

Errata: The third sentence of the abstract should read "Experiment 1, using ASL stimuli, found evidence for manual motoric coding (worse recall under articulatory suppression), and replicated findings of ASL-based phonological coding (worse recall for phonologically similar lists)."

Memory & Cognition
1997, 25 (3), 313-320

A visuospatial "phonological loop" in working memory: Evidence from American Sign Language

MARGARET WILSON and KAREN EMMOREY
The Salk Institute for Biological Studies, La Jolla, California

In two experiments, the question of whether working memory could support an articulatory rehearsal loop in the visuospatial domain was investigated. Deaf subjects fluent in American Sign Language (ASL) were tested on immediate serial recall. In Experiment 1, using ASL stimuli, evidence for manual motoric coding (worse recall under articulatory suppression) was found, replicating findings of ASL-based phonological coding (worse recall for phonologically similar lists). The two effects did not interact, suggesting separate components which both contribute to performance. Stimuli in Experiment 2 were namable pictures, which had to be recoded for ASL-based rehearsal to occur. Under these conditions, articulatory suppression eliminated the phonological similarity effect. Thus, an articulatory process seems to be used in translating pictures into a phonological code for memory maintenance. These results indicate a configuration of components similar to the phonological loop for speech, suggesting that working memory can develop a language-based rehearsal loop in the visuospatial modality.

Research on working memory has focused on two domains: verbal and visuospatial. The functioning of the verbal domain has been well mapped. A large body of research shows that representations of speech phonology and mechanisms of speech output form a partnership, called the "phonological loop," for maintaining verbal information (Baddeley, 1986; Gathercole & Baddeley, 1993). In contrast, the visuospatial domain of working memory has been less tractable. The relationship between visual and spatial working memory, the relationship between sensory storage and motoric processes, and the mechanisms by which retention or "rehearsal" might occur within visuospatial working memory are issues that remain unresolved. Thus, unlike the phonological loop, the "visuospatial scratch pad" has failed to reveal a clear structure. Why the disparity between these two domains?

One possibility is that the rehearsal-loop structure in verbal working memory is an inherent property of the auditory and vocal mechanisms in which speech is grounded. This could occur because of the relationship of these modalities to language in our evolutionary history or because auditory processing might lend itself to informa-

tion maintenance in ways that visual processing does not. On either of these accounts, differences between phonological working memory and visuospatial working memory are built into the system. A different possibility is that the phonological loop owes its structure, at least in part, to the impact of spoken language on the development of the individual. On this account, the phonological loop is shaped by a particular form of expertise.

To address this issue, we ask whether the visuospatial domain of working memory can support a language-based rehearsal-loop similar to the phonological loop. We do this by studying working memory in subjects who use a visuospatial language, American Sign Language (ASL). ASL is a natural human language possessing its own phonology (Coulter, 1993), morphology (Padden, 1983/1988), and syntax (Lillo-Martin, 1991). Like all languages, it is an abstract symbolic system capable of supporting the full range of linguistic expression of thought. However, it differs from spoken languages in the modality of transmission. It is produced with the hands and face rather than the vocal apparatus and is perceived through the eyes rather than the ears. In addition, it uses spatial relationships for conveying referential and semantic information (Emmorey, Corina, & Bellugi, 1995; Klima & Bellugi, 1979).

The present study asks whether ASL signers possess a visuospatial rehearsal loop that exhibits some or all of the structural properties of the phonological loop. This approach allows us to investigate which properties of working-memory subsystems are dependent upon the subsystem's relationship to language and which are more fundamental properties of the sensorimotor modalities involved.

The Phonological Loop

Verbal working memory in hearing subjects consists of a highly efficient rehearsal loop with two components:

This work was supported by NIH Grant DC-00128 awarded to M.W. and NIH Grant HD-13249 awarded to K.E. and Ursula Bellugi, as well as NIH Grants DC-00146, DC-0021, and HD-26022. We would like to thank Bonita Ewan, Steven McCullough, and Kathleen Say for their assistance in developing these experiments; Dennis Galvan, Bonita Ewan, and Steven McCullough for their assistance in obtaining and testing subjects; and Kevin Clark, Susan Gathercole, Steven McCullough, Michael Masson, Colleen Lee Smith, and one anonymous reviewer for their valuable suggestions and feedback. We are especially grateful to Gallaudet University, Washington D.C., California State University, Northridge, and the deaf subjects who participated in these studies. Correspondence should be addressed to M. Wilson, The Salk Institute for Biological Studies, 10010 North Torrey Pines Rd., La Jolla, CA 92037 (e-mail: wilson@crl.ucsd.edu).

a phonological storage buffer and an articulatory rehearsal process. The material in the buffer fades rapidly and must be repeatedly refreshed. This refreshing is done by the articulatory process. The evidence for the separation of these two components, as well as for their close interdependence, comes from a variety of experimental effects, including the *phonological similarity effect*, the *word-length effect*, and the effects of *articulatory suppression* and *irrelevant speech* (see Baddeley, 1986; Baddeley & Hitch, 1994; Gathercole & Baddeley, 1993, for reviews). The present study focused on phonological similarity and articulatory suppression, which interact in characteristic ways to reveal the structure of the phonological loop.

The phonological similarity effect occurs when items to be held in working memory contain many of the same phonemes. Such items yield worse memory performance than do dissimilar items (e.g., *mad, man, cad, mat, cap* vs. *pit, day, cow, sup, bar*). This effect shows phonological coding in working memory (e.g., Baddeley, 1966; Conrad & Hull, 1964).

Articulatory suppression occurs when subjects occupy their vocal apparatuses with an irrelevant activity, such as repeating "ta ta ta," during presentation or retention. This manipulation reduces memory span (e.g., Murray, 1968). The effect does not appear to be due to a general distraction or attention-load effect, but is instead a specific form of interference in the verbal domain (e.g., Brandimonte, Hitch, & Bishop, 1992; Logie & Salway, 1990; Smith, Reisberg, & Wilson, 1992). It appears that competing motor activity prevents the use of an articulatory strategy which normally assists memory performance. (We should note that this articulatory strategy need not involve actual execution of a motor program in the form of subvocalization. Instead, it is probably a more abstract level of representation, possibly of the early stages of motoric planning. See Baddeley & Wilson, 1985; Bishop & Robson, 1989; Vallar & Cappa, 1987.)

The relationship between phonological similarity and articulatory suppression is also revealing. When the stimuli are presented auditorily, the two effects do not interact. Specifically, suppression does not reduce the similarity effect, as one would expect if the two effects stemmed from the same process or representation. This indicates that two independent mechanisms are operating: a phonological store, which is responsible for the similarity effect, and an articulatory mechanism, which is responsible for the suppression effect. When materials are presented in auditory form, they gain automatic access to the phonological store and the similarity effect occurs whether or not articulatory rehearsal is available (Baddeley, Lewis, & Vallar, 1984).

However, the interdependence of these two mechanisms can be seen when stimuli are presented visually, as printed words or as namable pictures. Under these conditions, the similarity effect does occur, indicating that the visual materials are being phonologically recoded. Articulatory suppression, however, abolishes the similarity effect (e.g., Baddeley et al., 1984; Murray, 1968). Apparently, the ar-

tulatory mechanism is needed to recode visual materials into phonological form. Thus, this interaction provides evidence for the interdependence of the phonological buffer and the articulatory mechanism. The articulatory process can load materials into the buffer, thus providing a mechanism by which materials may be refreshed in a repetitive rehearsal process.

Visuospatial Working Memory

Can similar effects be found for visuospatial representations within working memory? Studies have found that spatial manual movements can disrupt visuospatial working memory (e.g., Baddeley, Grant, Wight, & Thomson, 1975; Quinn & Ralston, 1986; Smyth, Pearson, & Pendleton, 1988). In addition, studies have found effects of visual similarity and of irrelevant visual input on memory for nonnamable visual displays (e.g., Frick, 1988; Logie & Marchetti, 1991; Walker, Hitch, & Duroe, 1993). Thus, visuospatial working memory appears to rely, at least in some conditions, on sensory coding, and, at least in some conditions, on motoric processes. It is unclear, however, whether these mechanisms come together to form any kind of a rehearsal loop.

One reason for the confusion is that the relationship between visual and spatial working memory is unclear. On the one hand, there is evidence for a dissociation between visual and spatial working memory (Baddeley & Lieberman, 1980; Logie & Marchetti, 1991; Tresch, Sinnamon, & Seamon, 1993). On the other hand, visual and spatial working memory appear to interact. For instance, visual input can disrupt spatial working memory (Smyth & Scholey, 1994). However, the nature of this interdependence is unclear.

Another difficulty in drawing analogies between visuospatial and phonological working memory is that the two domains appear to differ in their relationship to executive control. Visuospatial memory tasks appear to involve more "central executive" resources than do verbal memory tasks (see Logie, 1995, for a review). Logie notes that visuospatial tasks appear to be particularly vulnerable to disruption during encoding, and argues that encoding is a function of the central executive.

A further difficulty is that it is unclear how visual representations may be maintained or "rehearsed" in working memory. Some authors have suggested that materials may be refreshed by generating representations of movements between locations, either by "sketching" materials motorically (Logie, 1995; Reisberg & Logie, 1993) or by scanning with an attentional mechanism (Smyth & Scholey, 1994). However, these proposals seem most applicable to the maintenance of spatial information, such as a sequence of locations. It is unclear whether such a mechanism could generate visual representations of complex pictures or shapes.

These various considerations seem to suggest that visuospatial working memory is not analogous to the phonological loop, and does not possess the same rehearsal structure. But, in fact, this appearance of dissimilar structure may be due to the lack of a clear relation-

ship between input and output representations in the visuospatial *materials* that have been studied. One of the striking features of phonological working memory is the direct mapping between articulated phonemes and heard phonemes. This mapping is characteristic of language, which provides an extremely overrehearsed "vocabulary" of sound representations, where the identity relationship between the heard form and the articulated form must be understood by both speaker and listener. Furthermore, the process of articulation results in immediate feedback in the form of auditory input, which serves to reinforce the identity relationship. This automatic mapping between input and output representations may be the basis of the partnership in which articulatory processes can be used to generate a phonological representation (cf. Reisberg & Logie, 1993, p. 58).

In contrast, it is unclear how one could "rehearse" most nonlinguistic visual materials. There is no obvious motoric representation that corresponds to the visual representation of a picture, leaving aside verbal labeling. Motor sequences that can actually be used to generate external pictorial materials (e.g., sketching) are time-consuming and usually not automatized. There is generally no standard or canonical format for how such sequences should be carried out. The difficulty, in short, is that there is no systematic relationship between input and output for such stimuli.

Furthermore, this difficulty with representing visuospatial materials in a form that can be rehearsed may explain the heavy executive load involved in the encoding phase with such materials. Subjects may be struggling during this phase to encode the material in a form that can be maintained and retrieved. This may lead to heavy use of the central executive, and also to wide variation in strategy between individuals. In contrast, if there is an obvious and overlearned means of encoding material for rehearsal, then the encoding phase may be effortless and automatic.

If, indeed, an overlearned representational system such as a language is critical for the emergence of a rehearsal loop structure, then, given the right conditions, the visuospatial domain of working memory may be able to support a rehearsal loop. A visuospatial language, such as ASL, might provide the necessary basis to allow such a visuospatial rehearsal loop to develop.

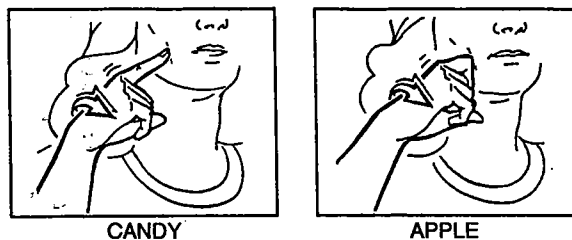
Sign Language and Working Memory

Could there be a "phonological loop" for sign language? Linguistic research over the past 30 years has revealed that sign languages contain sublexical structure analogous in many ways to spoken language phonology (Stokoe, 1960/1978; see papers in Coulter, 1993). This sublexical structure can be accounted for by using the same phonological constructs proposed for spoken languages (e.g., syllables, underspecified representations, feature geometry), and similar phonological processes apply to sign and spoken languages (e.g., assimilation, coarticulation, sonority; see Corina & Sandler, 1993; Sandler, 1995, for reviews). Given these findings, linguists

have broadened the term *phonology* to encompass the sublexical structure of both spoken and sign languages (Coulter & Anderson, 1993). Signs are classically analyzed into four basic phonological components: handshape, location, movement, and palm orientation (see Figure 1 for some examples of minimally contrasting sign pairs). These components are not meaningful in themselves but are combined in rule-governed ways to create signs—just as consonants and vowels are combined to create spoken words.

Clear evidence already exists that deaf subjects can and do use the phonology of ASL for coding materials in working-memory tasks. These phonological effects are primarily based on handshape. Studies have shown reduced span with handshape-similar materials (Hanson, 1982; Klima & Bellugi, 1979; Poizner, Bellugi, & Tweney, 1981) and intrusion errors that resemble the stimulus materials in terms of handshape (Bellugi, Klima, & Siple, 1974; Klima & Bellugi, 1979; Krakow & Hanson, 1985). These results are a promising indication that something like the phonological loop may exist in the visuospatial domain for deaf subjects. Will this ASL-based phonological coding show further effects and interactions that are characteristic of a rehearsal-loop structure? Just as the phonological loop uses representations of vocal ar-

SIGNS CONTRASTING ONLY IN HANDSHAPE



CANDY

APPLE

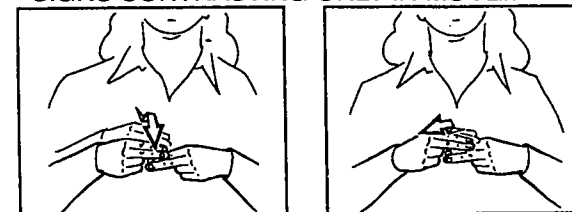
SIGNS CONTRASTING ONLY IN LOCATION



SUMMER

UGLY

SIGNS CONTRASTING ONLY IN MOVEMENT



NAME

SHORT

Figure 1. Illustration of part of the phonological system of American Sign Language. Each pair of signs contrasts a different phonological component.

tication, so ASL-based coding in working memory may use representations of manual articulation. And just as vocal articulation loads the phonological buffer, so too manual articulation may load an ASL-based phonological buffer.

In the present study, these possibilities are tested by looking for a suppression effect from irrelevant use of the hands and for appropriate interactions between suppression and phonological similarity. When stimuli are presented directly, as signs, the two effects should not interact, indicating the use of two separate components. However, when stimuli are presented indirectly (in the form of pictures) and must be recoded in order to use ASL-based phonological working memory, then suppression should abolish the phonological similarity effect.

EXPERIMENT 1

Method

Subjects. The subjects were 24 deaf individuals from Gallaudet University, California State University, Northridge, and the San Diego area. They ranged in age from 18 to 42 years, with a mean age of 27 years. All had been deaf either from birth ($n = 23$) or by age 2 ($n = 1$), and all had a hearing loss greater than 80 dB. All had learned ASL as their first language by age 6. Of these, 17 had deaf parents or older deaf siblings and had learned ASL at home from infancy. All considered ASL to be their primary language. The subjects were paid \$10 for 1 h of participation, which included the present experiment and various other experiments. Twelve subjects had participated in Experiment 2 of the present study prior to participating in Experiment 1. Two subjects who performed below 55% correct were replaced. This was done because, in hearing subjects at least, "phonological coding is often abandoned when the memory span is grossly exceeded" (Baddeley & Hitch, 1994). These 2 subjects are not counted in the descriptive statistics above, but did fit the selection criteria.

Materials. The stimuli were constructed from two sets of eight signs. One set consisted of phonologically similar signs, all sharing handshape (the "B" handshape) but differing in movement and/or place of articulation. These signs, glossed in English, were BROOM, PIE, BOOK, SOAP, BABY, DOOR, PAPER, BREAD.¹ The second set was dissimilar with respect to handshape, palm orientation, movement, and specific location features. These signs were KEY, EARTH, SOCKS, EGG, ROCK, PIANO, PUMPKIN, CUP. Examples from the phonologically similar and dissimilar sets are illustrated in Figure 2. The dissimilar signs were matched as closely as possible with the similar signs with respect to number of hands (all were two-handed), general location (neutral space), and types of movement (e.g., twists, brushes, short repetitions). In addition, to control for any possible word-length effect resulting from an English-based strategy, the two sets were closely matched for the number of syllables in their English translations. (In fact, it is the "dissimilar" set, the set that should be easier with ASL coding, that is slightly longer in terms of English syllables. Thus, any English-based word-length effect would actually work against our predictions.) All signs were concrete nouns that could be represented as pictures. This constraint was for the purposes of Experiment 2.

From each set, 16 lists of four signs were constructed, making for a total of 32 lists. Eight lists were used for practice, chosen so that each sign occurred exactly twice during practice. The remaining 24 lists were randomly mixed. A native signer was filmed producing each list at a rate of approximately one sign per second. This master videotape was edited to create an experimental tape, which contained a fixation cross prior to each list and black video between lists. The experimental videotape was edited using a Panasonic AG-A650

editor controller and Panasonic AG 6500 and 6300 videocassette recorders.

Design and Procedure. Two factors were varied in a 2×2 repeated measures design. The factors were similarity (dissimilar vs. similar) and suppression (not suppressed vs. suppressed). Similar and dissimilar trials were randomly mixed; not-suppressed and suppressed trials were blocked.

The subjects were tested individually, using a SONY PVM 1380 trinitron color video monitor. Each trial began with a fixation cross for 1 sec, followed by a list of four signs. At the end of the list, the screen went black and the experimenter paused the videotape while the subject responded. The subjects were instructed to repeat, in the correct order, the signs they had just seen. If they could not remember every sign, they were asked to indicate the serial positions of the missing signs.

The entire block of 32 lists was presented twice to each subject, once each for the suppressed and not-suppressed conditions. During the not-suppressed condition, the subjects were instructed to simply wait while each list was presented. During the suppressed condition, the subjects were instructed to move their hands, by changing handshape from a fist to an open hand (ASL S handshape

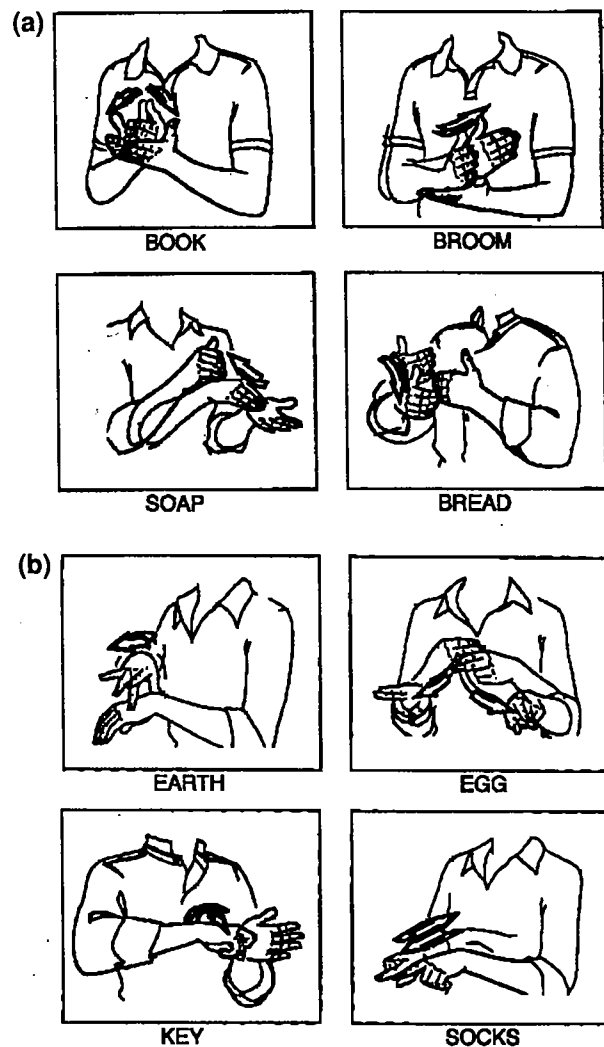


Figure 2. Illustration of (a) phonologically similar signs and (b) phonologically dissimilar signs.

to ASL 5 handshape), alternating the two hands. The order of the two suppression conditions was counterbalanced across subjects.

The suppression gesture was modeled for subjects by the experimenter at a rate of approximately four repetitions per second (two repetitions with each hand). Subjects were instructed that the gesture was to begin when the fixation cross appeared and to end when the last stimulus disappeared. Subjects who interrupted their gesturing during a trial, who were noticeably variable in speed, or who slowed to a rate of less than approximately two repetitions per second were corrected at the end of the trial.

The gesture used for suppression is of a simple, repetitive type that has been shown not to be generally disruptive of working memory but is capable of interfering with memory for hand configuration (Smyth & Pendleton, 1989). Furthermore, the gesture obeys the phonological constraints of ASL, and thus is a "pseudosign," analogous to the nonsense syllables used for articulatory suppression with hearing subjects in previous studies. (In fact, the gesture resembles an actual sign, and subjects may have treated it as a lexical item. However, given that similar results have been found in the hearing literature for suppression with words and nonwords, this distinction is not critical. In either case, the gesture we use here consists of phonological material.)

Scoring. The subjects were filmed during the experiment in order to subsequently score their responses. Each list was scored for the number of correct items reported in the correct serial position, and scores were totaled for each condition. In a few cases, trials had to be eliminated because of the ambiguity of the response or because a disruption had occurred during testing.

Results

Mean performance in the four conditions is shown in Table 1. An analysis of variance (ANOVA) showed that the main effect of similarity was significant [$F(1,23) = 8.12$, $MS_e = 64.9$, $p < .01$] and the main effect of suppression was also significant [$F(1,23) = 4.41$, $MS_e = 55.0$, $p < .05$]. There was no significant interaction between similarity and suppression [$F(1,23) = 0.003$, $MS_e = 57.5$, n.s.]. A power analysis was performed on the test for an interaction, with expected effect size taken from the significant interaction found in Experiment 2. Standard deviation was calculated from difference scores in Experiment 2. With an effect size of 5.1, standard deviation = 10.9, $N = 24$, and $\alpha = .05$ one-tailed, power to detect an interaction was .74. Because we are interested particularly in whether suppression eliminates the similarity effect, a pairwise comparison was made for the suppression condition, and a significant effect of similarity was found [$F(1,23) = 6.99$, $MS_e = 39.1$, $p < .02$].

Discussion

These results extend previous findings which have shown ASL-based coding in working memory (cf. Hanson, 1982; Klima & Bellugi, 1979; Poizner et al., 1981). In addition to finding coding based on the phonology of

ASL, as shown by the significant similarity effect, the significant suppression effect suggests that this ASL-based memory is an active rehearsal system employing an articulatory strategy. Furthermore, the similarity and suppression effects appear to be independent, just as in speech-based memory in hearing subjects. We found no significant interaction between the two effects, even though power to detect an interaction was moderately high. Under suppression, the similarity effect remained significant, and inspection of the means shows no trend toward suppression blocking the similarity effect. This pattern of data suggests that two separate components are contributing to performance, just as in speech-based working memory in hearing subjects.

One potential concern is that the sizes of the effects appear to be small compared with other results in the literature. However, this discrepancy can be primarily attributed to differences in scoring systems. With deaf subjects, both Hanson (1982) and Poizner et al. (1981) found that phonological similarity caused a drop in performance in the range of 15%–20%, but these results were based either on the number of lists repeated correctly (Poizner et al., 1981) or on the number of correct responses to a probe, with one probe per list (Hanson, 1982). In contrast, when Poizner et al. scored their data on the basis of total number of items reported in the correct position, they found drops in performance in the range of 5.8%–9.9%. Scoring by item rather than by list probably gives the more meaningful estimate of the magnitude of the effect, given that only 12 or 16 lists are used in most experiments in the literature. (In the present experiment, when number of correct lists is scored, the differences between means become larger but statistical significance is lost.)

With hearing subjects, again, most studies count number of completely correct lists, making comparison difficult. Furthermore, size of effect varies widely with list length and other variations in method (e.g., Baddeley et al., 1984). Thus, it is difficult to say whether the effects in the present Experiment represent weaker influences of phonological similarity and articulatory suppression for ASL than for spoken English. Nevertheless, the rather small effects found here suggest that such a disparity may indeed exist. If so, a number of factors might be contributing.

One possibility is that deaf subjects have a wider variety of strategies available to them than do hearing subjects. For instance, deaf subjects have an English-based code available to them in working memory as well as an ASL-based code (Campbell & Wright, 1989; Dodd, Hobson, Brasher, & Campbell, 1983; Hanson, 1982, 1990). In addition, some of our subjects reported using conceptual or elaborative strategies, such as imagining pairs of items related in some way. Deaf subjects may have a tendency to seek such alternate strategies because ASL coding appears not to support as long a memory span as does the phonological loop for speech. Hearing subjects, on the other hand, may be much more committed to the single strategy of phonological rehearsal. Thus, variation

Table 1
Mean Percent Correct With Sign Stimuli

	Dissimilar		Similar		<i>M</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Not suppressed	82.4	17.1	77.8	16.7	80.1
Suppressed	79.3	17.0	74.5	18.3	76.9
<i>M</i>	80.9		76.2		

