JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **15**, 17–31 (1976)

Acoustic Masking in Primary Memory

HERBERT A. COLLE AND ALAN WELSH

University of Chicago

Primary memory (or short-term memory), according to models such as Sperling's, maintains a substantial amount of its information by storing it in an auditory sensory memory. Since the auditory sensory memory is used to store memory information these models predict that concurrent auditory stimulation should destroy memory information and, hence, reduce recall performance. To test this hypothesis, a foreign language was presented over earphones while subjects performed a serial recall task with visual presentations and written recall. The subjects were told to ignore the noise. In Experiment I the presence of the irrelevant foreign language noise reduced recall performance on phonologically different lists but it did not reduce performance on phonologically similar lists. Passive articulatory restraint had little effect. In Experiment II this noise effect was eliminated after 30 sec of silent arithmetic, indicating that the noise effect is a primary memory phenomenon.

At least three different lines of research have suggested that short-term or primary memory estimates are particularly sensitive to the sound structure of the material to be remembered. The phonological similarity effect (Baddeley, 1972; Shulman, 1971), the phonological intrusion phenomenon (Conrad, 1964; Wickelgren, 1965), and the modality effect (Murdock & Walker, 1969; Crowder & Morton, 1969) all have been used to argue for the phonological nature of primary memory. The exact mechanisms by which these three effects are produced from a phonologically coded primary memory have not been fully explicated; the effects may not be attributable to a single mechanism (Watkins, Watkins, & Crowder, 1974). In addition, serial recall, probe recall, and free recall tasks may not all measure the same memory systems (Baddeley & Hitch, 1974). The following discussion is probably more conservatively generalizable only to serial and probe recall tasks which require the retention of order information.

A short-term memory system which uses phonological coding has a number of advantages. On the one hand, it is a code which is simple in comparison to the code which would seem to be necessary to store every one of the manifold kinds of distinctions that people can remember for extended periods of time. On the other hand, it can contain almost all of this information when it is used in conjunction with a linguistic decoding system. It can be decoded to make any distinction which spoken language permits.

The assumption of a separate memory store with phonological coding is attractive for another reason. It is possible that primary memory consists of a collection of components which exists primarily to serve other functions. Sperling's (1967) Model III utilizes components in this fashion. In his model, estimates of information in primary memory arise from information stored in either the auditory sensory memory or the R-buffer. The auditory sensory memory is a sensory memory analogous to its visual counterpart (Sperling, 1960). Its normal purpose would be to buffer

This research was supported by grants from the Social Sciences Divisional Research Fund, University of Chicago, and from the University Research Council, Wright State University. Requests for reprints should be sent to Herbert A. Colle, Department of Psychology, Wright State University, Dayton, Ohio 45431.

incoming auditory stimulation so that subsequent perceptual mechanisms have adequate time to perform their analyses. The R-buffer is a small memory which codes instructions for articulatory movements. The existence of a small memory for articulatory movements could exist as a part of the mechanism which produces context-sensitive speech. Sperling (1967) assumed that when the commands in the R-buffer are executed explicitly, speech would be produced. The resulting auditory stimulation would be stored temporarily in the auditory sensory memory. He also assumed that the commands in the R-buffer could be executed implicitly, an internal speech. Furthermore, it was assumed that this implicit execution placed auditorily coded information in the auditory sensory memory, utilizing the auditory memory without the presence of acoustic energy. With the addition of the capability to read information from the auditory memory into the R-buffer, the model attempted to describe the mechanism underlying the ability to maintain information in primary memory by the use of rehearsal. Thus these two components which are specialized for the reception and production of auditory information may be responsible for the phenomenon of primary memory, and the resulting short-term recall of auditory and visual linguistic stimuli.

The model proposed by Sperling makes a strong prediction. Since the information which supports the short-term recall estimates of primary memory must pass through the auditory sensory memory, it should interact with and be degraded by any other information also using that memory at the same time. If the auditory sensory memory works like the visual sensory memory, it is largely preattentive and many different kinds of acoustic stimuli should produce auditory memory effects. Hence, irrelevant acoustic stimuli should mask the memory representative in the auditory sensory memory, producing decrements in the primary memory estimates. Studies by Hintzman (1965) and Murray (1965) did not support the prediction. Both reported that the presence of gaussian noise did not impair short-term recall performance. There are at least two possible reasons why these two studies failed to find a noise effect. They will be called the *filter* and the *compensation* explanations.

The filter explanation assumes that not all incoming auditory information is processed in the same way by the auditory system, and not all of it is processed to the same extent. If an irrelevant noise source is filtered out before it reaches the level at which the recoded visual information is stored, then no interaction would take place, and hence, no degradation would be expected. In other words, not all masking stimuli are equally effective. For example, gaussian noise has been found to be a relatively poor backward masking stimulus for vowels (Dormen, Kewley-Port, & Turvey, Note 1) and also for tones (Holding & Loeb, Note 2). A similar distinction between patterned and unpatterned masks has been described for visual masking studies (Haber, 1970; Turvey, 1973). Evidence for differences in the mode of processing linguistic vs. nonlinguistic information also has been obtained from dichotic listening experiments (Kimura, 1961, 1964). Therefore, it is possible that Hintzman and Murray failed to find a noise effect because the gaussian noise they used was not an effective masking stimulus.

Another possible explanation of the ineffectiveness of the noise stimulus in Hintzman's and Murray's experiments is that the subjects were able to compensate for the presence of the noise. One of the phenomena that can be observed when people perform delayed short-term recall tasks with noise in their ears is that some of them will rehearse the list aloud (Sperling, 1967). One possible explanation for the appearance of this articulatory activity is that it compensates for the increased noise level. For example, executing the R-buffer explicitly produces speech. When a room is noisy, the loudness of speech increases and little loss of intelligibility results. If a similar phenomenon exists when the R-buffer is executed implicitly, then the absence of a performance decrement could occur because of a more vigorous implicit execution of the R-buffer. In Sperling's model this increase might lead to a clearer or stronger auditory sensory memory representation. As the noise is made more intense, explicit articulatory activity should start to occur, producing the speech which is observed.

Murray (1965) did vary the magnitude of articulatory involvement. He argued that it contributed nothing above that which would have been expected from the presence of the acoustic stimuli which were generated by it. However, the magnitude of articulatory involvement was manipulated by enhancing overt rehearsal above its normal levels. Since the overt rehearsal which is elicited by the presence of noise seldom gets very loud, the enhanced overt rehearsal which Murray used may have been increasing articulatory activity above the level which was already maximally effective. A more stringent test of the compensation explanation would be to try to reduce even normal articulatory involvement.

EXPERIMENT I

In this experiment an attempt was made to test the filter and the compensation hypotheses. To minimize the possibility that the noise source would be filtered out, linguistic material was used as a noise source. To minimize the likelihood that this linguistic material would be rehearsed, confused with the material to be remembered, or processed semantically, a foreign language was used. To further minimize these problems, the instructions emphasized that the foreign language was to be ignored. To test the compensation hypothesis more stringently by reducing normal articulatory activity, the subjects performed, on one half of the trials, while holding a tongue depressor in their mouths in a relatively restrictive way. The noise and tongue depressor conditions were combined factorially.

The filter and the compensation hypotheses predict two different patterns of results. If the previous failures to find a noise effect were due to the use of an inadequate masking stimulus, if the foreign language is an effective mask, and if compensation does not take place, then the data should display a simple noise effect. The presence of noise should depress recall performance without regard to the presence or absence of articulatory restrictions. However, if compensation takes place and if the articulatory restriction reduces its effectiveness, then correct performance should only be depressed when both the noise is present and articulation is restricted.

Since masking may depend on the phonological similarity of the mask to the target material, the similarity of the list members also was varied. The results suggest that the phonological similarity effect and the noise effect are closely related phenomena.

Method

Subjects. Forty-eight University of Chicago students volunteered to serve as subjects. The first 24 students responded to advertisements and to classroom requests and received no payment or class credit. The second 24 were paid $2.00 for the session. None of the students reported understanding any German or Yiddish. There were 20 males and 28 females.

Stimulis and materials. Two different sets of letters were used as stimuli to be recalled. The phonologically different set (PD) consisted of the letters F, K, L, M, Q, R, X, Y. The phonologically similar set (PS) consisted ot the letters B, C, D, G, P, T, V, Z. All letters were printed in upper case and were displayed visually. For each trial a list was generated by sampling randomly without replacement from the set of letters. Two sets of 48 lists were generated in this way, one for each set of letters. Lists containing familiar combinations of letters were discarded and replaced by others. Each of the letters were used about the same number of times at each serial position for the 40 lists used on the experimental trials.

For the noise conditions a passage from "A Hunger Artist" by Franz Kafka was played binaurally over earphones at roughly 85 dB sensation level. The passage was chosen because it had a low occurrence of English sounding words in it. It was recorded by a female speaker. The tape was obtained from the University of Chicago's language center.

Procedure. The subjects were randomly assigned to use either the similar or the dissimilar lists of letters. In each of these groups the subjects were tested under all four of the main experimental conditions. The four conditions were obtained by factorially combining the two noise conditions (noise, silence) and the two articulation conditions (tongue depressor, no tongue depressor). Each of the four experimental conditions was studied in a distinct block of 12 trials. The first two trials of each block were discarded to reduce practice effects. Each of the 24 permutations of the testing order was used with one of the subjects. Also for both paid and unpaid subjects, testing order was balanced so that each experimental condition occurred equally often in each of the four blocks of trials. The subjects were given five practice trials prior to any of the blocks. The practice trials were presented without noise and without articulatory restrictions.

During the noise conditions the German passage was presented binaurally over earphones for the entire block of trials. The subjects were told before the block of trials that the German would be presented. They were instructed to ignore the noise and to concentrate entirely on the serial recall task. For the silence conditions, the subjects wore the earphones, but no sound was presented. They were told that no noise would be presented.

The two articulatory conditions varied in the extent to which articulation was normal. For the restricted articulation condition (tongue depressor), each subject had a tongue depressor placed flat between his tongue and lower teeth. It was inserted until the corners of the mouth were slightly drawn back. The tongue was extended as fully as possible under these constraints and the upper teeth were brought to rest upon it. The tongue was held in this position by the subject. The position was an unusual one for the articulatory organs and also considerably restrained articulatory movement. The tongue depressor was lubricated to prevent it from adhering to the underside of the tongue.

The tongue depressor was inserted prior to the relevant block of trials and was held that way until the end of the block. In the normal articulation condition, the tongue depressor was not used. However, the subject never was allowed to speak aloud. An experimenter sitting about three feet in front of the subject monitored his speech. If the subject made audible sounds during a trial, that trial was discarded and the subject was asked to refrain from speaking. The discarded trial was replaced without the subject's knowledge by another at the end of the block of trials.

The following events took place on each trial. Following a warning signal, the letters from the eight-letter list were presented sequentially for 1.2 sec each. A 10-sec unfilled delay followed the termination of the last letter and the end of the delay was indicated by visually signaling the start of the recall period. Written serial recall was used. The subject wrote from left to right on a recall sheet which contained eight blank lines. The subjects were allowed to retrace leftward to change an answer but they were not allowed to fill in a blank on the right side until all of the preceding blanks to its left were filled. No blanks were allowed; guessing was forced. A cover sheet prevented the subject from inspecting his recall of previous lists during later trials.

Results

Phonologically different lists. The number of errors made at each serial position was obtained for the four experimental conditions and they were analysed with a mixed model analysis of variance. The factors containing correlated measures were (*a*) the presence or absence of the irrelevant German (noise vs. silence), (*b*) the presence or absence of articulatory restrictions (tongue depressor vs. no tongue depressor), and (*c*) serial position. The factor containing independent measures was the type of subjects (paid vs. unpaid). All of the treatment effects for correlated measures were evaluated by reducing the degrees of freedom to form conservative tests to protect against covariance heterogeneity (Greenhouse & Geisser, 1959).

Figure 1 presents the proportion of incorrect responses made at each serial position for the four treatment conditions. As the figure shows, there was a large effect produced by the irrelevant noise, $F(1, 22) = 104$, $p < .01$. The percentage of errors increased 12% when the noise was present (N) compared to when there was silence (S). This effect was approximately the same at all serial positions; no noise by serial position effect was obtained, $F(1, 22) = 2.11$, $p > .05$.

Clearly, the data in Fig. 1 provide no support for the compensation hypothesis. According to this hypothesis more errors should be made when the presence of noise is combined with the presence of the tongue depressor (TD), which artificially restricts articulatory movements. The other three conditions should produce the same level of better performance. Although it appears from the curves in Fig. 1 that the tongue depressor increased the number of errors in the silent condition more than in the noise condition, there was no interaction between the presence or absence of noise and the presence or absence of the tongue depressor, $F(1, 22) = 2.30$, $p > .05$. However, the overall effect of the presence of the tongue depressor was significant, $F(1, 22) = 5.96$, $p < .05$, and the interaction of the tongue depressor effect with the type of subjects also was significant, $F(1, 22) = 4.94$, $p < .05$. There was no suggestion of an error increase due to the presence of the tongue depressor in the paid subjects' data. For the unpaid subjects, 5.6% more errors were made with the tongue depressor present than were made when it was absent.

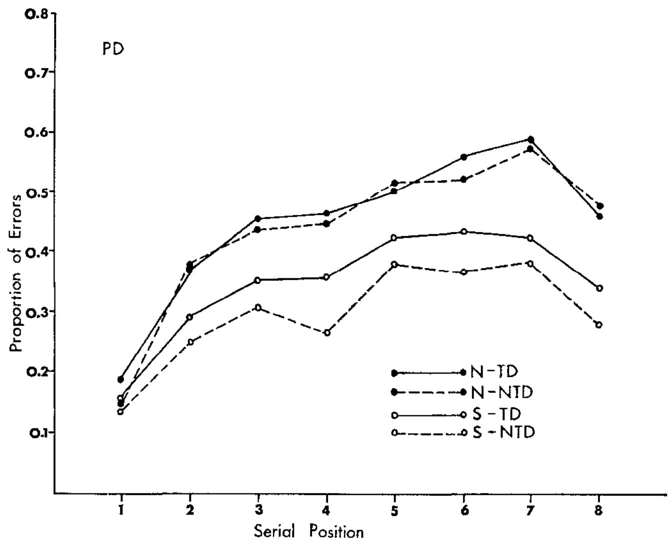

FIG. 1. A comparison of performance on the phonologically different lists (PD) with or without the noise (N vs. S) and with or without the tongue depressor inserted (TD vs. NTD).

As expected, performance varied considerably over serial position $F(1, 22) = 40.7$, $p < .01$. The curves are typical serial recall curves for visual stimuli (Morton, 1970). Serial position did not interact with any of the other treatments. None of the other interactions were significant at the 5% level of confidence.

Phonologically similar lists. The phonologically similar lists were analysed in the same way as the phonologically different lists were. Figure 2, which is directly comparable to Fig. 1, summarizes these data. As Fig. 2 shows, there was no noise effect, $F(1, 22) = 2.02$, $p > .05$. Also, neither the presence of the tongue depressor, $F(1, 22) = 1.11$, $p > .05$, nor its interaction with the noise conditions, $F(1, 22) < 1.0$, produced an increase in the frequency of errors. Serial position was the only variable which produced a major effect, $F(1, 22) = 42.3$, $p < .01$. All of the other main effects and interactions were not significant at the 5% level of confidence.

To compare performance on the similar and different lists the data were collapsed over the serial position and paid-unpaid factors. An analysis of variance was performed using list similarity as a between groups factor and noise treatment and articulatory restriction as factors with correlated measures. Degrees of freedom were reduced again to form conservative tests. The only effects which were significant at the 5% level of confidence were the noise effect $F(1, 46) = 23.0$, $p < .01$ and the list similarity by noise interaction $F(1, 46) = 11.8$, $p < .01$. This interaction can be seen easily in Table 1. When there was silence, the typical phonological similarity effect was demonstrated. Performance was better for the phonologically different lists than it was for the phonologically similar lists. However, with

TABLE 1

THE INTERACTION OF THE NOISE CONDITIONS WITH LIST SIMILARITY

Noise condition	Phonological similarity: Different	Phonological similarity: Similar
Silence	32.3	44.2
Noise	44.3	46.2

Note: Scores represent the percentage of incorrect responses.

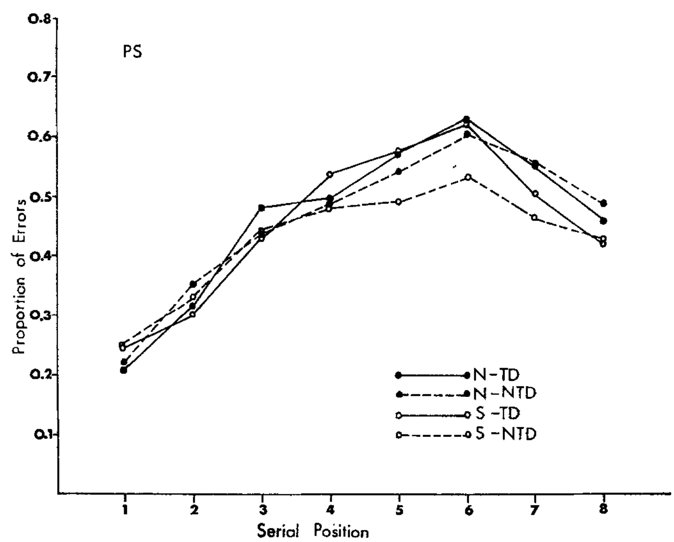

FIG. 2. A comparison of performance on the phonologically similar lists (PS) with or without the noise (N vs. S) and with or without the tongue depressor inserted (TD vs. NTD).

noise present, there was no difference in performance for the two types of lists. It is the pattern of results which would be expected if list similarity and irrelevant noise had a redundant influence on the memory representation.

Discussion

The discovery of a noise effect supports theories such as Sperling's (1967) which assume that primary memory estimates arise at least in part because of information that is implicitly sent to and maintained in an auditory sensory memory. During no part of the recall task was there any acoustic stimulation present which could have been masked by the irrelevant noise. The stimuli were presented visually and recall was written. In addition, the subjects were prevented from speaking aloud at any time during a trial. There was no opportunity for explicit sensory masking.

A problem that must be addressed is the selectivity of the noise effect. Perhaps the noise only distracted the subjects, making them miss the stimulus presentations. The data from the phonologically similar lists argue against this interpretation. Remembering phonologically similar lists is a more demanding task than remembering phonologically different lists, yet no noise effect was produced. If the noise was a simple distracting stimulus it should have produced an increase in errors for both situations. The performance of the groups which received the similar lists was still considerably above the chance level, so that ceiling effects cannot be evoked. The data suggest that it is the specific relationship between the sound source and the task being performed that is important. Gibson and Yonas (1966) found that a male voice speaking numbers had no effect on visual scanning performance. Hence, speech noises should not be expected to disrupt performance on all tasks; only those in which the auditory sensory memory component is used.

Since there was no evidence for the existence of compensation, the data suggest that Hintzman (1965) and Murray (1965) failed to find a noise effect because of an inadequate noise stimulus. However, there were other differences between those studies and the present one. Hence, the relative efficacy of various sound sources should be evaluated more systematically.

Experiment II

Although the presence of the irrelevant German reduced short-term recall performance in the first experiment, the locus of that effect was not established. A number of investigators (Baddeley, 1972; Waugh & Norman, 1965) have argued that immediate short-term recall confounds recall from primary memory with recall from secondary (or long-term) memory. In the first experiment the reduced performance could have been due to a loss of information from primary memory, as the above argument suggested. On the other hand, the presence of noise may have interferred with the subjects' ability to transfer information to secondary memory. If the noise impairs the acquisition of information by the secondary memory system, then the noise effect should not be reduced after primary memory information is eliminated by an unrelated intervening task. If, on the other hand, the noise effect occurs because information is lost from primary memory then the noise effect should be eliminated when recall follows an unrelated intervening task.

It is especially important to determine if the noise effect obtained in Experiment I is localized in primary memory because some investigators (e.g., Morton, 1970) have claimed that output interference in serial recall eliminates the primary component. If the claim is true, it would mean that the noise effect is a secondary memory phenomenon. Because very little is known about the dynamics of serial recall, the serial position curve cannot be used to determine the relative contributions of primary and secondary memory as it sometimes can for free recall or probed recall performance. Therefore the Brown–Peterson approach of using an unrelated arithmetic task was used to obtain an estimate of the secondary memory component. Although the distractor technique is not universally accepted as an adequate method for getting unbiased estimates of the primary and secondary memory components (Watkins, 1974; Wickelgren, 1973), it is a standard technique which has been used and which can be readily applied to serial recall. It should provide at least a rough estimate of the relative contributions of primary and secondary memory to the noise effect.

Method

Subjects. The subjects were 24 female undergraduate and graduate students at the University of Chicago. They were recruited by advertising and were paid $2.00 for participating. None of them reported knowing any German or Yiddish.

Procedure. The procedure used in Experiment I was retained. The phonologically different set of letters was always used. They were presented for 1.2 sec each and were followed by an unfilled 10 sec delay. Serial recall of the list was used again. On one half of the trials recall was allowed immediately after the 10 sec delay; on the other half the subject was given a three digit number and was required to subtract threes from it for 30 sec. A visual cue at the end of the 30 sec interval reminded the subject that she was to write down the number she was currently subtracting and then attempt to recall the list correctly. Unlike the procedure normally followed when arithmetic is used as intervening activity, the subjects did not subtract aloud. Instead, they subtracted silently and wrote down the final result before recalling.

There were four phases to each experimental session. In the first phase the subjects practiced serial recall for 10 trials in silence without any intervening arithmetic activity. In the next phase they practiced their arithmetic aloud. A three digit number was presented and the subjects counted backwards aloud by threes until told to stop. They were encouraged to subtract rapidly and accurately. There were five of these trials. Following the second phase, two experimental phases occurred. Each phase consisted of a block of 20 trials. One half of the subjects were randomly assigned to receive the noise during the first block and silence during the second block. The other half of the subjects received the noise during the second block and silence during the first block. For the noise blocks the irrelevant German was played continuously throughout the entire block. On one half of the trials in a block, recall was allowed immediately after the unfilled delay. On the other trials, 30 sec of arithmetic intervened between the unfilled delay and recall. The arithmetic and the no-arithmetic trials occurred in an irregular order. The first four trials of each block were discarded. They always consisted of two trials with arithmetic and two trials without. Subjects were not allowed to talk aloud during a trial and were instructed to ignore the irrelevant sounds.

Results

Planned mutually orthogonal comparisons were used to determine if the noise effect was still present when recall took place after an intervening arithmetic task. The analysis was performed on the total number of errors made at all serial positions, disregarding serial position information. The three main comparisons which were tested were (*a*) noise vs. silence without arithmetic, (*b*) noise vs. silence after arithmetic, and (*c*) arithmetic vs. no arithmetic, regardless of the noise conditions. The analysis included an order of testing effect to determine if the order of testing the noise and silence blocks influenced overall performance or interacted with any of the comparisons. Conservative tests, using reduced degrees of freedom, were used again to test each of the comparisons and their interactions with testing order.

Without arithmetic, the presence of the noise produced a substantial increase in the percentage of incorrect answers; the percentage of errors increased from 27.7 to 39.8%, $F(1, 22) = 15.7$, $p < .01$. This 12% increase approximates the increase observed in Experiment I. However, after the intervening arithmetic activity, recall performance during the presence of the noise was not significantly different from performance in silence, $F(1, 22) = 1.04$, $p > .05$. Figure 3 shows these data for each serial position. After the arithmetic, only 3% more errors were made during the noise than were made during the silence.

The intervening arithmetic task produced a considerable increase in the frequency of errors, $F(1, 22) = 107$, $p < .01$. Thus, the intervening unrelated activity successfully reduced the primary memory component (Baddeley, 1972; Waugh & Norman, 1965). Figure 3 shows that this increase in errors occurred equally at all serial positions. A factorial analysis revealed that the presence or absence of an intervening task did not interact with serial position, $F(1, 22) = 1.48$, $p > .05$. Thus the loss of information from primary memory does not influence the serial position curve selectively. Serial position, also, did not interact with any of the other factorial effects, although the main effect of serial position was again statistically significant, $F(1, 22) = 24.0$, $p < .01$.

To estimate the number of subtractions which were made during the silent arithmetic period, the number which was reported before recall was attempted was subtracted from the starting number and the difference was divided by 3. Five of these estimates were discarded as being anomalous.[1] The mean number of subtractions per trial was 13.2 and 12.9 for the silence and noise conditions, respectively. These estimates were not significantly different from each other, dependent $t(23) = 0.71$, $p > .05$.

There was a small effect of testing order, $F(1, 22) = 4.36$, $p < .05$. The people who received a block with silence before a block

[1] The five estimates were dropped because the number reported was larger than the starting number. Three of them were from the noise condition and two were from the silence condition. Subjects, even when subtracting aloud, make errors in the most significant digit, especially when they reduce the second most significant digit.

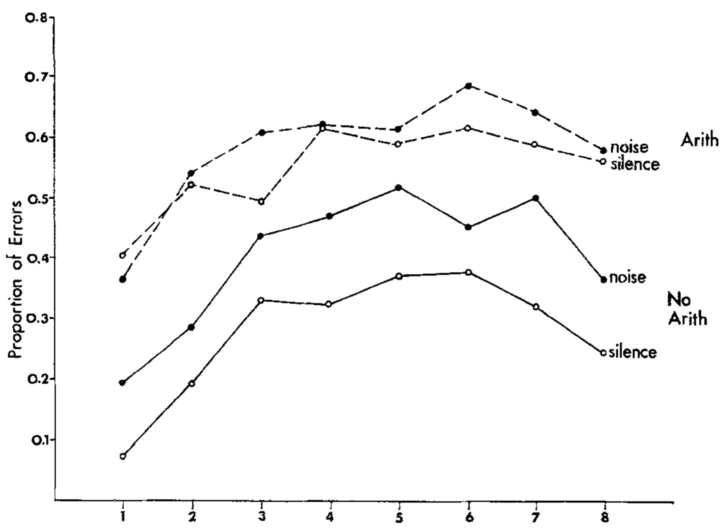

FIG. 3. The influence of noise on the recall of phonologically different lists when recall occurred after an arithmetic task was performed and when recall occurred before this filled delay.

with noise made more errors in both blocks than people who received the opposite order. However, testing order did not interact with any of the planned comparisons; all three F ratios were less than 1.0.

Discussion

The results support the conclusion that the noise effect is located in primary memory and not in secondary memory. The noise effect was no longer present after the irrelevant arithmetic activity. If it is assumed that the contribution of secondary memory can be estimated by recall performance after considerable irrelevant arithmetic (Baddeley, 1972; Waugh & Norman, 1965), then the absence of the noise effect implies that it is not acting upon secondary memory. On the other hand, the noise effect was obtained when the arithmetic task was not performed, replicating Experiment I. By exclusion, the effect must occur in primary memory unless a third memory store is postulated.

The absence of a noise effect after intervening activity also argues against the hypothesis that the noise only distracts people, causing them to fail to attend to the stimuli properly. Since the trials with arithmetic occurred irregularly and since the subject was not informed about whether or not she would be performing arithmetic until after the stimulus presentation, all stimulus presentations, on the average, should have been treated equivalently. Hence, failure to find a noise effect when recall followed the intervening activity argues against a simple distraction hypothesis.

Although it was not explicitly tested, it is apparent from Fig. 3 that even when the irrelevant German was present recall performance was substantially better without arithmetic than it was after the arithmetic, implying that the presence of the German did not eliminate all of the primary memory information. The inability of the German to mask all of the information could arise because only part of the primary memory estimate is obtained from information sent to the auditory system. Part of the information could be stored, say, as a visual or an articulatory code. On the other hand, not all masks are completely effective and, therefore, all of the information being sent through the auditory system may not have been masked. This alternative will be discussed more fully below.

Contrary to the assumption made by Morton (1970), as estimated by the distractor technique, considerable primary memory information is used to support the serial recall performance of lists which are just supraspan. Unlike the free recall and probed recall of long lists, the serial position curve for the serial recall of just supraspan lists provided little information about the relative amount of information supported by primary and secondary memory. Like the Brown–Peterson paradigm, an extraneous task must be used to obtain an estimate of the secondary memory component.

In both experiments an attempt was made to prevent the subjects from presenting themselves with acoustic information by vocalizing aloud. The results support the effectiveness of these procedures. There were no differential effects over serial position, effects which would be expected if the presence of vocalizing selectively produced acoustic stimulation for some of the experimental conditions. All of the curves had very small, if any, recency effects, indicating that the effective stimulation was not produced by self-generated acoustic stimuli (Crowder & Morton, 1969; Morton, 1970).

General Discussion

The results from both experiments are consistent with the hypothesis that one component of primary memory is an auditory memory and that even visual information can be recoded and sent there. The phonological similarity effect and the phonological intrusion phenomenon may owe their existence to this auditory component. Since the phonological similarity effect was eliminated by the presence of the irrelevant spoken German, some interrelationship between the two would seem to be indicated. If the similarity effect is produced by the destruction of nonauditory information, then it would be expected to be additive with the noise effect. Since the two effects acted redundantly, they both appear to be operating on the same information.

The term "auditory memory" has been used in a broad sense to indicate that a system normally used to process incoming acoustic stimuli somehow retains information momentarily. As discussed in the introduction all acoustic stimuli may not be processed in the same way or to the same extent. The procedure utilized to obtain the noise effect provides a method for isolating the part of the system in which the auditory information is held. By systematically varying the characteristics of the noise as well as the type of task performed, the interaction of the two can be determined. The method assumes that information can only be destroyed by other incoming information when they simultaneously utilize the same system processes. Unfortunately both auditory perception and selective attention are incompletely understood. However, the auditory system is known at least well enough to allow crude comparisons. For example, what is the difference between gaussian noise and spoken German which may cause only the latter to impair short-term recall performance? Could it be that distinctive features or phonemes are extracted from the German while they cannot be extracted from the gaussian noise? Or, it is the relative concentration of energy in the speech frequencies which accounts for the difference?

Although the similarity and noise effects appear to be closely related, a number of investigators also have shown that both the similarity effect and the intrusion phenomenon depend upon articulatory activity at the time of stimulus presentation. If an unrelated phrase or word is articulated with the material to be recalled is presented, then the similarity effect and the intrusion phenomenon are eliminated or reduced (Estes, 1973; Levy, 1971; Murray, 1967, 1968; Peterson & Johnson, 1971; Tell, 1972). In most of these studies the subjects vocalized aloud, confounding the presence of irrelevant acoustic stimulation with the articulatory involvement. However, Levy (1971) has found that similar results occur when people only mouthed an irrelevant phrase without saying it aloud. Using the logic applied above, these results would seem to argue that the similarity and intrusion effects depend upon an articulatory memory.[2]

A Classification System

In order to discuss the question of articulatory versus auditory encoding the status of three major assumptions must be clarified. The following terminology is introduced so that different memory models may be classified according to their coding assumptions, permitting a discussion of them as classes rather than as individual models. The classification is not to be construed as a theoretical proposal of a new model.

The three major coding assumptions made by primary memory models deal with (*a*) the number of codes utilized, (*b*) the interrelationship of these codes, and (*c*) the manner in which the codes can arise. A model may assume that information in primary memory can be represented by a *single code* such as a phonemic code (Sperling & Speelman, 1970). On the other hand, primary memory may be *dually* or *multiply coded* (Laughery, 1969);

[2] Cheng (1974) has reported that heard and mouthed similarity acted differently across serial positions, suggesting that different mechanisms underlie the two variables. Indeed, it could be interpreted as consistent with the separate auditory memory proposed by Crowder and Morton (1969). However, the study is difficult to interpret because all of the effect is due to one group which had the most unusual task to perform in under one second, making CV syllables out of normally pronounced letters. The type of articulation performed was confounded with the amount of processing that was required. The necessity for active processing is a problem that exists to some extent whenever forced articulation is used.

Sperling, 1967). For example, there may be auditory coding and no articulatory coding or articulatory coding and no auditory coding. Alternatively, both may exist together.

If more than one type of code exists, then assumptions about their interrelationships must be made. *Independent codes* are those in which the status of one code for a stimulus does not depend upon the present status of the other codes for that stimulus. That is, if one code is incorrectly encoded or distorted, the other memory codes will not change because of it. Information may be entered or lost from one without affecting the other. With *dependent coding* one code is assumed to change with changes of another one. The nature of the dependence would be specified by the model.

Coding mechanisms may be further classified as explicit or implicit. A coding process will be called *explicit* if it is assumed that the code can be present only when the physical energy which could normally produce it is present. Thus auditory coding only could occur when the appropriate acoustic energy had been present, and articulatory coding would depend upon the presence of actual articulatory movements, and so on. *Implicit* coding, on the other hand, will refer to the assumption that information received in one modality may be recoded into a form which would normally arise when energy is received by another modality without this energy actually being present. For example, the assumption that an auditory code can arise from visual presentations in the absence of acoustic energy is an implicit coding assumption.

Auditory versus Articulatory Coding

The existence of the noise effect is consistent with the assumption of implicit visual to auditory recoding. Whether there is also implicit visual to articulatory or auditory to articulatory recoding was not directly tested. However, since both irrelevant articulations and irrelevant noise have produced effects which are redundant with similarity, a redundancy between the noise effect and the effect of articulatory activity is suggested.

A redundancy between irrelevant noise and articulation is consistent with the close dependent relationship between the R-buffer and the auditory sensory memory assumed in Sperling's (1967) model. Since implicit execution of the R-buffer sends the recoded information to the auditory sensory memory, the recoded information should act as the same information would if it had been presented auditorily. The appearance of overt rehearsal during the presence of noise suggests that articulatory activity does occur implicitly. Although it could be argued that the articulatory activity does not occur without noise, arising only in response to the introduction of noise, this argument suggests a compensatory function for the activity and evidence for compensation was not found. Thus the articulatory activity probably occurs implicitly in silence, becoming explicit in noise because the absence of acoustic feedback eliminates the normal control process which held it in check.

The Modality Effect and Auditory Memory

As Morton (1970) has argued, the modality effect creates problems for component models such as Sperling's (1967) which assume that visual stimuli are implicitly recoded and stored in an auditory memory. Crowder and Morton (1969; Morton, 1970) have argued that the modality effect arises because recall of acoustically presented material has the advantage of being supported by a more stable auditory memory. Acoustic suffixes were shown to reduce the enhanced recency effect obtained when lists are presented acoustically. The destruction of the modality effect by acoustic suffixes strongly supports their argument that the increase is produced by information stored in an auditory memory, although alternative explanations do exist (Kahneman, 1973; Massaro, 1972). Morton's (1970) model assumes that auditory memory is coded explicitly. With acoustic presentations there is independent dual coding; with visual presentations an articulatory "response buffer" is singly coded. If the suffix effect supports the existence of a separate explicitly coded auditory memory, Sperling's (1967) assumption of implicit dependent dual coding is contradicted. If the utilization of auditory memory produces the advantage of auditory over visual stimuli then the auditory memory must not be available to visually presented stimuli. On the other hand, the noise effect is inconsistent with Morton's (1970) logogen model. If auditory memory is coded explicitly then how could irrelevant noise overwrite visually presented lists?

It may be possible to reconcile the two models. The above argument assumes that auditory memory is a unitary entity. However, if acoustic stimuli undergo differential processing, then a suffix may cause overwriting only if it is processed similarly. The magnitude of the suffix effect does depend upon a considerable number of relationships between the list members and the suffiix. For example, the suffix effect is reduced if a different voice speaks the suffix, if the suffix is not localized at the same place as the list, and if gaussian rather than speech is used as a suffix (Crowder, Morton, & Prussin, 1971). These differences would seem to imply that the magnitude of the effect depends upon the level or type of processing received by the list and the suffix. Similarly, the noise effect appears to depend upon the type of noise used. In previous studies no noise effect was found when gaussian noise was used (Hintzman, 1965; Murray, 1965). Perhaps the noise effect and the suffix effect utilize the same auditory memory system but use it in different ways. If visual presentations are recoded auditorily they would not necessarily be in the same voice as an experimentally presented acoustic suffix nor would they be localized in the same place. When people are asked to localize their silent speech they report that their "voice" appears to originate from a location inside the head about midway between the ears (Weber & Bach, 1969). If these reports are taken seriously, then visual and auditory presentations may be stored in different parts of an auditory memory. Only acoustic presentations that reach the part of auditory memory used by silent speech would affect recall performance.

This description of auditory memory could help to explain the relative ineffectiveness of certain kinds of stimuli. Crowder, Morton, and Prussin (1971) emphasized the importance of selective attention for explaining the results of different suffixes. The approach outlined above (which depending upon the nature of selective attention is not necessarily contradictory) emphasizes that separate storage locations may be activated by different kinds of stimuli. If two stimuli are stored in separate locations, overwriting would not occur. This could help to explain the differences between stop consonants and vowels (Crowder, 1973). Only when there were vowel differences did the modality and suffix effects occur. Vowels, because of their energy concentration at certain frequencies, differ in pitch as tones do. Roffler and Butler (1968) have shown that, in the midline, high pitches are localized higher than lower pitches. This effect occurs in children before they learn the names "high pitch" and "low pitch". Likewise von Békésy (1971) described how singers use the subjective vertical location of their voice (inside the head) as a technique for tone control. If these subjective localizations correspond to separate memory locations, then pitch information could act as a carrier, determining the vertical location of incoming auditory stimuli. Thus pitch changes which occur in vowels and in phonemic phrases may aid language processing by preventing overwriting, and, hence, expanding the capacity of auditory memory.

By utilizing different auditory memory locations the traces of activity left in the auditory system, taken in the broad sense, could appear to be a separate auditory sensory memory. The hypothesis can be tested by comparing the perceptual processing received by different classes of stimuli as determined by psychophysical criteria and comparing their ability to produce a noise effect in primary memory experiments. For example, the potential importance of localization for auditory memory is demonstrated by the effect of ipsilaterial vs. binaural suffixes on a monaurally presented list. Crowder et al. (1971) found that a binaural suffix was less effective than an ipsilateral one. This phenomenon is analogous to the masking level differences obtained with direct auditory masking (e.g., Egan, 1965). Localization[3] of the stimuli is one of the mechanisms which has been proposed as an explanation of these differences (Hafter, 1971). In the same vein, if the implicitly executed R-buffer creates auditory "voices" in the midline (Weber & Bach, 1969), then a monaural noise ought to produce a smaller noise effect than a binaural noise would.

References

Baddeley, A. D. Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin*, 1972, **78**, 379–385.

Baddeley, A. D., & Hitch, G. Working Memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8). New York: Academic Press, 1974.

von Békésey, G. Localization of visceral pain and other sensations before and after anesthesia. *Perception & Psychophysics*, 1971, **9**, 1–4.

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

Conrad, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, **55**, 75–84.

Crowder, R. G. Representation of speech sounds in precategorical acoustic storage. *Journal of Experimental Psychology*, 1973, **98**, 14–24.

Crowder, R. G., & Morton, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, **5**, 365–373.

[3] To simplify the above discussion the distinction between localized and lateralized stimuli was not utilized.

Crowder, R. G., Morton, J., & Prussin, H. A. Experiments with the stimulus suffix effect. *Journal of Experimental Psychology Monograph*, 1971, **91**, No. 1.

Egan, J. P. Masking-level differences as a function of interaural disparities in intensity of signal and of noise. *Journal of the Acoustical Society of America*, 1965, **38**, 1043–1049.

Estes, W. K. Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 360–372.

Greenhouse, W. W., & Geisser, S. On methods in the analysis of profile data. *Psychometrika*, 1959, **24**, 95–112.

Gibson, E. J., & Yonas, A. A developmental study of the effects of visual and auditory interference on a visual scanning task. *Psychonomic Science*, 1966, **5**, 163–164.

Haber, R. N. A note on how to choose a visual noise mask. *Psychological Bulletin*, 1970, **74**, 373–376.

Hafter, E. R. Quantitative evaluation of a lateralization model of masking-level differences. *Journal of the Acoustical Society of America*, 1971, **50**, 1116–1122.

Hintzman, D. L. Classification and aural coding in short-term memory. *Psychonomic Science*, 1965, **3**, 161–162.

Kahneman, D. *Attention and effort.* Englewood Cliffs: Prentice-Hall, 1973.

Kimura, D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 1961, **15**, 166–171.

Kimura, D. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 1964, **16**, 355–358.

Laughery, K. R. Computer simulation of short-term memory: A component-decay model. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 3). New York: Academic Press, 1969.

Levy, B. A. Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 123–132.

Massaro, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, **79**, 124–125.

Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

Murdock, B. B., Jr., & Walker, K. D. Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 665–676.

Murray, D. J. The effects of white noise on the recall of vocalized lists. *Canadian Journal of Psychology*, 1965, **19**, 333–345.

MURRAY, D. J. The role of speech responses in short-term memory. *Canadian Journal of Psychology*, 1967, **21**, 263–276.

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

PETERSON, L. R., & JOHNSON, S. T. Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 346–354.

ROFFLER, S. K., & BUTLER, R. A. Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America*, 1968, **43**, 1260–1266.

SPERLING, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, **74**, No. 11.

SPERLING, G. Successive approximations to a model for short-term memory. *Acta Psychologica*, 1967, **27**, 285–292.

SPERLING, G. & SPEELMAN, R. G. Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

SHULMAN, H. G. Similarity effects in short-term memory. *Psychological Bulletin*, 1971, **75**, 399–415.

TELL, P. M. The role of certain acoustic and semantic factors at short and long retention intervals. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 455–464.

TURVEY, M. T. On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, 1973, **80**, 1–52.

WATKINS, M. J. The concept and measurement of primary memory. *Psychological Bulletin*, 1974, **81**, 695–711.

WATKINS, M. J., WATKINS, O. C., & CROWDER, R. G. The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior* 1974, **13**, 430–447.

WAUGH, N. C., & NORMAN, D. A. Primary Memory. *Psychological Review*, 1965, **72**, 89–104.

WEBER, R. J., & BACH, M. Visual and speech imagery. *British Journal of Psychology*, 1969, **60**, 199–202.

WICKELGREN, W. A. Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology*, 1965, **70**, 102–108.

WICKELGREN, W. The long and the short of memory. *Psychological Bulletin*, 1973, **80**, 425–438.

(Received June 19, 1975)

REFERENCE NOTES

1. DORMAN, M. F., KEWLEY-PORT, D., & TURVEY, M. T. *Backward and forward masking of brief vowels*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November, 1973.
2. HOLDING, D. H., & LOEB, M. *Postcategorical masking by tones or noise*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November, 1973.