

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Is Lexical Retrieval in Speech Production like Recall or Recognition? The Effects of Word Frequency and Neighbourhood Size

Permalink

<https://escholarship.org/uc/item/0xn678xq>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors

MacAndrew, Siobhan B.G.
Harley, Trevor A.

Publication Date

2000

Peer reviewed

Is Lexical Retrieval in Speech Production like Recall or Recognition? The Effects of Word Frequency and Neighbourhood Size

Siobhan B. G. MacAndrew (s.macandrew@abertay-dundee.ac.uk)

Division of Psychology, School of Social & Health Sciences
University of Abertay Dundee
Dundee DD1 1NJ, UK

Trevor A. Harley (t.a.harley@dundee.ac.uk) & **Sheila Colgan**

Department of Psychology, University of Dundee
Dundee DD1 4HN, UK

Abstract

We investigate the effects of word frequency and lexical neighbourhood density on word recall and recognition. We found a three-way interaction between memory task, the size of lexical neighbourhood of a target word, and target word frequency. In particular, performance on low frequency words with many lexical neighbours was surprisingly good in the recognition condition. The results show that the number of lexical neighbours of the target moderates the word frequency effect in recognition. Large neighbourhood size always has a facilitatory effect upon performance. The findings are contrasted with those observed in lexical access in speech production.

Introduction

To what extent is retrieving a word when speaking like accessing a fact from long-term memory? In particular, how do the language processes involved in lexical access for spontaneous speech production relate to the memory processes involved in the retrieval of word lists from long term memory? On the one hand, our intuition is that lexical retrieval is like recall. Indeed, we even talk in these terms in every day use, using constructions such as “I cannot recall that word”. On the other hand, some models of lexical access in speech production involve search through a list of phonological forms (Butterworth, 1980, 1989; Fay & Cutler, 1977). Such a search might well involve an element of recognition when the appropriate form is reached. This paper looks at two psycholinguistic variables that are well known to influence lexical access (word frequency and lexical neighbourhood size), and investigates their effect on free recall and recognition performance. We compare their effects on a memory task with their effects on a language production task.

Word frequency is an important variable in all language tasks, including speech production (Harley, 1995). Frequency always has a facilitatory effect in speech production. For example, we are faster to name high frequency words and objects with high frequency names (see Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965). Harley and Bown (1998), using a laboratory-based “tip-of-the-tongue” (TOT) induction task (Brown &

McNeill, 1965), showed that we are more likely to experience a TOT state on less common words.

The second variable employed in this study is lexical neighbourhood size. Some words (e.g. “corpse”) are phonologically unique in that there are no other words that sound like them. Other words (e.g. “cage”) have a large number of phonological neighbours (“page”, “rage”, “sage”, and “cave”, among others). Obviously we need a suitable measure of lexical similarity; we discuss this below. It is well established that a word’s lexical neighbours play in an important role in word recognition (e.g. Glushko, 1979; Grainger, 1990). It is now also becoming apparent that they play some role in word production. Harley and Bown (1998) showed that the number of phonological neighbours a word has affects lexical retrieval in the tip-of-the-tongue state. In particular, they showed that when word length and frequency are controlled for, people are more likely to have difficulty with words that have few phonological neighbours. This result showed that a large set of potential responses can in fact increase the chances of successful retrieval of the target. Harley and Bown hypothesised that structurally similar items provide supporting activation for each other. This finding also supports the “insufficient activation” hypothesis for the origin of TOTs (Burke, MacKay, Worthley, & Wade, 1991).

Although research on the effects of neighbourhood density on lexical access in speech production is at an early stage, the pattern observed is that it is easier to produce frequent words that have many neighbours. Will this pattern be observed in memory tasks? Of course, the pattern observed might well differ depending upon the exact task used. In particular, we might observe different outcomes depending on whether we use a recognition or recall memory task.

The effect of word frequency on recognition is well known, if poorly understood. The “word frequency effect” is the finding that recognition memory is better for low frequency words than high frequency words (Baddeley, 1990; Gregg, 1976; see Guttentag & Carroll, 1998, for a recent review). In recognition, we make a judgement about whether or not we have recently seen a particular item. Is

the stimulus activated because of recent exposure, or is it activated just because of an intrinsic property, such as its high frequency? There is no such conflict in the case of low frequency words, where high activation of the stimulus representation is much more likely to have come from recent exposure in the study list. Put more colloquially, frequent words are less distinctive. This line of reasoning is commonly known as the *memorability hypothesis* (e.g. Brown, Lewis, & Monk, 1977).

Less is known about how word frequency affects performance in a recall task. If frequency operates by raising the activation levels of frequently used items (e.g. Morton, 1979), then the free recall of a high frequency word should be relatively easy and that of a less frequent word relatively difficult. In summary, frequency should facilitate recall but might hinder recognition. In the light of these hypotheses, the finding that in speech production high target word frequency always has a facilitatory effect suggests that lexical access is more like recall than recognition.

The effect of a large lexical neighbourhood is to increase the number of potential responses. The existence of plausible alternative responses may have different effects on recall and recognition. Recognition is more difficult when selecting from a large set of plausible responses than a smaller set (e.g. in the long-term memory version of Sternberg, 1966). Crucially, the similarity between targets and distractors affects recognition (Dale & Baddeley, 1962). Hence large lexical neighbourhoods should hinder recognition.

It is less clear how neighbourhood size will affect recall. If free recall acts like speech production, we would expect that words with many neighbours should be relatively easy to recall. One way of conceptualising this is that a word's neighbours should act as possible retrieval cues. On the other hand, in the two-process, generation-recognition account of free recall (Anderson & Bower, 1974; Kintsch, 1970), recall contains within it an element of recognition. In this case, the recall of words with many neighbours will be either hindered, or the effects of the two processes may cancel out so that no difference is observed. We attempt to explore these issues by examining the effect of the target word's phonological neighbourhood.

It is unclear how word frequency and lexical neighbourhood size will interact. If either recognition or recall resembles speech production while the other does not, we will obtain a three-way interaction with particularly poor performance on low frequency words with few neighbours. The simplest prediction is that speech production resembles recall, and that the pattern observed in speech production should therefore also be observed in the free recall task. It is less clear what should happen in the recognition task. One possibility is that the word frequency effect should overwhelm any effects of neighbourhood size, but any prediction here is prematurely speculative.

In summary, the aim of this paper is to examine the effects of word frequency and lexical neighbourhood size on measures of memory.

Method

Participants

We tested 30 volunteers, who had a mean age of 34 years. They were all psychology undergraduates of the University of Dundee, Scotland. Ten females and five males took part in each of the two experimental conditions.

Materials

All of the words used in the experiment were nouns of one or two syllables in length. The experiment required a printed list of target words used in the learning phase of the study, and a printed list of these words plus distractor words in the recognition condition.

The target items were the same as those of Experiment 2 of Harley and Bown (1998). There were 60 words in the target list, of which 30 were of high frequency (at least 100 instances per million words, with a mean of 163.7, as sampled in Francis & Kucera, 1982), and 30 of low frequency (under 9 instances per million, with a mean of 3.7). Within each list of 30, 15 of the target words had a dense lexical neighbourhood as evidenced by a mean N value of 15.1 (see Coltheart, Davelaar & Besner, 1977; our figures were taken from the MRC Database of Coltheart, 1981). The remaining 15 words had no close orthographic neighbours as evidenced by a mean N value of 0. The N value is a measure of a word's orthographic neighbourhood size: it is the number of other words that can be made from a particular word by changing one letter. Obviously the higher the N score, the larger the orthographic neighbourhood. Orthographic and phonological neighbourhood sizes are highly correlated. This issue is discussed in depth in Harley and Bown (1998), who found the same results whether orthographic or phonological neighbourhood size was used. The properties of the materials are summarised in table 1.

Table 1: Properties of materials

Condition	Frequency	N Value
High F, high N	246.7	15.1
High F, low N	225.5	0
Low F, high N	7.2	15.1
Low F, low N	5.6	0

This process yielded four sets of fifteen target words, balanced for frequency and orthographic neighbourhood size, comprising words of high frequency and high N value, high frequency and low N value, low frequency and high N value, and low frequency and low N value. The words were combined in random order to form one list.

The target words were printed in black ink, one beneath the other, in two columns on A4 paper in random order for use in the presentation phase of the experiment. Examples include “ball” and “date” (high frequency, high N), “cage” and “dove” (low frequency, high N), “growth” and “view” (high frequency, low N) and “corpse” and “tinsel” (low frequency, low N).

The recognition condition of the experiment consisted of the targets and 60 distractor items. In this particular experiment, the distractors were related in meaning to items from the target word set. Although it is clearly of interest to study other types of distractor, we wanted to make this task similar to speech production. Therefore the potential competing words were maximally plausible alternatives that were semantically similar to the targets. The target words were paired with close semantic associates. The items for the recognition task were also hand-printed in black ink, one beneath the other, in random order, on a single sheet of A4 paper.

Procedure

All participants were given 5 minutes in which they were told to read the presentation list of 60 words and to try to remember them. This was followed by an interval of 5 minutes during which participants engaged in conversation and listened to music. Participants in the recall condition were then given 5 minutes to write down as many words as they could recall from the presentation list. Participants in the recognition condition were told they had 5 minutes to read the recognition list of 120 words and underline in pencil any words that they thought they had previously seen.

Results

The experimental design comprised three factors. There was a between-subjects factor of memory task (with the two levels of free recall and recognition). There were two within-subjects factors, one of word frequency (with the two levels of high and low frequency) and one of lexical neighbourhood size (with the two levels of high N score and low N score).

A 2x2x2 ANOVA on the correct memory scores of the participants showed main effects of memory task ($F(1, 28) = 14.72, p < 0.001; MSE = 143.0$), word frequency ($F(1, 28) = 7.71, p < 0.025; MSE = 27.1$), and neighbourhood size ($F(1, 28) = 27.9, p < 0.001; MSE = 130.2$).

Importantly, there was a significant three-way interaction between memory task, word frequency, and lexical neighbourhood size ($F(1, 28) = 10.80, p < 0.01; MSE = 27.1$). There was also a significant two-way interaction between memory task and word frequency ($F(1, 28) = 12.96, p < 0.005; MSE = 49.4$). The interaction between memory task and neighbourhood size approached significance ($F(1, 28) = 3.95, p = 0.06, MSE = 18.4$), but there was no hint of any interaction between frequency and

neighbourhood size ($F(1, 28) = 1.2$). Figure 1 summarises these results.

As was expected, the level of recognition performance was better than that of free recall. Performance on words with dense lexical neighbourhoods was better than that on words with sparse neighbourhoods across both the recall and recognition conditions. Performance on high frequency words was generally better than on low frequency ones. The likely source of the three-way interaction, however, is that low frequency, dense-neighbourhood words perform unusually well in the recognition task (or unusually poorly in the recall condition). Words with many neighbours are significantly easier to recognise than those with few neighbours ($t[28] = 4.01, p < 0.001$). There is no difference between the corresponding conditions in the recall task ($t[28] = 1.10$). Indeed, recognition performance for the less frequent words with many neighbours was the best of all conditions. A consequence of this interaction is that there is no word frequency effect for words with few neighbours in the recognition task; performance on low frequency words is in fact worse than that on high frequency words, although not significantly so ($t[28] = 0.95$).

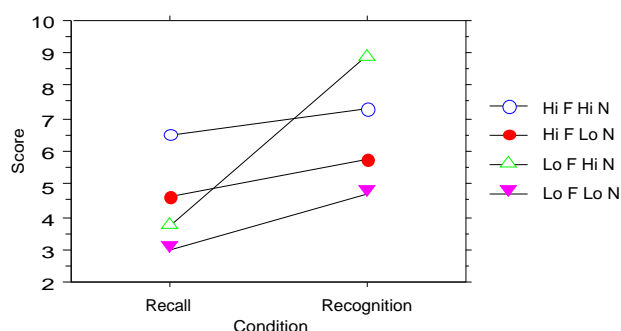


Figure 1: The effects of word frequency and lexical neighbourhood size on recall and recognition.

Discussion

In summary, we obtained a three-way interaction between memory task, word frequency, and phonological neighbourhood size, demonstrating that these two variables have significant differential effects upon recall and recognition. Large neighbourhood size always has a facilitatory effect on both recall and recognition performance, suggesting that the neighbours of target items act as a source of support rather than interference. In the recognition task, performance on words with few neighbours was better on high frequency words than low frequency words, reversing the usual frequency effect.

In the recall task, performance on high frequency words was uniformly better than on low frequency words, confirming our predictions based on the consideration of lexical activation levels. The amount of facilitation provided by dense lexical neighbourhoods is not large, but

words with many neighbours are easier to recall than words with few.

The findings in the recognition task are more complex. The word frequency effect in recognition, whereby less frequent items are easier to recognise than high frequency words, was replicated only for words with many lexical neighbours. There was no advantage (indeed, a slight disadvantage) for low frequency words with few neighbours. This suggests that any account of the word frequency effect must take into account the role of lexical neighbourhood size. The other conditions in the recognition task are in line with those of the recall task, bearing in mind the expected generally better performance in the recognition task.

Why should low frequency words with no or few neighbours be particularly difficult to recognise? The result appears contrary to the memorability hypothesis. A word which is orthographically and phonologically unique as well as uncommon should be more noticeable and therefore memorable than one with many neighbours.

There are at least two possible explanations. The first is that during the study phase, a target word primes the words in its neighbourhood. During the test phase of the experiment, the primed items then cue the target. The more neighbours there are to act as primes in the test phase, the more likely is a correct response. Words with few neighbours do not have this advantage.

Consideration of the attention-likelihood model of Glanzer and Adams (1990) suggests another explanation. They suggested that in the study phase of a recognition experiment, people pay more attention to some items than others. In general people might redistribute effort at encoding or in rehearsal towards troublesome items (see also Fritzen, 1975; Hastie, 1975; Murnane & Shiffrin, 1975). Low-frequency words with many neighbours may strike participants as odd. They therefore pay a disproportionate amount of attention to them, in particular ensuring that the low-frequency target is not in fact one of its own neighbours. On the other hand, it is possible that participants consider low-frequency words with few neighbours to be "obvious", and therefore pay little attention to them. In the recognition phase, performance will be poor on those items that had less attention allocated to them in the study phase (the low-frequency few-neighbours words). We cannot distinguish between these two possible explanations on the basis of our current data, and of course, they may not be incompatible.

Attention-likelihood theory is one explanation of the "mirror effect". Consider an experiment with two conditions (e.g. high and low frequency items) where the items in one are better recognized than items in the other. Then the superior condition will give better recognition of previously-seen items (i.e. targets) as being old but also better recognition of new items (i.e. distractors) as being new. (See Glanzer, Adams, Iverson, & Kim, 1993; Glanzer, Kim, & Adams, 1998; Stretch & Wixted, 1998; but see also Murdock, 1998.) Consideration of lexical

neighbourhood size may be helpful in giving an account of the mirror effect.

Another surprising finding is that, counter to our intuitions and prediction, the pattern of performance observed by Harley and Bown (1998) in the TOT task is here mirrored in the recognition task, and not in the recall task. In particular, Harley and Bown found a large difference between low frequency words with dense and sparse neighbourhoods. Here we only observed this difference in the recognition task. This suggests that lexical access in speech production contains an important recognition component. Of course, some caution is necessary in making this claim; it is necessary to reproduce our findings on a task more directly oriented to speech. There are at least two possible loci for a recognition component in lexicalization. First, lexical search models such as those of Butterworth (1980) and Fay and Cutler (1977) involve search through ordered lists of lexical entries. Selecting the correct entry might involve recognition. Second, speech production might contain an element of monitoring and editing. These processes might involve recognition. There is independent evidence for the existence of monitoring processes from self-repair of speech (see Levelt, 1989) while others (e.g. Baars, Motley, & MacKay, 1975; Butterworth, 1982) postulate that it is necessary to account for characteristics of speech errors.

An important caveat to any conclusion regarding the resemblance of speech production to other memory tasks concerns what happens in the tip-of-the-tongue state. The presumption in the literature is that a TOT state is an extended form of a hesitation in normal speech (see Harley, 1995, for a review; see also Levelt, 1989). Harley and Bown (1998) suggested that strategic factors might sometimes be operative in laboratory-induced TOT states. In particular, we suggested that there might be an editor responsible for monitoring the output of the interlopers, the words that often spontaneously come to mind when in a TOT state. Others have also proposed that our potential speech output can be edited by a late-acting monitor (e.g. Levelt, 1989). This editor might sometimes discard grossly implausible candidates. The editor must be far from perfect, however, as many implausible candidates are often output; and about a quarter of the time these interlopers bear no obvious relationship to the target. If and when it operates, this post-access monitor might plausibly contain an element of a recognition process. There is no reason to suppose that this applies to either spontaneous production or the strivings to retrieve the target word itself.

If this is the case, the recognition component observed in TOT states comes from the action of post-access strategic processes, rather than the processes of lexical retrieval themselves.

In summary, we have shown that lexical access in the tip-of-the-tongue state surprisingly resembles performance on a recognition task rather than on a free recall task. We have also shown that the word frequency effect in recognition is moderated by the size of the lexical neighbourhoods of the target items. The exact way in

which neighbours exert their effects in these tasks remains to be explored.

Acknowledgements

We are grateful to several anonymous reviewers for their helpful comments. Our data were collected by Sheila Colgan as part of an honours dissertation at the University of Dundee. We gratefully acknowledge financial assistance from the University of Abertay Dundee research development fund.

References

- Anderson, J. R., & Bower, G. (1974). A propositional theory of recognition memory. *Memory and Cognition*, 2, 406–412.
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 14, 382-391.
- Baddeley, A. (1990). *Human memory: Theory and practice*. Hove: Lawrence Erlbaum Associates.
- Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, word frequency, and negative recognition. *Quarterly Journal of Experimental Psychology*, 29, 461-473.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325-337.
- Burke, D., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, 30, 237-246.
- Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language production, Vol. 1: Speech and talk* (pp. 423-459). London: Academic Press.
- Butterworth, B. (1982). Speech errors: Old data in search of new theories. In A. Cutler (Ed.), *Slips of the tongue*. Amsterdam: Mouton.
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108-135). Cambridge, MA: MIT Press.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). London: Academic Press.
- Dale, H. C. A., & Baddeley, A. D. (1962). On the nature of alternatives used in testing recognition memory. *Nature*, 196, 93-94.
- Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505-520.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin company.
- Fritzen, J. (1975). Intralist repetition effects in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 1, 756-763.
- Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition*, 13, 8-20.
- Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, 100, 546-567.
- Glanzer, M., Kim, K., & Adams, J. K. (1998). Response distribution as an explanation of the mirror effect. *Journal of Experimental Psychology: Human Learning and Memory*, 24, 633-644.
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691.
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, 29, 228-244.
- Gregg, V. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recognition and recall*. Chichester: John Wiley and Sons.
- Guttentag, R., & Carroll, D. (1998). Memorability judgements for high- and low-frequency words. *Memory and Cognition*, 26, 951-958.
- Harley, T. A. (1995). *The psychology of language*. Hove: Psychology Press.
- Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, 89, 151-174.
- Hastie, R. (1975). Intralist repetition in free recall: Effects of frequency attribute recall instructions. *Journal of Experimental Psychology: Human Learning and Memory*, 1, 3-12.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 824-843.
- Kintsch, W. (1970). *Learning, memory, and conceptual processes*. New York: John Wiley and Sons.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics series 2: Structures and processes*. London: Paul Elek.
- Murdock, B. B. (1998). The mirror effect and attention-likelihood theory: A reflective analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 524-534.
- Murnane, K., & Shiffrin, R. M. (1991). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 855-874.

- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273-281.
- Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652-654.
- Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1379-1396.