Reading Time Evidence for Storage of Frequent Multiword Sequences

Colin Bannard[1] & Michael Ramscar[2]

Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany[1]
Stanford University, Stanford, USA[2]

The idea that frequent word sequences are stored by speakers and constitute a core part of their linguistic knowledge has been popular in recent linguistic theory (Goldberg, 2006; Culicover and Jackendoff, 2005). However, while processing advantages have been reported for frequent words and bigrams (e.g. Reali and Christiansen, in press; McDonald and Shillcock, 2003), there has been little work on the processing of longer sequences. In a lexical decision experiment using three word combinations, Bod (2000) found an inverse relationship between reaction time and frequency. However this study has been criticised (e.g. Jurafsky, 2003) for failure to control for component bigram frequencies. Our study looks at whether an advantage for frequent longer sequences can be found even when we control for component frequencies.

Our stimuli were 12 frequent sequences of between 4 and 7 words (e.g. the quality of life) paired (using the BNC) with 12 infrequent sequences (e.g. the quality of men) which were identical except for the last word. The final words and bigrams were matched for frequency ("men" had approximately the same frequency as "life" and "of men" the same as "of life"). The sequences were located in sentences which were presented as a series of chunks of between 3 and 8 words in a self-paced reading study. 30 undergraduates took part, in a counterbalanced within-subject design. Minor differences in word length and component frequencies were factored out using a regression model and the residuals were put into a repeated-measures ANOVA. We found effects of frequency in both by-subjects ($F1(1,29) = 12.667$; $p < .001$) and by-items analyses ($F2(1,11) = 5.489$; $p < .05$), with frequent sequences being read faster than infrequent ones. We interpret this as an expertise effect congruous with the domain-general tendency for well-rehearsed behaviours to be performed quickly. We briefly discuss the implications for models of language processing.

References

Bod, R. (2000). The storage and computation of frequent sentences. AMLAP-2000.
Culicover, P. and Jackendoff, R. (2005) Simpler Syntax. OUP.
Goldberg, A. (2006). Constructions at Work. OUP.
Jurafsky, D. (2003). Probabilistic Modeling in Psycholinguistics. In Bod et al, Probabilistic Linguistics (MIT Press).
McDonald, S. & Shillcock, R. (2003). Low-level predictive inference in reading. Vision Research, 43, 1735-1751
Reali, F. & Christiansen, M. (in press). Word-chunk frequencies affect the processing of pronominal object-relative clauses. Quarterly Journal of Experimental Psychology.