Journal of Memory and Language 64 (2011) 119–132

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The effects of word frequency and context variability in cued recall

Amy H. Criss *, William R. Aue, Larissa Smith

Department of Psychology, Syracuse University, United States

ARTICLE INFO

Article history:
Received 22 March 2010
Revision received 14 October 2010
Available online 16 November 2010

Keywords:
Cued recall
Episodic memory
Word frequency
Context variability
Mathematical models
Memory models

ABSTRACT

Normative word frequency and context variability affect memory in a range of episodic memory tasks and place constraints on theoretical development. In four experiments, we independently manipulated the word frequency and context variability of the targets (to-be-generated items) and cues in a cued recall paradigm. We found that high frequency targets were better recalled in both pure and mixed lists, even when context variability was held constant. High frequency cues were slightly more effective, but this benefit was eliminated when context variability was held constant. Low context variability cues were most effective while the context variability of the target had little effect on performance. The data suggest that words with fewer pre-experimental connections are better able to isolate the list and that generation of an item from memory benefits from frequency, perhaps due to the ease of generating common orthographic and phonological features. Implications for current models of memory and the prospects of future models are discussed.

Introduction

Properties of to-be-remembered items partially determine the results of later memory tasks. For example, studies have shown differences in memory between pictures vs. words, words vs. non-words, emotional vs. neutral items, high vs. low arousal items, etc. (e.g., Bradley, Greenwald, Petry, & Lang, 1992; Gillund & Shiffrin, 1981; Greene, 2004; Grider & Malmberg, 2008; Kapucu, Rotello, Ready, & Seidl, 2008; Nelson, Reed, & McEvoy, 1977; Onyper, Zhang, & Howard, 2010; Paivio, 1971; Snodgrass & McClure, 1975). Our goal here is to evaluate item properties (i.e., word frequency and context variability) that play a role in the successful generation of a target word and the effectiveness of a cue word.

Word frequency

Normative word frequency (WF) is one property of words that has received much empirical and theoretical attention. In single item recognition, uncommon low frequency (LF) words are remembered better than common high frequency (HF) words. Typically this manifests as a mirror pattern where hit rates (HR) are higher and false alarm rates (FAR) are lower for LF than HF words (e.g., Glanzer & Adams, 1985; Schulman, 1967). This word frequency mirror effect is a benchmark finding that is accounted for by most models of recognition memory, albeit with different underlying mechanisms.

Critically, the pattern of accuracy for HF and LF words changes when the paradigm by which memory is evaluated changes. In a free recall task where participants are asked to generate as many target words from the study list as possible without being provided any explicit memory cue, more HF than LF words are successfully recalled (DeLosh & McDaniel, 1996; Gregg, 1976; Hall, 1954). This pattern holds when the study list is composed of a single frequency (either all HF or all LF). The HF benefit in recall is less reliable, sometimes absent, or even reversed when

* Corresponding author. Address: Department of Psychology, Syracuse University, Syracuse, NY 13244, United States.
E-mail address: acriss@syr.edu (A.H. Criss).

0749-596X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2010.10.001

the encoded list is mixed in WF composition, containing both HF and LF words. This is referred to as the mixed list paradox (Gillund & Shiffrin, 1984; Gregg, Montgomery, & Castaño, 1980; May, Cuddy, & Norton, 1979; Watkins, LeCompte, & Kim, 2000). The mixed list paradox is also present in immediate serial recall (Hulme, Stuart, Brown, & Morin, 2003). This has been interpreted as evidence that pre-experimental associations (assumed to be more plentiful for HF pairs) play an important role in a recall task. This and other dissociations between free recall and single item recognition are difficult to account for with a single model. Such dissociations have contributed to the current state of the field where models tend to be applied to either single item recognition or free recall, but not both.

Context variability

Nearly all studies of WF confound WF with context variability (CV)due, in part, to the high correlation between the two measures (Dennis & Humphreys, 2001; Steyvers & Malmberg, 2003). CV is a measure of the number of *different* contexts in which a word appears in a corpus. Consider the words *soccer* and *wrist*, both low frequency words. *Soccer* almost always appears in the context of sports while *wrist* may appear in a variety of contexts such as anatomy, medicine, jewelry, sports, etc. Soccer is an example of a low CV (LCV) word and wrist is an example of a high CV (HCV) word. Steyvers and Malmberg (2003) were among the first to empirically address the confound between WF and CV in episodic memory. They did so by constructing stimulus sets where the mean WF was approximately the same for both the HCV and LCV word sets and the mean CV was approximately the same for both the HF and LF word sets (this same stimulus set is used in Experiments 3 and 4 of the current paper, see Table 3). In other words, they orthogonally manipulated WF and CV. In their single item recognition experiment, Steyvers & Malmberg found an advantage for LF and LCV words in the form of simultaneous mirror effects for both WF and CV. In experiments where subjects were asked to provide a subjective report of their recognition decisions, LCV words had higher reports of recollection in the HRs and HCV words had higher reports of familiarity in the FARs; the same pattern has also been observed for WF (Cook, Marsh, & Hicks, 2006). To summarize, in single item recognition experiments, independent manipulations of WF and CV (e.g., not confounded with one another) affect performance in the same way. Specifically, LCV and LF words are better recognized than HCV and HF words, respectively.

The effects of WF and CV do not show the same pattern in a free recall task. Recall that pure lists result in a HF advantage in free recall (e.g., Gregg, 1976; Hall, 1954). Both between- and within-subject manipulations of CV reveal better performance for LCV compared to HCV words regardless of WF (Hicks, Marsh, & Cook, 2005). The LCV advantage in free recall holds for the same stimulus set used by Steyvers and Malmberg (2003) and when the stimulus set is further constrained so that concreteness is controlled (Marsh, Meeks, Hicks, Cook, & Clark-Foos, 2006).

The overall pattern across these studies shows separate and independent effects of CV and WF on episodic memory performance. In recognition, LF and LCV targets are better remembered and LF and LCV foils more likely to be rejected than HF or HCV counterparts. In free recall, LCV and HF words are more likely to be recalled than HCV or LF words, respectively.

Models of single item recognition

Mathematical models that account for the word frequency mirror effect in single item recognition are plentiful, as are the proposed underlying mechanisms (e.g., Glanzer & Adams, 1990; Dennis & Humphreys, 2001; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997; Reder et al., 2000). We will consider two examples that are most relevant for this work, acknowledging that there are several alternatives. First, consider the Retrieving Effectively from Memory model (REM; Shiffrin & Steyvers, 1997) which attributes the WF mirror effect to the high diagnosticity provided by uncommon features of LF words. For example, LF words are composed of atypical letters and letter combinations relative to HF words (e.g., Cleary, Morris, & Langley, 2007; Criss & Malmberg, 2008; Freeman, Heathcote, Chalmers, & Hockley, 2010; Landauer & Streeter, 1973; Malmberg & Nelson, 2003; Malmberg, Steyvers, Stephens, & Shiffrin, 2002; Zechmeister, 1969). Uncommon words are composed of uncommon features in REM and therefore tend to not match other words by chance, resulting in a lower FAR for LF words. However, matching an uncommon feature during retrieval provides more evidence in favor of that item than does matching a common feature, leading to a higher HR for LF words. In other words, on average LF foils are a poor match to *other* words stored in memory reducing the FAR and LF targets are a good match to their own memory trace increasing the HR.

Second, consider models that attribute the WF effect to interference caused by the large number and variability of prior contexts in which HF words have been encountered (Dennis & Humphreys, 2001; Reder et al., 2000). The Bind Cue Decide Model of Episodic Memory (BCDMEM; Dennis & Humphreys, 2001) operates via a single process where the features of the reinstated study context are matched against all prior contexts in which the test item had been encountered. According to BCDMEM, HF words tend to be experienced in more pre-experimental contexts and thus have more interference and lower accuracy compared to LF words. In the Source of Activation Confusion (SAC; Reder et al., 2000) model, HF words have higher baseline familiarity at the concept node due to the larger number of times they have been previously encountered, resulting in a higher FAR for HF words. Further, LF words are better recollected due to the relatively smaller number of prior contexts in which they appeared. This LF benefit in recollection overcomes the higher baseline familiarity for HF words, resulting in a higher HR for LF words. Thus, both the SAC and BCDMEM models predict that items with many pre-experimental associations are more difficult to remember because the other contexts to which they are associated interfere with remembering the association of the item and the experimental context.