Memory & Cognition
2001, 29 (5), 774-776

Phonological similarity in working memory

DAVID E. COPELAND and GABRIEL A. RADVANSKY
University of Notre Dame, Notre Dame, Indiana

That phonologically similar words in a short-term memory test are more difficult to recall than phonologically dissimilar words is a well-known phenomenon. This effect is the *phonological similarity decrement*. In the present study, we examined whether this phonological similarity decrement is present when additional semantic information is available, as in a reading span test, as compared with a standard presentation, or in the context of an operation span test. The results revealed a phonological similarity facilitation. Phonologically similar words were remembered better than phonologically dissimilar words.

Most people have read the passage, "I do not like green eggs and ham, I do not like them Sam-I-am." (Geisel, 1960). The phonological similarity of rhymes like this can be used to aid in retention (e.g., Searleman & Herrmann, 1994). However, a prominent finding in short-term memory research is that phonological similarity leads to poorer memory (Conrad & Hull, 1964).

Short-term memory, or working memory (Baddeley, 1986), is used for the transient retention of information. One of the most popular ways to study working memory is to measure its capacity. There are various ways that this can be done, but this paper will only address the more popular of these. One of the simplest measures is the word span test. In this test, a person is presented with a list of words and then must recall them in the order in which they were presented. Two other, more complex tests are the reading (Daneman & Carpenter, 1980) and the operation (Turner & Engle, 1989) span tests. These span tests include processing (reading sentences or verifying math problems) and storage (list of words) components. In general, the more words that can be recalled, the greater a person's working memory span is thought to be.

One of the most prominent effects in working memory is a phonological similarity decrement (Conrad & Hull, 1964). This effect is characterized as worse performance when the words in a list are phonologically similar (e.g., *bat*, *flat*, *mat*) than when they are phonologically dissimilar (e.g., *dirt*, *pen*, *hug*). The phonological similarity decrement has been demonstrated repeatedly in word span tests. Our goal was to determine whether a phonological similarity decrement is present in more complex span tests, such as the reading and the operation span tests.

This research was supported in part by Grant ARMY-DASW01-99-K-0001 from the Army Research Institute, awarded to the second author. We thank Diane Berish, Anthony Bishara, Kris White, and Sarah Wieber for their assistance in collecting the data. Copies of the stimuli are available from the authors. Correspondence concerning this article should be addressed to D. E. Copeland, 118 Haggar Hall, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556 (e-mail: copeland.9@nd.edu).

EXPERIMENT 1

The aim of Experiment 1 was to compare performance for phonologically similar and dissimilar words in the word span and the reading span tests.

Method

Subjects

Fifty-six undergraduates, 27 males and 29 females, from the University of Notre Dame participated in exchange for partial course credit. The data from 2 additional subjects were not analyzed owing to experimenter error. The subjects were all native English speakers.

Materials and Procedure

The materials consisted of a list of words and a list of sentences. Each of these sets was constructed for a similar and a dissimilar condition. The similar condition contained sets of phonologically similar target words. The dissimilar condition used the same words and sentences, but reordered so that the target words in any one set were not phonologically similar.

The subjects were randomly assigned to one of the two conditions. The first consisted of the word span test with similar words and the reading span test with dissimilar words. The second condition was the opposite. The order of presentation (word vs. reading span) was randomly assigned for each subject. Note that across subjects, the same words were used in both conditions within a span test. Thus, the conditions are equivalent for word frequency and concreteness. Also, the words used in each span test did not differ in terms of frequency and concreteness (all ts < 1).

Word span. For the word span test, the subjects were presented with a word on the computer screen for 1 sec, followed by a 1-sec pause before the next word. The subjects read the words aloud as they were presented. They were presented with three sets of two words for practice. Following this, they worked through three sets of three, four, five, six, and seven words, in that order. At the end of each set, the subjects recalled the words from that set out loud while the experimenter typed their responses into the computer.

Reading span. The reading span test was conducted in a similar manner. The subjects were presented with sets of sentences. They had to read the entire sentence out loud, and the experimenter pressed the space bar to advance to the next sentence. At the conclusion of each set, the subjects recalled the last word from each sentence in

that set. Again, the experimenter typed in the words. There were three sets of two sentences for practice, followed by three sets of two, three, four, five, and six sentences in the reading span test itself.

Scoring. For both tests, if a set was recalled in the correct order, the subscore was equal to the number of words in that set (following LaPointe & Engle, 1990). These subscores were totaled for a final score. The maximum possible scores were 75 and 60, for the word and the reading span tests, respectively.

Results and Discussion

The data for the two groups in Experiment 1 are presented in Table 1. For the word span test, the results revealed the phonological similarity decrement (Conrad & Hull, 1964). The subjects scored higher in the dissimilar condition than in the similar condition [$F(1,54) = 3.87$, $MS_e = 37$, $p < .06$]. Importantly, the opposite pattern occurred for the reading span test. The subjects scored higher in the similar condition than in the dissimilar condition [$F(1,54) = 7.18$, $MS_e = 48$, $p < .05$]. This is consistent with the idea that the additional contextual information in the reading span test may aid recall performance when combined with the knowledge of a rhyming cue.

EXPERIMENT 2

One possible reason for the result of Experiment 1 is that between-group differences may be responsible for the pattern of data. A within-subjects design was used in Experiment 2 to control for this.

Method

Subjects

Thirty-four undergraduates, 16 males and 18 females, from the University of Notre Dame participated in exchange for partial course credit. The data from 9 additional subjects were discarded owing to either experimenter error or failure to finish all four tasks. The subjects were all native English speakers.

Materials and Procedure

Experiment 2 used the same words and sentences as those from Experiment 1. In addition, a second set of words and sentences were created and organized into phonologically similar and dissimilar lists. This allowed each person to complete all four span tests. The order of the tests was randomized for each subject. The subjects were also randomly assigned to one of two groups. These groups differed in that one group saw a particular list in the similar condition, whereas the other group saw that same list in the dissimilar condition. Again, as a reminder, the same words were used across subjects, and the words in the two span tests did not differ in terms of frequency and concreteness. Scoring was the same as that for Experiment 1.

Table 1
Means and Standard Deviations for Similar and Dissimilar Conditions in the Three Experiments

	Similar		Dissimilar	
Test	*M*	*SD*	*M*	*SD*
Word span (Experiment 1)	22	7	26	5
Reading span (Experiment 1)	21	7	16	6
Word span (Experiment 2)	19	6	25	7
Reading span (Experiment 2)	24	10	17	9
Operation span (Experiment 3)	22	9	25	10

Note—The maximum score for the word, reading, and operation spans were 77, 60, and 60, respectively.

Results and Discussion

The data for Experiment 2 are listed in Table 1. These results replicated those for Experiment 1. The subjects scored higher in the dissimilar condition than in the similar condition for the word span test [$F(1,33) = 27.04$, $MS_e = 23$, $p < .001$]. In contrast, for the reading span test, scores were higher in the similar condition than in the dissimilar condition [$F(1,33) = 11.53$, $MS_e = 72$, $p < .01$].

EXPERIMENT 3

The aim of Experiment 3 was to reinforce the idea that the content of the sentences in the reading span test contributed to the phonological similarity facilitation. Two alternative explanations were tested. First, it is possible that the reversal was due to the delays between the to-be-remembered words (Nairne & Kelley, 1999).[1] Second, articulatory suppression has been shown to eliminate the phonological similarity decrement (Murray, 1968). The reading span test involves a delay filled with articulatory suppression between word acquisition. Thus, it could be that the delay, articulatory suppression, or both is contributing to the phonological similarity facilitation. Experiment 3 used the operation span test (Turner & Engle, 1989) with either phonologically similar or dissimilar words. The operation span test, like the reading span test, includes a delay filled with articulatory suppression between word acquisition. However, the operation span test uses math problems instead of sentences.

Method

Subjects

Thirty-four undergraduates, 22 females and 12 males, from the University of Notre Dame participated in exchange for partial class credit. No subjects needed to be replaced. All the subjects were native English speakers.

Materials and Procedure

Experiment 3 used the operation span test. This task consists of a two-step math problem with an answer provided [e.g., $(5 \times 2) - 3 = 7$]. The subjects read the problem out loud, then indicated whether the answer was correct by pressing one of two mouse buttons. Following this, a word was displayed on the screen for 1 sec. The subjects read the word out loud as well. At the end of a set, the subjects recalled the words from that set in the appropriate order while the experimenter typed the words into the computer. The operation span consisted of the same number of sets, set sizes, and practice trials as the reading span test. Also, the words used in the reading span test in Experiment 2 were used here. Scoring was done in the same manner as in Experiments 1 and 2.

Results and Discussion

The results of Experiment 3 (see Table 1) were analyzed using a within-subjects ANOVA. The subjects scored higher in the phonologically dissimilar condition

than in the phonologically similar condition [$F(1,33) = 3.82$, $MS_e = 42.405$, $p < .07$]. This is the standard phonological similarity decrement.

We also entered the data into a 2 × 2 analysis of variance, with span test (operation span/reading span from Experiment 2) as a between-subjects factor and list type (phonologically similar/dissimilar) as a within-subjects factor. The interaction was significant [$F(1,66) = 15.09$, $MS_e = 57.339$, $p < .001$]. Despite using the same words as the reading span test, the operation span test showed the standard phonological similarity decrement. Proportion correct on the math verification task was high and did not differ significantly between the similar (.97) and the dissimilar (.96) conditions ($F < 1$).

GENERAL DISCUSSION

In the first two experiments, for the word span test, the phonological similarity decrement was replicated (Conrad & Hull, 1964). The subjects performed better for phonologically dissimilar words than for phonologically similar words. However, for the reading span test, the opposite pattern was observed. The subjects performed better for sentences that ended with phonologically similar words, as compared with phonologically dissimilar words. In the reading span test, the subjects were not relying solely on a phonological code. The sentence contexts in the reading span test, coupled with the knowledge that the target words rhyme, provided retrieval cues that aided recall of the words.

This idea is consistent with the concept of long-term working memory (Ericsson & Kintsch, 1995). According to this view, retrieval cues are used to make information in long-term memory more available. In the present study, the reading span test is the only one that includes a retrieval cue that is semantically related to the target word. In the similar condition, this retrieval cue can be coupled with the knowledge that the words in the set rhyme. However, in the dissimilar condition, the rhyme cue is not available, leading to worse performance. For the word and operation span tests, although knowledge that the words rhymed was available, it was not sufficient to provide an effective cue.

Alternative explanations for the phonological similarity facilitation based on articulatory suppression (Murray, 1968) and delay (Nairne & Kelley, 1999) were explored in Experiment 3 by using an operation span test (Turner & Engle, 1989). Importantly, this test provided articulatory suppression and delay, but not contextual clues. Nonetheless, the standard phonological similarity decrement was observed. Thus, the context provided in the reading span test must have led to the phonological similarity facilitation.

REFERENCES

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, **55**, 429-432.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, **102**, 211-245.

Geisel, T. S. [Dr. Seuss] (1960). *Green eggs and ham*. New York: Random House.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1118-1133.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, **78**, 679-684.

Nairne, J. S., & Kelley, M. R. (1999). Reversing the phonological similarity effect. *Memory & Cognition*, **27**, 45-53.

Searleman, A., & Herrmann, D. (1994). *Memory from a broader perspective*. New York: McGraw-Hill.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, **28**, 127-154.

NOTE

1. The reversal of the phonological similarity decrement occurred for Nairne and Kelley (1999) when different items were used on every trial. The present experiments used this procedure as well.

(Manuscript received October 6, 2000;
revision accepted for publication April 10, 2001.)