

BRIEF REPORTS

List composition and the word length effect in immediate recall: A comparison of localist and globalist assumptions

NELSON COWAN

University of Missouri, Columbia, Missouri

ALAN D. BADDELEY

University of Bristol, Bristol, England

and

EMILY M. ELLIOTT and JENNIFER NORRIS

University of Missouri, Columbia, Missouri

Lists of short words usually are recalled better than lists of longer words in immediate recall tasks. Such word length effects might be explained by *localist* accounts, in which the length of each word in a list affects the recall of that word only, or by *globalist* accounts, in which the lengths of at least some words affect the recall of other words (e.g., Baddeley, 1986). In a recent localist account, Neath and Nairne (1995) proposed that the recall of each word depends on the likelihood that features within the word are contaminated within the memory representation. We tested this by presenting not only homogeneous lists of short and long words, but also mixed lists, and by including articulatory suppression on some trials. The short-word advantage depended on the composition of the list, ruling out a strictly localist approach. There appear to be several globalist influences on recall, including distinctiveness factors as well as phonological storage and articulation.

The aim of models of short-term memory is to describe and explain limits in the ability of humans to recall recently presented sets of stimuli. Baddeley, Thomson, and Buchanan (1975) contributed to a model of verbal short-term memory by pointing out the importance of word length. Lists composed of short words were recalled better than lists composed of longer words. This was true even when the shorter and the longer words had the same number of phonemes and syllables. Baddeley (1986) has described a *phonological loop* model in which the duration of the phonological information that could be held in an individual's short-term memory was said to approximate the amount that the individual could rehearse covertly in about 2 sec. This was supported by evidence that the amount recalled approximated the maximal amount that could be pronounced aloud in 2 sec (Baddeley et al., 1975). The theoretical reason for the

limit was, presumably, that information resided in a phonological buffer and could be refreshed in memory through rehearsal but would be lost from the buffer, becoming unavailable for recall on that trial, if it decayed away before it could be rehearsed.

The phonological loop model has been challenged in a number of ways. However, most of the alternative theoretical accounts that have been suggested do not necessarily challenge what we term the *globalist* assumption, according to which the recall of any particular word in a list depends on the list context—that is, on the set of words in which the word in question is embedded. In the phonological loop model, this occurs because rehearsal of a long word in the list delays the rehearsal of other words in the list and allows them to decay more than is the case when a short word is rehearsed. It may also occur because long words delay the overt recall of other words (Cowan et al., 1992; Cowan, Wood, Nugent, & Treisman, 1997) or because long words typically create more phonological interference than do short words (Caplan, Rochon, & Waters, 1992; Cowan et al., 1997), sometimes making it possible to minimize word length effects by controlling phonological complexity (Lovatt, Avons, & Masterson, 2000; Service, 1998).

This work was initiated while N.C. was supported by a Benjamin Meaker Visiting Professorship at the University of Bristol. The research was conducted with funding from NIH Grant HD-21338 to N.C. Correspondence concerning this article should be addressed to N. Cowan, Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, Columbia, MO 65211 (e-mail: cowann@missouri.edu).

An alternative type of model of the word length effect (Neath & Nairne, 1995) relies on what we call a *localist* assumption, according to which the recall of any one word depends not on the list context, but on the characteristics of the word itself. In the model of Neath and Nairne, the short-word advantage is assumed to be due not to the depletion of the available buffer space by long words, but rather to the poorer recallability of each long word per se. In their model of the word length effect, each word was considered to be, in effect, a miniature *list* of segments. Forgetting of part of this list of segments was said to jeopardize recall of the word. Long words contain more segments and, therefore, present more opportunities to forget enough phonemes so as to make the word unrecallable. Ordinarily, one might equate a segment with a phoneme, except for an assumption in the model that, when the duration of phonemes is increased, the number of segments per phoneme may increase also. (Neath and Nairne pointed out on p. 439 that “even if these segments are time based, this does not logically imply decay,” inasmuch as there can be degradation from interference that accumulates over time.) That characteristic allows the model to account for effects of both temporal duration and phonological length in recall. Cowan, Nugent, Elliott, and Geer (2000) demonstrated both types of effect; the number of syllables and the rate at which items were to be spoken both mattered, with advantages for words that were shorter in either way.

Models based on both the globalist and the localist assumptions make some similar predictions. Both types of models predict that, for lists of a fixed length, as the proportion of long words increases, the overall proportion correct should decrease. This occurs, in the globalist view of Baddeley (1986), because the more long words there are, the fewer the number of words that can be included in a time-limited, repeating rehearsal loop. It occurs, in the slightly different globalist view of Cowan et al. (1992), at least partly because the more long words there are, the more forgetting there is during spoken recall. It also occurs, though, in the localist view of Neath and Nairne (1995), because the more long words there are, the more the overall average across all words in the list is influenced by performance on those long words.

The globalist and the localist views differ in the manner in which this overall effect occurs, however. According to the particular globalist view based on the phonological loop model, performance levels on both long and short words should decrease as the proportion of long words in the list increases. According to globalist views more generally, the performance levels on short words, long words, or both must change *somehow* as a function of the proportion of long words in the list (e.g., it is possible that performance on long words, but not on short words, would decrease as a function of the proportion of long words in the list). In contrast, according to a strict localist view, performance on neither short words nor long words is influenced by the makeup of the list; the

overall average changes only because the proportion of long words contributing to that average changes.

We compared these views, using six-item, printed lists comprising between zero and six monosyllabic words and a complement of five-syllable words, arranged in a random order. Similar to Baddeley, Lewis, and Vallar (1984), the participants were to record only the first three letters of every word to help control output interference effects. As in most previous studies of short-term serial recall, small sets of items were used over and over in order to minimize the usefulness of unique lexical identities as cues to recall, making it more likely that phonological representations of the words and, hence, the phonological loop (Baddeley, 1986) would be necessary for recall. The main question was whether the mean performance on short and long words would be affected by the list composition, as is predicted by the globalist views, but not by the localist view. If so, another important question is which globalist view is supported by the data. We added the requirement of articulatory suppression to half of the trial blocks to help us distinguish between different storage mechanisms, inasmuch as articulatory suppression effectively blocks the contribution of covert articulatory processes in recall (e.g., Baddeley et al., 1984).

METHOD

Design

Every list was six words long. Each participant completed 98 list recall trials, including a block of 49 trials with articulatory suppression and a block of 49 trials without it, in a counterbalanced order. Each trial block included seven subblocks, each of which included 1 trial of each possible list composition, with these trials in a different random order for every participant. The seven list compositions included lists with zero, one, two, three, four, five, or six short words and a complement of long words in each case, with short and long words ordered randomly within the list.

Participants

The participants were 40 undergraduate students at the University of Missouri, who received course credit for their participation.

Apparatus and Stimuli

The study was conducted 1 participant at a time in a sound-attenuated room with a DOS-based program. Words appeared on the screen one at a time in a normal screen font. The short and long sets of words (all nouns) were designed to maximize the word length difference and to have as little phonological similarity between items as possible and were matched for word frequency across sets, using several different standard sources of word frequency information (Carroll, Davies, & Richman, 1971; Fearnley, 1997; Wilson, 1988) because no one set included all of the words. The short, monosyllabic words included *birch*, *clam*, *hoof*, *myth*, *pest*, and *trait*, and the long, five-syllable words included *administration*, *criminology*, *enthusiasm*, *hypochondriac*, *photosynthesis*, and *undergraduate*.

Procedure

The participant initiated each trial with a buttonpress. In the no-suppression condition, the phrase “Get ready” appeared at the center of the screen, and in the suppression condition, the phrase “Count from 1 to 8” appeared instead. After the beginning phrase had been

presented for 2.5 sec, it was replaced with the first list item. Each item stayed on the screen for 1.5 sec and was replaced by the next one until the final list item was replaced by a recall prompt ("?????") that stayed on the screen as the participant recorded a response on paper. The instruction was to write just the first three letters of each word in the preestablished slot on the paper, although some participants sometimes wrote in four letters. It was emphasized that the words were to be recorded in the same order as that in which they were presented. When the trial was completed, the participant pressed several computer keys in succession to proceed to the next trial.

Just before the articulatory suppression condition, the participants were trained to pronounce the numbers 1 through 8 repeatedly at the rapid rate of about 1.5 sec per cycle. The experimenter sat with the participant throughout the session and made sure that articulatory suppression was carried out adequately. Articulatory suppression was to be carried out from the moment of the presentation of the ready prompt to the moment of the presentation of the recall prompt, but not during the recall period. Previous research had shown that type of suppression to be sufficient to eliminate the standard word length effect for printed words (Baddeley et al., 1984).

RESULTS AND DISCUSSION

Means for every list composition for the no-suppression and the suppression blocks of trials are shown in Table 1. It should first be pointed out that the standard word length effects (Baddeley et al., 1975) were obtained. Examining only the homogeneous lists in a length \times suppression condition analysis, there was a main effect of articulatory suppression [$F(1,39) = 82.48$, $MS_e = 0.02$, $p < .001$] and of word length [$F(1,39) = 4.60$, $MS_e = 0.02$, $p < .04$], as well as an interaction between these factors [$F(1,39) = 20.36$, $MS_e = 0.01$, $p < .001$]. Separate analyses confirmed that there was a word length effect for the no-suppression condition, but not for the suppression condition.

If one examines average list performance across both short and long words within a list, there was an orderly, simple pattern of results. Not only was there a large effect of articulatory suppression [$F(1,39) = 90.47$, $MS_e = 0.06$, $p < .001$], but also an interaction of the suppression condition with the list composition [$F(6,234) = 4.71$, $MS_e = 0.01$, $p < .001$]. As is shown in Figure 1, in the no-suppression condition, increases in the number of long words in the list produced a nearly monotonic decrease in

overall performance on the list, whereas in the suppression condition, there was little change in overall performance as a function of list composition. Separate analyses for the two suppression conditions revealed a significant effect of list composition for the no-suppression condition only [no-suppression, $F(6,234) = 7.30$; suppression, $F(6,234) < 1$].

As was explained above, either globalist or localist theories of word length effects could account for the systematic change in performance as a function of list composition in the no-suppression condition, shown in Figure 1. It may be that either type of theory also could account for the suppression means shown in that figure (inasmuch as articulatory suppression can cause overwriting of features). Where the theories more clearly differ is in the prediction of the localist theory that the proportion correct for short words and long words, examined separately, should not change as a function of the list composition; by that theory, the change in overall performance can be attributed only to the change in the mixture of short and long words contributing to the overall average. Separate analyses of the list composition for short and long words disconfirmed that expectation of the localist view, however.

In the no-suppression condition, for the short words, there was an effect of list composition [$F(5,195) = 2.38$, $MS_e = 0.02$, $p < .04$; see Table 1]. The pattern was not monotonic, however, and post hoc Newman-Keuls tests failed to turn up any significant pairwise differences. For long words, there was a more systematic effect of list composition [$F(5,195) = 12.10$, $MS_e = 0.01$, $p < .001$]. According to Newman-Keuls tests, performance on long words within lists that had only one long word was superior to performance on long words within any other list composition. (In addition, performance on long words within lists with two long words was superior to that for lists with four long words; the other list compositions did not differ significantly.) Thus, there seems to be some sort of interference between long words presented within the same list. This might be attributed to participants' attempts to rehearse the long words selectively and to a limit in the capacity of the phonological loop to do so.

Table 1
Proportion Correct on Short and Long Words in Trials With Every Possible Word Length Mixture

Word Length	Number of Long Words Out of Six Words													
	0		1		2		3		4		5		6	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Lists With No Articulatory Suppression														
Short	.61	.03	.58	.03	.54	.03	.58	.03	.54	.04	.61	.04	—	—
Long	—	—	.66	.04	.57	.03	.52	.04	.48	.03	.52	.03	.50	.04
Word-length effect, <i>t</i> (39)	—		-3.09*		-1.06		2.76*		2.10*		2.35*		3.57*†	
Lists With Articulatory Suppression														
Short	.34	.02	.36	.02	.34	.02	.36	.02	.36	.03	.43	.04	—	—
Long	—	—	.36	.04	.34	.03	.34	.02	.31	.02	.33	.02	.35	.02
Word-length effect, <i>t</i> (39)	—		-0.09		0.27		0.86		1.74		2.79*		-0.93†	

Note—Positive *t* values indicate the standard word length effect (better performance on short than on long words); negative values indicate the opposite. * $p < .05$ or better, two-tailed test. †The last item in the row shows the between-list word length effect using homogeneous lists.

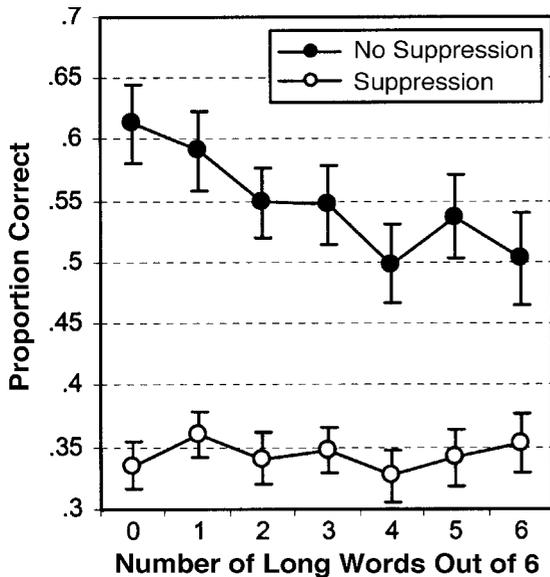


Figure 1. Mean proportion correct as a function of the list composition for trial blocks with an articulatory suppression task during list encoding (open points) and without articulatory suppression (solid points), collapsed across long and short words within a list. Error bars reflect standard errors.

For the mixed-length lists only, a list composition \times word length analysis was possible. The overall within-list word length effect, collapsed across all compositions of mixed lists, did not approach significance. The analysis did reveal not only a significant list composition effect [$F(4,156) = 6.87, MS_e = 0.02, p < .001$], but also a significant list composition \times word length interaction [$F(4,156) = 7.75, MS_e = 0.01, p < .001$]. As Table 1 shows, adding more long words to the list greatly decreased the recall of long words but did little to alter the recall of short words, in keeping with the suggestion that participants may have tried to use the phonological loop selectively for long words.

For the suppression condition, there was still an effect of list composition among the short words, examined separately [$F(5,195) = 2.37, MS_e = 0.02, p < .05$; see Table 1]. According to Newman-Keuls tests, performance was superior when there was only one short word than when there were six, four, or two short words, and no other conditions differed significantly. Thus, here, it appears to be the isolated short word among long words that stood out for recall. For long words in the suppression condition, there was no significant effect of list composition. For the mixed-length lists, a list composition \times word length analysis produced no significant effects.

Although these findings support a globalist interpretation, they also illustrate that the particular globalist interpretation that has become standard, the phonological loop theory (Baddeley, 1986), cannot account for performance unassisted by other factors. It cannot explain why an isolated short word among long words is recalled

better than short words within lists including other short words. It also has no basis for explaining word length effects for printed lists that occur in the presence of articulatory suppression during list presentation (Baddeley et al., 1984). This did occur, within lists with one short and five long words, as is illustrated by the t tests in Table 1.

Figure 2 clarifies the effects of word length in this experiment. It plots the within-list word length effect for each list composition in the no-suppression (top panel) and suppression (bottom panel) conditions. Positive numbers indicate a short-word advantage. The figure shows that, with no suppression, the word length effect depended on the list composition. With only one long item mixed with five short items, there was an advantage for the long item. As the number of long items increased, performance shifted steadily to a short-word advantage. The fact that the word length effect rises above the diagonal line of symmetry shows that there was a slight advantage for short words overall. The discrepancy from symmetry was greatest when there were three short and three long words, when a significant short-word advantage occurred according to a t test.

With suppression, as is shown in the bottom panel of Figure 2, long words no longer had an advantage over short words in any list composition, but short words sometimes still had an advantage. This could be explained on the grounds that, with articulation tied up by the suppression task, long words were not phonologically encoded in a manner sufficient to allow them to stand out well even if there were very few such words in the list.

The horizontal dashed lines in Figure 2 indicate the word length effect for homogeneous lists (which must be compared across trials). It appears that this between-list effect occurred on a basis different from the within-mixed-list effects. Thus, when there was articulatory suppression, the between-list effect was abolished, as in previous research (Baddeley et al., 1984), even though the within-list effect to some extent remained, as is shown in Table 1 and in the bottom panel of Figure 2. This remaining advantage for isolated short words fits the principle that isolates do tend to be recalled well (von Restorff, 1933). It also must be considered that visuo-spatial information storage apparently contributes to recall and is easier for short words. Thus, Chincotta, Underwood, Ghani, Papadopoulou, and Wresinki (1999) found an advantage for the recall of Arabic numerals (e.g., 5, 8, 3, etc.) over digit names (e.g., *five*, *eight*, *three*, etc.) that was not eliminated by articulatory suppression.

At first glance, one might attempt to account for performance in the no-suppression condition through a modification of the localist account in which a von Restorff effect is added. (Strictly speaking, that still would be a type of globalist account.) However, our data cannot be described well in this manner. As one can see in Table 1, as the number of long words in the list decreased from six to one, there was a systematic increase

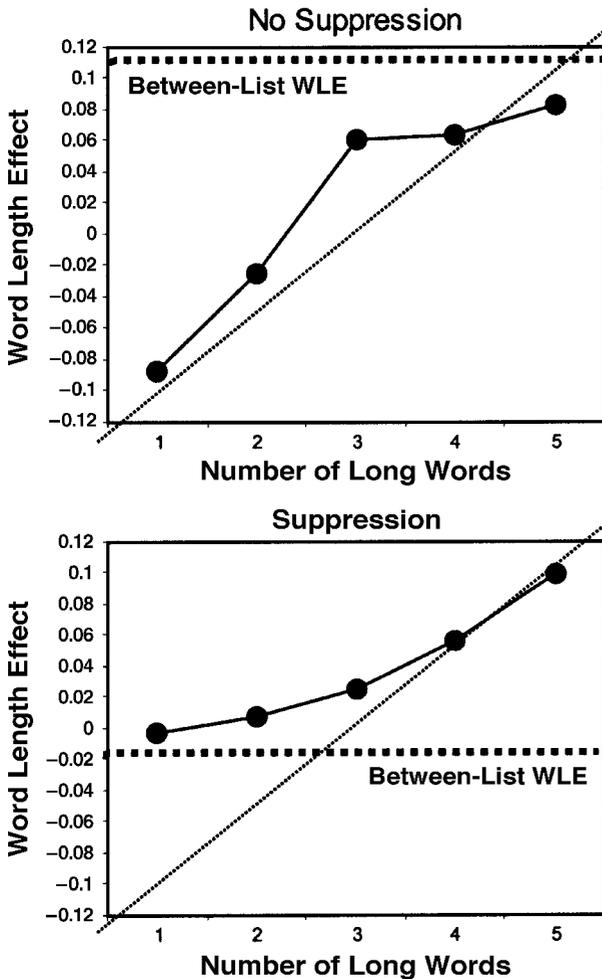


Figure 2. Magnitude of the word length effect (WLE) for trial blocks with no suppression (top panel) and with articulatory suppression (bottom panel) as a function of the list composition. Positive effects reflect an advantage for short words. In each panel, the diagonal line is a line of symmetry that approximates a pattern of WLEs that hypothetically might be obtained within mixed lists if short and long words formed two distinct sets of equal status; note that some of the actual WLEs favor short words more than would be indicated by this line of symmetry. The horizontal dashed line in each panel is the between-list WLE between homogeneous lists of short versus long words.

in the proportion of long words recalled (from .50 to .66, reading from right to left in the table), in keeping with a von Restorff effect. The linear trend was significantly different from 0. In contrast, as the number of short words in the list decreased from six to one (reading from left to right), there was no systematic change in the performance on short words; the linear trend was not significantly different from 0. The linear trends for short versus long words differed significantly [$F(1,39) = 13.13, p < .001$]. Although this pattern of responses could reflect a von Restorff effect restricted to the long words, that would be only a description, and not a theoretical explanation. Theoretically, in cases of mixed lists,

participants may use the rehearsal loop (when available) to retain long words and may be able to use some other means to keep the short words and their serial locations in mind. Therefore, only the recall of long words would suffer when their number within the list increased.

The present findings clearly support a globalist interpretation of word length effects, inasmuch as the direction, as well as the magnitude, of the effects depended on the list context. One way to account for the findings is on the basis of several principles in combination. (1) When a list is homogeneous, the phonological loop principle provides a good description of performance. (2) When a list includes both short and long words, organizational factors are important; recall tends to be better for words of the length that is underrepresented in the list. Within the mnemonic representation of the list, that underrepresented word length may stand out as a figure against the background of the other words. (3) Such organizational factors may operate in combination with the phonological loop, which may be used selectively to encode and rehearse the more difficult, longer words within mixed lists. Thus, the suppression task prevented the participants from demonstrating a long-word advantage regardless of the list composition, as is shown in Figure 2, presumably by preventing the application of covert articulation to the full encoding of such words. (4) Visuospatial storage may selectively preserve one or more short words under such circumstances.

An account such as the one we have offered fits into a recent tradition in which a phonological loop factor is combined with other factors in an account of immediate serial recall (e.g., Hulme, Maughan, & Brown, 1991). Within mixed lists, it appears that the distinction between short and long words can provide a basis that makes the less frequent type of word within the list stand out, in working memory, as a figure against the background of the more frequent type of word.

This notion, that a minority type of item in a mixed list forms a group that stands out, begs the question of where or how that privileged information is held between the time of presentation and the time of recall. Recently, two of the authors have discussed aspects of verbal short-term memory storage that are nonphonological in nature, are severely limited in capacity, and are likely to require attention for adequate storage—namely, a capacity-limited focus of attention (Cowan, 2001) or a temporary episodic buffer (Baddeley, 2000, 2001). It is possible that these aspects of working memory, dating back to earlier work such as that of Miller (1956), will come into play in combination with the phonological loop to provide a fuller account of immediate, serial recall of verbal materials. Participants may store some information in phonological form and some in a more conceptual form, combining the sources of information at the time of recall. The available sources of storage together may define the participant's capacity on a particular trial, whereas the breakdown of the list into subsets may help to determine which items come to predominate and occupy that capacity.

As an alternative theoretical approach to the one we have suggested, it is also possible that the general theoretical framework in which Neath and Nairne (1995) worked could be used to postdict the present results on the basis of some sort of principle other than the one in which each word is viewed as a miniature list to be recalled. Indeed, Neath and Nairne viewed their model as an extension of the feature model of Nairne (1990), and that model included what we would view as a type of buffer storage—namely, an overwriting process based on a temporary store of feature properties susceptible to such overwriting. It may be more immediately achievable to help determine what “flavor” of an investigator’s favored theoretical model would be needed to account for extant data, as opposed to a definitive determination of which model is best in absolute terms. Contrasting within- and between-trial effects, such as those of word syllabicity demonstrated here, can promote this refinement of models.

REFERENCES

- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- BADDELEY, A. [D.] (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-423.
- BADDELEY, A. [D.] (2001). The magic number and the episodic buffer. *Behavioral & Brain Sciences*, **24**, 117-118.
- BADDELEY, A. [D.], LEWIS, V., & VALLAR, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, **36A**, 233-252.
- BADDELEY, A. D., THOMSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 575-589.
- CAPLAN, D., ROCHON, E., & WATERS, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology*, **45A**, 177-192.
- CARROLL, J. B., DAVIES, P., & RICHMAN, B. (1971). *The American Heritage word frequency book*. Boston: Houghton Mifflin.
- CHINCOTTA, D., UNDERWOOD, G., GHANI, K. A., PAPADOPOULOU, E., & WRÉSINKI, M. (1999). Memory span for Arabic numerals and digit words: Evidence for a limited-capacity, visuo-spatial storage system. *Quarterly Journal of Experimental Psychology*, **52A**, 325-351.
- COWAN, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-185.
- COWAN, N., DAY, L., SAULTS, J. S., KELLER, T. A., JOHNSON, T., & FLORES, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory & Language*, **31**, 1-17.
- COWAN, N., NUGENT, L. D., ELLIOTT, E. M., & GEER, T. (2000). Is there a temporal basis of the word length effect? A response to Service (1998). *Quarterly Journal of Experimental Psychology*, **53A**, 647-660.
- COWAN, N., WOOD, N. L., NUGENT, L. D., & TREISMAN, M. (1997). There are two word length effects in verbal short-term memory: Opposed effects of duration and complexity. *Psychological Science*, **8**, 290-295.
- FEARNLEY, S. (1997). MRC Psycholinguistic Database search program. *Behavior Research Methods, Instruments, & Computers*, **29**, 291-295.
- HULME, C., MAUGHAN, S., & BROWN, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory & Language*, **30**, 685-701.
- LOVATT, P., AVONS, S. E., & MASTERSON, J. (2000). The word-length effect and disyllabic words. *Quarterly Journal of Experimental Psychology*, **53A**, 1-22.
- MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.
- NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.
- NEATH, I., & NAIRNE, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, **2**, 429-441.
- RESTORFF, H. VON (1933). Analyse von Vorgängen im Spurenfeld. I. Über die Wirkung von Bereichsbildung im Spurenfeld [On the effect of field formations in the trace field: I. Analysis of processes in the trace field]. *Psychologische Forschung*, **18**, 299-342.
- SERVICE, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *Quarterly Journal of Experimental Psychology*, **51A**, 283-304.
- WILSON, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, **20**, 6-10.

(Manuscript received December 19, 2000;
revision accepted for publication February 22, 2002.)