

A strict test of the phonological loop hypothesis with Libyan data

MUSTAFA F. A. SHEBANI
Al-Fateh University, Tripoli, Libya

FONS J. R. VAN DE VIJVER
Tilburg University, Tilburg, The Netherlands
and North-West University, Potchefstroom, South Africa

and

YPE H. POORTINGA
Tilburg University, Tilburg, The Netherlands
and Catholic University of Leuven, Leuven, Belgium

In Arabic, there are two ways, differing in length, of pronouncing each digit. This feature of word pairs that are conceptually identical but of different length allows for a stricter test of the phonological loop hypothesis than has been reported previously. Libyan schoolchildren, both boys and girls, of two grades took part in a quasi-experimental study in which shorter and longer digits and shorter and longer words were presented in recall and pronunciation tasks. The results confirmed the validity of the phonological loop model in that memory span was longer for shorter stimuli. Analysis of the data with structural equation modeling showed that only 17% of the individual-level variation in memory span could be explained by pronunciation speed. In an analysis of covariance, it was determined that pronunciation speed could account for a substantial proportion of the variation found across age (8-year-olds vs. 10-year-olds), stimulus modality (digits vs. words), and stimulus length (short vs. long), but not for gender differences.

Baddeley and Hitch (1974, 1977; see also Baddeley, 1997) developed a model of working memory that distinguished between a central executive, supported by two subsystems, one visual and one auditory or verbal. The latter subsystem has been called the *articulatory* or *phonological loop*. The phonological loop, the focus of the present study, is assumed to consist of two components, the phonological store and control processes. Since the present study focuses on the phonological loop as a subsystem, these components are not further considered or discussed separately here. Memory traces in the store decay rapidly unless refreshed by rehearsal. The notion of the phonological loop can help to explain a variety of findings about short-term memory span. First and foremost, memory span depends on word length; more short words than long words can be recalled after presentation (Baddeley, Thomson, & Buchanan, 1975). Recall is influenced by articulation rate, which is often operationalized as the speed of reading or pronunciation (Baddeley, 1997). The role of rehearsal has been demonstrated by the suppression of articulation (Gupta & MacWhinney, 1995).

A host of other findings have been interpreted as supporting the notion of a phonological loop. For example, phonological similarity affects speed and accuracy in mental calculations, suggesting temporary storage of addends (Noël, Désert, Aubrun, & Seron, 2001). Overall, the findings suggest that temporal factors are important in memory span; the phonological loop amounts to a temporary phonological store of speech-based material with a duration of approximately 1.5 sec.

The phonological loop model has also been applied to ontogenetic development (e.g., Case, Kurland, & Goldberg, 1982; Hitch, Halliday, Dodd, & Littler, 1989; Hulme, Thomson, Muir, & Lawrence, 1984; Nicolson, 1981; Schneider & Pressley, 1997). Younger children have slower speech rates and rehearse more slowly. With age, increase in speech rate implies faster rehearsal, leading to better recall. Schweickert and Boruff (1986; see also Schweickert, 1993) have found similar outcomes with a variety of different materials. Differences in immediate memory across individuals, different materials, ages and languages can be accounted for largely on the basis of differences in the rate at which items can be pronounced (Schweickert, 1993).

According to the phonological loop model, memory span should vary across cultures and languages according to articulation time for a given set of items. Ellis and Hennessey (1980) tested English and Welsh bilinguals on

We thank Jan de Kuijer for help with collection and processing of speech data. Correspondence should be addressed to F. J. R. van de Vijver, Department of Psychology, Tilburg University, 5000 LE Tilburg, The Netherlands (e-mail: fons.vandevijver@uvt.nl).

a memory span and reading speed task. Children were required to read aloud 20 lines of digits as fast as possible eight times, altering the language on each trial. The mean reading time for the digits was significantly longer in Welsh than in English, whereas, as would be predicted from the phonological loop hypothesis, memory span was larger for English digits than for Welsh digits. The correlation between reading speed and memory span was positive and significant.

The word length of Chinese numerals has been mentioned as a factor in the superior performance of Chinese children in arithmetic, compared with children in the USA. Stigler, Lee, and Stevenson (1986) administered a digit span test to children from Taiwan, Japan, and the United States in kindergarten, first grade, and fifth grade. The mean digit span for the Chinese children was at least a full digit more than that of the American children at all three age levels, and the difference between the Chinese and Japanese samples was even larger. Hoosain and Salili (1987) reported a faster articulation rate for Chinese digits than for English digits. Zhang and Simon (1985) found that bilingual Chinese scholars had better memory span of Chinese and English digits.

Naveh-Benjamin and Ayres (1986) conducted a study with adult English, Spanish, Arabic, and Hebrew speakers. Their ages were between 20 and 30 years. Participants were tested on digit span, speeded digit reading, and story reading speed. Average memory span in the Arab-, English-, Hebrew-, and Spanish-speaking samples was 5.77, 7.21, 6.51, and 6.37, respectively. Speeded digit reading resulted in averages of 370, 256, 309, and 287 msec per digit, respectively. These findings suggest a close relationship between memory span and reading speed. The normally paced story reading task showed results in the same direction as those obtained for digit span and speeded digit reading, with English the fastest and Arabic the slowest.

The evidence in favor of the phonological loop hypothesis is fairly consistent. Still, there is a methodological problem: It is difficult to test the hypothesis within a single individual. Most tests of the model are based on comparisons of age or cultural groups; comparability of participants on outcome-relevant factors (such as education), then, must be assumed. The use of longitudinal designs could help to partly solve these problems. Some studies have attempted to circumvent incomparability problems by using bilinguals (e.g., Cheung, Kemper, & Leung, 2000; Zhang & Simon, 1985). However, a similar level of mastery of both languages has to be assumed among the participants, which is rarely the case.

By studying the phonological loop hypothesis in Libya, we had the opportunity to do a "natural experiment" (e.g., Scheier, 1959) and bypass the methodological problems just described. In Arabic, there are two ways to pronounce digits. One is long, and the other is relatively short; the long forms have a syllable added to the short form. Primary school textbooks in arithmetic document both types. When Libyan children learn to count, they use long forms. In schools, Libyan children (as well as other Arabic-

speaking children) use both ways of pronunciation. In everyday mathematical tasks, such as counting in verbal games, Libyan children use both forms of number words. Long versions of digit names are more often used by the children when they recite text from the Koran. In general, both types are commonly used by schoolchildren.

The present study set out to examine memory span development in Libya (the data reported here are part of a larger cross-cultural project in which Dutch children are also involved). Two hypotheses were tested, both of which derived from the phonological loop model. First, the children were expected to show a longer memory span for the short form of the digits than for the long form of the digits. Second, it was expected that individual differences in short-term memory performance could be accounted for by differences in pronunciation speed.

METHOD

Participants

Sixty-four third- and fifth-grade children were recruited from two elementary schools in Garian, a semiurban area in western Libya. There were 14 boys and 18 girls of 8 or 9 years of age, and 16 boys and 16 girls of 10 years of age. All children spoke (only) their native language, which is Arabic. They came from middle-class families.

Materials and Procedures

Memory span. There were two subtests for memory span, one dealing with digits and one with words (non-numerals). The digits were further subdivided into tasks with short and long pronunciation. The word span test was also subdivided into a task with short words and a task with long words.

Digit span test. For each trial, digits were selected randomly from the set of 1 to 9. The trials varied in length from two to eight digits. The following restrictions applied: No digits were repeated in a trial, no successive numbers followed each other, and sequences of numbers with mnemonic cues (e.g., phonological similarity) were avoided. The digits were individually spoken into a tape recorder by a female teacher and were then digitized and stored. A trial was constructed by randomly selecting digits with the described restrictions and then copying the chosen set of digitized digits to an audiotape. The digits were presented at a rate of approximately one per second. The end of every trial was marked by a tone. The same set of trials was heard by each child (up to the stopping point), so the items composing a trial were randomly selected, but everyone heard the same tape.

Each child was tested individually. The child had to recall the digits in the same order as presented. A brief training session was given by presenting 3 trials with a length of three digits and 3 trials with four digits per trial.

Three trials of each length were presented from audiotape, beginning with two digits, followed by three digits, and so on. Presentation of trials continued as long as the child recalled correctly in serial order the digits of all 3 trials with a given length. If the child failed to recall the digits on all 3 trials, the test was terminated, and the score was equal to the number of digits at the previous level. If 1 of the 3 trials was correct, the test was stopped, and the score was the number of digits recalled correctly three times, plus a bonus of 0.5. With 2 out of 3 trials correct, the test was stopped, and the score was the number of digits in these trials.

Word span test. Two sets of words were assembled from large pools of words, derived from textbooks and reading books for children. The first set had three-syllable words, and the second set had four-syllable words. Both sets contained eight words. Libyan teach-

ers evaluated the words for their familiarity to children. The following short words were used: door (باب), cat (قط), sand (رمل), apples (تفاح), camel (جمل), beans (فول), elephant (فيل), and oven (فرن). The long words were: green beans (فاصوليا), honey duke (بطيخ), lamb (خروف), tomato (طماطم), book (كتاب), apricots (انجاص), olive tree (زيتون), and couscous (كسكس). These words were read aloud by a female teacher at a rate of approximately one word per second. Stimulus trials were prepared in the same way as described for the digits.

In the first of two sessions, the words were presented one by one, and the child was asked to repeat each word until this could be done without difficulty. In the second session, memory span was tested. Three trials were presented of a given length, beginning with two words. Each time the trial length increased, the child was told so. Stopping rules and scoring were the same as for digits.

In order to get a (conservative) estimate of the reliability of the span tests, correlations were computed between short and long versions of the same type of stimulus. The correlation was .48 for the digit tests and .61 for the word tests (both $ps < .001$). Somewhat comparable figures have been found for the test-retest stability of digit span subtests of both the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale (Wechsler, 1991).

Pronunciation speed. There were two tests for pronunciation speed, for digits and words, respectively. Both the digit and word tests consisted of a short and a long form.

Digit pronunciation speed. Pronunciation speed was operationalized here as the number of items a child could say in 1 sec. There were eight pairs of digits typed in medium-size fonts, each pair on a separate sheet. A tape recorder was used for recording.

The experimenter asked the child to read and repeat again and again, aloud and as quickly as possible, the two digits presented at a trial. As with the memory span test, counterbalancing was applied in order to compensate for order effects. Each pair of digits was to be repeated 10 times; the counting was done (silently) by the experimenter.

No instruction was given to the child about the number of repetitions. It should be noted that with this procedure, in which the child repeats from memory, individual differences in reading skills do not influence speed.¹

Word pronunciation speed. Pronunciation speed was operationalized here as the number of words a child could say in 1 sec. The same two sets of short and long words that were used in the memory span test were grouped in pairs, resulting in four short-word pairs and four long-word pairs. Each of these pairs of words was typed in medium-size fonts on a separate sheet. The administration procedure for this subtest was the same as for the digit pronunciation subtest.

The reliability estimates of the pronunciation measures were again based on the correlations of the short and long forms of a type of stimulus; the correlation was .69 for digits and .91 for words (both $ps < .001$).

RESULTS

Means and standard deviations obtained under the different conditions for both memory span and pronunciation speed are presented in Table 1. The differences for word length and age were all in the expected direction. The crucial difference in memory span between longer and shorter forms of the same digits clearly emerged [mean for short digits = 3.91; mean for long digits = 3.67, $t(63) = 3.21, SE = .07, p < .001$]; Cohen's (1988) effect size is .41, which is considered to constitute a small- to medium-size effect. The length of the digit words has an influence on recall, even if the meaning or conceptual identity is kept constant. This finding con-

Table 1
Means and Standard Deviations of Memory Span and Pronunciation Speed Tests for Age and Gender

| | 8 Years | | | | 10 Years | | | | Overall | |
|-------------------------------|---------|-----|-------|-----|----------|-----|-------|-----|---------|-----|
| | Boys | | Girls | | Boys | | Girls | | | |
| | M | SD | M | SD | M | SD | M | SD | M | SD |
| Memory Span | | | | | | | | | | |
| Digits | | | | | | | | | | |
| Short | 3.46 | .41 | 4.00 | .57 | 3.97 | .53 | 4.12 | .50 | 3.91 | .56 |
| Long | 3.39 | .63 | 3.67 | .51 | 3.66 | .60 | 3.97 | .53 | 3.67 | .59 |
| Words | | | | | | | | | | |
| Short | 3.25 | .47 | 3.44 | .45 | 3.44 | .63 | 3.56 | .63 | 3.43 | .55 |
| Long | 3.07 | .33 | 3.19 | .35 | 3.34 | .57 | 3.50 | .48 | 3.28 | .46 |
| Pronunciation Speed | | | | | | | | | | |
| Digits | | | | | | | | | | |
| Short | 2.17 | .32 | 2.13 | .51 | 2.63 | .36 | 2.44 | .44 | 2.34 | .46 |
| Long | 1.28 | .19 | 1.32 | .17 | 1.58 | .14 | 1.57 | .19 | 1.44 | .22 |
| Words | | | | | | | | | | |
| Short | 1.49 | .28 | 1.47 | .23 | 1.83 | .24 | 1.79 | .25 | 1.65 | .30 |
| Long | 1.30 | .20 | 1.31 | .19 | 1.60 | .23 | 1.58 | .25 | 1.45 | .26 |
| Mean Residual of Memory Span* | | | | | | | | | | |
| Digits | | | | | | | | | | |
| Short | -.32 | .39 | .18 | .56 | -.04 | .52 | .12 | .51 | .00 | .53 |
| Long | -.19 | .59 | .02 | .51 | -.09 | .60 | .23 | .54 | .00 | .57 |
| Words | | | | | | | | | | |
| Short | -.10 | .47 | .11 | .41 | -.11 | .59 | .08 | .61 | .00 | .52 |
| Long | -.14 | .34 | -.02 | .39 | -.03 | .53 | .17 | .45 | .00 | .44 |

*Mean residual of a cell (after correction for the pronunciation measures); within each row, the (weighted) residuals have an average of zero.

firms expectations derived from the phonological loop model and corroborates the first hypothesis.

The collected data enable us to estimate the difference in time needed to pronounce short and long digits. The difference was found to be substantial; the participants took on average 448 msec ($SD = 110$) to pronounce a digit in its short form and 715 msec ($SD = 121$) to pronounce a long-form digit. This difference is highly significant [$t(63) = -20.96, SE = 12.89, p < .001$].

Differences were further analyzed by means of analysis of variance (ANOVA). A first analysis (repeated measures) was carried out on the memory span scores. The between-subjects variables were age (8 vs. 10 years) and gender (boys vs. girls), whereas word length (short vs. long) and type of stimulus (digits vs. words) were the within-subjects variables. The dependent variables were the digit span score for short digits, long digits, short words, and long words. The F values with significance levels and effect size (η^2 , the proportion of the total variance accounted for by the effect) are presented in the left part of Table 2. The age effect was significant, indicating that older children had a longer memory span than did younger children ($M = 3.70$ and $M = 3.45$, respectively). Gender also showed a significant effect, with girls scoring higher than boys ($M = 3.67$ and $M = 3.46$, respectively). Stimulus length was highly significant; shorter stimuli were easier to remember than were longer stimuli ($M = 3.67$ and $M = 3.48$, respectively). Digits were easier to remember than were words ($M = 3.79$ and $M = 3.36$, respectively). Since the latter effect is mainly due to the choice of stimuli, it is irrelevant for examining the phonological loop hypothesis and is not further discussed here. Finally, none of the interaction effects were significant.

The same repeated measures design was used to analyze pronunciation speed scores. Differences between boys and girls were not significant [$F(1,60) = 0.35, MS_e = 0.21, \eta^2 = .01$]; grade differences were highly significant [$F(1,60) = 31.26, MS_e = 0.21, p < .001, \eta^2 = .34$]. Finally, the interaction of grade and gender was not significant [$F(1,60) = 0.32, MS_e = 0.21, \eta^2 = .01$].

Structural Equation Modeling

The relationship between pronunciation speed and memory span measures was examined using structural equation modeling. The model involved all eight tests. The covariance matrix to be analyzed was based on deviance scores, because significant effects had been found for age and gender, and the data set was too small for separate analyses per subsample. Theoretically, the speed measures can be taken to compose a single latent speed factor, whereas the span measures define a (latent) memory span factor. According to the phonological loop model, the speed factor should impact on the memory span factor. In terms of a structural equation model, the theoretically assumed structure implies that the pronunciation speed measures are the observed input variables that together compose a latent factor; similarly, a latent memory span factor is defined by the observed output measures. Finally, there is a path from the speed factor to the span factor.

An inspection of the skewness and kurtosis showed that only the span test for long words showed a significant value; the test showed a negative skewness of -0.95 ($p < .001$). The results are presented in Figure 1. A good model fit was found [$\chi^2(19, N = 64) = 28.18, p = .08; \chi^2/df = 1.48, RMSEA = 0.09$ (90% interval: $0.00 - 0.15$), CFI = 0.97, NNFI (also known as TLI) = 0.96]. As could be expected, all speed measures showed a positive and significant loading on the latent pronunciation speed factor ($p < .01$). Similarly, all span measures showed a positive and significant loading on the memory span factor ($p < .01$). The (standardized) path coefficient of pronunciation speed on memory span was also significant ($\gamma = 0.42, p < .01$). Pronunciation speed explained 17% of the variance of memory span, which is a moderate but significant value.

The analyses underscored the feasibility of the phonological loop model. Both the pronunciation speed and the memory span measures defined their own underlying factors, as assumed in the phonological loop model, and these two factors showed a significant relationship. Children who have a higher pronunciation speed tend to show a longer memory span.

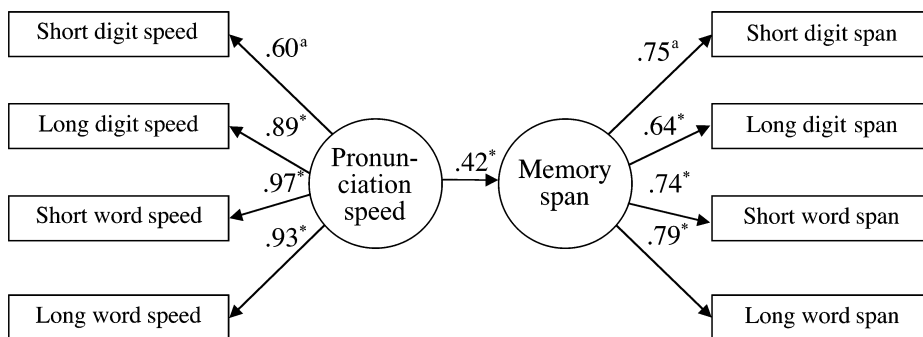


Figure 1. Structural equation model linking pronunciation speed and memory span (standardized solution). ^aParameter fixated in the nonstandardized solution at a value of 1 to set the scale of the latent variable. * $p < .01$.

Table 2
Analysis of Variance (Left) and Covariance (Right) of Memory Span With Pronunciation Speed as Covariate: F Ratios, Their Significance, and Effect Sizes (η^2)

| Source | Before Correction for Pronunciation Speed | | After Correction for Pronunciation Speed | |
|-------------------------|---|----------|--|----------|
| | <i>F</i> (1,60) | η^2 | <i>F</i> (1,56) | η^2 |
| Between Subjects | | | | |
| Age | 7.32** | .11 | 1.87 | .03 |
| Gender | 5.64* | .09 | 6.69* | .11 |
| A × G | 0.20 | .00 | 0.08 | .00 |
| <i>MS_e</i> | 0.60 | | 0.59 | |
| Within Subjects | | | | |
| Stimulus length | 19.30*** | .24 | 0.86 | .02 |
| A × L | 0.49 | .01 | 2.22 | .04 |
| G × L | 0.26 | .00 | 0.36 | .01 |
| A × G × L | 2.63 | .04 | 2.47 | .04 |
| Type of stimulus | 56.73*** | .49 | 1.49 | .03 |
| A × T | 0.56 | .01 | 0.62 | .01 |
| G × T | 2.07 | .03 | 1.24 | .02 |
| L × T | 0.61 | .01 | 0.24 | .00 |
| A × G × T | 0.37 | .01 | 0.38 | .01 |
| A × L × T | 0.58 | .01 | 0.59 | .01 |
| G × L × T | 0.05 | .00 | 0.01 | .00 |
| A × G × L × T | 0.72 | .01 | 0.79 | .01 |
| <i>MS_e</i> | 0.16 | | 0.17 | |

* $p < .05$. ** $p < .01$. *** $p < .001$. A, age; G, gender; L, stimulus length; T, type of stimulus.

It could be argued that the current sample size is too small to warrant structural equation modeling. However, the ratio of cases to observed variables is 8:1, which is not inappropriate; furthermore, the fit measures that are sensitive for sample size, such as the chi-square statistics, yield the same conclusions as the other fit measures, which underlines the suitability of the analysis.

Analysis of Covariance

The present data also provide an opportunity to test whether pronunciation speed explains age and gender differences in memory span, as implicitly claimed by the phonological model (cf. Schweickert, 1993). Such claims seem to run counter to our finding in the previous analyses that only 17% of the variation in memory span could be accounted for by pronunciation speed.

A repeated measures analysis of covariance (ANCOVA) was carried out, with the four memory span measures as the dependent variables, constituting two dichotomous within-subjects factors: type of stimulus (digits vs. words) and stimulus length (short vs. long). The four pronunciation speed measures (i.e., short digits, long digits, short words, and long words) were the "time-dependent" covariates; like in the previous ANOVA, age and gender were the (between-subjects) independent variables. Effect sizes are presented in Table 2 (right side). Only main effects are discussed, since none of the interactions were significant. The influence of pronunciation speed on age differences was substantial, reducing the initial effect size from .11 to a nonsignificant value of .03 after correction. However, a different pattern emerged for gender differences. The effect size went up slightly from .09

($p < .05$) to .11 ($p < .05$). An inspection of Table 1 shows how this finding came about. Whereas girls tended to score somewhat better on memory span, pronunciation speed differences showed a small difference in favor of boys. This pattern is confirmed in the lower part of Table 1, where residual scores are presented (defined as the average residual score per cell after correction for the pronunciation measures). Boys showed a negative mean residual score (indicating that their obtained score is smaller than their expected score), whereas girls showed the opposite pattern.

The most salient result was obtained for both within-subjects factors. The effect size of stimulus type (digits vs. words) was reduced from .49 to .03 after correction for pronunciation speed; similarly, the effect size of stimulus length went down from .24 to .02. The large reduction of both effect sizes can be seen as strong evidence in favor of the phonological loop hypothesis.

DISCUSSION

The main interest of this study was to test the word-length effect using short and long digits, and short and long words. In particular, the difference in performance on short and long digits was relevant, because the presence of short and long forms of the digits in Arabic provided the conditions of a "natural experiment" to test the phonological loop model in a within-subjects design with stimulus materials of identical meaning but of different word length. The data provided support for the model; Libyan children showed a longer memory span for short digits than for long digits. Other factors were also supportive of Baddeley's model (Baddeley, 1997; Baddeley & Hitch, 1974). In a structural equation model, memory span constituted a single underlying factor, with four span tests (short digit span, long digit span, short word span, and long word span) as indicators; similarly, pronunciation speed constituted a single factor, with four pronunciation measures (short digit speed, long digit speed, short word speed, and long word speed) as indicators. As would be predicted from Baddeley's model, pronunciation speed was a significant predictor of memory span, although the proportion of explained variance was modest (17%). A final line of support for Baddeley's model came from an ANCOVA in which it was found that pronunciation time differences could successfully explain both between-subjects span differences (i.e., between 8-year-olds and 10-year-olds) and within-subjects differences (between words and digits, as well as between short and long stimuli).

These results were in line with the phonological loop model, replicating and extending previous findings. The effects of word length can be explained by the greater articulation time needed to pronounce long words in the phonological store (Broadbent, 1984; Hitch, Halliday, Schaafstal, & Heffernan, 1991; Hulme, Silvester, Smith, & Muir, 1986). The longer span for words than for digits could be accounted for by the shorter pronunciation

time of the latter (Landauer, 1962). The developmental increase in memory span between the ages of 8 and 10 follows a pattern similar to that found in previous investigations (Hitch & Halliday, 1983; Hulme et al., 1984; Nicolson, 1981), with a clear increase in memory span across the two age groups.

Although the current data provide impressive support for Baddeley's model (Baddeley, 1997; Baddeley & Hitch, 1974), three unexpected findings emerged. First, the structural equation model yielded a modest 17% of explained variance, demonstrating that important sources of individual span differences are unaccounted for by pronunciation time. It is unlikely that the relatively low reliability estimates could account for this modest percentage. If the four subtests making up the speed factor are treated as separate items, the internal consistency of the speed factor is .86, and the value for the span factor is .81. After correction for attenuation, the correlation between speed and span becomes .50 (the unattenuated correlation is .42). The proportion of explained variance increases then from 17% (unattenuated) to 25% (attenuated). There is no reason to expect that with higher reliability coefficients, the percentage of explained variance would have been substantially higher. In most tests of Baddeley's model, there has been a close similarity between the stimuli employed to measure reading or pronunciation speed and memory span. The relationship between speed and memory span can be boosted by stimulus particulars that are shared by their respective measures. The similarity of stimuli in speed and memory span measures could induce mode-specific strategies. A similar argument has been developed by Gupta and MacWhinney (1995), who have argued that the control processes in Baddeley's model are presumably both articulatory and auditory; the first is close to Baddeley's model and involves "the refreshing of phonological representations by an output phonological process" (p. 67), whereas the second involves the "replaying of auditory images" (p. 67) and is nonarticulatory. In this line of reasoning, a measure of speed pronunciation could only partially explain individual differences in working memory performance.

Second, the short digit span was overestimated by the pronunciation speed for short digits, whereas all other span scores were underestimated. Moreover, the difference in memory span for short and long stimuli is smaller than would be expected on the basis of the speed pronunciation scores. Mode-specific strategies as discussed in the previous paragraph could also play a role here. Our data would then suggest that nonverbal strategies (perhaps visualization) that are less time dependent than rehearsal become more prominent in the storage of longer stimuli.

Third, Baddeley's model cannot account for the observed gender differences, which were not the same for the speed and span measures. Compared with girls, boys may have shown a higher motivation on the relatively active speed pronunciation test than on the more passive memory tasks. It is also possible that the often-observed slightly higher verbal skills of girls (as compared with

boys) only involve the mental processing of stimuli, whereas pronunciation speed is influenced by noncognitive aspects, such as the level of advancement of the speech muscles.

CONCLUSION

Our study has provided support for Baddeley's model (Baddeley, 1997; Baddeley & Hitch, 1974) in what appears to be the strictest test so far—namely, a within-subjects test with familiar stimulus materials that have the same meaning and only differ in the crucial variable of word length. The model could explain differences in performance across age, type of stimulus (digits vs. words), and stimulus length (short vs. long) but could not account for gender differences. It can be concluded that our data support a weak version of the word-length effect, according to which memory performance is influenced by pronunciation speed, but a strong version, according to which memory performance is fully determined by pronunciation speed, is not supported. Gender and task particulars are among the moderators with a bearing on memory span that are not completely accounted for in the phonological loop hypothesis.

REFERENCES

- BADDELEY, A. [D.] (1997). *Human memory: Theory and practice* (rev. ed.). Mahwah, NJ: Erlbaum.
- BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 742-775). New York: Academic Press.
- BADDELEY, A. D., & HITCH, G. J. (1977). Recency re-examined. In S. Dornic (Ed.), *Attention and performance VI* (pp. 647-667). Hillsdale, NJ: Erlbaum.
- 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.
- BROADBENT, D. E. (1984). The Maltese Cross: A new simplistic model for memory. *Behavioral & Brain Sciences*, **7**, 55-94.
- CASE, R., KURLAND, D. M., & GOLDBERG, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, **33**, 386-404.
- CHEUNG, H., KEMPER, S., & LEUNG, E. (2000). A phonological account for the cross-language variation in working memory processing. *Psychological Record*, **50**, 373-386.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- ELLIS, N. C., & HENNELLY, R. A. (1980). Bilingual word length effect: Implication for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, **71**, 43-51.
- GUPTA, P., & MACWHINNEY, B. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory & Language*, **34**, 63-88.
- HITCH, G. J., & HALLIDAY, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 325-340.
- HITCH, G. J., HALLIDAY, M. S., DODD, A., & LITTLER, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, **7**, 347-362.
- HITCH, G. J., HALLIDAY, M. S., SCHAAFSTAL, A. M., & HEFFERNAN, T. M. (1991). Speech, "inner speech" and the development of short-term memory: Effects of picture labelling on recall. *Journal of Experimental Child Psychology*, **51**, 220-234.

- HOOSAIN, R., & SALILI, F. (1987). Language differences in pronunciation speed for numbers, digit span, and mathematical ability. *Psychologia: An International Journal of Psychology in the Orient*, *30*, 34-38.
- HULME, C., SILVESTER, J., SMITH, S., & MUIR, C. (1986). The effects of word length on memory for pictures: Evidence for speech coding in young children. *Journal of Experimental Child Psychology*, *41*, 61-75.
- HULME, C., THOMSON, N., MUIR, C., & LAWRENCE, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, *38*, 241-253.
- LANDAUER, T. K. (1962). Rate of implicit speech. *Perceptual & Motor Skills*, *15*, 646.
- NAVEH-BENJAMIN, M., & AYRES, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology*, *38A*, 739-751.
- NICOLSON, R. (1981). The relationship between memory span and processing speed. In M. P. Friedman, J. P. Das, & N. O'Conner (Eds.), *Intelligence and learning* (pp. 179-184). New York: Plenum.
- NOËL, M. P., DÉSSERT, M., AUBRUN, A., & SERON, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory & Cognition*, *29*, 34-42.
- SCHEIER, I. H. (1959). The method of natural variation in applied educational research. *Journal of Educational Research*, *52*, 167-170.
- SCHNEIDER, W., & PRESSLEY, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ: Erlbaum.
- SCHWEICKERT, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, *21*, 168-175.
- SCHWEICKERT, R., & BORUFF, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory & Cognition*, *12*, 419-425.
- STIGLER, J. W., LEE, S. W., & STEVENSON, H. W. (1986). Digit memory span in Chinese and English: Evidence for a temporary limited store. *Cognition*, *23*, 1-20.
- WECHSLER, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: Psychological Corporation.
- ZHANG, G., & SIMON, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory & Cognition*, *13*, 193-201.

NOTE

1. The children's recorded performance was digitized. Subsequently, individual trials were edited and converted to a graphical representation that could be examined visually. The mean pronunciation speed per trial was determined and copied to an SPSS database for further statistical analysis.

(Manuscript received May 14, 2003;
revision accepted for publication June 3, 2004.)