1976

Verbal short-term storage and analysis-by-synthesis of speech : evidence for common mechanisms

James W. Aldridge
Binghamton University--SUNY

VERBAL SHORT-TERM STORAGE AND ANALYSIS-BY-SYNTHESIS OF SPEECH: EVIDENCE FOR COMMON MECHANISMS

BY
JAMES WINSTEAD ALDRIDGE, JR.

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in State University of New York
at Binghamton
1976

Accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in State University of New York at Binghamton.

NAME	DEPARTMENT	SIGNATURE	DATE
C. James Scheirer	Psychology	C. James Scheirer	1/15/76
Richard G. Burright	Psychology	Richard Burright	1/15/76
Richard E. Pastore	Psychology	Richard E. Pastore	1/15/76
Jane M. Connor	Psychology	Jane M. Connor	1/15/76
Jerrold Aronson	Philosophy	Jerrold Aronson	1/15/76

TO
JERE BEAUCHAMP
AND
ROBERT BJORK,
WHO TAUGHT ME
THE MEANING
OF
PATIENCE

CONTENTS

TABLES

FIGURES

Since the late 1950's a great deal of research has been directed toward elucidating the process(es) by which humans retain verbal material for short intervals of time. The purpose of the present paper will be to explore the extent to which such memorial processes can be identified with processes which may be employed in information processing tasks in general, and speech perception and/or production in particular.

"Limited Capacity" Conceptions

The problem of the "immediate memory span" or the "span of apprehension" has concerned experimental psychologists for a very long time (see Woodworth & Schlosberg, 1954, for a review). The research to be considered in this section is based on a procedure introduced by Pillsbury and Sylvester in 1940. The general procedure involves a single presentation of material to be remembered, requiring subjects to engage in an apparently unrelated "distracting" task during a short retention interval, and then testing the effect of the intervening task on subsequent recall of the memorized material. Pillsbury and Sylvester found that 10 seconds of intervening rest led to better recall of eight items than 10 seconds of arithmetic or of question-answering.

More recently, the distractor method has been reintroduced in two forms. In one procedure, (Brown, 1958; Peterson & Peterson, 1959), subspan lists of items are used. In the other procedure, longer lists are presented for free recall, and the effect of subsequent distraction on the recency portion of the retention curve is examined (Postman & Phillips, 1965). The most frequent forms of distraction used in subsequent research have been those used by Brown (reading subsequent items aloud) and Peterson

and Peterson (counting backward, usually by threes). Numerous experiments have manipulated the difficulty of verbal distracting tasks (e.g., reading numbers vs. arithmetically transforming them; counting backward by threes vs. counting backward by sevens), and a consistent finding has been that increasing apparent difficulty in such ways results in a retention decrement (Posner & Rossman, 1965; Posner & Konick, 1966; Bruning, Shappe, & O'Malley, 1966; Talland, 1967; Merikle, 1968; Scheirer & Voss, 1970; Lowe & Merikle, 1971; Kroll & Kellicutt, 1972)[1].

Other investigators have found that difficult <u>non</u>verbal retention interval tasks may result in a retention decrement. Crowder (1967 a,b) found that manipulating stimulus lag in a sequential multiple-alternative reaction time task reduced recall of previously presented word trigrams, and that decreasing stimulus-response compatibility reduced recall of noun quintagrams. Watkins, Watkins, Craik, & Mazuryk (1973) found that manually "shadowing" four-alternative sequences of piano notes produced substantial forgetting of noun quintagrams in a Brown-Peterson paradigm, with the amount of forgetting increasing over one- to twenty-second retention intervals. In a second experiment, they found that 20 seconds of pursuit rotor activity following a free recall list eliminated the serial position curve's end peak if presentation had been visual, and substantially reduced the peak if presentation had been auditory.

Results such as those outlined above have led several theorists to conclude that active maintenance of material in memory requires the use of some mechanism which is of limited processing capacity. "Limited capacity" explanations of short-term memory performance have taken two basic forms.

One form of limited capacity conception places the limitation in a "channel". This hypothesis was proposed by

Broadbent (1958,1963, 1971), and has been invoked by other authors (Murdock, 1964, 1965 a,b; Rabbitt, 1966,1968). Although Broadbent (1971) has recently expanded the model greatly, the basic idea has remained as originally proposed in 1958. The model involves a short-lived sensory buffer, followed by a limited capacity channel through which material must pass before further processing. Short-term memory phenomena are said to appear when items are repetitively passed from a buffer through the channel, returned to a buffer, and repassed through the channel for as long as the material needs to be retained. Retention loss occurs when an intervening task occupies the channel and material in the buffer is allowed to decay. The feature of Broadbent's system which differentiates it from other limited capacity notions is the idea that the limitation is on a particular mechanism which is necessarily involved in processing incoming information beyond a very peripheral level. That is, there is a fixed processing "bottleneck" at some point in the flow of information through the organism.

The second form of limited capacity hypothesis involves a limited amount of "attention", "mental energy", or "central processing capacity" which is freely allocable to diverse kinds of processing (Posner & Rossman, 1965; Posner & Keele, 1967; Moray, 1967; Posner, 1967,1973; Posner, Boies, Eichelman, & Taylor, 1969; Posner & Boies, 1971; Kerr, 1973; Craik & Lockhart, 1972; Kahneman, 1973). Although it is sometimes stated that certain tasks require this central capacity while others do not (Posner & Boies, 1971; Kerr, 1973), the limited capacity processor itself appears to be considered relatively undifferentiated. That is, tasks are said to draw upon it as a function of their general difficulty or attentional demands, and the suggestion has been made that the processor may be identifiable with consciousness (Posner & Klein, 1974; Craik &

Lockhart, 1972). Within this second framework, short-term memory phenomena are said to appear when attention required to keep remembered material "active" or "in mind" is demanded by the intervening distraction task.

Although the limited capacity hypotheses seem attractive in light of the results reviewed above, two recent sets of experiments seem difficult to interpret in terms of a unitary processor in which capacity limitations are determined by attentional requirements. Reitman (1971), using noun trigrams in a Brown-Peterson paradigm, required subjects to engage in one of three auditory signal detection tasks during a 15 second retention interval. In one condition, subjects were required to detect the presence of 100 msec threshold-level square wave segments embedded in noise. In the other conditions, they engaged in a verbal discrimination task, monitoring a string of repeated synthetic syllables ("doh") for the occurrence of another, similar syllable ("toh") while both vocalizing the target syllable and pressing a key upon detection or while only key pressing. Reitman found that while the discrimination task produced substantial forgetting (with vocalization adding a slight further decrement), the tonal detection task produced almost no forgetting. Shiffrin (1973) subsequently found very little forgetting of consonant quintagrams when subjects engaged in a tonal detection task during retention intervals up to 40 seconds.

In order to deal with Reitman's and Shiffrin's results within a general "attention" or "central capacity" framework (see above), one would be forced to conclude that either the memory task or the detection task (or both) required little attentional capacity. This conclusion is complicated, however, by Reitman's findings with her verbal discrimination conditions. It would follow from the general capacity notion that the verbal task required "mental energy",

while the nonverbal task did not. It is not apparent why this should be the case, since Reitman equated her filler tasks for difficulty (as measured by performance). It also seems intuitively reasonable that conscientious performance on any threshold-level detection task would demand a great deal of "concentration" or attention.

Broadbent's conception seems potentially more capable of explaining Reitman's and Shiffrin's results, if the limited capacity channel can be identified with a process which would not have been involved in the tonal detection conditions. Note that in all of the studies reviewed in previous paragraphs, activities which have been found to be effective in producing short-term memory loss have involved complex motor responses, either verbal or nonverbal. In contrast, while Reitman's and Shiffrin's tasks required difficult perceptual discriminations, the response requirements (pressing a single key based on a yes-no decision) seem minimal. It may be possible that the limited capacity channel is in fact a response selection or organization mechanism, which was not required in the tonal task. Such an interpretation, however, still does not explain the interference produced by the verbal discrimination task, unless there is reason to believe that analyzing the syllables required the use of a response selection mechanism. In the following section, evidence will be reviewed from speech processing research which suggests that there _is_ reason to believe that in some situations, analysis of speech stimuli does involve processes which are also involved in response production. Following this, a system for storage of verbal material will be outlined which is neither incompatible with Broadbent's position nor with Reitman's and Shiffrin's results.

Analysis-by-synthesis

Analysis-by-synthesis theories have enjoyed considerable popularity among speech researchers in recent years, primarily as a result of efforts by theorists at M.I.T. (Halle & Stevens, 1962; Stevens & House, 1972; Stevens, 1973) and at Haskins Laboratories (e.g., Liberman, 1957; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Cooper, 1973). The particular theories differ in detail, but all share the idea that speech is somehow analyzed perceptually by reference to the production process. For example, the Halle and Stevens (1962) model suggests that after feature extraction, the perceiver forms an hypothesis as to the identity of incoming verbal material. This hypothesized item is then generated by the perciever, and if the synthesized material matches the original item, the item is specified by the procedures used to generate the match. The generation process does not necessarily involve the peripheral articulatory apparatus, but the speech production system is assumed to be involved, at some level. The Haskins group appears to lean more toward stimulus specification by means of articulatory commands, and the term "motor theory of speech perception" seems more closely associated with their conception. The term "analysis-by-synthesis" seems more closely associated with the M.I.T. group, which appears to lean toward specification at a more abstract phonetic level. These distinctions, however, need not concern us here, and "analysis-by-synthesis" as a generic term will be used in reference to both conceptions[2].

The case for perception of speech by reference to articulation has been presented in detail by several authors (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Studdert-Kennedy, Liberman, Harris, & Cooper, 1970;

Stevens & House, 1972) and will not be repeated here. Suffice it to say that some of the evidence considered by the above authors to provide the strongest support for the conception has recently been called into question (Kirman, 1973; Miller, Pastore, Wier, Kelley, & Dooling, 1974; Cutting & Rosner, 1974; see also Lane, 1965). However, other evidence suggests that while perception of speech may not differ qualitatively from the perception of other acoustic stimuli in many situations, the processing of speech may involve a motor component in some circumstances. Specifically, research with apparent cerebral laterality effects suggests motor involvement when (1) the stimulus information is reduced or ambiguous; and/or (2) when the stimulus is encoded phonetically. Research supporting these conclusions will now be briefly reviewed.

When engaging in a dichotic speech identification task, right-handed people generally identify stimuli presented to the right ear more readily than stimuli presented to the left ear (e.g., Kimura, 1961a). The ear showing the dichotic advantage seems almost always to be that contralateral to the hemisphere controlling speech production, as determined by injection of sodium amytal (Kimura, 1961b). Subjects showing a right ear advantage for speech, however, typically show either a left ear advantage or no advantage for music (Kimura, 1964; Gordon, 1970; Spreen, Spellacy, & Reid, 1970; King & Kimura, 1972), or nonspeech human sounds such as coughing or crying (King & Kimura, 1972) or intonation contours of sentences (Blumstein & Cooper, 1974). Thus, speech processing seems sometimes to be lateralized to the same hemisphere as speech production. One factor governing the magnitude of the laterality effect seems to be the degree to which identification of a speech stimulus depends on surrounding context. Fricatives may be specified either by their

initial steady state friction, which is invariant across vowel context, or by the transition to a following vowel formant, which varies with vowel context. Darwin (1971) found that when synthetic fricatives were specified by context-dependent transitions, a right ear advantage was obtained in a dichotic identification task. The advantage was not obtained when the same fricatives were specified by context-invariant initial friction. Both Darwin (1971) and Haggard (1971) found a right ear advantage for vowels whose identification required consideration of the speaker's identity, but not for vowels which could be identified without such consideration. It therefore seems possible that in situations where the stimulus is made ambiguous by dependency on context, hypothesis testing of the sort proposed by Halle & Stevens (1962) takes place. That is, when a stimulus is only partially specified by some of its acoustic characteristics, the observer may generate possibilities by means of the production system.

Haggard & Parkinson (1971) argued that laterality effects may depend more on the subject's task than on the stimulus itself. This conclusion was based largely on an experiment in which dichotic identification of the emotional tone of sentences yielded a left ear advantage. Subsequent results have strongly suggested the same conclusion. Muraski & Sharf (1973) found a dichotic right ear advantage in the identification of /p/,/t/, and /k/ when subjects were instructed to identify the stimuli as speech, but obtained a non-significant left ear advantage when the same stimuli were identified as being of high, low, or medium pitch. Spreen, Spellacy, & Reid (1970) obtained a left ear advantage for tonal patterns in a recognition paradigm, but Halperin, Nachson, & Carmon (1973) found a right ear advantage for the same types of stimuli in a paradigm requiring subjects to name the patterns.

Other investigators have obtained similar results measuring auditory evoked potentials from the right and left hemispheres. Wood, Goff, & Day (1971) and Wood (1975) presented the same consonant-vowel stimulus (e.g.,/ba/) to subjects in differing task contexts. In one task, subjects were required to discriminate between the experimental syllable and a phonemically different stimulus (e.g.,/da/). The potential evoked by the experimental syllable at the left hemisphere differed reliably between the two task contexts, but the potential at the right hemisphere did not. Matsumiya, Tagliasco, Lombroso & Goodglass (1972) found little, if any, lateralization of evoked potential magnitude for either sound effects or words when subjects were instructed to simply record the frequency of sound occurrence without regard for sound identity. The magnitude lateralization increased reliably toward the left hemisphere when subjects were instructed to determine the number of sound effects that could be named, and was greatest when meaningful sentences were presented.

Finally, Sussman (1971) and Sussman, MacNeilage, and Lumbley (1975) have reported a right ear advantage in a task which is completely nonverbal, but which clearly involves stimulus analysis with reference to the state of articulatory structures. In an auditory pursuit task, when the frequency of the cursor tone was controlled by movements of the tongue or jaw, it was found that performance was better when the cursor was delivered to the right ear than the left. The right ear advantage was much diminished and unreliable when cursor frequency was controlled by hand movements.

In summary, it appears that some component of the analysis of auditory stimuli is lateralized to the speech production hemisphere when (1) the sound is closely related to articulatory movements; (2) the stimulus is context-dependent and therefore ambiguous; or (3) the sound is

encoded as a linguistic unit. While the evidence that these three sources of right ear advantages reflect the same underlying process is circumstantial, it is certainly tempting to interpret them within a single framework involving similar use of speech production processes. It has already been suggested that ambiguous information may elicit hypothesis testing of the sort proposed by Halle and Stevens (1962). Since a name is in fact an articulatory response, it also seems tenable that encoding a stimulus as a linguistic unit (naming it) involves reference to articulation.

A System For Verbal Storage

Consider three processes which are likely to be involved in performance of any information processing task involving an overt response. First, the incoming material must undergo some initial perceptual analysis. Secondly, an appropriate response to the analyzed material must be selected. Finally, the response, once selected, must be executed. Next, consider how the same three processes might be used for short term retention of verbal material. First, incoming items are analyzed perceptually. Secondly, a decision must be made involving the selection of an appropriate response to the material; i.e., an appropriate name (articulatory response) is synthesized (executed) by use of <u>some level</u> of the speech production system. The product of the synthesis is then reanalyzed, the name reselected, the response resynthesized, etc. The analysis and synthesis loop can be executed iteratively for as long as the material needs to be retained. This is not to say that the response in question will necessarily be a fully actualized, overt articulation, but merely some component of the speech motor organization/ execution process (cf. Liberman, 1957; Cooper, 1973).

The evidence most strongly suggesting that storage can involve a production component like that outlined above is exceedingly straightforward: people sometimes tend to overtly repeat items that they are trying to remember (Sperling, 1963; Hintzman, 1965). Even when the repetition is not overt, EMG recordings indicate that the speech musculature is active while verbal material is being held in memory (e.g., Locke & Fehr, 1970, 1972). In addition, it is well known that short-term memory confusions are most frequently based on phonemic similarity (Conrad, 1962, 1964), and that the precise nature of the confusions is well predicted by linguistic "distinctive feature" systems (Wickelgren, 1965, 1966; Klatt, 1968; Cole, Haber & Sales, 1968, 1973; Cole, Sales, & Haber, 1969, 1974; Sales, Cole, & Haber, 1969, 1974; Sales, Haber, & Cole, 1969; Cole, 1973 a,b, 1974).

When a system which can be used for short-term storage of verbal material is identified with operations involved in information processing tasks in general, a conceptual framework results within which specific kings of interference may be localized. Consider the initial analysis component of the model. Craik and Lockhart (1972) have suggested that modality-specific short-term memory phenomena (e.g., Kroll, Parkinson, & Parks, 1972; Crowder & Morton, 1969) appear when incoming information is held at a perceptual level of processing, rather than being analyzed by post-perceptual processes. Although the present system lends itself well to such an analysis, this paper will be primarily devoted to situations in which the entire perception-selection-synthesis sequence seems involved in storage.

Existing data are ambiguous with regard to the question of whether subsequent perceptual processing interferes with the retention of verbal material that has been processed

beyond a perceptual level. The results of several investigations suggest that intervening perception does not result in memory loss. As mentioned in the previous section, neither Reitman (1971) nor Shiffrin (1973) found any storage loss produced by signal detection tasks which, while extremely difficult in terms of the perceptual discrimination, involved very simple responses. Other experimenters have found that the simple occurrence of noise in a retention interval (with no required response) had no effect on memory (Sperling, 1963; Sloboda, 1969; Sloboda & Smith, 1968).

Others have reported results which might be interpreted as suggesting that interpolated sounds can produce some interference with retention. Using fifth grade children as subjects, Newell (1968) found that an unexpected scream during a retention interval reduced the number of letters recalled from a consonant quintagram by more than 50%, relative to recall after unfilled intervals. Watkins, Watkins, Craik, & Masuryk (1973) presented noun quintagrams followed by a retention interval which was either unfilled, contained sequences composed of four notes, or filled with a task in which subjects were required to "shadow" the note sequences by pressing four telegraph keys. After 20 second retention intervals, mean recall scores (estimated from their Figure 1) were 4.25 words after unfilled intervals, 4.0 words after intervals during which the tone was presented but responding was not required, and 2.0 words after intervals during which the tones were "shadowed". Although the decrement due to tone presence alone was small, it was statistically reliable. Reitman (1974) and Leshowitz, Zurek, & Robbins (1974) have recently found some forgetting in a Brown-Peterson situation using tonal distractors. Reitman used five-word memory items and tonal detection as distraction. Under nonrehearsal instructions, recall scores after 15 seconds of distraction were only

71% of the values obtained for immediate recall. Leshowitz, Zurek, & Robbins (1974) found forgetting of word trigrams only by some subjects after detection or intensity discrimination tasks up to 30 seconds in duration.

Although the above results may suggest that some interference with verbal retention may occur as a result of purely perceptual distraction, it is not completely clear that this is the case. The surprise scream may have elicited any number of covert or overt reactions from the children Newell (1968) used as subjects. Watkins, et. al. (1973) presented the tonal stimuli to be ignored on trials randomly mixed with "shadowing" trials, and signalled the subjects to avoid responding only after the memory material had been presented. It is quite possible that this procedure led to response selection by the subjects for the first few tones, even though the responses were not overtly executed. This interpretation is supported by their findings that the ignored tones produced as much interference during 3 second retention intervals as they did during 20 second retention intervals.

Apart from problems of interpretation raised by post-hoc subject classification in the Leshowitz, et. al. (1974) study, subjects received a monetary bonus for rapid responding to interpolated stimuli. It is possible that requiring rapid responding causes even the decision involved in a simple yes-no response to produce measurable interference. Reitman's results may have been entirely due to nonrehearsal instructions, rather than to the detection task. Recall scores after the detection task with instructions to concurrently rehearse were 98% of immediate recall values. In any case, Reitman (1974) found 44% more forgetting when a letter discrimination task was used during retention of the five-word items.

Thus, the results of the above experiments do not suggest interference with the perceptual component of the

memory loop as strongly as might be surmised at first glance. It is possible that after initial perceptual processing of the memory material has been completed, the processing of other material at a perceptual level does not interfere with storage. Whatever perceptual-like processes are involved in storage may not be susceptible to interference by concurrent perceptual processing of external stimuli. The possibility cannot be excluded, however, that interference may occur at the perceptual level, if either (1) the sounds to be discriminated are acoustically more complex than the nonverbal sounds used thus far; or (2) the sounds to be discriminated are speech stimuli, and thus are highly perceptually similar to the memory items. Reitman found large interfering effects of an intervening speech discrimination task in both the 1971 and 1974 studies. These possibilities will be explored further in Experiments 1 and 2 of the present paper.

From the data considered thus far, it appears clear that tasks which involve complex verbal or nonverbal responses produce forgetting in short-term memory experiments. In terms of the model proposed above, it seems possible that the response selection component is a very general mechanism used to some degree by any motor response, whether verbal or nonverbal. Thus, tasks which involve complex responses interfere with memory. It is not at all clear that perceptually demanding tasks produce interference. Why, then, have investigators consistently found that difficult interpolated discrimination tasks involving speech stimuli produce large recall deficits (Rabbitt, 1966, 1968; Reitman, 1971, 1974)? The possibilities that the interference resulted from the acoustic complexity of speech or from its similarity to the memory material have already been mentioned. There are other possible explanations, however. Evidence has been reviewed suggesting

that there is reason to believe that, in certain circumstances, the processing of verbal material involves a motor response component. In fact, the input analysis, response selection, and response execution processes which were proposed for the storage system seem readily identifiable with the feature extraction, hypothesis selection, and synthesis processes of analysis-by-synthesis (e.g., Halle & Stevens, 1962). If the storage and analysis-by-synthesis models are valid and refer to the same underlying system, they suggest that interference between remembering old and analyzing new verbal material might occur at all three stages of processing. In addition to requiring the same input or feedback analyzers, the two tasks might compete for use of the response selection mechanism, and/or for use of whatever component of the execution process might be involved in the two tasks. In this context, Experiments 2,3, and 4 were undertaken in the attempt to localize the interference with verbal memory produced by interpolated verbal processing.

Experiment 1

Experiment 1 examines the effect on short-term memory of some extraverbal differences between Reitman's (1971, 1974) tonal stimuli and her syllable stimuli. Unlike her tonal detection condition, her discrimination condition involved processing approximately 30 discrete syllables per trial and making a yes/no decision about each item. It is possible that this discrete-item decision task engages a qualitatively different analytic process from that involved in waiting for a signal's occurrence, unrelated to the fact that the items Reitman used happened to be verbal. For example, the occurrence of the stimuli at regular rates would seem to result in defined detection intervals, a procedure which is known to affect signal

detection performance (Egan, Greenburg, & Schulman, 1964). Secondly, the verbal stimuli were acoustically more complex than the stimulus in the tonal task. If perceptual and memorial processing involve capacity limitations determined by the number of "features" composing a stimulus (Massaro, 1972), one might expect an interpolated stimulus more complex than the tones used by Reitman (1971) and Shiffrin (1973) to produce a larger retention decrement.

Method

Subjects. Subjects were 20 undergraduate and 16 graduate psychology students at the State University of New York at Binghamton. Sixteen of the undergraduates received course credit for their participation. None of the subjects were trained psychophysical observers, but previous participation in memory research was permitted.

Design. The 36 subjects were randomly assigned to four groups of nine, with the restriction that the graduate students were equally distributed among the groups. A Brown-Peterson procedure was used with a 16 second retention interval. Groups differed in terms of retention interval filler activity.

Retention interval tasks. The signal detection group engaged in a simple detection task similar to those used by Reitman (1971, 1974) and Shiffrin (1973). Signals were 70 msec segments of a 1000 Hz. square wave which were to be detected in broad-band noise.

Subjects in the frequency discrimination group monitored a series of discrete 70 msec 1000 Hz square wave pulses for occurrences of a pulse whose fundamental was slightly above 1000 Hz in frequency. Pulses were presented at the rate of 1/sec. Although performance data indicate that the discrimination was difficult (see below), pulses were presented at a comfortable intensity without

masking noise and were clearly audible. This condition was intended to assess the effect on memory of a task involving decisions about each member of a series of discrete nonverbal items.

The <u>clatter discrimination</u> group listened for changes in an acoustically complex stimulus. The stimulus was the clatter produced by the rotation of plexiglass boxes containing sheet metal screws. This stimulus has been previously found to elicit investigation by rats (Aldridge & Burright, 1974). Six such boxes were constructed, each mounted on the shaft of a separate 10 rpm electric motor. The peak intensity of the sound produced by each box during a 1 min period was measured on the A scale of a General Radio 1551A sound level meter set at "fast", with a General Radio 1560-P6 microphone located approximately 30 cm from the center of the array of boxes. Peaks of 51 db were obtained for each of three of the boxes, 52 db for each of two others, and 54 db for the remaining box. When all six boxes were operated for 1 min, the peak measured intensity was 57 db. During a retention interval, either one, two, or three of the boxes rotated continuously. The signals to which the subjects responded were two second intervals during which the remaining five, four, or three boxes were also rotated. Thus, "signal" always consisted of the operation of a variable number of boxes. By increasing the number of boxes producing "noise" (thereby decreasing the amount of difference between the "noise" and the "signal"), task difficulty could be varied to obtain an approximately equal performance level for all subjects. The boxes were rectangular (4.5 X 4.0 X 7.5 cm) and were rotated about the long axis. Since a complete rotation required six seconds, the sound produced by one box was fairly regular with intensity peaks as the screws struck the small face of the box after sliding along the long face. When the number of

boxes rotating was changed, the sound changed along several dimensions, the most noticable of which were temporal characteristics, quality, and intensity. This condition was intended to assess the effect on memory of processing a complex nonverbal stimulus during retention.

Finally, in the letter discrimination group, subjects listened to a string of spoken "Bs" in noise for occurrences of spoken "Ps". The letters were spoken in a monotone at a rate of 1/sec by a male native speaker of English.

Materials and Apparatus. Memory items were single trigrams with a Witmer association value of less than 30% (Underwood & Schultz, 1960). Presentation of stimuli was accomplished using digital logic and a set of solid state timers connected to reed switches. The equipment was operated by control signals recorded on two track magnetic tape and played on a Wollensak 6250 stereo recorder. Each channel of the recorder was connected to the logic and timers via independent Schmitt triggers.

The experiment required three tapes. For the letter discrimination condition, memory items and discriminanda were recorded on the same tape as the control signals. For the other three conditions, only memory items and control signals were recorded. Because of differences in the nature of the control signals required, the clatter condition and letter condition required separate tapes. The same control tape was used for the signal detection condition as for the frequency discrimination condition. In the clatter condition, the clatter boxes were located approximately 10 cm behind the subject's chair and .5m above the floor. All other stimuli were presented binaurally over high-fidelity Koss Ko 727 B stereo headphones after mixing, either from the tape or from tone or noise generators. In the clatter condition, subjects heard the memory items over the headphones. The headphones were then automatically disconnected when the clatter was operating.

In the letter discrimination condition, discriminability was manipulated by adjusting the intensity of the white noise before mixing with the letters. White noise was produced by an Elgenco 624A Gaussian Noise Generator. Each letter was independently spoken: stimuli were not repetitions of one another.

For the detection and frequency discrimination groups, a timer-controlled reed switch repetitively closed for 70 msec and opened for 930 msec. In the signal detection condition, the tone was delivered to the headphones when the switch closure occurred during a 930 msec control pulse from the tape recorder. With this apparatus, a few intended signals failed to be delivered, when the control pulse began just as the switch opened and ended just as the switch closed. So that the experimenter would be aware of the number of signals actually delivered, a signal light located out of the subject's view was illuminated for the duration of any signal actually delivered to the headphones. If a signal failed to be delivered, or was seriously shortened, the light failed to reach its full brightness. When this happened, the event was considered a nonsignal. Although this method of signal presentation is somewhat crude, signals which were not delivered or were shortened were rare, and the apparatus seems adequate for present purposes. In the signal detection condition, discriminability was manipulated by adjusting the intensity of the signal by means of a Hewlett- Packard 350D Attenuator Set before mixing with the noise.

The instrumentation for the frequency discrimination group was similar to that for the signal detection group, and the same control tape was used. For both conditions, the square waves were produced by a Wavetek 131A VCG Generator. For the frequency discrimination group, a 1000 Hz tone was delivered to the headphones whenever

the 70 msec switch was closed during a retention interval. A control pulse from the tape recorder caused the voltage to be increased at the VCG input of the Wavetek, which in turn increased the frequency of the signal. The problems with signal shortening discussed above were also present in the frequency discrimination condition. Discriminability in this group was manipulated by adjusting the "signal" VCG input voltage. This was accomplished using the Hewlett-Packard attenuator.

In all conditions, subjects indicated a detection of the appropriate retention interval signal by pressing a telegraph key. Responses were manually recorded by the experimenter, who monitored a display indicating occurrences of signals, hits, and false alarms.

Procedure. Each trial began with a recorded warning burst of white noise one sec in duration. A recorded trigram then occurred during a 2 sec period, and was preceded and followed by .5 sec of silence. The 16 sec retention interval then began, followed by a 5 sec recall interval. Seven sec elapsed between the end of the recall interval and the next warning burst. The recall interval was signaled to the subject by a small light next to the telegraph key which remained on for the 5 sec. No feedback was given regarding either memory performance or interpolated task performance.

Interpolated task difficulty was roughly equated across subjects by a pretest with each subject which determined the discriminability level producing a 50% hit rate with a minimal false alarm rate. The pretest involved 21 16-sec intervals during which either three, four, or five signals occurred. Pretest trials were presented in the same manner as memory trials, except that no memory items were presented. Since some pilot subjects had been unable to perform the clatter discrimination, subjects used in the clatter condition of the experiment were allowed

to watch the boxes for three 16 sec sample trials before the pretest. This visual pre-exposure allowed performance on the clatter task to reach a level comparable to that in the other conditions.

Following difficulty level adjustment, subjects were instructed regarding memory trials. The instructions emphasized that during a retention interval the subject should "concentrate completely" on the filler task and not "think about" the memory letters in any way. A verbatim sample set of instructions (that used for the clatter group) appears in Appendix A.

After instructions there were 38 trials. The first six and last six trials contained no memory items. These trials were used to measure performance on the filler task with no memory load. During a retention interval, 0-5 signals could occur. The number of signals per trial was randomly determined, with the restriction that each block of six trials contain exactly one trial with each of the possible frequencies of signal occurrence (although due to apparatus limitations discussed above, it was impossible to rigidly adhere to this restriction for the frequency discrimination and signal detection groups). The time of occurrence for each signal within a retention interval was randomly determined, with the restriction that signal onsets be at least 3 sec apart. Subjects were not informed of these restrictions. Each subject experienced, in the same order, the same trigrams, signal frequencies, and signal times of occurrence.

Recall of the trigrams was verbal, and any responses made after the 5 sec recall interval were scored as incorrect. On the clatter discrimination task, the first response made during a two sec signal interval was scored as a hit, with other responses scored as false alarms. For the other conditions, a hit was a response occurring in an interval 140-1140 msec after onset of a signal.

Results and Discussion

The first six memory trials were considered practice and were not included in the data analysis, so mean recall scores are based on 20 trials. Graduate student means were nearly identical to undergraduate means, so the data from the two populations were pooled. Individual recall scores from all experiments are included in the appendices.

Trigram recalls were scored as correct only if all letters were recalled in their original order. Subjects recalled all three letters in an inappropriate order on less than 1% of the trials. This scoring procedure, while typically used in Brown-Peterson experiments, is considerably more stringent than the procedure used by Reitman (1971) and Shiffrin (1973). A scoring procedure analogous to that used by these investigators would involve reporting the proportion of total letters recalled. For example, if a subject erred only on a single letter on each of two trials in the present experiment, his recall score would be 90% using the present procedure, but 98% using the procedure of Reitman and Shiffrin. The former procedure is used in all experiments reported in this paper to facilitate comparison with most other short-term memory research.

Memory performance. The results of the recall test were quite unambiguous. Mean recall was 91.2% for the signal detection group; 89.8% for the frequency discrimination group; 94.0% for the clatter discrimination group; and 68.9% for the letter discrimination group. An overall analysis of variance applying specific comparisons to the groupings of interest indicated a highly reliable difference between the letter discrimination group and the other groups, $F(1,32) = 22.31$, $p < .001$, while the nonverbal groups did not differ from one another, $F < 1$. The overall standard error was $\pm$ 4.2%.

Interpolated task performance. The main reason for recording interpolated task performance scores was to insure that subjects conscientiously performed the tasks during concurrent retention. It was expected that the untrained observers would produce data too variable to be of any further theoretical use. Although there was a fair bit of "slop" in the data, there was enough consistency to suggest that the verbal memory load affected the letter discrimination task differently than it affected the nonverbal tasks.

Mean hit and false alarm probabilities on the interpolated tasks are shown in Table 1. False alarm rates were computed on the basis of responses in one second intervals during which no signal occurred. The groups differed reliably in both hit rate, $F(3,32) = 3.09$, $p < .05$, and false alarm rate, $F(3,32) = 8.14$, $p < .001$. The frequency discrimination group had the highest hit rate, and the letter group had the highest false alarm rate. Neither of these variables was reliably associated with recall within any group. Pearson product-moment correlations between hit rate and recall were -.51 for the signal detection group; +.54 for the frequency discrimination group; -.49 for the clatter discrimination group; and -.23 for the letter discrimination group, all $p > .10$. Correlations between false alarm rate and recall were -.39 for the signal detection group; +.15 for the frequency discrimination group; -.26 for the clatter discrimination group; and -.46 for the letter discrimination group, all $p > .10$. The differences in interpolated task performance indicate that the attempted equation of interpolated task difficulty in this experiment was only approximately successful. This was expected, given the differences in tasks, the use of completely untrained observers, and the relatively few trials which could be devoted to establishment of difficulty level. The overall levels

Table 1

Mean Hit and False Alarm Probabilities On the Interpolated Tasks for the Four Groups in Experiment 1

	Pre-memory	Memory	Post-memory
Signal Detection			
P(Hit)	.55	.49	.45
P(F/A)	.01	.02	.01
Frequency Discrimination			
P(Hit)	.71	.67	.55
P(F/A)	.07	.08	.06
Clatter Discrimination			
P(Hit)	.43	.41	.44
P(F/A)	.07	.08	.06
Letter Discrimination			
P(Hit)	.56	.47	.62
P(F/A)	.16	.16	.10

of performance, however, indicate that all tasks were quite difficult, and within-group differences on interpolated task performance do not seem related to recall in any consistent fashion.

Interpolated task performance on trials without a memory load ("pre-memory" and "post-memory") was compared with performance when a memory load was present. For hits, this comparison was nonsignificant, $F<1$, and did not interact significantly with groups in the overall analysis, $F(3,32) = 1.4$, $p>.10$. When each group was analyzed individually, the comparison between performance on memory and non-memory trials reached marginal significance only for the letter group, $F(1,8) = 4.4$, $.05<p<.10$. Overall, there was no reliable difference between hit rates on the six pre-memory and six post-memory trials, $F(1,32) = 1.7$, $p>.10$, and no interaction of this comparison with groups, $F(3,32) = 2.0$, $p>.10$. In individual group analyses, the only pre- vs. post-memory difference approaching reliability was that shown by the frequency discrimination group, $F(1,8) = 20.1$, $p<.001$. This group's hit rate declined (from .71 to .55) over trials, while that of the other groups remained essentially unchanged. The reason for this differential decline in performance is not obvious, but seems unrelated to the questions at hand.

Similar comparisons were made for false alarm rate. The overall memory load vs. no load comparison was significant as a main effect, $F(1,32) = 8.2$, $p<.01$, but did not interact with groups, $F(3,32)= 1.07$, $p>.10$. The overall pre-memory vs. post-memory comparison was reliable, $F(1,32) = 4.8$, $p<.05$, as was the interaction of this comparison with groups, $F(3,32) = 4.0$, $p<.05$. However, when these comparisons were made individually for each group, the only effects which approached reliability were those found in the letter group. For this group, both

memory load vs. no load, $F(1,8) = 8.9$, $p<.05$, and the pre- vs. post-memory comparison , $F(1,8) = 12.8$, $p<.01$, were reliable.

It appears that there may have been some effect of a verbal memory load (the trigrams) on performance of the letter discrimination task. Hit rate for this group dropped when a memory load was added and false alarm rate increased. However, the effect on hit rate is weak in this experiment, and the interpretation of the false alarm data is rendered ambiguous by the finding that false alarm rate dropped reliably between the pre-memory and post-memory trials. Experiment 2 contains conditions which are similar to the letter detection condition of Experiment 1, and will furnish additional data on the effect of a memory load on interpolated task performance.

Recall did not seem to vary reliably as a function of number of detection responses (hits + false alarms) made during a retention interval. The numbers of responses made on individual trials were divided into Vincent thirds for each subject. Table 2 shows the mean recall for each response third for each group. There was no main effect of response third on recall, $F<1$. No effect appeared when each group was analyzed alone, all $p>.10$.

Figure 1 shows the mean recall for each group as a function of number of signals occurring during a retention interval. A groups X signals analysis of variance (with trend analysis) on memory performance yielded the following reliable effects: group, $F(3,32) = 7.0$, $p<.001$; quadratic trend, $F(1,32) = 19.6$, $p<.001$; quadratic X group, $F(3,32) = 3.7$, $p<.05$; cubic, $F(1,32) = 7.9$, $p<.01$; and quintic, $F(1,32) = 5.1$, $p<.05$. The most potentially interesting effect appearing in Figure 1 is the sharp increase in recall shown by subjects in the letter discrimination condition on trials during which no "Ps" (signals) were presented during the retention interval. Unfortunately, the procedure used in Experiment

Table 2
Mean Percent Recall for Vincent Third
Number of Detection Responses in
Experiment 1

	Third		
	1st	2nd	3rd
Signal Detection	91	89	95
Frequency Discrimination	90	91	89
Clatter Discrimination	94	94	96
Letter Discrimination	68	70	67

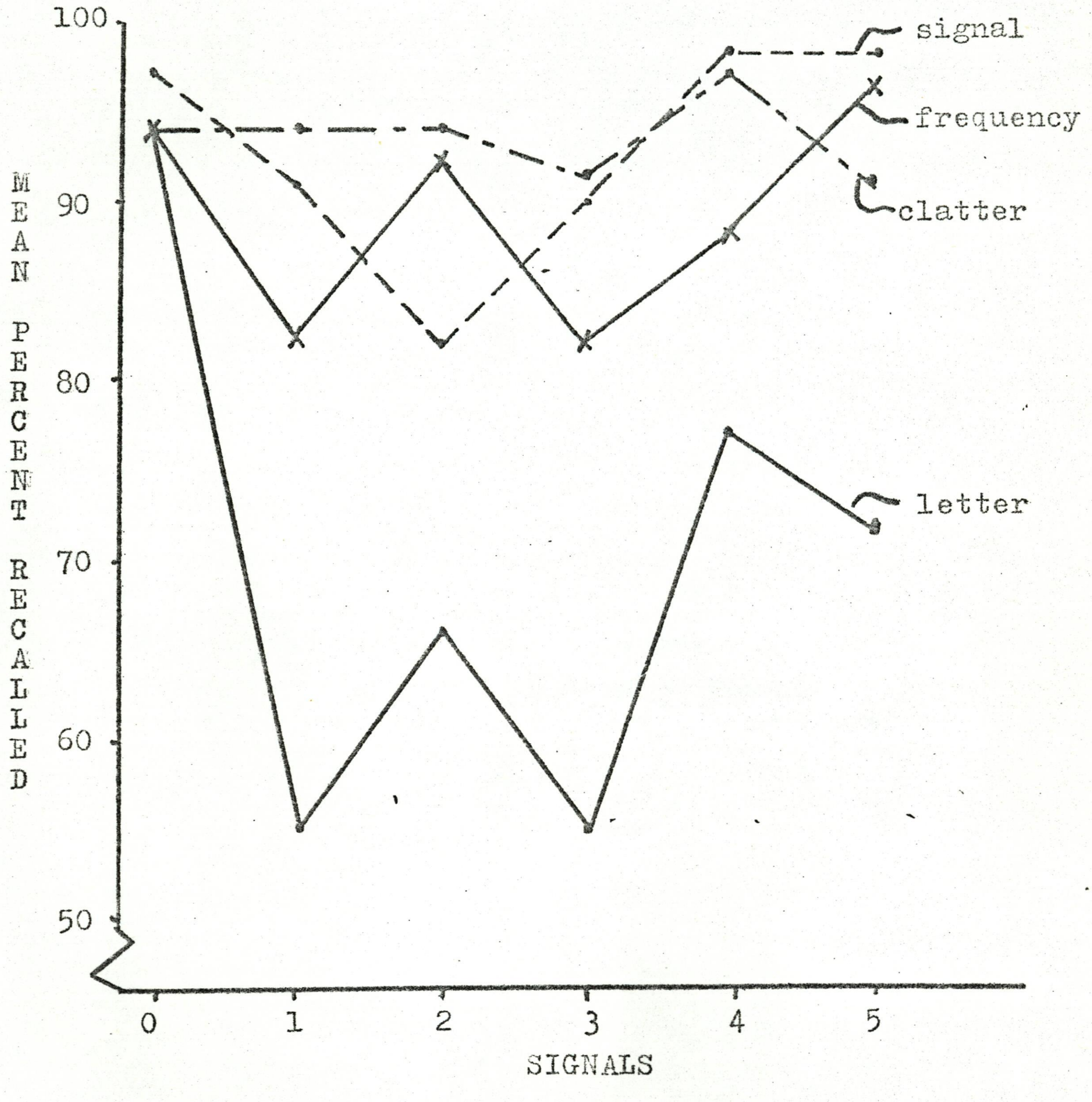

Figure 1. Mean recall as a function of number of detection interval signals in Experiment 1.

1 necessitated presenting the same trigrams, in the same order, to all subjects. Therefore, all zero signal trials in all groups followed a particular set of trigrams, single signal trials followed another particular set, etc. The presence of cubic and quintic trends suggests that particular trigram sets may have contributed significantly to the apparent effect of number of signals. Although the same stimuli are used in Experiment 2 that were used in Experiment 1, Experiment 3 examines the effect of presence vs absence of signals in interpolated letter discrimination tasks with the confound of particular trigram sets partially eliminated. The design of Experiment 4 again incorporates the confound, but does so using a different set of trigrams.

The result of Experiment 1 which is of primary importance is the finding that only the letter discrimination condition produced appreciable interference with storage of the consonant trigrams. In addition, there is some suggestion in the data that the memory task interfered with performance on the letter discrimination task, but not with performance on the nonverbal tasks. In terms of the model proposed in the introduction, it appears that occupation of the perceptual component, per se, does not produce substantial interference with storage, at least with the memory load used in Experiment 1. The possibility remains, however, that the letter discrimination task interfered at the perceptual level as a result of the similarity between the verbal memory items and the verbal discriminanda. An interpolated task is now needed for comparison which requires processing of letters at an acoustic feature level, but which would not be expected to involve processing by reference to production. A task approximating these conditions is used in Experiment 2.

Experiment 2

Subjects. Subjects were 30 undergraduate and six graduate psychology students at the State University of New York at Binghamton. Most of the undergraduates received course credit for their participation. None of the subjects had participated in Experiment 1, although other expereince in memory experiments was permitted. None of the subjects were trained psychophysical observers.

Design. Four independent groups of nine subjects were again used. Two of the groups differed from the letter group of Experiment 1 only in instructions. The noise-articulating group was instructed to pronounce each "B" and "P" aloud during the retention interval. This condition was intended to provide data on the interference produced by an interpolated task which clearly involved a verbal production component tied to analysis of incoming verbal stimuli. In the noise-nonarticulating group, subjects were told that they might be tempted to pronounce the "Bs" and "Ps", but that they should avoid doing so. It was expected that this group would be less likely to analyze the incoming material by reference to production than the noise-articulating group.

Subjects in the monitoring condition listened to the same stimuli as subjects in the other discrimination conditions, but without the masking noise, so that the letters were clearly discriminable. No instructions were given this group regarding articulation. The monitoring group had two functions. It was suggested in the introduction that one condition for eliciting analysis by reference to production is the use of stimuli which are deficient in information. If this suggestion is valid, it would be expected that the undegraded monitoring stimuli would be more likely to be analyzed solely at an acoustic feature level than letters in the noise conditions. If interference occurs only at post-perceptual levels, the

monitoring task should therefore interfere less with memory than the tasks involving letters in noise. On the other hand, if interference is due to similarity of retention interval material to the memory material, the monitoring condition should produce more interference than any other condition. The "Bs" and "Ps" in the monitoring condition were spoken in the same voice and at approximately the same intensity as the memory letters, and were not degraded by white noise.

A final group of subjects in an emission condition did not engage in a discrimination task during the retention interval, but were merely instructed to repeat aloud the letter sequence "ABCDEFGHI" as rapidly as possible. This group was used to test the effect on storage of a simple "emission" task (Peterson, 1969) that required the use of the articulatory apparatus. If both analysis by reference to production and the verbal storage system require the use of the articulatory system, interference may occur at the level of execution of the articulatory response. If "tying up" the peripheral articulatory effector mechanism is the only condition responsible for storage interference, recall after the emission task should be lower than recall after the noise-articulating task. The articulatory apparatus is nearly completely occupied in the former condition. If other, higher-level verbal production processes are responsible for the interference, higher recall would be expected after the simple emission task, which Peterson (1969) found produced virtually no interference with a simultaneous anagram-solving task.

Apparatus and procedure. The apparatus and procedure were generally the same as for Experiment 1. Stimulus presentation was controlled by the tape used in the letter condition of Experiment 1, except in the case of the emission group, where stimulus presentation was controlled by the tape used to control the clatter in

Experiment 1. This latter tape caused the headphones to be automatically disconnected during the retention interval. Subjects were instructed to begin speaking when the headphones "went dead" and to stop when the recall light came on.

For the noise conditions, 21 initial trials without memory trigrams were again used for determining a 50% threshold for each subject. For the emission and monitoring groups, these trials were presented as "practice". After the initial 21 trials, subjects were given memory task instructions. As in Experiment 1, there were six pre-memory, 26 memory, and six post-memory trials. The first six memory trials were again excluded from the data analysis so that mean recall was based on the last twenty memory trials. Subjects in all conditions were instructed to "concentrate completely" on the retention interval activity to the exclusion of the memory items.

Results and Discussion

Mean and median recall scores are presented in Table 3. Recall scores for the monitoring condition were sharply negatively skewed, with most of the skewness contributed by one subject whose recall score was 12.5%. This subject also had the highest false alarm rate and lowest hit rate of the monitoring group. Analysis of variance using the specific comparisons of *a priori* interest indicated that recall in the monitoring condition was reliably better than in the other conditions, $F(1,35) = 4.68$, $p < .05$. The noise-articulating condition resulted in poorer recall (49%) than the noise-nonarticulating condition (69%), $F(1,17) = 5.28$, $p < .05$, and the difference between the emission and noise-articulating groups in amount recalled was not reliable, $F < 1$.

Mean hit and false alarm rates for the three letter

Table 3
Percent Recalled in Experiment 2

	Group: Monitoring	Noise-Non-articulating	Noise-Articulating	Emission
Mean	78.2	69.4	48.6	57.4
Median	91.7	66.7	50.0	58.3

discrimination groups are shown in Table 4. Performance of the emission group on their retention interval task will be considered later. There was a highly reliable effect of group on both hit, $F(2,24) = 44.3$, $p<.001$, and false alarm rate, $F(2,24) = 11.3$, $p<.001$. These effects are almost entirely attributable to the high level of performance by subjects in the monitoring condition. The noise-articulating and noise-nonarticulating groups did not differ from one another either in hit rate, $F<1$, or in false alarm rate, $F(1,16) = 1.8$, $p>.10$.

Interpolated task performance on memory trials was again compared with trials on which no load was present ("Pre-memory" + "Post-memory" vs. "Memory"). For hit rate, this comparison was highly reliable as a main effect, $F(2,24) = 21.4$, $p<.001$. The interaction of the comparison with groups, which primarily reflects the relatively greater "Memory" drop shown by the noise groups, was only marginally reliable, $F(2,24) = 2.6$, $p<.10$. Neither the "Pre-memory" vs. "Post-memory" comparison, nor the interaction of this comparison with groups, approached reliability, $Fs<1$, for hit rate. When the same comparisons were made using false alarm rate, the only effect to reach even marginal reliability was the "Pre-memory" vs. "Post-memory" X group interaction, $F(2,24) = 3.3$, $p<.10$.

The decline in hit rate due to a memory load for the noise groups (Table 4) is approximately of the same magnitude as the decline shown by the letter group of Experiment 1 (Table 1). The marginally reliable effect on hit rate found in Experiment 1 became highly reliable when consideration of three groups of subjects, rather than only one group, increased statistical power. On the other hand, the ambiguous increase in false alarm rate due to a memory load shown by letter-group subjects in Experiment 1 disappeared in Experiment 2. In any case, it generally appears that a verbal memory load reduced

Table 4

Mean Hit and False Alarm Probabilities On the Interpolated Tasks for the Discrimination Groups in Experiment 2

	Pre-memory	Memory	Post-memory
Noise-Articulating			
P(Hit)	.59	.47	.61
P(F/A)	.17	.15	.12
Noise-Nonarticulating			
P(Hit)	.58	.38	.58
P(F/A)	.21	.21	.18
Monitoring			
P(Hit)	1.00	.94	.98
P(F/A)	.01	.03	.06

hit rate for verbal interpolated tasks, while the false alarm rate either remained constant or increased slightly as a result of the load.

As in Experiment 1, recall did not differ as a function of number of retention interval detection responses for the three letter discrimination groups. Mean responses per Vincent third for these three groups is shown in Table 5. There was no main effect of response third on recall, $F<1$, and no interaction of response third with group, $F(4,48) = 1.4, p > .10$.

Figure 2 shows mean recall as a function of the number of retention interval signals. Analysis of variance with trend analysis on these data yielded the following reliable effects: group, $F(2,24) = 4.0$, $p<.05$; quadratic trend, $F(1,24) = 9.5$, $p<.01$; quintic trend, $F(1,24) = 18.1$, $p<.001$; quintic trend X group, $F(2,24) = 8.1$, $p<.01$. Although there is some suggestion in Figure 2 that the memory performance of the monitoring and noise-nonarticulating groups was better after detection intervals which did not contain "Ps" (0 signals), the presence of higher-order trends again suggests a strong influence of particular trigram sets. The effect of signal presence vs. absence will be considered under more suitable circumstances in Experiment 3.

For the most part, the data presented thus far are consistent with the interpretation that the degree of interference with the storage system by the interpolated tasks is solely a function of the probability that the task will cause articulatory responses to be executed. Of the perceptual tasks used in Experiment 2, the monitoring condition stimuli seemed least likely to be analyzed by reference to production. This condition produced less interference with memory than the other conditions of the experiment. Similarly, subjects in the noise-non-articulating condition recalled a greater number of items

Table 5

Mean Percent Recall for Vincent Third
Number of Responses in Experiment 2:
Letter Discrimination Groups Only.

	Third		
	1st	2nd	3rd
Noise-Articulating	.45	.48	.50
Noise-Nonarticulating	.65	.73	.61
Monitoring	.87	.68	.79

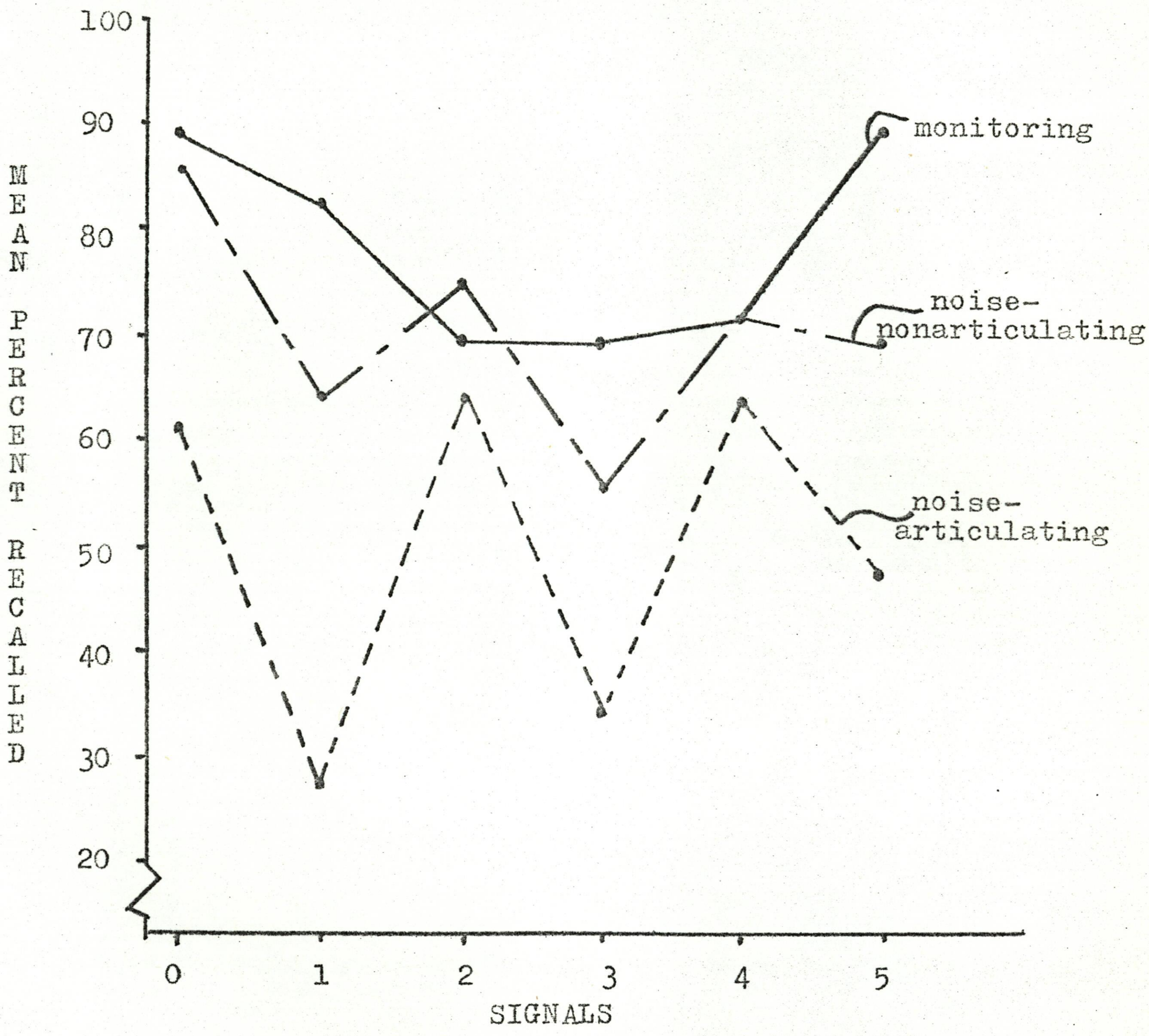

Figure 2. Mean recall as a function of number of detection interval signals for the letter discrimination groups of Experiment 2.

than subjects in the noise-articulating condition, with subjects in the latter group not differing in recall performance from those engaging in an articulatory emission task. However, further analysis of data from the emission group suggested that interference may occur at a level of response processing prior to peripheral response execution.

Filler task performance of subjects in the emission group was measured in two ways. First, the experimenter recorded the number of A-I sequences completed on each trial, and this value was multiplied by nine to obtain an estimate of the number of letters emitted per trial. Mean number of letters pronounced during each 16 second retention interval was primarily determined by the number and duration of hesitations between letters. Fewer letters were pronounced on memory trials, mean = 83.7, than on trials without a memory load, mean = 90.0; $F(1,8) = 6.0$; $p < .05$, but recall on memory trials increased markedly with number of letters emitted. Numbers of letters spoken per trial were divided into Vincent thirds for each subject. Recall was 34.2% on trials involving the fewest articulations, 47.6% on trials involving an intermediate number of articulations, and 63.3% on trials involving the largest number of articulations. Analysis of variance on these data showed a highly reliable linear effect, $F(1,8) = 15.1$, $p < .01$, with no reliable deviation from linearity, $F < 1$. That is, the greater the number of letters a subject was able to produce on a trial, the better his recall on that trial. The number of repetitions completed on a given trial was apparently not determined by overall practice or fatigue effects, since blocks of six trials each did not differ reliably in numbers of letters emitted, $F(3,24) = 1.5$; $p > .10$. Means for the four blocks were 82.8, 82.8, 83.7, and 86.4 letters per trial. Recall also did not differ over the four blocks, $F < 1$. Mean recall scores for the four blocks were 63%,

57%, 59%, and 55%.

The direct relation between number of letters emitted and recall is a most interesting finding. The effect is in precisely the opposite direction from that expected on the hypothesis that interference with storage is solely a function of frequency of articulation on the intervening task. Rather, it appears that interference may also occur at a level of response processing prior to the final execution of an articulatory response.

The second measure of interpolated task performance used with the emission group yielded parallel results. For the second analysis, the experimenter noted the occurrence of articulatory errors on any trial. Errors included mispronunciation of a letter, omission of a letter from the sequence, or pronunciation of a letter more than once in the same sequence. The number of letters emitted on errorless trials, mean = 84.4, did not differ reliably from the number of letters emitted on trials containing articulatory errors, mean = 83.5, $F<1$; nonetheless, the number of errors on a trial was divided by the number of letters emitted on that trial to obtain an error rate. Mean number of errors/letter/trial (error rate) was .02 on pre-memory trials, .06 on memory trials, and .02 on post-memory trials. Comparison of error rate on memory trials with error rate on the blocks containing no memory load yielded a highly reliable difference, $F(1,8) = 15.5$, $p<.01$. Mean recall on memory trials involving no articulatory errors was 61%, as opposed to 43.8% on trials during which errors were committed, although this difference is only marginally reliable, $F(1,8) = 4.5$, $p<.10$.

It would appear that a large proportion of the interference produced by the emission task occurred at a level prior to the final articulation of the interpolated verbalizations. In terms of the model currently under consideration, it may be suggested that repetitive articulation

of the relatively short, highly overlearned A-I sequence was usually a ballistic, "open loop" process. The decision involved was simply to repeat the sequence or to stop. When, for whatever reason, execution was interrupted in mid-sequence, the model would require the selection mechanism to intervene and select the next response in the sequence from the nine alternatives. While this decision was being made, the selection mechanism was not available for the storage loop and memory loss resulted.

Experiment 3

Although the results of Experiments 1 and 2 seem fairly clear in implicating verbal response production as the source of a large amount of interference with concurrent storage, they do not speak to possible interactions between stimulus discriminability and mode of analysis. That is, it is possible that the highly discriminable stimuli in the monitoring condition of Experiment 2 would produce little interference with storage, even if analyzed with reference to articulation. The primary purpose of Experiment 3 was to examine this possibility. A second purpose of the experiment was to study the effect of signal absence vs signal presence during a retention interval on recall. Any possible effect of signal nonoccurrence in the previous experiments was completely confounded with particular trigram sets, as discussed above.

Subjects. Subjects were 12 undergraduate students from an introductory psychology class at the State University of New York at Binghamton. All subjects received course credit for their participation. None had participated in previous experiments in this series, and none were trained psychophysical observers. Participation in other previous memory research was permitted.

Design. All conditions in Experiment 3 were within-subject. The monitoring-nonarticulating condition was the same as the monitoring condition of Experiment 2, except that the subjects were explicitly instructed to avoid saying the interpolated "Bs" and "Ps" even to themselves. In the monitoring-articulating (B-P) condition, subjects were required to pronounce each interpolated letter aloud, simultaneously pressing a key when a "P" occurred. The noise-articulating condition was the same as the noise-articulating condition of Experiment 2.

If all of the interference produced by the noise-articulating task is due to the fact of articulation -- i.e., stimulus discriminability, per se, has no effect -- recall after the monitoring-articulating (B-P) task should be as low as recall after the noise-articulating task. It is also possible, however, that the forgetting produced by the noise-articulating task results from the requirement that the selection of the verbal interpolated responses be made on the basis of stimuli which are difficult to discriminate. The extent to which stimulus discriminability interacts in this way with articulation should be reflected in a recall difference between the noise-articulating and monitoring-articulating conditions.

The fourth condition, monitoring-articulating (B-F), was presented after subjects had been tested under each of the above three conditions. This task was the same as the monitoring-articulating (B-P) condition, except that the target letters were "Fs" rather than "Ps". This condition was included against the possibility that, even without noise, the "Bs" and "Ps" were sufficiently acoustically similar to be somewhat difficult to discriminate.

Apparatus and Procedure. The apparatus was the same as that used in previous letter discrimination conditions. A new control tape was constructed using a new randomly selected set of consonant trigrams with a Witmer value of

less than 30%.

Each subject first received ten trials without memory items for the purpose of adjusting the noise level to produce a 50% hit rate with minimal false alarms in the noise-articulating condition. Following difficulty level adjustment, instructions were given regarding the memory tasks, including an admonition to avoid "thinking about" the memory items during retention intervals. All four of the conditions were outlined, and the subjects told that they would be informed regarding what to do before each of the four conditions. The monitoring-articulating (B-F) condition was always presented last. Two subjects received each of the six possible orderings of the other conditions. Six practice trials were presented before the ten experimental trials in the first condition. There was no break between the practice and the initial experimental trials, and from the subjects' point of view the practice trials were indistinguishable from the first experimental block. Conditions were administered in blocks of 10 trials each. The time parameters on each trial were the same as in the previous experiments. Between blocks, brief instructions (e.g., "Now there'll be no noise, say the "Bs" and "Ps".) were given during the 7 sec intertrial interval without breaking the usual pattern. All subjects responded appropriately to all instructed condition changes.

The number of "Ps" presented on a given trial was randomly determined, with the restriction that each experimental block of ten trials contain exactly two trials with no signals ("Ps"). Time of occurrence of each signal within a trial was randomly determined. From one to four signals could occur on a given signal trial, with no attempt to equate total number of signals between blocks. The same control tape, and therefore the same trigrams in the same order, was used for all subjects. Since subjects

received conditions in different orders, however, any given condition from the first three was presented with a particular set of trigrams to only four subjects. The B-F condition, always being presented last, involved the same set of trigrams for all subjects, the trigram set being different from that used in any B-P condition.

In order to be certain that clear differences in discriminability existed between the monitoring and noise conditions, it was decided that any subject would be replaced whose hit rate on a monitoring block was below .8, or whose hit rate on the noise block was lower than .25 or higher than .75. This combination of criteria was rather rigid, especially considering the small number of trials available for noise level adjustment, and 19 subjects had to be tested before 12 useable sets of data were obtained. All subjects whose data were dropped were excluded on the basis of noise-articulating hit rate. Five of those excluded had hit rates below .25, and two had hit rates above .75.

Results and Discussion

Mean recall scores for the four conditions were 45.8% after the noise-articulating task, 59.2% after monitoring-articulating (B-P), 57.5% after monitoring-articulating (B-F), and 85.8% after monitoring-nonarticulating. Analysis of variance with specific comparisons applied to the groupings of interest yielded a reliable difference between the monitoring-nonarticulating condition and the other conditions, $F(1,11) = 30.3$, $p < .001$; a reliable difference between the noise-articulating and monitoring-articulating conditions, $F(1,11) = 7.4$, $p < .05$; and no difference between the two monitoring-articulating conditions, $F < 1$.

Mean hit rate on interpolated task performance was .46 for the noise-articulating condition, .97 for the monitoring-articulating (B-P) condition, .96 for the monitoring-articulating (B-F) condition, and .93 for the

monitoring-nonarticulating condition. The noise condition hit rate was reliably lower than that in the other conditions, $F(1,11)=111.4; p<.001$, which did not differ reliably from one another, $F(2,22)=2.3, p>.10$. False alarm rates were .10 in the noise-articulating condition, .009 in the monitoring-articulating (B-P) condition, .002 in the monitoring-articulating (B-F) condition, and .007 in the monitoring-nonarticulating condition. The noise condition false alarm rate was reliably higher than in the other conditions, $F(1,11)=26.0$, $p<.001$. The monitoring nonarticulating false alarm rate did not differ from the combined rates of the two monitoring-articulating condition, $F(1,11) = 1.3$, $p>.10$.

Data from the monitoring-articulating (B-F) condition will not be included in further analyses. Since this condition involved a different set of memory trigrams from the other conditions, was always presented last, and all subjects in the condition received the same trigram set, interpretation of results from the B-F condition must be tempered with consideration of possible effects of order and/or of specific trigrams. The purpose of the B-F condition was to determine whether any obvious changes in memory performance would occur if the noiseless B-P stimuli were made even more discriminable. Even though the false alarm data indicate that the B-F stimuli may have been even more discriminable than the monitoring condition B-P stimuli, no obvious effect on concurrent retention was obtained.

Although the noise-articulating task led to poorer recall than the monitoring-articulating task, the difference was entirely due to performance on trials during which no signals were presented. For the noise-articulating condition, recall was 41.7% on trials containing no "Ps", and 47% on trials on which at least one signal occurred. In the monitoring-articulating condition, recall was 79.2% on trials containing no signals, but was 53.1% on trials

containing at least one signal. Eight of the 12 subjects had perfect recall on zero-signal monitoring-articulating trials. Although the recall scores after noise-articulating and monitoring-articulating tasks are reliably different when compared on zero-signal trials, $F(1,11) = 9.0$, $p<.025$, the difference disappears when only trials containing signals are considered, $F>1$. For the monitoring-<u>non</u>-articulating condition, no recall errors were made by any subject on zero-signal trials: mean recall was exactly 100%. On trials with signals present, mean monitoring-nonarticulating recall was 82.0%.

Analysis of variance on the three B-P conditions using conditions and signal presence as within factors yielded reliable effects of signal presence, $F(1,11) = 13.8$, $p<.01$, conditions, $F(2,22) = 22.0$, $p<.001$, and a reliable signal presence X condition interaction, $F(2,22) = 5.3$, $p<.05$. The effect of signal presence was reliable in the monitoring-articulating group, $F(1,11) = 20.0$, $p<.001$, and in the monitoring nonarticulating group, $F(1,11) = 40.0$, $p<.001$, but not in the noise-articulating group, $F<1$.

The results are rather striking. It would appear that discriminability of the to-be-articulated stimuli, per se, had little or no effect on concurrent retention. The monitoring-articulating condition produced less interference with memory than the noise-articulating condition only on those trials containing no interpolated "Ps". This difference may be more related to decision variables than to stimulus variables. It is possible that on zero-signal trials in the monitoring-articulating condition, the subjects tended to analyze the interpolated stimuli at an acoustic feature level, repetitively emitting "Bs" with minimal involvement of response selection processes. When acoustic analysis yielded data indicating a possible occurrence of the signal, the response selector was required to decide

whether a "B" or "P" should be emitted, and greater memory loss occurred. By contrast, in the noise-articulating condition, zero-signal trials were not clearly differentiable from signal trials. Because of the degradation by noise, decisions had to be made much more often as to which response was appropriate in the noise-articulating condition. Thus, the effect of stimulus discriminability can be interpreted as having an indirect detrimental effect on retention. Lowering stimulus discriminability may be hypothesized to adversely affect recall only if the lowered discriminability tends to increase the probability of a decision between alternative articulatory responses.

There was also a large effect of signal presence vs. absence in the monitoring-nonarticulating condition. In fact, recall was perfect for every subject on zero-signal monitoring-nonarticulating trials. In terms of the hypothesis outlined in the previous paragraph, these trials were more likely than any condition considered thus far to involve analysis of the interpolated stimuli at a purely acoustic level. A subject was required only to listen for a cue signalling a possible "P" without necessarily involving response processes in any way. Instructions, in fact, encouraged the subjects to avoid pronouncing the interpolated letters. The occurrence of a signal, however, may have had some tendency to elicit analysis by reference to articulation, producing the relatively lower recall after trials containing "Ps".

Experiment 4

It has been hypothesized that stimulus discriminability will have an effect on concurrent retention only if lowering discriminability increases the probability of analysis by reference to production. One test of this hypothesis would be to compare memory performance of subjects set to analyze interpolated material by synthesis with those

set to analyze the same material at an acoustic level. Experiment 4 provides such a comparison.

Subjects. Subjects were 30 undergraduate students from an introductory psychology class at the State University of New York at Binghamton. All subjects received course credit for their participation. Although participation in previous memory research was permitted, none of the subjects had participated in previous experiments in this series, and none were trained psychophysical observers.

Apparatus. The apparatus and control tape were the same as those used for Experiment 3. Thus, all subjects in all groups received the same trigrams in the same order.

Design and procedure. The subjects were divided into three groups of ten. All subjects first engaged in a pretest consisting of ten "B"-"P" discrimination trials without memory items for the purpose of establishing a noise level producing a 50% hit rate with minimal false alarms. Following difficulty adjustment, the subjects were instructed regarding memory trials, and told to avoid "thinking about" the memory material during the retention interval. No instructions were given regarding articulation of the interpolated letters. The noise-first group received six practice trials and 15 initial experimental trials listening to the interpolated letters in noise and key-pressing for "Ps". On experimental trials 16-30, the interpolated letters were presented without noise. The subjects had been told before the onset of the practice trials that the noise would be removed on "the second set of trials", but no further warning was given. The noise-second group received the initial practice and experimental trials without interpolated noise. These subjects were told that noise would be added on "the second set of trials", but, again, no further warning was given. Finally, an all-noise control group

received all trials with the interpolated letters in noise.

Based on the results of the previous experiments, it was expected that on the initial block of trials recall probability in the noise-second group (which received no noise on the initial block) would be considerably higher than in the all noise and noise-first groups. The difference, it has been hypothesized, will occur because the stimuli degraded by noise will be more likely to be analyzed by reference to production. If subjects in the noise-first group continue analyzing by synthesis when the noise is removed, recall should not improve as a function of trial blocks. Similarly, it subjects in the noise-second group continue to analyze at an acoustic level even when noise is added, recall for this group should not decline over trial blocks. More likely, however, whether analysis takes place at an acoustic level or by reference to production is probabilistically determined. If this is the case, it may be more realistic to expect only partial carryover of processing strategy from the first to the second block of trials. That is, the differences between the noise-first and noise-second groups which should appear on the first block may be much diminished on the second block.

Trial blocks and subject selection. For the major analysis, the thirty experimental trials were blocked as follows. For the noise-first and noise-second groups, the change in conditions occurred between the 15th and 16th trials. Trials 16-20 are considered separately from trials 21-30 to permit separate consideration of transient and longer-lasting effects of the change in conditions. Experimental trials 1-10 are used to measure pre-change effects, since this block contains the same number of zero-signal trials as the block comprised of trials 21-30.

The same subject selection criteria were applied to trials 1-10 and 21-30 as were used in Experiment 3.

Using these criteria, 37 subjects had to be tested to obtain 30 sets of acceptable data. Two subjects were excluded from the all-noise control group for low hit rates on trials 21-30, and one was excluded for a high hit rate on trials 1-10. Three subjects were excluded from the noise-second group for low hit rates on trials 21-30, and one was excluded for a high hit rate on trials 21-30.

Results and Discussion

Mean recall scores for the three blocks being considered are shown in Table 6. Block 1 in the table is composed of trials 1-10. Block 2 is composed of trials 16-20, and Block 3 is composed of trials 21-30. Overall blocks X groups analysis of variance on the data of Table 6 yielded a blocks X groups interaction as the only reliable effect, $F(4,54) = 6.9$, $p<.001$. On Block 1, the noise-second group (then receiving no noise) showed a reliably higher recall probability than the noise-first group and all-noise control group, $F(1,27) = 10.3$, $p<.01$. The two groups receiving noise on Block 1 did not differ from one another, $F<1$. On Block 3, the noise-first group (then receiving no noise) did not differ in recall from the noise-second group (then receiving noise), $F<1$. The difference between these two experimental groups and the all-noise control on Block 3 falls just short of statistical significance, $F(1,27) = 3.9$, $p<.10$. Transient effects (Block 2) will be considered shortly.

It appears that the large effect of stimulus discriminability on concurrent storage shown in Block 1 is completely eliminated in Block 3. The differential effectiveness of the noise on the first and last blocks was not due to differential performance on the noise tasks between Blocks 1 and 3. Table 7 shows the mean hit and false alarm rates for each group for each of the three

Table 6

Percent Recall Scores for the Experimental Trial Blocks of Experiment 4. Addition or Deletion of Noise Immediately Preceded Block 2

	Block 1	Block 2	Block 3
(Trial)	(1-10)	(16-20)	(21-30)
Noise second	86 --	54 (noise)	74 (noise)
Noise first	60 (noise)	70 --	76 --
All noise	53 (noise)	64 (noise)	56 (noise)

blocks after the exclusions were made. The noise-first group does not differ from the control group in either hit or false alarm rate on Block 1, $F's<1$. On Block 1, the noise-second hit rate is reliably higher than the rate of the groups then receiving noise, $F(1,27) = 104.5$, $p<.001$, and the false alarm rate is reliably lower, $F(1,27) = 31.5$, $p<.001$. Similarly, on Block 3, the hit rate is reliably higher for the noise-first group than for the groups then receiving noise, $F(1,27) = 127.6$, $p<.001$, and the false alarm rate is reliably lower, $F(1,27) = 22.7$, $p<.001$. Thus, even though noise had markedly <u>different</u> effects on memory on Blocks 1 and 3, the noise had the <u>same</u> effect on interpolated task performance level on Blocks 1 and 3. The memory effect would seem to have been manipulated independently of interpolated task difficulty.

The transient effects revealed by consideration of Block 2 (Table 6) merit consideration. When noise was removed, the noise-first group's recall performance reached the level of Block 3 almost immediately. The improvement from Block 2 to Block 3 shown by this group (70-76%) is small and unreliable, $F<1$. By contrast, the immediate effect of noise addition for the noise-second group was a plunge in recall rate (86-54%) followed by a sharp recovery to 74% recall in Block 3. For the noise-second group, the improvement in performance from Block 2 to Block 3 is highly reliable, $F(1,9) = 11.2$, $p<.01$. There were no reliable changes in recall across blocks in the all-noise control, $F<1$.

Interpolated task performance showed reliable increases in hit rate from Block 2 to Block 3 for both the noise-second group, $F(1,9) = 12.5$, $p<.01$ and the noise-first group $F(1,9) = 8.3$, $p<.05$, while the control group showed no change, $F<1$. No accompanying changes appeared in false alarm rate between Blocks 2 and 3 (Table 7).

While the transient effect of noise addition shown

Table 7

Interpolated Task Performance in Experiment 4

Hit Probability

	Block 1	Block 2	Block 3
Noise-second	.96	.35	.51
Noise-first	.48	.91	.97
All-noise	.50	.49	.50

F/A Probability

	Block 1	Block 2	Block 3
Noise-second	.01	.13	.14
Noise-first	.21	.01	.01
All-noise	.18	.16	.16

by the noise-second group was not predicted, an ad hoc explanation of it is easily reconcilable with the model under consideration. It is as though the noise-second subjects had been utilizing acoustic features for the discrimination task before the noise was added. The noise appears to have rendered the features originally used ineffective, requiring subjects to find a new set of features which were effective in the noise. This they did rather quickly, but for the first few noise trials were forced to analyze the interpolated material primarily by synthesis, producing the initial sharp decrement and subsequent rapid improvement.

In summary, it was predicted that the effect of discriminability of interpolated stimuli on concurrent retention could be reduced by a noninstructional manipulation intended to affect subjects' mode of analysis, independent of performance level. In fact, such a manipulation completely eliminated the effect of stimulus discriminability on concurrent retention. The transient effect of noise addition on both memory and interpolated task performance, while not predicted, is readily explained.

The data from Experiment 4 allowed two other analyses related to prior findings which will be briefly considered. Experimental trials 1-15 contained four zero-signal trials, and were analyzed for the effect of signal presence vs signal absence. These trials occurred before the change in conditions on trial 16. For the noise-second group, 85% of the trigrams were recalled on zero-signal trials, as compared with 82% on trials containing signals. The corresponding scores were 58% and 62% for the noise-first group, and 58% and 55% for the all-noise control. Analysis of variance on these data indicated no main effect of signal presence, $F<1$, and no signal presence X groups interaction, $F<1$. The failure to obtain a recall advantage for zero-signal trials in the noise-second group was surprising,

since not one of the same trigrams was forgotten by any subject in the monitoring-nonarticulating condition of Experiment 3. The major difference between the two groups appears to be the precise nature of the instructions. The monitoring-nonarticulating subjects of Experiment 3 were expressly cautioned against "saying the letters to themselves". The noise-second subjects of Experiment 4 received no instructions regarding articulation of the interpolated stimuli. Consequently, it appears that the Experiment 4 subjects may have been more likely to analyze some of the interpolated stimuli by reference to articulation, even on zero-signal trials.

Secondly, there were sufficient data from the all-noise control group to permit analysis of the effect of interpolated signal frequency on recall. In Experiments 1 and 2, such an effect was obtained, but was confounded with the particular trigrams associated with each signal frequency. The same confound appears in the all-noise group of the present experiment, but confounds signal frequency with a completely different set of trigrams. In the present experiment, with data analyzed over 30 trials, subjects in the all-noise control group recalled 62% of the material after trials with no signals, 50% after trials with one signal, 59% after two signals, 62% after three, and 54% after four signals. Analysis of variance revealed neither an overall effect of signal frequency, $F<1$, nor any orthogonal polynomial trend, maximum $F(1,9) = 1.7$, $p<.10$. Apparently the effect of signal frequency observed in the first two experiments was indeed due to effects of the particular trigrams with which each signal frequency was associated.

General Discussion

In the introduction, it was hypothesized that stor-

age of verbal material for brief intervals could be considered to be based on perceptual, response selection, and response execution processes common to information processing tasks in general, and speech processing in particular. The viability of such a model in accounting for the present results seems quite clear.

I have been unable to find an interpolated task which seems to interfere with concurrent verbal storage at a perceptual level. Neither defined-interval discrete nonverbal stimuli nor acoustically complex nonverbal stimuli produced more interference than a simple tonal detection task in Experiment 1. The interference produced by the processing of interpolated verbal material was markedly reduced when high stimulus discriminability allowed analysis at an acoustic level (Experiment 2) or when subjects were set to process noisy letters at an acoustic level (Experiment 4). When high discriminability, instructions to avoid articulation, and lack of any overt response requirement were combined on zero-signal trials in the monitoring-nonarticulating condition of Experiment 3, literally <u>no</u> measurable interference with storage was produced, even though the high level of performance on this task indicated that the interpolated letters were being perceptually processed.

There is evidence, however, for interference with storage by tasks involving the selection of a verbal response from one or more alternatives. Perhaps the most striking example of interference due to response organization occurred in the emission group of Experiment 2. In this condition, recall <u>improved</u> markedly with the number of letters a subject was able to pronounce during a retention interval. If hesitations during the emission sequences reflected response selection, the degree to which such selection processes were required on a particular trial was a major determinant of forgetting on that

trial. One immediately suggested further test of this interpretation would involve the use of emission sequences of varying length. As sequences become shorter, demands on response selection and organization mechanisms should be reduced, and concurrent retention should be facilitated. Recently gathered data clearly indicate that this is the case[3]. Another indication that involvement of response selection processes produces interference was obtained with the monitoring-articulating condition of Experiment 3. More than 25% greater recall was obtained in this condition on zero-signal trials (when all interpolated items were "Bs") than on trials when at least one interpolated letter was a "P", even though the same number of overt articulations were required on both types of trial.

Every manipulation in the present series of experiments which was intended to increase the probability of stimulus analysis by reference to production also increased the probability of forgetting of verbal material which was being concurrently retained. The manipulations included a requirement of overt stimulus articulation (Experiments 2 and 3), degradation of the stimuli by noise (Experiments 1,2,3, and 4), and manipulation of subjects' mode of analysis by means of a set produced by a prior task (Experiment 4). It cannot be determined at this point whether such interference results solely from response selection processes or whether an execution component is involved in the interference. Both of the results discussed in the previous paragraph as evidence for decision-produced interference were taken from conditions where recall was less than 80% at maximum. It is possible that storage places some requirement on response execution processes, and intervening tasks which may also require these processes will interfere. A useful task for examining this question would be one which involved minimal response selection,

but which kept the peripheral articulatory apparatus occupied during a retention interval. It is possible that the repetitive overt articulation of a single letter would be useful in this regard.

In summary, no evidence for interference with the perceptual component of the hypothesized storage loop has been found. Evidence has been found for interference at the level of response selection, but it is unclear whether additional interference takes place at the level of response execution. The storage system outlined in the introduction was quite useful in generating predictions tested in the present experiments, and is equally useful in ordering the results of those experiments. It is unclear how a general attentional capacity notion can be reconciled with the present results. Three very difficult nonverbal tasks were found to produce almost no interference with storage. On the other hand, very easy verbal tasks were made either very interfering or non-interfering, depending on the extent to which instructions (Experiments 1,2, and 3) or other manipulations (Experiment 4) could be said to elicit analysis-by-synthesis. That is, interference with memory does not seem to vary with interpolated task difficulty per se, but rather with the probability that the interpolated task engages response production processes.

A Note on the Nature and Reportability of "Rehearsal"

Since the storage loop described in the present paper seems to bear some similarity to a process usually described as "rehearsal", some mention should be made of the way rehearsal has been previously measured and interpreted. Numerous investigators seem to equate rehearsal with an appearance of some kind of an image of the material being retained in a subject's consciousness. In

fact, several investigators have recently used systematic introspection to measure the frequency of rehearsal, with results which seem highly internally consistent (Montigue, Hillix, Kiess, and Harris, 1970; Kroll and Kellicutt, 1972; Kroll, Kellicutt, and Parks, 1975; Reitman, 1971, 1974; Shiffrin, 1973). There is an obvious assumption in these studies that functional rehearsal is a conscious process, and Reitman (1974, p.375) has gone so far as to characterize subjects not reporting functional rehearsals as "liars".

It therefore came as something of a surprise when inquiries following Experiments 1 and 2 about success at avoiding rehearsal were most often met with blank stares from the subjects. Specifically, each subject was verbally asked "Did you find it possible to keep your mind off of the letters you were remembering while you were listening to the beeps/clatter/letters ?". There was then a pause, followed by "If they came to mind, how did they appear? Was it as though you saw them, heard them, said them to yourself, or are you not sure?" At this point, most subjects opted for one of the three suggested forms of rehearsal, although a few chose to say that the letters had been "completely out" of their minds[4]. The experimenter had a definite feeling that he was being humored. During the pause after the first question, most subjects seemed to have a great deal of difficulty in deciding whether their minds had been off of the letters. Responses such as "Well... Sort of." and "I guess so." were extremely common. Even after the second question, eight of the subjects from the two experiments insisted that the questions could not be answered: that there were no visual, auditory or articulatory images, but a very strong sense that the memory trigrams were "just there".

To get some first-hand information on the experiences being reported by the subjects, the author tested himself

on a number of trials from the signal detection condition of Experiment 1. The cause of the subjects' consternation became immediately obvious. After the first trial or two, it was possible to prevent any distinct image of the memory items from entering consciousness, albeit with some difficulty. There was, however, very frequently a strong sense that the items were, indeed, "just there" in some indescribable form. The experience seemed very close to the Bewusstheiten ("awarenesses") during imageless thought reported by Ach (Boring, 1950). To quote Boring (p.405-406),

> A Bewusstheit is,... a vague, intangible, conscious content that is not image or sensation. The descriptive word for it is unanschaulich, which Titchener has translated 'impalpable'. The action or thought consciousness lacks enough content adequately to clothe itself, but neverless systematic experimental introspection reveals something more than palpable contents; the consciousness has impalpable moments, Bewusstheiten.

That such an experience should occur is not really surprising. There is evidence that percepts gradually attain clarity as processing progresses (e.g., Haber and Hershenson, 1965), and the entire methodology of signal detection was devised to deal with situations in which subjects are unsure whether a phenomenal experience has occurred. In the typical signal detection paradigm, however, control by the experimenter of stimulus events permits separation of sensitivity and decision components of detection. With stimuli being held in memory, the experimenter has no way of knowing whether an item is actually being circulated through the storage loop. It is therefore difficult to know whether reported rehearsals reflect all functional rehearsals, or only those rehearsals during

which an associated image exceeds some criterion of clarity. Consequently, introspective reports about rehearsal were not taken from subjects in Experiments 3 and 4, and no assumptions are made concerning the correlates of the presently hypothesized storage processes in conscious experience.

References

Aldridge, J.W., & Burright, R.G. Auditory stimulus information change: reinforcing or investigation-eliciting? Paper read at the Forty-Fifth Annual Meeting, Eastern Psychological Association, Philadelphia, Pennsylvania, April, 1974.

Blumstein, S., & Cooper, W.E. Hemispheric processing of intonation contours. Cortex, 1974, 10, 146-158.

Boring, E.G. A history of experimental psychology (2nd. Ed.). New York: Appleton-Century-Crofts, 1950.

Brown, J.A. Some tests of the decay theory of immediate memory. Quarterly Journal of Experimental Psychology, 1958, 10, 12-21.

Bruning, J.L., Schappe, R.H., & O'Malley, J.S. Active vs. passive activity in STM. Psychological Reports, 1966, 19, 126.

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

Broadbent, D.E. Flow of information within the organism. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 34-39.

Broadbent, D.E. Decision and stress. New York: Academic Press, 1971.

Cole, R.A. Different memory functions for consonants and vowels. Cognitive Psychology, 1973a, 4, 39-54.

Cole, R.A. Listening for mispronunciations: a measure of what we hear during speech. Perception and Psychophysics, 1973b, 13, 153-156.

Cole, R.A., Coltheart, M., & Allard, F. Memory of a speaker's voice: reaction time to same- or different-voiced letters. Quarterly Journal of Experimental Psychology, 1974, 26, 1-7.

Cole, R.A., Haber, R.N., & Sales, B.D. Mechanisms of aural encoding: I. Distinctive features for consonants. Perception and Psychophysics, 1968, 3, 281-284.

Cole, R.A., Haber, R.N., & Sales, B.D. Mechanisms of aural encoding: VI. Consonants and vowels are remembered as subsets of distinctive features. Perception and Psychophysics, 1973, 13, 87-92.

Cole, R.A., Sales, B.D., & Haber, R.N. Mechanisms of aural encoding: II. The role of distinctive features in articulation and rehearsal. Perception and Psychophysics, 1969, 6, 361-365.

Cole, R.A., Sales, B.D., & Haber, R.N. Mechanisms of aural encoding: VII. Differences in consonant and vowel recall in a Peterson and Peterson short-term memory paradigm. Memory and Cognition, 1974, 2, 211-214.

Conrad, R. An association between memory errors and errors due to acoustic masking of speech. Nature, 1962, 196, 1314-1315.

Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-83.

Cooper, F.S. How is language conveyed by speech? In J.F. Kavanaugh & I.G. Mattingly (eds.), Language by ear and by eye. Cambridge: The M.I.T. Press, 1973.

Craik, F.I.M., & Lockhart, R.S. Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Crowder, R.G. Reciprocity of retention and interpolated task scores in short-term memory. Perceptual and Motor Skills, 1967a, 24, 903-909.

Crowder, R.G. Short-term memory for words with a perceptual-motor interpolated activity. Journal of Verbal Learning and Verbal Behavior, 1967b, 6, 753-761.

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

Cutting, J.E., & Rosner, B.S. Categories and boundaries in speech and music. Perception and Psychophysics, 1974, 16, 564-570.

Darwin, C.J. Ear differences in the recall of fricatives and vowels. Quarterly Journal of Experimental Psychology, 1971, 23, 46-62.

Egan, J.P., Greenberg, G.Z., & Schulman, A.I. Interval of time uncertainty in auditory detection. In J.A. Swets (ed.), Signal detection and recognition by human observers. New York: Wiley, 1964.

Glanzer, M., Gianutos, R. & Dubin, S. The removal of items from short-term storage. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 435-447.

Gordon, H.W. Hemispheric asymmetries in the perception of musical chords. Cortex, 1970, 6, 387-398.

Haber, R.N., & Hershenson, M. Effects of repeated brief exposures on the growth of a percept. Journal of Experimental Psychology, 1965, 69, 40-46.

Haggard, M.P. Encoding and the REA for speech signals. Quarterly Journal of Experimental Psychology, 1971, 23, 34-45.

Haggard, M.P., & Parkinson, A. Stimulus and task factors as determinants of ear advantages. Quarterly Journal of Experimental Psychology, 1971, 168-177.

Halle, M. & Stevens, K.M. Speech recognition: a model and a program for research. IRE Transactions on Information Theory, 1962, IT-8, 155-159. Also in J.A. Fodor & J.J. Katz (eds.), The structure of language. Englewood Cliffs: Prentice Hall, 1964.

Halperin, Y., Nachson, I., & Carmon, A. Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. Journal of the Acoustical Society of America, 1973, 53, 46-49.

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

Hochberg, J. Attention, organization, & consciousness. In D.I. Mostofsky (ed.), Attention: contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970.

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

Kerr, B. Processing demands during mental operations. Memory and Cognition, 1973, 1, 401-412.

Kimura, D. Some effects of temporal-lobe damage on auditory perception. Canadian Journal of Psychology, 1961a, 15, 156-165.

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

King, L., & Kimura, D. Left ear superiority in dichotic perception of vocal nonverbal sounds. Canadian Journal of Psychology, 1972, 26, 111-115.

Kirman, J.H. Tactile communication of speech: a review and an analysis. Psychological Bulletin, 1973, 80, 54-74.

Klatt, D.H. Structure of confusions in short-term memory between English consonants. Journal of the Acoustical Society of America, 1968, 44, 401-407.

Kroll, N.E.A., & Kellicutt, M.A. Short-term recall as a function of covert rehearsal and of intervening task. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 196-204.

Kroll, N.E.A., Kellicutt, M.H., & Parks, T.E. Rehearsal of visual and auditory stimuli while shadowing. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 215-222.

Kroll, N.E.A., Parkinson, S.R., and Parks, T. Sensory and active storage of compound visual and auditory stimuli. Journal of Experimental Psychology, 1972, 95, 32-38.

Lane, H. The motor theory of speech perception: a critical review. Psychological Review, 1965, 72, 275-309.

Leshowitz, B., Zurek, P.M., & Robbins, D. Forgetting of visually presented words after retention intervals filled with detection of acoustic signals. Bulletin of the Psychonomic Society, 1974, 3, 211-213.

Liberman, A.M. Some results of research on speech perception. Journal of the Acoustical Society of America, 1957, 29, 117-123.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. Perception of the speech code. Psychological Review, 1967, 74, 431-461.

Locke, J.L., & Fehr, F.S. Subvocal rehearsal as a form of speech. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 495-498.

Locke, J.L., & Fehr, F.S. Subvocalization of heard or seen words prior to spoken or written recall. American Journal of Psychology, 1972, 85, 63-68.

Lowe, D., & Merikle, P.M. Interpolation effects in short-term memory. Psychonomic Science, 1971, 22, 89-91.

Massaro, D.W. Perceptual processes and forgetting in memory tasks. Psychological Review, 1970, 77, 557-567.

Matsumiya, Y., Tagliasco, V. Lombroso, C.T., & Goodglass, H. Auditory evoked response: meaningfulness of stimuli and interhemispheric asymmetry. Science, 1972, 175, 790-792.

Merikle, P.M. Unit size and interpolated-task difficulty as determinants of short-term retention. Journal of Experimental Psychology, 1968, 77, 370-375.

Miller, J.D., Pastore, R.E., Wier, C.C., Kelley, W.J., & Dooling, R.J. Discrimination and labeling of noise-buzz sequences with varying noise lead times. Journal of the Acoustical Society of America, 1974, 55, 390.

Montague, W.E., Hillix, W.A., Kiess, H.O., & Harris, R. Variation in reports of covert rehearsal in STM produced by differential payoff. Journal of Experimental Psychology, 1970, 83, 249-254.

Moray, N. Where is capacity limited? A survey and a model. Acta Psychologica, 1967, 27, 84-92.

Muraski, A.A., & Sharf, D.J. The effect of varying labels on the identification of speech and nonspeech stimuli presented dichotically. Journal of the Acoustical Society of America, 1973, 54, 285.

Murdock, B.B. Proactive inhibition in short-term memory. Journal of Experimental Psychology, 1964, 68, 184-189.

Murdock, B.B. A test of the "limited capacity" hypothesis. Journal of Experimental Psychology, 1965a, 69, 237-240.

Murdock, B.B. Effects of a subsidiary task on short-term memory. British Journal of Psychology, 1965b, 56, 413-419.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Newell, M. The effects of acoustic disruption on short term memory. Psychonomic Science, 1968, 12, 61.

Norman, D.A. Toward a theory of memory and attention. Psychological Review, 1968, 75, 522-536.

Peterson, L.R. Concurrent verbal activity. Psychological Review, 1969, 76, 376-386.

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

Pillsbury, W.B., & Sylvester, A. Retroactive and proactive inhibition in immediate memory. Journal of Experimental Psychology, 1940, 27, 532-545.

Posner, M.I. Cognition: an introduction. Glenview, Ill.: Scott Foresman, 1973.

Posner, M.I., and Boies, S.J. Components of attention. Psychological Review, 1971, 78, 391-408.

Posner, M.I., Boies, S.J., Eichelman, W.H., and Taylor, R.L. Retention of visual and name codes of single letters. Journal of Experimental Psychology Monographs, 1969, 79, No. 1, part 2.

Posner, M.I., and Keele, S.W. Decay of visual information from a single letter. Science, 1967, 158, 137-139.

Posner, M.I., and Klein, R. On the functions of consciousness. In S. Kornblum (ed.), Attention and performance, vol. 4. New York: Academic Press, 1973.

Posner, M.I., & Konick, J. On the role of interference in short-term retention. Journal of Experimental Psychology, 1966, 72, 221-231.

Posner, M.I., and Rossman, E. Effect of size and location of informational transforms upon short-term retention. Journal of Experimental Psychology, 1965, 70, 496-505.

Postman, L., & Phillips, L.W. Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology, 1965, 17, 132-138.

Rabbitt, P. Recognition: memory for words correctly heard in noise. Psychonomic Science, 1966, 6, 383-384.

Rabbitt, P.M.A. Channel-capacity, intelligibility and immediate memory. Quarterly Journal of Experimental Psychology, 1968, 20, 241-248.

Reitman, J.S. Mechanisms of forgetting in short-term memory. Cognitive Psychology, 1971, 2, 185-195.

Reitman, J.S. Without surreptitious rehearsal, information in short-term memory decays. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 365-377.

Sales, B.D., Cole, R.A., & Haber, R.N. Mechanisms of aural encoding: V. Environmental effects of consonants on vowel encoding. Perception and Psychophysics, 1969, 6, 361-365.

Sales, B.D., Cole, R.A., & Haber, R.N. Mechanisms of aural encoding: VIII. Phonetic interference and context-sensitive coding in short-term memory. Memory and Cognition, 1974, 2, 596-600.

Sales, B.D., Haber, R.N., & Cole, R.A. Mechanisms of aural encoding: IV. Hear-see, say-write interactions for vowels. Perception and Psychophysics, 1969, 6, 385-390.

Scheirer, C.J., & Voss, J. Short-term association recall as a function of number and type of retention interval tasks. Psychonomic Science, 1970, 19, 223-224.

Shiffrin, R.M. Information persistence in short-term memory. Journal of Experimental Psychology, 1973, 100, 39-49.

Sloboda, W. The disturbance effect of white noise on human short-term memory during learning. Psychonomic Science, 1969, 14, 82-83.

Sloboda, W., & Smith, E.E. Disruption effects in human short-term memory: some negative findings. Perceptual and Motor Skills, 1968, 27, 575-582.

Sperling, G. A model for visual memory tasks. Human Factors, 1963, 5, 19-31.

Spreen, O., Spellacy, F.J., & Reid, J.R. The effect of interstimulus interval and intensity on ear asymmetry for nonverbal stimuli in dichotic listening. Neuropsychologia, 1970, 8, 245-250.

Stevens, K.N. Segments, features, and analysis by synthesis. In J.F. Kavanaugh and I.G. Mattingly (eds.), Language by ear and by eye. Cambridge: The M.I.T. Press, 1973.

Stevens, K.N., & House, A.S. Speech perception. In J. Tobias (ed.) Foundations of modern auditory theory, Vol. 2. New York: Academic Press, 1972.

Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., Cooper, F.S. Motor theory of speech perception: a reply to Lane's critical review. Psychological Review, 1970, 77, 234-249.

Sussman, H.M. The laterality effect in lingual-auditory tracking. Journal of the Acoustical Society of America, 1971, 49, 1874-1880.

Sussman, H.M., MacNeilage, P.F., & Lumbley, J.L. Pursuit auditory tracking of dichotically presented tonal amplitudes. Journal of Speech and Hearing Research, 1975, 18, 74-81.

Talland, G.A. Short-term memory with interpolated activity. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 144-150.

Underwood, B.J., & Schultz, R.W. Meaningfulness and verbal learning. Philadelphia: Lippincott, 1960.

Watkins, M.J., Watkins, O.C., Craik, F.I.M., & Mazuryk, G. Effect of nonverbal distraction on short-term storage. Journal of Experimental Psychology, 1973, 101, 296-300.

Wickelgren, W.A. Distinctive features and errors in short-term memory for English vowels. Journal of the Acoustical Society of America, 1965, 38, 583-588.

Wickelgren, W.A. Distinctive features and errors in short-term memory for English consonants. Journal of the Acoustical Society of America, 1966, 39, 388-398.

Wood, C.C. Parallel processing of auditory and phonetic information in speech discrimination. Perception and Psychophysics, 1974, 15, 501-508.

Wood, C.C. Auditory and phonetic levels of processing in speech perception: neurophysiological and information-processing analyses. Journal of Experimental Psychology: Human Perception and Performance, 1975, 104, 3-20.

Wood, C.C., Goff, W.R., & Day, R.S. Auditory evoked potentials during speech perception. Science, 1971, 173, 1248-1251.

Woodworth, R.S., & Schlosberg, H. Experimental psychology. New York: Holt, Rinehart, & Winston, 1954.

NOTES

1. Glanzer, Gianutos, and Dubin (1969) found no effect of interpolated task difficulty on the degree to which a free recall end peak was reduced, and interpret other results as reflecting an effect on transfer between two storage systems.

2. Other theorists have used components of the analysis-by-synthesis conception, and their theories have been given the apellation "analysis-by-synthesis", although the theories are much broader in scope than those presented here. Neisser (1967) used the active percept construction notion, and Norman (1968) and Hochberg (1970) used similar conceptions phrased in terms of "expectancies". Broadbent (1971) extensively used the categorization process (e.g., naming) implicit in the present description of analysis-by-synthesis.

3. Aldridge, in preparation.

4. There was no apparent relation between memory performance and reported rehearsal. Over both experiments, rehearsers ("said to self, heard, or saw", N=40) produced a mean recall score of 75%. The mean was 78% for nonrehearsers ("just there or out of mind", N=23) and was 68% for subjects responding "don't know", N=9. These differences are not reliable, $F<1$.

APPENDICES

A. Sample Instructions: Clatter Group of Experiment 1

B. Individual Recall Scores

C. Analyses of Variance on Recall Scores

Appendix A

Sample Instructions: Clatter Group of Experiment 1.

I. Prior to difficulty level determination:

For this part of the experiment, I want to test your ability to detect changes in a clattering sound. On each trial, you will first hear a noise burst as a warning signal. About two seconds later the headphones will go dead and you will hear a clattering sound behind you that sounds like this. There are several devices behind you which make the same sound, and while the steady clatter is on, other devices of the same kind will come on occasionally for two second periods, creating extra clatter. Listen carefully, and whenever you're sure the extra devices have come on, quickly give the telegraph key a firm tap so that it goes all the way down. You can respond once during each two second period of extra clatter. Taps at any other time count as errors. When the steady clatter goes off and the light by the telegraph key comes on, wait a few seconds for the next trial to begin. The amount of difference between the steady clatter and the extra clatter will change from trial to trial. Performance on this kind of task is greatly impaired by momentary lapses of attention, so please concentrate completely on listening for extra clatter while the steady clatter is on. To show you what to listen for, you will watch the clatter boxes for a few trials. However, after these trials, don't turn around until the experiment is finished, since to do so could give you a clue as to when to expect extra clatter on the following trial.

II. Prior to the experimental trials:

Next comes the main part of the experiment. The first few trials will be exactly like the ones you've just had, except the extra clatter will be equally difficult to distinguish from the steady clatter on all of the trials. On trials after the first few, you will hear a set of three letters between the warning signal and the onset of the steady clatter. Commit the three letters to memory completely, as you hear them. When the steady clatter comes on, concentrate completely on listening for extra clatter. It is very important that you not think about the letters in any way, since to do so will impair your performance on detecting the extra clatter. When the steady clatter goes off and the light comes on, you have five seconds to remember the letters. Say what you remember clearly and loudly so I can record your responses. After the light goes off, your time is up, so relax until you hear the next warning burst. The last few trials will be exactly like the first few, with no letters presented.

Appendix B

Individual Recall Scores

Experiment 1. Percent recalled: one score per subject

Signal Detection	Clatter	Frequency	Letter
87.5	95.8	70.8	43.5
83.3	91.7	91.7	67.0
87.5	95.8	95.8	87.5
95.8	91.7	95.8	91.7
91.7	83.3	95.8	41.7
95.8	100.0	75.0	92.7
100.0	95.8	100.0	79.1
95.8	91.7	83.3	41.7
83.3	100.0	100.0	75.0

Experiment 2. Percent recalled: one score per subject

Monitoring	Noise-Articulating	Noise-non-Articulating	Emission
87.5	54.2	100.0	79.2
95.8	16.7	79.2	54.2
70.8	79.2	79.2	58.3
91.7	66.7	58.3	50.0
100.0	45.8	45.8	79.2
12.5	50.0	58.3	79.2
100.0	70.8	66.7	66.7
50.0	25.0	58.3	8.3
95.8	29.2	79.2	41.7

Appendix B (continued)

Experiment 3. Percent recalled: completely within design

Subject	Monitoring-Articulating (B-P)	Monitoring-Articulating (B-F)	Noise-Articulating	Monitoring-Non-Articulating
1	90	60	60	90
2	60	50	20	90
3	20	30	40	80
4	50	90	70	100
5	90	50	50	80
6	90	80	50	70
7	0	40	0	90
8	50	50	50	80
9	30	50	50	80
10	80	60	50	90
11	60	70	50	90
12	90	60	60	90

Experiment 4. Percent recalled: noise first group

	Trial Block		
Subject	1	2	3
1	100	100	100
2	0	20	50
3	40	60	20
4	70	60	80
5	60	80	90
6	70	80	60
7	30	80	80
8	80	60	100
9	70	100	90
10	80	60	90

Appendix B (continued)

Experiment 4 (cont.) Percent recalled: noise second group

Subject	Trial Block 1	2	3
11	100	60	90
12	90	60	90
13	70	60	80
14	70	0	40
15	80	40	30
16	90	60	100
17	100	60	60
18	80	80	80
19	90	60	100
20	90	60	70

Percent recalled: All noise control group

Subject	Trial Block 1	2	3
21	40	20	50
22	40	40	40
23	100	100	100
24	60	100	60
25	30	40	10
26	10	60	40
27	80	80	80
28	80	80	50
29	40	40	70
30	50	80	60

Appendix C

Analyses of Variance on Recall Scores

Experiment 1

Source	df	SS	MS	F
Groups	3	3583.56	1194.52	7.42
Error	32	5149.36	160.92	

Experiment 2

Source	df	SS	MS	F
Groups	3	4599.05	1533.02	2.86
Error	32	17141.67	535.68	

Experiment 3

Source	df	SS	MS	F
Monitoring-non-articulating vs. others	1	0.9025	0.9025	30.31
Error	11	0.3275	0.0298	
Noise-articulating vs. Monitoring-articulating	1	0.1250	0.1250	7.43
Error	11	0.1850	0.0168	
Monitoring-articulating: B-P vs. B-F	1	0.0017	0.0017	0.05
Error	11	0.3883	0.0353	

Appendix C (continued)

Experiment 4

Source	df	SS	MS	F
Groups	2	0.32	0.16	1.22
Error	27	3.50	0.13	
Trial Blocks	2	0.05	0.02	1.15
Trial Blocks X Groups	4	0.66	0.17	6.9
Error	54	1.29	0.02	