Are word sequences like words? The role of phrase frequency in recognition memory.

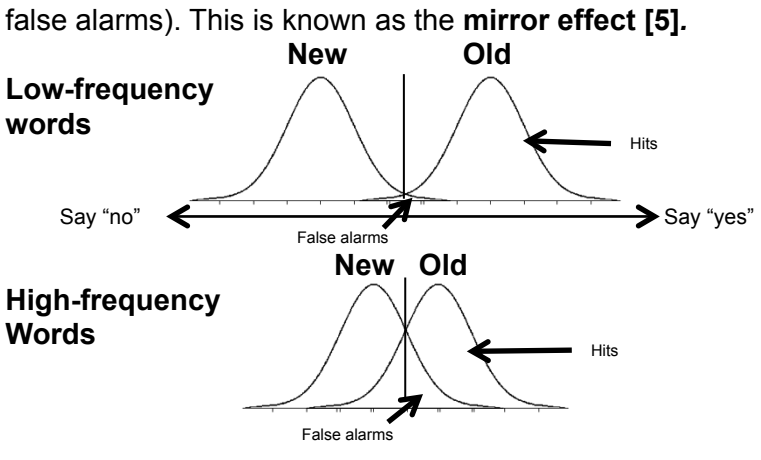
Cassandra L. Jacobs[1], Gary S. Dell[1], Aaron S. Benjamin[1], Colin Bannard[2]
1: University of Illinois at Urbana-Champaign; 2: University of Texas at Austin
This research was supported in part by NIH grant DC000191. The first author is also a recipient of an NSF Graduate Research Fellowship.

The wordlikeness of multi-word sequences

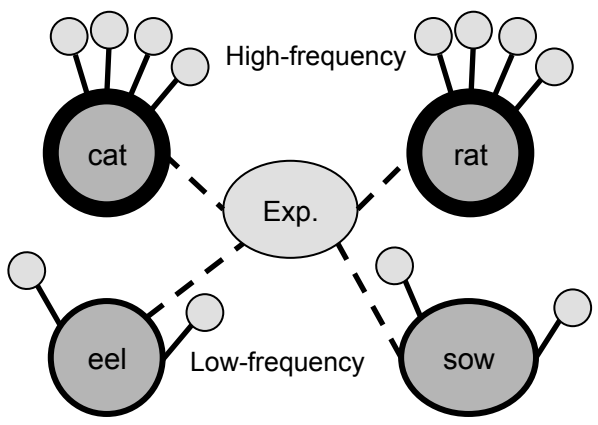
	Frequency advantages			
	Words		Phrases	
	Low	High	Low	High
Task	*anvil*	*tree*	*flaming bounds*	*undue hardship*
Production	High > Low [1,2]		High > Low [1,2]	
Comprehension	High > Low [3]		High > Low [3]	
Acquisition	High > Low [4]		High > Low [4]	

A word frequency paradox in recognition memory

Low-frequency words do better (more hits and fewer false alarms). This is known as the **mirror effect [5].**

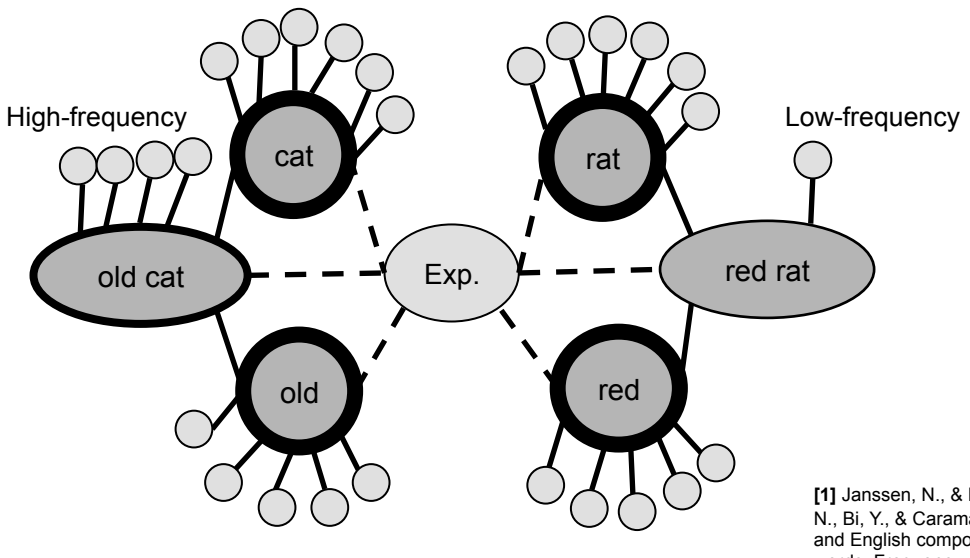


A proposed model for single words (Reder et al., 2002)



Recognition requires retrieval of an experimental episode (**Exp.**). With high-frequency words, there are more *other* episodes to interfere with access of **Exp.** for old words and to promote false alarms for new words.

A proposed model for multi-word phrases



Experiment 1

Participants study 44 nouns and are tested on 88 (44 targets, 44 distractors) in a yes-no recognition test (e.g. *wizard*, *anvil*, *tree;* from [6]). We should replicate the mirror effect here.

Figure 1: Accuracy for words by word frequency. Higher accuracy for low-frequency words

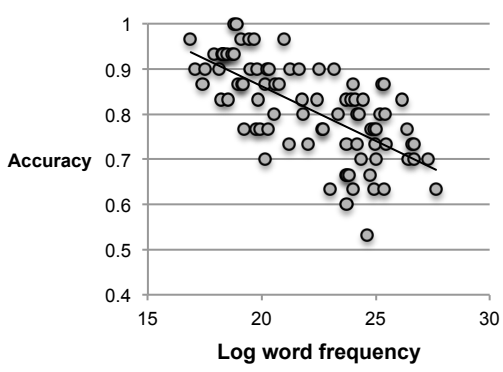
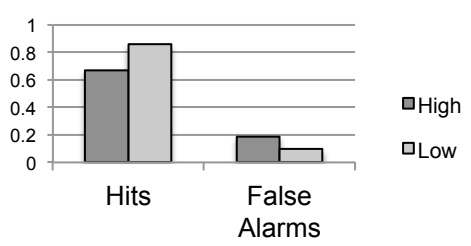


Figure 2: Replication of the mirror effect. More hits and fewer false alarms for low-frequency words



15 in $\log_2$ ~ 3 per 100 million words
30 in $\log_2$ ~ 1000 per million words

Experiment 2

Materials: Adjective-noun pairs from the Google 1T n-gram corpus, with varying phrase frequency but not noun or adjective frequency (e.g. *alcoholic beverage*, *psychic nephew*, *undue hardship;* **Figure 3**).

Procedure: Participants passively study 26 phrases, are given a 30 minute puzzle, and are tested on those 26 plus 26 distractors in a yes-no recognition test.

Figure 3: Phrases have a wide range of frequencies

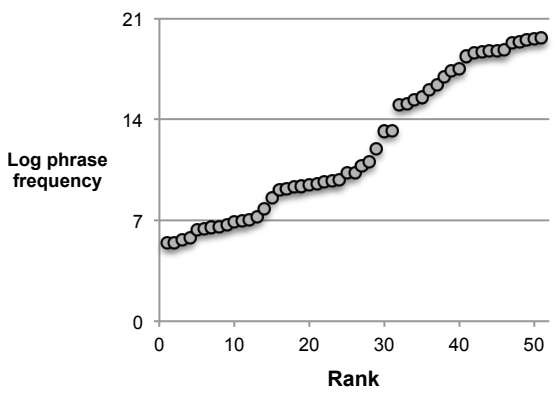
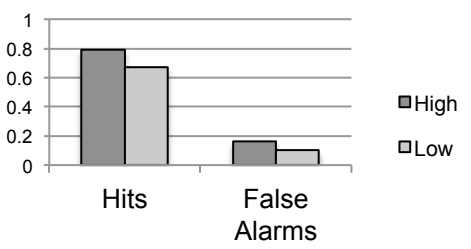


Figure 4: Greater hits *and* false alarms for high-frequency phrases.



6 in $\log_2$ ~ 1 per 100 billion words
20 in $\log_2$ ~ 1 per million words

Figure 5: No benefit on accuracy for low-frequency phrases

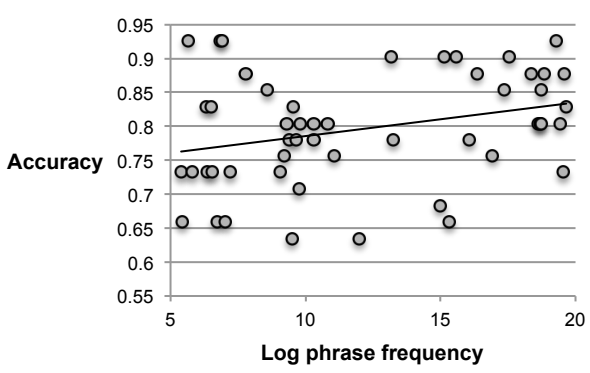
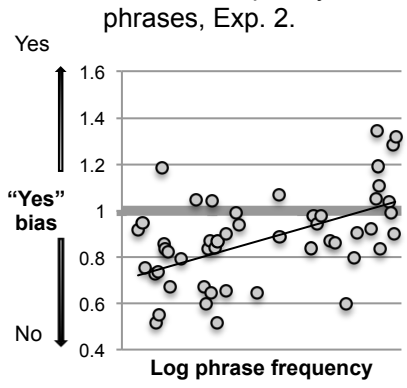


Figure 6: Strong bias to say "no" for low-frequency phrases, Exp. 2.



Discussion

Why don't high-frequency phrases suffer from more interference in finding the experimental episodic memory? Memories store meaning, and a phrase's meaning is mostly the sum of the words within it. Each phrase accesses memories associated with each word and the whole phrase but there are many more memories for the individual words, overwhelming any phrase frequency differences. The bias to say "yes" to high-frequency phrases demonstrates an influence of familiarity on recognition judgments. That the relevant familiarity is *phrasal* is evidence that multi-word sequences are stored and accumulate frequency.

References

[1] Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS ONE*, *7*, e33202. **[2]** Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes*, *23*, 1191–1223. **[3]** Arnon, I., Snider, N., (2010) More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*, 67–82. **[4]** Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning : The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*, 241-248. **[5]** Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13*, 8–20. **[6]** Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*, 199–226.