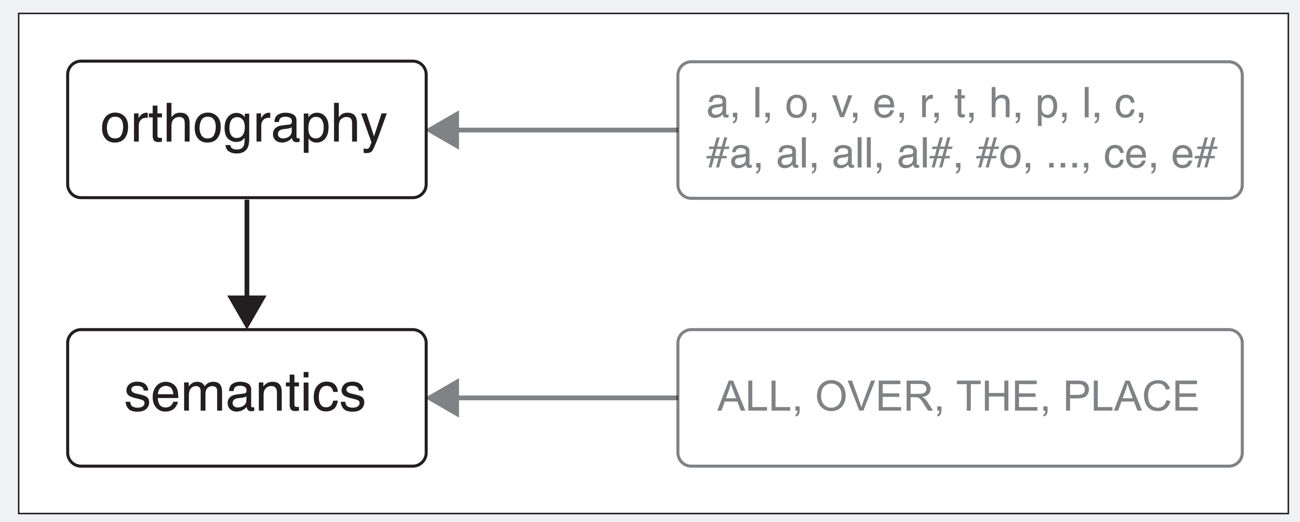
Baayen et al. (2010) present the Naive Discriminative Reader (NDR). The NDR is a discriminative learning model, based on the Rescorla-Wagner equations (Wagner & Rescorla, 1972). It is a full-decomposition model, in which there are no separate representations for morphologically complex words.

Nonetheless, the NDR correctly predicts whole-word frequency effects for complex words. In addition, it successfully replicates a range of morphological effects, including morphological family size effects, inflectional entropy effects, constituent frequency effects and paradigmatic entropy effects.

Here, we show that the NDR also replicates the whole-phrase frequency effect observed by Arnon & Snider (2010). It does so in the absence of any representations beyond the basic word level.

Naive Discriminative Reader

The NDR is a two-layer symbolic network model. The first layer is an orthographic layer, which consists of letters and letter bigrams. The second layer is a semantic layer. The units in this layer are elementary meanings.



The weights on the connections between the orthographic and semantic units are calculated using the Rescorla-Wagner equilibrium equations provided by Danks (2003). The estimated weights are optimal in the least-squares sense.

Simulation

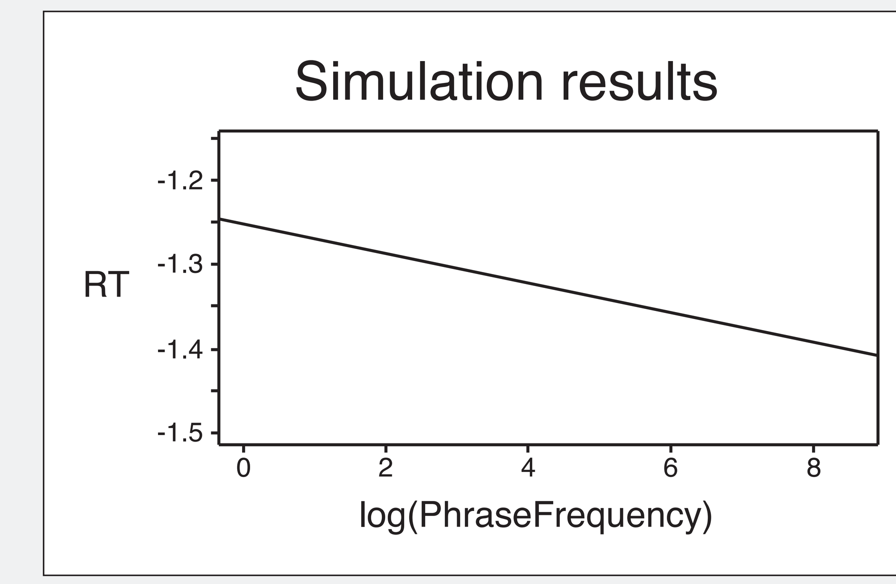
We trained the model on 337,069 phrase types extracted from the British National Corpus.

Given an input phrase (e.g., "*all over the place*"), we defined the activation for a word in that phrase (e.g., "*all*") as the sum of the weights on the connections between all orthographic units in the phrase (e.g., "*a, l, o, v, ..., #a, al, ll, l#, #o, ov, ...*") and that word's meaning (e.g., "ALL"). We then obtained the activation of the whole phrase by summing over the activations of its associated meanings (e.g., "ALL, OVER, THE, PLACE"). We modelled reaction times as proportional to the log of the reciprocal of this summed activation.

The simulation we report here is parameter-free and completely driven by the corpus input.

Results

We found a main effect of phrase frequency, with shorter latencies for more frequent phrases. We did not observe an effect of the frequency of the fourth word or the frequency of the final bigram.



This is an exact replication of the pattern of results found by Arnon & Snider (2010).

Conclusions

The NDR replicates the whole-phrase frequency effect without having to assume that hundreds of millions of n-grams have whole-phrase representations in the mental lexicon. Instead, the whole-phrase frequency effect emerges from co-occurence patterns of orthographic input units and elementary semantic output units. As such, it is quite literally *all over the place*.

References

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62* , 67-82.
Baayen, R. H., Milin, P., Filipovic Durdevic, D., Hendrix, P., & Marelli, M. (2010). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Under revision.*
Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology* , *47* (2), 109–121.
Wagner, A., & Rescorla, R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II* (pp. 64-99). Appleton-Century-Crofts.

Department of Linguistics
University of Alberta, Edmonton

R. Harald Baayen (baayen@ualberta.ca)
Peter Hendrix (hendrix@ualberta.ca)