Journal of Experimental Psychology:
Learning, Memory, and Cognition
1990, Vol. 16, No. 2, 258–269

0278-7393/90/$00.75

Properties of Memory for Unattended Spoken Syllables

Nelson Cowan, Wemara Lichty, and Tim R. Grove
University of Missouri

Whereas previous studies on memory for unattended speech have inadvertently included acoustic interference, the present study examines memory for unattended syllables during a silent period of 1, 5, or 10 s. The primary task was to read silently (Experiments 1–3) or whisper the reading (Experiment 4). Occasionally, when a light cue occurred, the subject was to recall the most recent spoken syllable, as well as the recent reading material. Memory for both the vowels and consonants of the syllables decreased across 10 s, confirming that auditory memory does decay in the absence of acoustic interference. However, the specific patterns of memory decay for vowels versus consonants depended on task demands, including the allocation of attention and the opportunity for subvocal coding. We suggest an account of performance that includes auditory sensory and phonetic memory codes with different properties, used in combination.

An important assumption within cognitive psychology is that memory is partly independent of attention and partly dependent on it (e.g., see Broadbent, 1958; Cowan, 1988; Norman, 1968; Shiffrin, 1988). Although some features of each stimulus presumably are retained for a short time automatically, without attention to that stimulus, memory for other features presumably can be vastly improved by mnemonic strategies or other attention-demanding processes.

Despite the long endurance of this assumption, however, both the amount that can be remembered automatically and the duration for which it can be remembered remain unknown. To be sure, these questions were addressed early on in cognitive psychology (e.g., Broadbent, 1958), but some of the lines of research that seem potentially the most relevant were abandoned without wholly satisfactory answers having been obtained.

Research with unattended stimuli delivered through headphones may be especially useful to address this question; subjects cannot influence the perceptual input stage by redirecting the effector organs as they can in visual experiments. In a well-known, relevant line of research, two competing messages are presented dichotically. The subject's task is either to report both messages (e.g., Broadbent, 1957; Bryden, 1971) or to continually monitor one message and report the other one only when a special cue is delivered (e.g., Glucksberg & Cowen, 1970; Norman, 1969; Treisman, 1964). In a related procedure, different sets of auditory stimuli are presented simultaneously at different apparent spatial locations, and the subject is informed of the spatial location to be recalled only after a poststimulus delay period (Darwin, Turvey, & Crowder, 1972; Massaro, 1976; Rostron, 1974; Treisman & Rostron, 1972). All of these procedures were designed to examine memory for sounds that were at least partly unattended at the time of presentation. On the basis of such research, it has generally been concluded that auditory memory persists for several seconds in the absence of attention.

There are two issues that remain unresolved, however. First, to what extent was the attention manipulation successful? Attention often has not been checked very diligently in this type of research. Second, how much can subjects remember of unattended sounds when acoustic interference is minimized? This question cannot be answered at all from the studies cited above, given that contralateral stimuli were always present. Contralateral stimuli appear to cause amounts of interference that cannot be neglected, both in perceptual encoding (Harris, 1974; Kallman & Morris, 1984) and in postperceptual auditory memory (Hawkins & Presson, 1977; Morton, Crowder, & Prussin, 1971). Further, in the studies in which subjects shadowed an attended message, the subject's voice would, of course, act as an additional source of auditory interference.

Eriksen and Johnson (1964) developed a method that helps to answer some of these concerns. Near-threshold-level tones occasionally were presented within a continual white-noise background as subjects silently read passages from a novel. At varying intervals following the tone (or in a control condition with no tone), a light was switched off to signal the subject to stop reading and recall whether or not a tone had been presented. It was found that tone detection was at better-than-chance levels but decreased monotonically as the delay between the tone and light cue increased to 10 s. In the attention manipulation check, subjects were to stop reading and respond to the tone as soon as it was presented. They responded to the tones only very rarely in this condition.

Although this single study was unquestionably innovative, further research along these lines still is needed. It is not at all clear if the memory decay function observed for the detection of near-threshold tones embedded in noise would generalize to the recognition of clearly suprathreshold sounds. Further, with speech sounds as the stimuli, an automatic phonetic

This project was supported in part by National Institutes of Health Grant 2-R23-HD21338 awarded to Nelson Cowan. Several of the present experiments (1a, 2a, and half of 3) were briefly described by Cowan, Lichty, and Grove (1988).

We thank J. Scott Saults, Robert Crowder, John Gardiner, and an anonymous reviewer for commenting on an earlier draft of this article, and we thank Cristy Cartwright and Eric Hille for assisting in data collection and analysis.

Correspondence concerning this article should be addressed to Nelson Cowan, Department of Psychology, 210 McAlester Hall, University of Missouri, Columbia, Missouri 65211.

encoding process could supplement a purely auditory process for some or all sounds. As one means of examining the generality of the Eriksen and Johnson (1964) findings, the present investigation involved a modification of their procedure with the use of unattended syllables of speech presented at a suprathreshold level of intensity.

In our study, the phonemic identities of the consonant and vowel of each spoken syllable were varied independently to enable separate scoring of memory for consonants versus vowels. Research with attended speech sounds has indicated that auditory sensory memory typically is less useful for retaining stop consonants, which are relatively complex acoustically, than for retaining vowels, which are acoustically simpler (Cole, 1973; Crowder, 1971, 1973; Darwin & Baddeley, 1974; Pisoni, 1973). On the other hand, there is no evidence for a vowel advantage within a more abstract form of memory for speech, such as a phonetic memory, that can include information from the visual modality as well as from audition (e.g., see Massaro, 1987). The presence or absence of a vowel superiority effect in the data would serve as a clue to whether or not auditory sensory memory was used, either alone or in combination with another form of memory for speech.

Experiment 1a

Method

Subjects. Thirty undergraduate students (10 male, 20 female) who received course credit for their participation served as subjects. They had no known hearing losses and spoke English as their native language.

Apparatus. Subjects were run one or two at a time in a sound-attenuating chamber, with the equipment and experimenter controls located outside of the chamber. Each subject sat in a chair with desktop attached, and, when 2 subjects were run concurrently, their seats faced away from one another. An intercom system was used to ensure that subjects were quiet during the session. The stimulus channel of an audiocassette tape was delivered to each subject binaurally through TDH-39 audiological headphones, with each syllable set at approximately 54 dB(A) with a GenRad Model 1565-B sound level meter equipped with a 9-A Type Earphone Coupler. That level is quiet, but is well above the detection threshold. There were two 100-W lights in the chamber facing upward to provide diffuse lighting, and one of them was turned off whenever the subject was to stop reading and begin a memory trial.

Stimuli. The reading material was the beginning of the novel *2001* (Clarke, 1968). The auditory stimuli were natural tokens of nine consonant–vowel (CV) syllables spoken in a male voice. These syllables, which are shown in Table 1, consist of three different consonant phonemes combined with three different vowel phonemes, so that recognition of phonemes in the consonant versus the vowel categories could be scored independently. The stimulus channel of the audiotape was constructed by reading the syllables aloud at predetermined intervals. Every block of nine stimuli on the tape included each of the nine possible syllables once, in some random order. The silent interval between syllables on the tape was always about 2 s following three of the syllables in each block, 6 s following three syllables, and 13 s following three syllables, with intervals randomly ordered. The purpose of most of these stimuli was only to promote habituation to the auditory channel. Only nine syllabic tokens, dispersed among all of the others, were used as memory targets.

Table 1
Phonemic Composition of the Nine Consonant–Vowel Syllables Used in Experiment 1

	Vowel		
Consonant	i	I	ɛ
b	[bi]	[bI]	[bɛ]
d	[di]	[dI]	[dɛ]
g	[gi]	[gI]	[gɛ]

Note. Key to pronunciations: consonants, [b] as in *by*, [d] as in *do*, and [g] as in *go*; vowels, [i] as in *beet*, [I] as in *bit*, and [ɛ] as in *bet*.

The nine targets included a single token of each different syllable. On the average, they were 264 ms long (*SD* = 21.50), with a fundamental frequency of 128 Hz (*SD* = 1.58). The target syllables always were followed by 13 s of silence on the tape, which was long enough to allow the same nine syllabic tokens to be used for trials at any of three posttarget delay intervals: 1, 5, or 10 s. There were three groups of subjects (A, B, and C), all of whom heard the same stimulus tape, but with a different assignment of the three delay intervals to particular test trials. Each subject received three test trials at each delay.

On the second channel of the audiotape there were cues that went to the experimenter but not to the subject. These cues included the word "ready" to alert the experimenter just before each target syllable, and a cue at the end of the posttarget delay for the experimenter to begin the memory trial. The experimenter's channel was designed to handle any of the three groups of subjects. For example, after the target syllable for the third trial, the experimenter's channel contained the cues *C* (after 1 s), *A* (after 5 s), and then *B* (after 10 s), corresponding to the assignments of groups to delays that happened to be selected for Trial 3.

Procedure. Subjects read and heard the following instructions:

> In this experiment we are interested in how well you can concentrate on a reading passage when there is a distracting source of noise presented through headphones. Allow yourself to become involved in the reading. We will interrupt the session several times to ask you to report on what you have been reading. We will signal you to stop reading by turning off a light. We also would like to know whether or not you remember the last thing that was said before you stopped reading. Whenever we turn the light off, first please circle a syllable on your answer sheet indicating which syllable was said *last*.

These instructions were followed by a list of English transliterations of the nine syllables (e.g., *bee, dih, geh*), each with an English word as an example. The subject then looked over the answer sheet, on which the 9 syllables were arranged in a 3 × 3 matrix form for each trial. The subject learned that he or she was to circle the correct syllable, indicate what page of the novel he or she was on, and write a "couple of short sentences" on lines provided to the right of the syllable matrix, to summarize the content of what was read just before the light cue. The subject also learned that there would be a multiple choice test on the reading at the end of the session.

Before the reading-and-memory test period began, however, there was a syllable-familiarization period and then a practice period. In these periods, an audiotape similar in format to the test tape was used. In the familiarization period, the subject listened to each of the nine syllables once while following along on an already-marked answer sheet. Then, in the practice period, the subject heard a sequence of 27 syllables (lasting just over 3 min), and 9 of these syllables were used as memory trials with no posttarget delay. In this practice period, unlike the test period, the subject attended to the syllables and had no concurrent task. Therefore, the practice data could be used as a baseline for the identifiability of the syllables when subjects were free to attend to them.

Subjects began reading immediately when the test tape began. However, the first memory and reading test trial on the tape did not occur until the subject had been reading for approximately 5 min. As soon as the experimenter heard the trial cue, he or she turned off a power switch that operated both the audiocassette deck and one of the lights in the sound-attenuating chamber, signaling the subjects to begin the recall trial. Subjects had 30 s to identify the syllable and then write briefly about what they were reading. The brief written description was not scored; it was included in order to encourage subjects to maintain their attention to the reading.

The total reading time was approximately 33 min. Consecutive memory trials were spaced several minutes apart ($M = 3.58$ minutes, $SD = 1.37$) to allow subjects sufficient time to habituate to the sounds and become involved in the reading before a memory test trial occurred.

At the end of the session, each subject removed the headphones and then took a multiple choice test on factual material from the reading (what a particular character said, did, was carrying, etc.) Each question was marked with the text page on which the event occurred, and each subject completed the test only up to the point where he or she stopped reading. A typical example is the following question (#5): "What abstract deduction did Moon-Watcher initially make about how the monolith appeared? (a) It grew the way mushrooms grow; (b) It was from another tribe; (c) It fell from the sky; or (d) An enormous animal had dragged it." Subjects were instructed to guess if they did not know the answer to a question corresponding to a portion of the novel that they had read. They had no problem following these instructions.

Results and Discussion

Syllabic identification. Each syllabic identification response was scored in two ways: for the correctness of the consonant and for the correctness of the vowel. These scores were averaged across trials to yield a proportion correct for consonants and for vowels at each delay for each subject. As an inspection of Table 1 suggests, chance performance would be .33 for either of these identification scores.

In the practice session, subjects correctly identified 94.4% of the consonants and 93.9% of the vowels. In the test session, however, a very different result emerged, as shown in Figure 1. Performance after a 1-s delay was similar to performance in the practice session, but memory decayed across the longer delays. Further, this decay of memory was much more pronounced for the consonants than for the vowels. Confirming this pattern, an analysis of variance (ANOVA) of proportion correct with delay (1, 5, or 10 s) and phoneme type (consonant or vowel) as within-subjects factors produced a significant effect of delay, $F(2, 58) = 23.82$, $p < .001$, $MS_e = 0.048$, and phoneme type, $F(1, 29) = 6.37$, $p < .02$, $MS_e = 0.061$, as well as a Delay × Phoneme Type interaction, $F(2, 58) = 6.16$, $p < .005$, $MS_e = 0.035$.

Reading task. Subjects read an average of 22.77 pages of the novel ($SD = 6.46$). On the final multiple choice test, in which subjects were to go only as far as they had read, they answered an average of 13.23 questions ($SD = 3.35$), 10.77 of them correctly ($SD = 3.17$). The mean proportion correct was .81 ($SD = 0.15$).

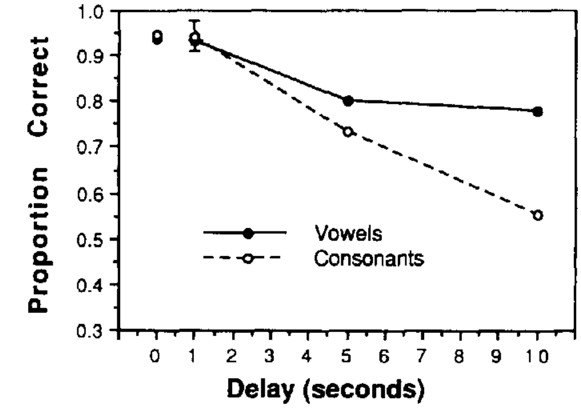

Figure 1. Mean proportion correct recognition of vowels and consonants at each poststimulus delay in Experiment 1a. (Unconnected points at a 0-s delay = single-task practice session; error bar = SE_M based on the Phoneme Type × Delay MS_e.)

These data suggest that subjects were attending to the reading fairly well. Therefore, it seems likely that subjects devoted at most a part of their attention to the encoding and retention of syllables on the stimulus tape. Nevertheless, more information is needed on the allocation of attention and on stimulus factors before these results can be interpreted theoretically.

Experiment 1b

This was a control experiment to demonstrate the validity of the multiple choice test on reading comprehension by providing an estimate of guessing rate.

Method

Nine college students served as subjects (5 male, 4 female). They completed the multiple choice test without having read any part of the novel and without participating in any other experimental task. Each subject was asked to complete 19 questions of the multiple choice test in whatever way seemed to make the most sense.

Results and Discussion

Of the 19 questions that subjects were to answer, they were correct on an average of 5.89 questions ($SD = 1.27$), and the mean proportion correct was .31 ($SD = 0.07$). Even though subjects in Experiment 1a answered far fewer than 19 questions on the average, the mean number answered correctly in that experiment, 10.77 ($SD = 3.17$), was significantly greater, $t(47) = 4.24$, $p < .001$. If one looks at just the first 13 questions, which is the integer closest to the mean number answered in the first experiment, the control subjects correctly answered only 4.00 of them ($SD = 1.00$), and the mean proportion correct was again .31 ($SD = 0.08$). The much poorer reading comprehension test scores in this control experiment confirm that subjects in Experiment 1a must have comprehended a substantial amount of what they read in the presence of irrelevant speech sounds.

Experiment 1c

The purpose of this control experiment was to determine how well subjects would do on the reading comprehension test if they did the reading but were not placed in a dual task situation.

Method

Twelve college students (5 male, 7 female) who had not participated in Experiment 1a or 1b served as subjects. They were tested concurrently in a large, quiet room. After instructions that included mention of the final multiple choice test, subjects read the novel in the same time schedule as the subjects in Experiment 1a. However, they did not wear headphones or hear any spoken stimuli during this reading. Only the experimenter listened to the audiotape in order to know when the room lights should be flashed off and on as a signal to recall. These recall trials were timed to match a group of subjects in Experiment 1a. Whenever the signal was received, subjects wrote brief passages indicating what was currently happening in the novel, but they naturally skipped the syllable-identification test items. At the end of the session, the reading comprehension test was administered.

Results and Discussion

Subjects read an average of 21.08 pages of the novel (SD = 4.87). They answered an average of 12.67 multiple choice questions (SD = 2.15), and were correct on an average of 10.50 of them (SD = 3.03). The mean proportion correct, .81 (SD = 0.15), is the same as what was obtained in Experiment 1a. None of the means is greater than what was observed in Experiment 1a, and there are no significant differences between them.

These results suggest that the attention manipulation of Experiment 1a was successful. Given that subjects seem to have read as efficiently in the presence of the speech stimuli as in their absence, attention apparently was devoted primarily to the reading rather than to the spoken stimuli in Experiment 1a.

Summary of Experiment 1 Findings

When presented with CV syllables while attending to a reading passage, subjects retained vowels (V) better than consonants (C), with the difference between the two types of phoneme becoming progressively larger across a silent, 10-s poststimulus period. This result is generally consistent with previous studies of memory for vowels and consonants (Cole, 1973; Crowder, 1971, 1982; Pisoni, 1973), but the time scale over which memory decay was observed (at least 10 s) is longer than the estimates that were offered in those previous studies (generally 3 s or less). The duration of decay we observed, however, is more consistent with all of the evidence on the duration of sensory memory (see Cowan, 1988).

Experiment 2a

There are several possible accounts of why an advantage for vowel memory occurred within Experiment 1a. One is that vowels are acoustically simpler than consonants. A second possibility is that the vowel advantage occurred because vowels were the final phonemes in the syllables that were used. Third, both phoneme complexity and syllabic composition may matter.

In order to investigate such stimulus factors, in this experiment VC syllables were used instead of CV syllables. If the vowel memory advantage of Experiment 1a resulted totally from the greater simplicity of vowels, the same pattern of performance should emerge with VC syllables. On the other hand, if factors of syllabic composition such as phoneme order are important, there could be an advantage for consonants. If both acoustic simplicity and order play a role, these factors should tend to cancel one another out in this experiment, and the difference between vowel and consonant memory should be much smaller than in Experiment 1.

Method

Subjects. The subjects were 30 students (12 male, 8 female) who did not participate in any of the previous experiments.

Apparatus, stimuli, and procedure. A new stimulus tape was constructed according to the same temporal schedule as in the first experiment, but with VC syllables instead of CV syllables. The syllables contained the same phonemes as in the stimuli of Experiment 1, but in the reverse order within a syllable (instead of [bI], [Ib]; instead of [gi], [ig]; and so on).

In the pronunciation, the consonants were released (i.e., were succeeded by a brief period of voicing) to ensure their clarity, although the releases were produced in a near-whisper, and unrealistic exaggerations of these releases were avoided. The nine target stimuli were 317 ms long (SD = 20.32), with a vowel fundamental frequency of 137 Hz (SD = 2.20). The released portions of these syllables were 87 ms long on the average (SD = 16.60), with a fundamental frequency of only 94 Hz (SD = 14.52). In every other respect, the method was the same as that of Experiment 1a.

Because the intervals between syllables on a tape were timed silently during a live recording session, there proved to be a small difference between the timing of stimuli in Experiments 1a versus 2a. On the average, a 1.000-s intersyllabic interval in Experiment 2 corresponded to a 0.936-s intersyllabic interval in Experiment 1. We assume that this small difference in timing had no more than a trivial effect on subjects' level of habituation to the sounds. Note that the timing of trial delay cues, on the experimenter's channel of the stimulus tape, was not affected by this difference in intersyllabic intervals. Unlike those intervals throughout the tape, the trial delays were timed with a stopwatch in a separate dubbing session and could be redubbed until they were accurate.

Results and Discussion

Syllabic identification. In the practice session, subjects correctly identified 96.0% of the vowels and 100% of the consonants. The proportions of correct recall in the test session, for each combination of delay interval and phoneme type, are shown in Figure 2. Performance decreased regularly across delay intervals, but, unlike in Experiment 1a, there was very little difference between the decay function for vowels versus consonants. An ANOVA with the same factors as in Experiment 1a produced a significant effect of delay, $F(2, 58)$

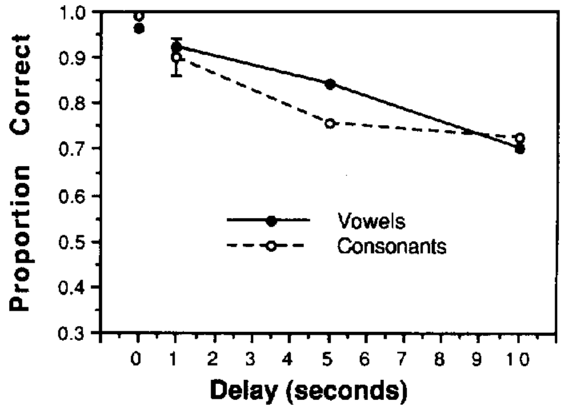

Figure 2. Mean proportion correct recognition of vowels and consonants at each poststimulus delay in Experiment 2a. (Unconnected points at a 0-s delay = single-task practice session; error bar = SE_M based on the Phoneme Type × Delay MS_e.)

= 13.98, $p < .001$, $MS_e = 0.043$. However, neither the phoneme type nor the interaction of Delay × Phoneme Type approached significance. This suggests that the relative simplicity of vowels and the position of the vowel in the syllable both make a difference for memory performance. In Experiment 1a, both of these factors favored memory for vowels, but in Experiment 2a, the effect of the simplicity of vowels apparently was counteracted by the placement of consonants after the vowels.

The results make sense with respect to the conventional concept of auditory sensory memory. This form of memory generally is more useful for the preservation of vowels than of consonants. However, it cannot be used well when the memory target is followed by an acoustically similar suffix item (Crowder, 1971; Morton et al., 1971). The syllable-final consonants in the present experiment apparently contained acoustic components that were similar enough to the vowels to cause suffix-like interference with memory for the vowels. This is not necessarily surprising, given that the vowels and consonants of a syllable were connected by continuous transitions in the formant frequencies. Nevertheless, the finding seems somewhat at odds with several previous studies that seem to suggest that the relative availability of consonant and vowel information may be equivalent in CV and VC syllables (Cole, 1973; Crowder, 1973; Darwin & Baddeley, 1974).

Cole (1973) presented sets of CV and VC syllables for recall and found an advantage for vowels in both types of syllable. However, a common problem in some VC syllables is that the consonant may not be pronounced clearly. This is especially true of stop consonants if they are not released. It is quite possible that the clarity of consonants was better equated across CV and VC syllables in the present study than in Cole's study and that the factor of phoneme recency is obscured unless this level of phoneme clarity is achieved. Also, the releases in the present study could have directly interfered with memory for the vowels, although they are not very similar acoustically (see above). A final possibility is that the difference in results could have occurred because the syllables were unattended in the present study and attended in Cole's (1973) study.

Crowder (1973) found no suffix effect for stop consonants within VC syllables, similar to what has been found for stop consonants within CV syllables. Darwin and Baddeley (1974) obtained comparable modality effects for lists composed of the syllables [ʃa], [ma], and [ga] and for lists composed of the syllables [aʃ], [am], [ag]. However, because memory for vowel information within VC syllables was not examined in these studies, there is a paucity of evidence directly relevant to our test situation.

Because of the factors mentioned above, we cannot be certain that phoneme recency within a syllable plays a role in memory. However, the data from Experiments 1 and 2 do appear to warrant the conclusion that there is superior memory for vowels overall and that this vowel superiority effect was cancelled by some acoustic aspect of the VC stimuli. In this respect, the data are consistent with the assumption that auditory sensory memory was used, either alone or with other types of memory.

Reading task. On the average, subjects read 25.13 pages of the novel (SD = 8.87). They answered 14.9 multiple choice questions on the average (SD = 3.60), and were correct on an average of 12.23 of these questions (SD = 4.19). The mean proportion correct on the multiple choice test was .82 (SD = 0.15), which is nearly identical to the proportion obtained in Experiments 1a and 1c. However, another reading control experiment was conducted to correspond to the VC tape.

Experiment 2b

As mentioned above, the VC stimulus tape was inadvertently recorded at a slightly slower pace than the CV tape used in Experiment 1. Although the differences in intersyllabic intervals were quite small, they do add up across the session; the VC tape was 2.25 min longer than the CV tape. Because subjects in Experiment 2a thus had longer to read, another reading-only control experiment was needed in which the reading period was timed according to the VC tape.

Method

The subjects were 19 students (5 male, 14 female) who had not participated in previous experiments. The procedure was the same as that of Experiment 1c, except that the reading period was timed according to the VC tape rather than the CV tape.

Results

Subjects read an average of 27.92 pages of the novel (SD = 7.47). They answered an average of 15.47 questions of the multiple choice test (SD = 3.22) and were correct on an average of 12.63 questions (SD = 3.39). The mean proportion correct was .82 (SD = 0.14), as in Experiment 2a. None of the differences between reading measures in Experiments 2a versus 2b were significant. This provides further confirmation that the attention manipulation has been successful.

Summary of Experiment 2 Findings

Memory for unattended vowels and consonants within VC syllables was found to decay at comparable rates across 10 s. Although the superiority of vowel memory over consonant memory that had been obtained with CV syllables in Experiment 1a was not obtained with VC syllables in Experiment 2a, the difference was canceled, not reversed. Apparently, phoneme recency or some other aspect of the VC syllables counteracted the vowel superiority effect. The data are still at least consistent with the view that an auditory sensory memory was used, given that this is the only type of memory in which a vowel superiority effect has been observed.

Experiment 3

In their manipulation check for attention, Eriksen and Johnson (1964) found that subjects usually could not spontaneously detect near-threshold tones unless they had received a cue to stop reading and attend to the auditory channel. This finding strengthened the assumption that attention was devoted to reading. Our Experiment 3 was a similar manipulation check for the present study, using CV syllables. Subjects had to detect occurrences of the phoneme [dI] while reading. They also received the occasional cues to stop reading and recall the last syllable presented, as in Experiments 1a and 2a. It is not at all a foregone conclusion that this uncued detection of a particular clearly audible syllable will be as difficult as uncued detection of a barely audible tone was in the Eriksen and Johnson study.

Method

The subjects were 30 students (7 male, 23 female) who did not participate in the previous experiments. The apparatus, stimuli, and procedure were the same as in Experiment 1a, with one major exception. Whenever the subject heard the syllable [dI], he or she was to flip a switch that turned on an indicator light outside of the chamber for the experimenter. This uncued syllabic monitoring was included in the practice sequence in which subjects were to make syllabic identifications, as well as in the test session. Only 1 subject was tested at a time in this experiment.

Results and Discussion

Monitoring for detection of [dI]. Throughout the test session, each subject heard the target syllable [dI] 28 times and spontaneously detected it an average of 16.8 times, for a hit rate of 60.0%. There were 224 other syllables, which were incorrectly identified as [dI] 4.17 times on the average, for a false alarm rate of 1.9%. Combining these percentages, signal detectability (d') for the group in this syllabic recognition task was 2.32. In contrast to this result, one can derive $d' = 0.86$ from the comparable condition in Eriksen and Johnson's (1964) study, a much poorer level of signal detectability than in the present experiment.

One interpretation of this difference between studies is that subjects can switch attention from reading to suprathreshold syllables more readily than they can switch attention from reading to near-threshold-level tones. Perhaps the ability to switch attention to the auditory channel in this situation depends on the target sound's tendency to automatically elicit attention. Cowan (1988) reviewed evidence that automatic attentional responses and deliberate attending do operate together to control the direction of the subject's attention at each moment.

Syllabic identification. In the memory test that occurred after a postsyllabic delay of 1, 5, or 10 s, there was no indication of the severe, monotonic decline across delay intervals that appeared in the first experiment (cf. the present results, in Figure 3, and those of Experiment 1a, in Figure 1). This difference between experiments suggests that subjects could carry out the spontaneous [dI]-detection task only by shifting attention to the speech stimuli more than when no such monitoring task was included. This confirms that the severe decline across delays occurs only when attention is allocated primarily to the reading.

Despite the much lower rate of speech memory decay observed in this experiment, there was a significant effect of delay, $F(2, 58) = 3.47$, $p < .04$, $MS_e = 0.033$, in an ANOVA with the same factors as in the previous speech memory experiments. Performance was slightly lower after a 5-s delay (.83) than it was after either a 1-s delay (.91) or a 10-s delay (.89), although the explanation of this small difference is unclear. No other effects of the analysis were significant.

In order to statistically verify that the apparent difference between experiments was real, an analysis including the memory data from both Experiment 1a and Experiment 3 was performed. It resulted in an overall effect of delay, $F(2, 116) = 18.37$, $p < .001$, $MS_e = 0.04$, and Phoneme Type × Delay, $F(2, 116) = 5.82$, $p < .005$, $MS_e = 0.03$, but there also was a main effect of experiment, $F(1, 58) = 11.32$, $p < .002$, $MS_e = 0.06$, and interactions of Experiment × Phoneme Type, $F(1, 58) = 6.98$, $p < .02$, $MS_e = 0.05$, and Experiment × Delay, $F(2, 116) = 12.45$, $p < .001$, $MS_e = 0.04$. These results further confirm that subjects in the first experiment were not attending fully to the speech stimuli while they were reading.

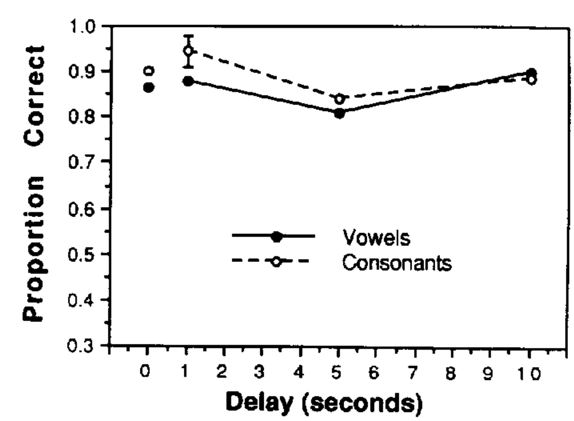

Figure 3. Mean proportion correct recognition of vowels and consonants at each poststimulus delay in Experiment 3. (Unconnected points at a 0-s delay = single-task practice session; error bar = SE_M based on the Phoneme Type × Delay MS_e.)

Detection and memory. It is possible to learn more about performance by examining correspondences between detection and memory. On the one trial in which the memory target was [dI], subjects sometimes spontaneously detected the target but then failed to correctly identify it. The probability of correct [dI] identification conditional upon that target having been spontaneously detected was .86 for trials with a 1-s delay, but it was only .60 with a 5-s delay and .57 with a 10-s delay. (The rates of spontaneous detection of these targets, upon which the memory results have been conditionalized, were 7 of the 10, 5 of the 10, and 7 of the 10 subjects for whom the [dI] target was presented at the 1-s, 5-s, and 10-s delays, respectively.) This pattern of performance could simply reflect guessing biases elicited by the particular task situation. Subjects may have incorrectly guessed that another syllable was presented in the 5- or 10-s delay period following [dI]. If they also had noticed (e.g., in the practice phase) that the same syllable rarely was presented twice in a row, there would have been a bias against using the label [dI] after having spontaneously detected [dI] several seconds earlier.

Theoretically, a subject might fail to detect a [dI] target spontaneously but then identify it correctly in the memory test. However, this happened in only 2 cases out of 30. Apparently, subjects allocated some attention to the spoken syllables rather consistently, given these task demands, and they rarely failed to detect [dI] simply because they were attending to the reading. Thus, when the syllable percept was accurate enough to permit the correct delayed identification, this percept usually had permitted spontaneous detection also.

Reading task. On the average, subjects read 23.87 pages of the novel (SD = 7.86). Subjects answered an average of 13.77 questions (SD = 2.90) and were correct on 10.57 of them (SD = 3.15). None of these means was significantly different from those of the appropriate reading control group (Experiment 1c) or Experiment 1a. The mean proportion correct, .77 (SD = 0.16), was only slightly lower than it was in Experiment 1a (.81), and the difference was not significant. Thus, contrary to our expectations, subjects appear to have learned to coordinate the reading and syllabic detection tasks fairly well. In this regard it should be noted that previous research also has shown that subjects can learn to coordinate pairs of tasks that at first would appear incompatible (e.g., Hirst, Spelke, Reaves, Caharack, & Neisser, 1980).

Nevertheless, anecdotal evidence suggests that this coordination of tasks was achieved only at a cost. Whereas subjects in Experiment 1a generally reported that the task was a pleasant one, subjects in the present experiment often mentioned that the task was stressful. These subjects apparently had to exert more effort in order to perform adequately in the dual task situation.

Summary of Experiment 3 findings. This experiment has confirmed in another way that the decay of memory for speech sounds observed in Experiments 1a and 2a occurred because subjects did not attend to the auditory channel. Apparently, though, subjects can learn to increase their attention to the auditory channel with very little deleterious effect on their reading performance as measured by the number of pages read and the scores on a multiple choice test of comprehension.

The present results differ from those of Eriksen and Johnson (1964), who found that subjects could not successfully monitor for quiet tones while reading. Nevertheless, the two sets of results both indicate that subjects ordinarily do not attend to the auditory channel while reading. When they did divide attention in this way in the present experiment, speech memory performance improved substantially.

Experiment 4

There are at least two limitations of the silent reading task that was used in Experiments 1a through 3. First, it is possible that subjects might have shifted their attention when a spoken syllable occurred and then quickly shifted it back to the reading. If so, memory performance conceivably might reflect partly attended rather than totally unattended speech, not only in Experiment 3, but to a lesser extent in the previous experiments as well, despite subjects' ability to read with comprehension. The relative insensitivity of the reading measure to the addition of an auditory monitoring task in Experiment 3 amplifies this concern.

A second concern is that subjects could have used covert articulation or "subvocalization" to rehearse the speech sounds while reading. They could do this because subvocalization might require only a minimum of attention and because subjects can read fairly well, although not perfectly well, without using subvocalization (Slowiaczek & Clifton, 1980).

In some previous studies of selective attention (e.g., Dawson & Schell, 1982; Norman, 1969; Treisman & Geffen, 1967), shadowing of prose has been used as the attended task. This procedure has several assets. It engages both attention and the articulatory apparatus, which should effectively prevent rehearsal of the supposedly unattended material. It also provides an index of possible shifts of attention away from the reading task. Specifically, slowed or erroneous shadowing responses are taken to indicate that an attention shift may have occurred. The method achieves these goals only at the expense of auditory interference from the subject's voice, however.

In the present experiment, a reading task was designed to circumvent this problem. Subjects read the text in a whispered voice into a microphone. Because whispering is not transmitted well through bone conduction, subjects generally cannot hear themselves whisper when they wear audiological headphones. Whispered speech was found to greatly attenuate covert rehearsal in a previous study (Cowan, Cartwright, Winterowd, & Sherk, 1987). Also, the whispering responses could be recorded well enough to allow an analysis of errors and hesitations indicative of subtle shifts of attention. The whispered speech waveforms were digitally stored in order to more carefully assess these shifts.

Method

Subjects. The final sample of subjects included 30 college students (14 male, 16 female) who did not participate in any of the previous experiments. Twelve additional subjects were excluded from the sample because they could not consistently produce an audible whisper.

Apparatus, stimuli, and procedure. The apparatus, stimuli, and procedure were similar to those of Experiment 1a, except that subjects were to whisper the reading into a microphone. The whispering was recorded on one channel of a response tape, and the stimulus sequence was simultaneously recorded from the stimulus tape deck to the second channel of the response tape deck using a connecting wire.

Subjects received 2 min of practice whispering the text, after the second phase of the practice session but before the stimulus tape for the text session began.

Analysis of whispering responses. Our analysis to observe possible subtle shifts of attention was a modification of a method used by Dawson and Schell (1982) within their study of semantic encoding of unattended material. They marked shadowing errors (omitted, mispronounced, or extraneous words) that occurred either concurrently with, or within two words following, the unattended target word. "Long hesitations" within this measurement interval also were considered indicative of possible shifts of attention, although Dawson and Schell offered no exact definition of a long hesitation.

To operationally define hesitations, we played the tape recordings of whispering surrounding each memory target into a computer equipped with an analog-to-digital conversion unit and an acoustic waveform analysis program. Hesitations were defined as the production of less than one full syllable of speech either in the exact 1-s period preceding the target onset or in the 1-s period following the target onset. The rationale for including hesitations differed for these two periods. In the pre-onset period, a hesitation in whispering could provide an opportunity for the subject to subvocalize or attend to the upcoming target syllable. In the postonset period, on the other hand, a hesitation could indicate that the subject had shifted attention to the target syllable.

Results

Syllabic identification. The proportion of correct recognition of consonants and vowels in both the practice and test conditions can be observed in Figure 4. The proportions in practice were .92 for both consonants and vowels, which is comparable to Experiment 1a. However, the pattern of results in the test phase was different from that obtained in Experiment 1a. Specifically, both consonant and vowel memory performances appeared to decline across delay intervals at similar rates, but with a large advantage for vowels, whereas in the first experiment the decline was much steeper for consonants than for vowels. In other words, whispered reading was much more detrimental than silent reading for the retention of consonants at short delays.

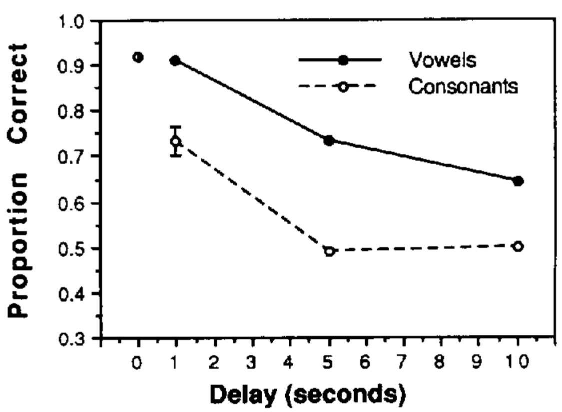

Figure 4. Mean proportion correct recognition of vowels and consonants at each poststimulus delay in Experiment 4. (Unconnected points at a 0-s delay = single-task practice session; error bar = SE_M based on the Phoneme Type × Delay MS_e.)

Confirming this pattern, an analysis comparable to the one carried out for the first experiment resulted in significant main effects of phoneme type, $F(1, 29) = 25.07$, $p < .001$, $MS_e = 0.064$, and delay, $F(2, 58) = 17.64$, $p < .001$, $MS_e = 0.062$, but the interaction of these factors did not approach significance, $F < 1.0$, unlike Experiment 1a.

There are several possible reasons why whispered reading was more devastating than silent reading for the retention of consonants at short delays. One possibility is that consonants were retained for short periods in Experiment 1a through the use of a residual amount of attention not devoted to the reading, and that whispered reading controls attention more carefully than silent reading does. Alternatively, the critical difference could be that consonants were retained in Experiment 1a through articulatory coding processes, and whispered reading could suppress this articulatory coding more completely than silent reading. In either case, if we assume that sensory memory includes neither attentive processes nor articulatory coding, the present experiment would yield a clearer view of sensory memory decay unconfounded by attentive or articulatory coding.

Syllabic identification with whispering performance taken into account. Whispering errors or hesitations occurred on 17.4% of all test trials. It was possible to recalculate each subject's mean for each poststimulus delay with these error trials omitted. The results of this analysis were quite similar to the results of the overall analysis. Vowels were correctly recalled on .85, .74, and .67 of the trials at the 1-, 5-, and 10-s delays, respectively, whereas consonants were correctly recalled on .66, .42, and .49 of the trials at those delays. There was once more a significant advantage for vowels over consonants, $F(1, 29) = 20.24$, $p < .001$, $MS_e = 0.12$, and a significant effect of delay, $F(2, 58) = 6.00$, $p < .003$, $MS_e = 0.10$, but once again the interaction of these factors did not approach significance. Thus, this pattern of responding appears to be indicative of the decay of memory for truly unattended speech sounds over a 10-s period.

It is also possible to examine memory performance on trials in which an error or hesitation in whispering *did* occur. These data are summarized in Table 2. They appear remarkably different from the nonerror data at the 1-s delay (although a statistical comparison is not possible because of empty cells in the error data for most subjects). Unlike the error-free data, the table shows little forgetting of either vowels or consonants at the 1-s delay. Apparently, subjects did shift attention to the sounds on at least a majority of these whispering-error-trials. The memory decay obtained at the longer delay intervals could have occurred because subjects' attention shifted back to the reading at some time after 1 s.

Reading task. Subjects read an average of 20.60 pages ($SD = 2.13$), answered an average of 13.27 multiple choice questions ($SD = 1.17$), and were correct on an average of 11.0 of these questions ($SD = 2.29$). The mean proportion correct was .83 ($SD = 0.14$). These means are similar to those of

Table 2
Proportions of Correct Recognition of Vowels and Consonants in Experiment 4 Following an Error or Pause in Whispered Reading

Type of phoneme	Delay		
	1 s[a]	5 s[b]	10 s[c]
Vowels	.93	.80	.56
Consonants	.93	.67	.50

Note. The contribution of each subject with data suitable for this table was the mean of all trials in which an error occurred in the two words whispered immediately after the target syllable's onset or in which a 1-s pause occurred just before or just after this onset. The numbers of subjects with two whispering errors at a delay were 3 (at 1 s), 2 (at 5 s), and 4 (at 10 s); the remaining subjects had only one such trial per delay.
[a] $N = 14$; [b] $N = 15$; [c] $N = 9$.

Experiment 1a and are not significantly different from that experiment.

Summary of Experiment 4 findings. The performance on the reading task and the analysis of whispering responses together go a long way toward ensuring the purity of the measure of memory for unattended syllables of speech. According to this measure, unattended consonants are not retained as clearly as unattended vowels in memory (i.e., performance levels were much higher for vowels across all three delays). Nevertheless, both unattended consonant and vowel memory traces appear to decay at comparable rates, at least between 1-s and 10-s delay intervals.

Of course, this experiment does not prove that the memory decay rates are identical for vowels and consonants; it would be impossible to prove this null hypothesis. However, it does seem that memory for consonants and vowels decayed at roughly comparable rates from 1 to 10 s. As a first estimate of decay rates, based on the means for trials in which no whispering error or hesitation occurred, the percentage of decrement in performance at the 10-s delay relative to the 1-s delay was 21.2% for the vowels and was a similar 25.8% for the consonants. This is in striking contrast to Experiment 1a, in which the percentages of decrement were 16.7% for the vowels and 41.1% for the consonants. The effect of whispering on speech memory decay thus was quite profound.

It is not completely clear what processing occurred in the first second after each target syllable. The added task of whispering could have rapidly interfered with phonetic memory of the consonants. In this case, the decay functions for consonants and vowels would actually be quite different if the function could be observed from the 0-s point rather than the 1-s point. On the other hand, the whispering task could have prevented subjects from encoding the target syllables phonetically in the first place. In this case, the theoretical decay functions for consonants and vowels would be similar, but the attended-speech data plotted at 0 s in the figure would not serve as a reasonable approximation of unattended performance at a 0-s delay, as it might have in the previous experiments.

Other research suggests that the initial phonetic coding of unattended speech sounds occurs automatically. Specifically, unattended speech has been found to interfere with processing of visually presented verbal information (Salamé & Baddeley, 1982), and interference from unattended spoken words in a cross-modal version of the Stroop effect also has been obtained (Cowan, 1989; Cowan & Barron, 1987). Therefore, it seems most likely that whispering of text during the poststimulus delay period caused consonant information to be lost from memory quickly.

In future work, it will be important to examine memory for VC syllables under the conditions otherwise similar to the present experiment, in order to determine if the vowel superiority effect would be preserved or if, instead, it would be nullified (as in Experiment 2a) or even reversed by the factor of phoneme order.

General Discussion

The present study included four experiments on memory for unattended speech (1a, 2a, 3, and 4), as well as three control experiments that confirmed the validity of the reading task to which subjects attended during the session (1b, 1c, and 2b). In the first speech memory experiment (1a), subjects heard CV syllables through headphones while reading silently. Memory for consonants was found to decline rapidly across 10 s, whereas memory for vowels declined more slowly. In Experiment 2a, a different pattern of memory results was obtained for VC syllables. The difference between vowel and consonant memory decay was not reversed from Experiment 1a—it was eliminated. This suggests that the differential decay rates observed in Experiment 1a might have resulted from an advantage for the most recent phoneme in combination with an intrinsic advantage for the vowels over the stop consonants. Experiment 3 was conducted like the first experiment, except that subjects also had to monitor the audiotape for occurrences of one particular syllable. With this additional need to attend to the sounds, memory performance remained at a relatively high level across delay intervals. This finding further demonstrated that the considerable memory decay across 10 s in the first two speech memory experiments occurred because attention was directed away from the auditory channel. Finally, in Experiment 4, subjects read in a whisper rather than reading silently. The result was that vowel and consonant memory decayed at roughly comparable rates from 1 to 10 s, but with a consistently lower performance level for consonants. This difference from Experiment 1a suggests that task demands (either attention or covert articulation) contributed to the difference between vowel and consonant memory decay rates that were observed in Experiment 1a.

The most fundamental conclusion to be drawn from this research is that the memory traces of both consonants and vowels within unattended syllables of speech undergo considerable decay. This decay lasted at least 10 s, which is similar to what Eriksen and Johnson (1964) observed when using tonal stimuli. It is important to bear in mind that this memory decay occurred during silent postsyllabic periods, which goes against the assumption implicitly or explicitly made by many theorists (e.g., Massaro, 1970; Nairne, 1988; Penney, 1989) that auditory memory decay results only from the cumulative effects of poststimulus interference.

It is also clear from this research that attention-demanding processes or articulatory coding may improve memory for speech. Although this finding in itself is certainly not surprising (e.g., see Baddeley, 1986), it was interesting that subtle differences in the allocation of attention yielded important differences in memory. Speech memory performance was poorest when subjects had to quietly articulate the reading and ignore the speech sounds (Experiment 4), somewhat better when subjects read silently (Experiment 1), still better when subjects had to split their attention between the reading and the speech sounds (Experiment 3), and best when the speech sounds were attended (practice phase in all experiments). Generally, memory for consonants was most affected by these attentional factors. In future research, it might prove helpful to study task factors more directly by manipulating task demands within a single experiment.

An unanswered question in this research is the exact extent to which the observed memory functions were based on auditory sensory or "echoic" memory as opposed to a more abstract code (e.g., a phonetic memory that can accept input from either visual or auditory information about speech; for related reviews, see Cowan, 1984; Massaro, 1987; Nairne, 1988; Penney, 1989). The finding most relevant to this point is that vowel memory was superior to consonant memory when the stimuli were CV syllables, whereas this symmetry was not reversed in VC syllables; consonant and vowel memory did not differ.

The finding of overall vowel superiority is consistent with what would be expected if subjects used an acoustic memory, but it is not consistent with what would be expected if subjects used only a phonetic memory (given that consonants and vowels were identified about equally well when they were attended, in the practice phase). In support of these statements, for example, Pisoni (1973) found roughly equivalent memory performance for pairs of stop consonants as for pairs of vowels when subjects were to compare speech sounds in a pair drawn from different phonemic categories (which presumably can be done on the basis of phonetic memory), whereas vowel memory was much superior when subjects were to compare speech sounds drawn from the same category (which presumably requires an acoustic form of memory). The advantage for vowels over stop consonants may occur because vowels are steady-state sounds and are therefore acoustically simpler than stop consonants, which change rapidly over time (Crowder, 1973; Darwin & Baddeley, 1974).

Although these considerations suggest that subjects used auditory sensory memory, this does not imply that subjects used only that form of memory; they could have used auditory and phonetic memory codes in combination. Previous research has suggested that, in fact, subjects do tend to use both sources of information about speech in combination (Cheng, 1974; Cowan & Morse, 1986; Crowder, 1982; Repp, Healy, & Crowder, 1979).

The decay of auditory sensory memory is assumed to be unaffected by task demands in a silent poststimulus environment (Anderson & Craik, 1974; Broadbent, 1958; Cowan, 1984; Eriksen & Johnson, 1964; Greenberg & Engle, 1983; Morton et al., 1971), whereas nonsensory forms of speech coding can be greatly affected (Cowan, 1988; Peterson & Peterson, 1959; Vallar & Baddeley, 1982; Watkins & Todres, 1980). Therefore, we suspect that the differences in the memory decay functions obtained with different task demands in the present study reflected changes in the availability of a nonsensory form of memory, such as phonetic memory. If this is the case, then the present data need not contradict the assumption that the availability of auditory sensory memory is unaffected by task demands.

We do not know the relative weights that subjects in the present study would have attached to auditory sensory information versus phonetic information, but Figure 5 illustrates one way in which these types of information could have been combined. The intent of the figure and accompanying explanation is simply to show that an account of the present results based on auditory sensory and phonetic codes used in combination is feasible. The left half of the figure shows hypothetical functions for the decay of auditory sensory and phonetic information about the speech sounds in each memory experiment. The functions were drawn arbitrarily rather than according to any actual mathematical model (which would have been premature, given the limited evidence available). The right half of the figure contains reproductions of the actual data. In order to predict the data from the hypothesized information sources underlying performance (auditory sensory and phonetic memory), one must derive a weighted

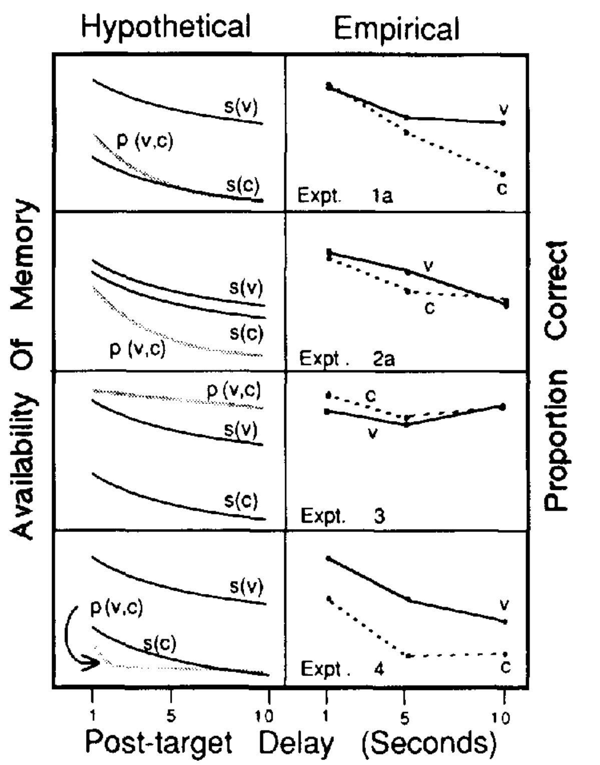

Figure 5. Hypothetical functions for the loss of auditory sensory memory and phonetic memory across delay intervals (left panels) and obtained memory functions (right panels), for Experiments 1a (top row), 2a (second row), 3 (third row), and 4 (bottom row). (V = vowels, c = consonants; s(v) = auditory sensory memory for vowels; s(c) = auditory sensory memory for consonants; and p(v, c) = phonetic memory for both vowels and consonants.)

average of these sources, attaching greater weight to whichever source is more available at a given delay. The assumption that information from multiple sources is combined, with greater weight attached to the more available or vivid source of information, is consistent with evidence from a wide variety of perceptual tasks reviewed by Massaro (1987).

In our experiments, the relative weights of auditory sensory and phonetic information would have shifted across delays in a manner that depended on the different decay rates of these memory codes. The task demands that differed from experiment to experiment would have affected primarily the decay of phonetic memory because it is assumed that auditory sensory memory decays at a fixed rate in silence (but at a level that can be affected by acoustic factors in the stimuli). In Experiment 1a, silent reading presumably resulted in a moderate rate of phonetic memory loss that would have allowed a shift in the primary source of information from phonetic memory at short delays to auditory memory at longer delays. This shift in the relative availability of the two memory codes would account for the divergence of the performance functions for vowels versus consonants because the vowel advantage presumably exists in auditory sensory memory only. In Experiment 2a, the shifts across delays would have been similar to Experiment 1a, but the lack of divergence between vowel and consonant memory decay would be attributed to the absence of a vowel advantage in sensory memory for our VC stimuli. In Experiment 3, monitoring the sounds while reading presumably stabilized phonetic memory at a level so high that auditory sensory memory never predominated at any delay. This would account for the consistently high performance across delays that was obtained for both vowels and consonants in that experiment. Last, in strong contrast to Experiment 3, the whispering task used in Experiment 4 presumably reduced the availability of phonetic memory so severely that auditory sensory memory already predominated by the 1-s delay, as in the longer delays.

In summary, we can account for the observed pattern of memory for unattended syllables in all four experiments with a model (see Figure 5) that includes at least the following assumptions: (a) that there is an auditory sensory memory with a decay rate across 10 s or more that is unaffected by task demands during a silent delay period; (b) that this auditory sensory memory, however, is affected by acoustic factors, which included a vowel advantage in our CV stimuli but not our VC stimuli; (c) that there is a phonetic memory for both consonants and vowels with a decay rate that is greatly affected by task demands, unlike auditory sensory memory; and (d) that the auditory sensory and phonetic memory codes are used in combination, with greater weight attached to whichever memory code is more available at the time of the syllabic memory test.

Out of all of our experiments, the difference between the availability of sensory memory versus phonetic memory presumably was greatest for vowel memory in Experiment 4. In that condition, sensory memory was maximized because the relevant stimuli were syllable-final, steady-state segments, but phonetic memory was minimized because of phonetic interference from the whispering task. Therefore, the vowel data of Experiment 4 would provide the purest index of auditory sensory memory decay in the present study. If the decay function observed in this condition can, in fact, be attributed to the decay of auditory sensory memory alone, then sensory memory apparently outlasts by a substantial margin the estimate that is conventionally given (i.e., several seconds). The assumption that sensory memory lasts 10 s or more is consistent, however, with previous evidence (Cowan, 1984, 1988; Eriksen & Johnson, 1964). Additional work on the decay of various features of unattended speech sounds, perhaps using computer-synthesized speech to achieve greater acoustic control, would help to clarify the role of auditory and phonetic codes in memory for speech.

References

Anderson, C. M. B., & Craik, F. I. M. (1974). The effect of a concurrent task on recall from primary memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 107–113.

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Clarendon Press.

Broadbent, D. E. (1957). Immediate memory and simultaneous stimuli. *Quarterly Journal of Experimental Psychology, 9,* 1–11.

Broadbent, D. E. (1958). *Perception and communication.* New York: Pergamon Press.

Bryden, M. P. (1971). Attentional strategies and short-term memory in dichotic listening. *Cognitive Psychology, 2,* 99–116.

Cheng, C. (1974). Different roles of acoustic and articulatory information in short-term memory. *Journal of Experimental Psychology, 103,* 614–618.

Clarke, A. C. (1968). *2001.* New York: New American Library.

Cole, R. A. (1973). Different memory functions for consonants and vowels. *Cognitive Psychology, 4,* 39–54.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104,* 163–191.

Cowan, N. (1989). The reality of cross-modal Stroop effects. *Perception & Psychophysics, 45,* 87–88.

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics, 41,* 393–401.

Cowan, N., Cartwright, C., Winterowd, C., & Sherk, M. (1987). An adult model of preschool children's speech memory. *Memory & Cognition, 15,* 511–517.

Cowan, N., Lichty, W., & Grove, T. (1988). Memory for unattended speech during silent reading. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 2, pp. 327–332). New York: Wiley.

Cowan, N., & Morse, P. A. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America, 79,* 500–507.

Crowder, R. G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behavior, 10,* 587–596.

Crowder, R. G. (1973). The sounds of speech in precategorical acoustic storage. *Journal of Experimental Psychology, 93,* 14–24.

Crowder, R. G. (1982). Decay of auditory memory in vowel discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 153–162.

Darwin, C. J., & Baddeley, A. D. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology, 6,* 41–60.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology, 3,* 255–267.

Dawson, M. E., & Schell, A. M. (1982). Electrodermal responses to attended and unattended significant stimuli during dichotic listening. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 315–324.

Eriksen, C. W., & Johnson, H. J. (1964). Storage and decay characteristics of nonattended auditory stimuli. *Journal of Experimental Psychology, 68,* 28–36.

Glucksberg, S., & Cowen, G. N., Jr. (1970). Memory for nonattended auditory material. *Cognitive Psychology, 1,* 149–156.

Greenberg, S. N., & Engle, R. W. (1983). Voice changes in the stimulus suffix effect: Are the effects structural or strategic? *Memory & Cognition, 11,* 551–556.

Harris, J. D. (1974). *Psychoacoustics.* Indianapolis: Bobbs-Merrill.

Hawkins, H. L., & Presson, J. C. (1977). Masking and preperceptual selectivity in auditory recognition. In S. Dornic (Ed.), *Attention & Performance VI* (pp. 195–211). Hillsdale, NJ: Erlbaum.

Hirst, W., Spelke, E. S., Reaves, C. C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General, 109,* 98–117.

Kallman, H. J., & Morris, M.E . (1984). Backward recognition masking as a function of ear of mask presentation. *Perception & Psychophysics, 35,* 379–384.

Massaro, D. W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology, 83,* 32–39.

Massaro, D. W. (1976). Perceptual processing in dichotic listening. *Journal of Experimental Psychology: Human Learning & Memory, 2,* 33–339.

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry.* Hillsdale, NJ: Erlbaum.

Morton, J., Crowder, R. G., & Prussin, H. A. (1971). Experiments with the stimulus suffix effect. *Journal of Experimental Psychology, 91,* 169–190.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition, 16,* 43–352.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review, 75,* 522–536.

Norman, D. A. (1969). Memory while shadowing. *Quarterly Journal of Experimental Psychology, 21,* 85–93.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17,* 398–422.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics,* 253–260.

Repp, B., Healy, A. F., & Crowder, R. G. (1979). Categories and context in the perception of isolated, steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 129–143.

Rostron, A. B. (1974). Brief auditory storage: Some further observations. *Acta Psychologica, 38,* 471–482.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150–164.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (pp. 793–811). New York: Wiley.

Slowiaczek, M. L., & Clifton, Jr., C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior, 19,* 573–582.

Treisman, A. M. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior, 3,* 449–459.

Treisman, A. M., & Geffen, G. (1967). Selective attention: Perception or response? *Quarterly Journal of Experimental Psychology, 19,* 1–17.

Treisman, M., & Rostron, A. B. (1972). Brief auditory storage: A modification of Sperling's paradigm. *Acta Psychologica, 36,* 161–170.

Vallar, G., & Baddeley, A. D. (1982). Short-term forgetting and the articulatory loop. *Quarterly Journal of Experimental Psychology, 34A,* 53–60.

Watkins, M. J., & Todres, A. K. (1980). Suffix effects manifest and concealed: Further evidence for a 20-second echo. *Journal of Verbal Learning & Verbal Behavior, 19,* 46–53.

Received June 7, 1989
Revision received July 27, 1989
Accepted July 31, 1989 ■