Journal of Memory and Language 97 (2017) 1–16

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Phrase frequency effects in free recall: Evidence for redintegration

Cassandra L. Jacobs [a,*], Gary S. Dell [a], Colin Bannard [b]

[a] University of Illinois at Urbana-Champaign, United States
[b] University of Liverpool, United Kingdom

ARTICLE INFO

Article history:
Received 22 February 2017
Revision received 19 June 2017
Available online 14 July 2017

Keywords:
Phrase frequency
Redintegration
Memory
Language production
Recall

ABSTRACT

Three experiments examined the effects of word and phrase frequency on free recall. Word frequency did not affect word recall, but when participants studied and recalled lists of compositional adjective-noun phrases (e.g. *alcoholic beverages*), phrase frequency had a consistently beneficial effect: both words from frequent phrases were more likely to be recalled than for infrequent phrases, providing evidence that long-term memory for phrases can aid in pattern completion, or redintegration. We explain these results and those of a previous study of phrase frequency effects in recognition memory (Jacobs, Dell, Benjamin, & Bannard , 2016) by assuming that the language processing system provides features that are linked to episodic contexts. Recall tasks map from these contexts to linguistic elements, and recognition maps from linguistic elements to contexts. Word and phrase frequency effects in both memory tasks emerge both within the language processing system and from multiple stored episodes, and the fact that the representations of phrases are tied to knowledge of their component words, rather than being representational islands.

Published by Elsevier Inc.

Introduction

In many linguistic tasks, phrase frequency effects mirror word frequency effects. Common words (e.g. woman) and phrases (e.g. alcoholic beverage) are easier to acquire, understand and produce than uncommon words and phrases (Arnon & Cohen Priva, 2013, 2014; Arnon & Snider, 2010; Bannard & Matthews, 2008; Bybee, 2006; Janssen & Barber, 2012; Morgan & Levy, 2016; Siyanova-Chanturia, Conklin, & Van Heuven, 2011). The existence of phrase frequency effects demonstrates that the language processing system pays attention to multiword linguistic units. Frequency effects for individual words have typically been accounted for by either positing a lexical entry that keeps track of something like the count of times a person has encountered a linguistic category, or individual memories (exemplars, episodes, or instances) for each of those experiences. Because phrases include a temporal or grammatical relationship between multiple words, it is less clear how phrases might be represented in long-term memory. The present study addresses this question.

One way to explain phrasal frequency effects and phrase representation in general is to propose the existence of a lexically-specific but usage-event-independent representation of the phrase, such as a "node" (e.g. MacKay, 1982) or "superlemma" (e.g. Sprenger, Levelt, & Kempen, 2006) that contains information about its category (e.g. noun phrase, for an adjective-noun combination) and connects to representations of its component words (e.g. Copestake et al., 2002). The frequency of a phrase could be stored with this lexical entry, or it could arise from the number of stored episodes that contain or point to it. Alternatively, phrases could lack explicit discrete representations entirely, in line with theories and computational models that encode all words and phrases implicitly in network weights (Baayen, Hendrix, & Ramscar, 2013; Baayen, Milin, Đurđević, Hendrix, & Marelli, 2010; Seidenberg & McClelland, 1989).

Some recent work has looked into whether phrase frequency effects arise from speakers of a language tracking the episodic representations of events in which a phrase is experienced. Jacobs, Dell, Benjamin, and Bannard (2016) tested whether people showed the same sensitivity to phrase frequency in recognition memory as they are known to have for words. In single-word recognition memory experiments, words that a participant has rarely experienced over the course of their life (low frequency words) have fewer episodic memories in long-term memory, and yet are more accurately discriminated from lures than high frequency words are (Glanzer & Adams, 1985; Hintzman, 1988; Reder et al., 2000). This paradoxical effect of word frequency can be explained by noting that to judge a test word as "old" in a recognition task, the participant may retrieve the episode in which the word was studied.

* Corresponding author at: 603 E. Daniel St., Department of Psychology, University of Illinois Urbana-Champaign, Champaign, IL 61820, United States.
E-mail address: cljacob2@illinois.edu (C.L. Jacobs).

http://dx.doi.org/10.1016/j.jml.2017.07.003
0749-596X/Published by Elsevier Inc.

When that word is low frequency, there are fewer other episodes of it to hinder the search for the crucial experimental episode. Jacobs et al. reasoned that, if adjective-noun phrases have their own episodic memories that contribute to memory in the same manner, then low frequency phrases like *psychic nephew* should also be more accurately recognized than high frequency phrases like *alcoholic beverages*. Surprisingly, they found that high and low frequency phrases were recognized equally well, but that recognition memory improved when the noun in a phrase was uncommon (e.g. *wizard* improves memory for *handsome wizard*). That is, the ability to discriminate new from old phrases, as reflected in a higher hit rate and a lower false alarm rate, was unaffected by phrase frequency, but it was benefited by low-frequency nouns within phrases. They concluded that recognition judgments for phrases are more influenced by the number of episodes containing particular words within the phrase, as opposed to the entire phrase. This is so because individual words are necessarily much more common than phrases. Thus, the many episodes sharing a word with a test phrase are more potent sources of interference in the recognition process than the few episodes containing the entire phrase.

This finding from Jacobs et al. (2016) provides evidence that phrasal processing is at least partially compositional, in that judgments about *psychic nephew* are influenced by memories of events of psychic things that are not nephews and nephews that are not psychic. However, the study also found that participants tended to say they had studied the more common phrases (e.g. *alcoholic beverages*), as evidenced by a bias to respond "yes" with increasing phrase frequency. This suggests that phrase frequency *is* represented in long-term memory, either as a single coherent representation or as individual episodes.

Recognition memory data provide a perspective on how speakers of a language map between linguistic material and a context. A canonical view of recognition is that, at test, speakers are given the linguistic content, the test items, and have to retrieve the experimental context in which they were experienced in order to endorse the items as old (Reder et al., 2000). The demands of a recognition task are therefore more comprehension-like than production-like. The other major memory task, recall, works in the opposite way. An act of recall starts with an existing temporal, discourse, or situational context representation ("recall all of the words *on the list you just saw*") and maps to the linguistic material that was experienced in this context (Criss, Aue, & Smith, 2011; Howard & Kahana, 2002). Recall is an explicit language generation task. In this respect, the demands of recall are more akin to production than comprehension. The current studies therefore examine phrase frequency effects in recall, rather than recognition, to provide a different perspective on the question of the source of such effects and what they tell us about phrasal representation.

Studies of language production demonstrate that frequent words and phrases are easier to say. Word and phrase frequency effects are apparent in a number of production measures including faster onset times (Janssen & Barber, 2012) and shorter word durations in frequent phrases (Arnon & Cohen Priva, 2013; Bannard & Matthews, 2008). Janssen and Barber assessed whether phrase frequency as measured by hits on the Google search engine predicted how easily speakers provided modified noun phrase picture descriptions like *blue car* or *red house* and noun-noun pairs like *bus car* in Spanish as well as noun-adjective pairs in French. They measured speech onset latencies as a function of phrase frequency, the frequency of the first word, and the frequency of the second word in each pair. When Janssen and Barber controlled for word frequency, phrase frequency explained the speedup in speech onset latencies, showing that high frequency phrases are easier to produce. Generally, the higher the phrase frequency, the earlier speakers began talking. Because they found phrase frequency effects, Janssen and Barber argued that phrases are stored holistically and that these representations lack a relationship between the component words and the phrase.

The results of Janssen and Barber were surprising because a previous study by Alario, Costa, and Caramazza (2002) had identified separable contributions of adjective and noun frequency to speech onset latencies, where high frequency adjectives and nouns sped up noun phrase production. Janssen and Barber argued that the results of Alario et al. could have also been due to variations in phrase frequency confounded with word frequency, as high frequency phrases tend to be made up of high frequency words, which have well-known frequency effects.

Additional evidence from child production data corroborates the hypothesis that the production system retrieves multiword units, perhaps in addition to individual words. Bannard and Matthews (2008) used a phrase imitation task in which children repeated phrases that an experimenter said to them. Children made fewer errors, and took less time to produce the overlapping words, when repeating more common phrases (e.g. "a drink of milk") than less common ones that shared the same first three words (e.g. "a drink of tea"). This suggests that long-term memory for multiword sequences has an effect on children's language production.

Theories of language production have not had a great deal to say about the production of phrases, with the possible exception of idiomatic phrases. The notion of a superlemma referred to earlier was developed by Sprenger et al. (2006) to allow for the model of Levelt, Roelofs, and Meyer (1999) to be able to produce idiomatic phrases. For non-idiomatic or compositional phrases, models have not assumed the existence of stored representations of multiword sequences (MacKay, 1982, is an exception in this respect). Because of the need for the production system to be able to assemble completely novel phrases (e.g. "an ugly beauty" cited by Chang, Dell, & Bock, 2006), models have emphasized that structural frames (e.g. adjective-noun) are retrieved, and then individual words UGLY and BEAUTY are retrieved and linked to slots in the frame (e.g. Chang et al., 2006; Dell, 1986; Dell, Oppenheim, & Kittredge, 2008; Garrett, 1975). Finding that production processes are sensitive to phrase frequency (e.g. Bannard & Matthews, 2008; Janssen & Barber, 2012) forces an amendment to these models.

To better understand phrase frequency effects, we consider the task of immediate free recall, which is an episodic memory task that engages the production system. We ask how phrase frequency supports retrieval for production. We will contrast phrase recall performance with recall of individual words. The first experiment (Experiment 1) explores the effects of word frequency on single-word (noun) recall, while Experiment 2 and Experiment 3 examine the influence of phrase frequency on recall of adjective-noun phrases.

Experiment 1

Frequency effects on free recall of nouns

The purpose of Experiment 1 is to examine whether a set of single words that show strong frequency effects in recognition in favor of the low frequency items (Balota, Burgess, Cortese, & Adams, 2002; Jacobs et al., 2016) exhibit similar frequency effects in a free recall task. Some studies have found no effect of frequency on recall (Clark & Burchett, 1994; MacLeod & Kampe, 1996; Hulme, Stuart, Brown, & Morin, 2003), while others have found an advantage for high frequency words (Balota & Neely, 1980; Criss et al., 2011).

When the words that we test for recall here were tested in yes-no recognition, the frequency effects were dramatic: the most