

Effects of imagining speakers' voices on the retention of words presented visually

RALPH E. GEISELMAN and JANET GLENNY
University of California, Los Angeles, California 90024

Subjects were introduced to one male and one female voice by a tape recording with instructions to attend to characteristics of the voices. Then 18 pairs of words were presented visually on slides. The subject's task during each 10-sec interslide interval was to repeat silently the pair of words over and over again in the male voice, in the female voice, or in the subject's own voice. A surprise recognition test for the words indicated that the words were more likely to be recognized if they were spoken in the same voice at test as was used to repeat them during presentation. Recognition of the words repeated in the subject's own voice was not affected by the sex of the speaker at test. In Experiment 2, different speakers were used at test than those used by the subjects to repeat the words. The interaction between the sex of voice used at encoding and at test was again significant, but recognition was generally lower than in Experiment 1. It was concluded that it is not necessary to assume that subjects have literal copies of spoken words in memory but speaker's voice does form an integral part of the verbal memory code and its influence is specific to a given speaker as well as to a given class of speakers (male or female).

Craik and Kirsner (1974) reported that the recognition of words presented auditorily was facilitated if the words were spoken by the same person at test as at presentation. In their experiments, some items in a series of unrelated words were spoken by a male and the remaining words were spoken by a female. The sex of the speaker at initial presentation was then reversed for half of the words at test. Craik and Kirsner attributed their results to a physical-similarity effect that is based on the assumption that subjects have access to non-linguistic aspects of auditory events in literal form for at least 2 min. Hintzman, Block, and Inskip (1972) suggested an experiment similar to that of Craik and Kirsner and agreed with the literal-copy explanation of such results. More recently, Crowder (in press) has also concluded that Craik and Kirsner's data reflect the influence of "literal episodic records" on later recognition.

The present Experiment 1 was designed to determine if the literal-copy assumption is necessary to account for the effect of speaker's voice on word recognition. None of the authors cited above has clearly defined the concept of an auditory literal copy; but, by any definition, it certainly must be the case that an episodic record of a covertly fabricated "auditory" event would not be comparable to a literal copy of the actual auditory event. A literal copy of an auditory event could not be stored in memory unless the event had in fact occurred. Therefore, in the present experiment, the subjects were

instructed to imagine that words presented visually were being spoken in terms of recently familiar voices. Then, the subjects were tested unexpectedly for occurrence recognition of the words using auditory presentation. It would no longer be necessary to account for Craik and Kirsner's (1974) data with the literal-copy assumption if the probability of recognizing a word was found to be greater when the voice at test matched the voice used to form the auditory image at presentation. Such results would not rule out the possibility that literal copies of auditory events are stored in memory, but it would only be necessary to attribute a mediating role to the original auditory events in Craik and Kirsner's experiments. That is, a speaker's voice at presentation might influence the coding process that forms the representation for an item in memory. At test, the target item would be compared to this representation rather than to a literal copy of the original auditory stimulus. This degree of freedom would be useful since the literal-copy notion is inconsistent with other memory concepts, such as Crowder and Morton's (1969) notion of precategorical acoustic storage.¹

There are at least two ways in which a speaker's voice might influence the code for an item in memory aside from being part of a literal copy. First, Carterette and Barnebey (1975) have proposed that a small set of sensory attributes of an unfamiliar voice may be stored in parallel with verbal attributes of the stimulus. These attributes include fundamental frequency, intonational pattern, and intensity. Carterette and Barnebey leave open the question as to whether or not such attributes are stored in their sensory form, but for purposes of the present discussion, it will be assumed that these attributes are stored in propositional form. Otherwise, the

This research was funded in part by a grant from the National Institutes of Health (1 F32 MH 05783-01). Requests for reprints should be sent to the first author, Department of Psychology, University of California, Los Angeles, California 90024.

Carterette and Barnebey hypothesis becomes a weak form of the literal-copy notion where only a few sensory attributes are stored in memory due to some filtering process. By referring to the representation of speaker's voice in memory as being propositional, we concur with Pylyshyn (1973) that the representation "(a) does not correspond to a raw sensory pattern, rather, is already highly abstracted and interpreted, (b) is not different in principle from the kind of knowledge asserted by a sentence, and (c) depends on the classification of sensory events into a finite set of concepts and relations" (p. 7).

Another way in which a speaker's voice might influence the code for the item in memory was proposed by Geiselman and Bellezza (1976, 1977). They suggested that an unfamiliar voice may be remembered because of the influence of its connotation of the verbal code for the item in memory. In other words, a sentence spoken by a male may not have the same meaning as the same sentence spoken by a female. These two views are, of course, highly compatible, in that a small set of voice attributes that is stored in parallel in propositional form might influence the connotative meaning of the item.

In the present experiments, the subjects were asked to construct "auditory" events at presentation; this constructive process might be based on a recollection of values for a small set of attributes of the intended voice, such as pitch (high or low), or the constructive process might be based on a recollection of connotative aspects of the voice (masculine vs. feminine). In either case, one framework for discussing the nature of the constructive process is provided by the analysis-by-synthesis model of speech recognition proposed by Halle and Stevens (1962). Halle and Stevens suggest that an input signal is matched with internally generated speech patterns and the analysis-by-synthesis procedure is said to adapt to the characteristics of different speakers through modification of internal speech parameters. Perhaps the constructive process in the present experiments can occur if, for purposes of word repetition, the subject matches the values of the internal speech parameters with his recollection of those values, in propositional form, for a given speaker. As Norman (1976) has argued, the format in which information is stored need not be the same as the form in which the information is used. The constructive process should influence the later recognition of the words unless it is necessary to have literal copies of auditory events in memory for the recognition effect to be observed.

A secondary purpose of Experiment 1 was to compare the effects of repeating words in another person's voice on retention vs. repeating words in one's own voice. First, is it more difficult to repeat words in another person's voice and does this difference affect recognition performance? Also, for words that are repeated in one's own voice, is it easier to recognize the words later if the sex of the speaker at test matches the

sex of the subject? Or is one's "inner voice" neutral with respect to any influence on the code for an item in memory?

EXPERIMENT 1

Method

Subjects. The subjects were 15 male and 15 female undergraduate volunteers from the introductory psychology course at the University of California at Los Angeles. The testing was done in groups of 10 subjects each (5 males, 5 females).

Materials and procedure. First, the subjects were presented a tape recording of a 14-sentence passage about psychoanalysis in which a randomly chosen 7 sentences were read by a male speaker and the remaining 7 sentences were read by a female. The subjects were told, "Rather than paying attention to what these two people are saying, I want you to listen carefully to characteristics of their voices. In other words, I want you to familiarize yourself with their voices. What they are saying is not important." The fundamental frequency for the male speaker was estimated to be 117 Hz and for the female speaker was 219 Hz.

Following the introductory recording, the subjects were shown a series of 18 slides of word pairs. The items were constructed such that all words were common four-letter non-homophonic nouns and no pairs contained words beginning with the same letter, rhymed, or had obvious associations. Each word pair was shown for 2 sec and a blank slide inserted after each word pair was shown for 10 sec. In addition the background of each slide containing a word pair was blue, pink, or yellow. The color of each slide was counterbalanced across sessions such that each slide was presented as blue, pink, or yellow for some subjects. The subjects were told, "If a slide with words on it is blue, I want you to imagine that the male speaker that you just heard is saying the words over and over again. In other words, if the slide is blue, you should repeat the words to yourself for the 10 sec in terms of the male's voice—not in your own inner voice. If the slide with words on it is pink, you should repeat the words to yourself for the 10 sec in terms of the female's voice. If the slide is yellow, I want you to repeat the words over and over again in your own voice for the 10 sec." The subjects were not told about the ensuing recognition test but rather, "When we are finished, I will ask you some questions about the experiment to see how well you were able to repeat the words in others' voices and in your own voice." Before presenting the slides, the recording about psychoanalysis was re-presented for the subjects to insure that they were familiar with the two voices.

After presenting the slides, the subjects were given a questionnaire that asked the following questions: (1) How well do you think that you were able to repeat the words to yourself in the male's voice when the slides were blue? (2) How well do you think that you were able to repeat the words to yourself in the female's voice when the slides were pink? (3) How well do you think that you were able to repeat the words to yourself in your own voice when the slides were yellow? The possible responses for each question were: very poorly, poorly, fairly well, and very well. The questionnaire took approximately 30 sec to complete.

Then the subjects were presented a tape recording of 72 individual four-letter words with 5-sec blank interword intervals. Thirty-six of the 72 words had been shown on the slides and 1 word from each slide was spoken by the male and the remaining word was spoken by the female. Half of the 36 distractors were also spoken by the male and the remaining distractors were spoken by the female. The subject's task during each 5-sec interitem interval was to respond on an answer sheet with "yes" or "no," depending on whether the word had been

shown on one of the slides. The color of the slide and the sex of the speaker on the test tape were irrelevant to the subject's response decisions because the subjects were told, "Your task is simply to indicate whether or not each word was presented to you on one of the slides."

Apparatus. The tape recorder was a Sony TC-200 and the slide projector was a Kodak Carousel 860H. The projector was controlled by two Davis Scientific Instruments timers, Model BS01.

Design and analysis. The data matrix for the recognition of the words presented on the slides was a 2 by 3 by 2, with the specific factors being sex of subject, repetition mode (male's voice or female's voice, or own voice), and test mode (male's voice or female's voice). The data matrix for the distractor items was a 2 by 2, with the specific factors being sex of subject and test mode. The sex-of-subject factor was the only between-subjects factor in either analysis.

In addition, the questionnaire data were analyzed using a 2 by 3 analysis of variance, with the factors being sex of subject and repetition mode. This was done to determine (1) whether repeating words in one's own voice is easier than repeating words in another person's voice and (2) whether this relationship interacts with the sex of the subject. The dependent variable was each subject's ratings of the ease of repetition on the three 4-point scales.

Results

The analysis of variance conducted on the recognition data for the words presented on the slides showed that the main effects of repetition mode and test mode were not significant ($F_s < 1$). However, the Repetition Mode by Test Mode interaction effect was significant [$F(2,48) = 7.63$, $MSe = .13$, $p < .005$]. These results are shown in Figure 1. A Cicchetti test (Cicchetti, 1972) indicated that (1) the words that were repeated in the male's voice were recognized more frequently if they were spoken by the male on the test tape ($p < .05$), (2) the words that were repeated in the female's voice were recognized more frequently if they were spoken by the female on the test tape ($p < .05$), and (3) the words that were repeated in the subject's own voice were not affected by the sex of the speaker on the test tape. With respect to the third result, for words that were repeated in the subject's own voice, it was not easier to recognize the words later if the sex of the speaker at test matched the sex of the subject. The male subjects recognized 58% of the words that were repeated in their own voices regardless of whether the test voice was male or female. The female subjects recognized 61% of the words that were repeated in their own voices if the test voice was male and 60% if the test voice was female.

The analysis of variance conducted on the distractor words showed no significant effects and, therefore, the subjects did not exhibit a response bias associated with the male speaker (false alarm rate = .19) or with the female speaker (false alarm rate = .20).

The analysis of variance conducted on the questionnaire data indicated a main effect of repetition mode [$F(2,48) = 4.96$, $MSe = .34$, $p < .025$]. A Tukey's HSD test (Kirk, 1968) showed that the subjects found it easier to repeat the words in their own voice than in the female's voice ($p < .05$) or in the male's voice

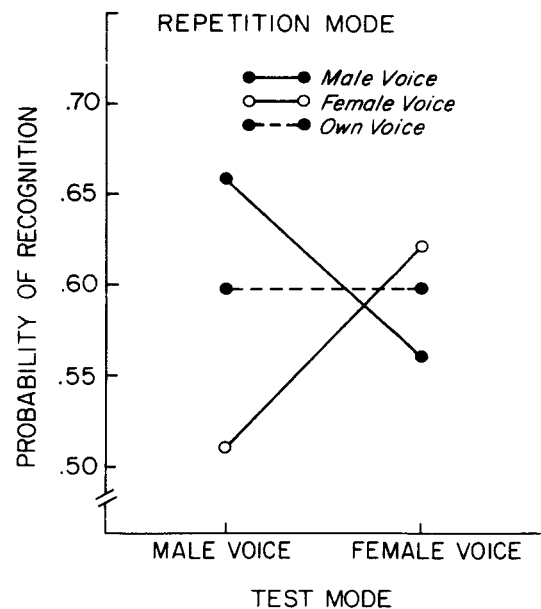


Figure 1. Probability of word recognition as a function of repetition mode and test mode in Experiment 1. The intrusion rate for new words was .19.

($p < .01$). Despite these differences, it is instructive to note that all three mean ratings were near or above the "fairly well" category. Hence, most of the subjects felt that they were able to perform the intended repetition task. The test-mode factor did not interact with the sex-of-subject factor, therefore, the subjects found it just as easy to image a voice of the opposite sex as to image a voice of their own sex.

Discussion

The results of Experiment 1 indicate that it is certainly not necessary to suggest that subjects have access to a literal copy of an auditory event in memory at test. The subjects in the present experiment did not receive external auditory stimuli, but rather constructed auditory images for visually presented words from their recollections of two unfamiliar voices. In spite of this dramatic procedural difference, the recognition results were found to be of the same pattern as that observed by Craik and Kirsner (1974). Words that were repeated in a male's (female's) voice were recognized more frequently if they were presented in the male's (female's) voice at test. The overall level of recognition was much lower in this experiment than in the Craik and Kirsner study, but the present test measured incidental learning, whereas Craik and Kirsner used a continuous recognition paradigm. Also, the magnitude of the same-voice effect was much larger in the present experiment than the approximate 4% difference reported by Craik and Kirsner. However, the present subjects processed the words for 10 sec under repetition instructions, whereas Craik and Kirsner's subjects studied each word

for later recognition for 4 sec. An extended repetition process may be more conducive to speaker's voice becoming an integral part of the memory codes.

It has been concluded elsewhere that long-term literal copies probably do not exist in the visual mode (Light & Berger, 1976) and the present data suggest that there is also no compelling reason to assume that long-term literal copies exist in the auditory mode. The precise nature of the influence of a speaker's voice on the code for an item in memory was not addressed by Experiment 1, but two possibilities were mentioned earlier (Carterette & Barnebey, 1975; Geiselman & Bellezza, 1976, 1977).

In addition, it appears that if an item is repeated in one's own voice, the sex of the speaker at test does not interact with the sex of the subject to influence the recognition of the item. This result is consistent with previous findings that the sex of the speaker of an item at presentation does not interact with the sex of the subject to influence the later retention of the item (Hintzman et al., 1972). One's own "inner voice" appears to be neutral with respect to any influence on the memory code. It is certainly not clear why this should be the case, but possibly one's inner voice is neutral because it acts as a partial interface between perception and higher order coding processes. That the subjects reported the repetition task to be easier if the repetition was conducted in their own voice is probably the result of the extra effort involved in maintaining the auditory image in an altered inner voice. This extra effort, however, did not affect the overall probability of incidental recognition of the words because the main effect of repetition mode was not significant.

EXPERIMENT 2

An interesting question that can be raised in reference to either the Carterette and Barnebey (1975) hypothesis or the Geiselman and Bellezza (1976, 1977) hypothesis is how specific are the representational propositions in memory corresponding to a speaker's voice? In other words, does one unfamiliar male's (female's) voice have a significantly different influence on the memory code than another unfamiliar male's (female's) voice? The purpose of Experiment 2 was to answer this question. The procedure for Experiment 2 was the same as that in Experiment 1 except different male and female voices were used for the recognition test than were used by the subjects to repeat the words during presentation. If the facilitative effect on recognition of preserving the sex of the speaker at test was observed to be as large as that in Experiment 1, then it could be assumed that the initial repetition process influenced the memory code in essentially a dichotomous manner rather than in a more specific way. If the same-sex voice effect on recognition was found to be greatly reduced or eliminated, then it could be assumed that the initial repetition process influenced the memory code in a manner specific to a given male or female voice.

Method

Subjects. The subjects were 15 male and 15 female undergraduate volunteers from the introductory psychology course at the University of California at Los Angeles.

Materials and procedure. The materials and procedure were the same as in Experiment 1 except that the test tape recording was constructed using different male and female speakers than those used to construct the introductory and test tape recordings in Experiment 1. In other words, in Experiment 2 the voices at test did not match the voices that the subjects used to auditorily image the words at presentation. One word from each slide was spoken by the new male speaker and the remaining word was spoken by the new female speaker on the test tape. The fundamental frequency for the new male speaker was estimated to be 135 Hz and was 233 Hz for the new female speaker. These voices were higher in pitch than those used to make the introductory tape recording.

Results

The analysis of variance conducted on the recognition data for the words presented on the slides showed that the Repetition Mode by Test Mode interaction was again significant [$F(2,48) = 4.39$, $MSE = .15$, $p < .025$]. However, unlike Experiment 1, the main effect of repetition mode was also significant [$F(2,48) = 4.36$, $MSE = .14$, $p < .025$]. A Tukey's HSD test indicated that the words that were repeated in the subject's own voice were recognized more frequently than the words that were repeated in another person's voice ($p < .05$). These results are shown in Figure 2. As in Experiment 1, for words that were repeated in the subjects's own voice, it was not easier to recognize the words later if the sex of the speaker at test matched the sex of the subject. Also as in Experiment 1, the analysis of variance conducted on the distractor words showed no significant effects, with the false-alarm rate for distractors spoken

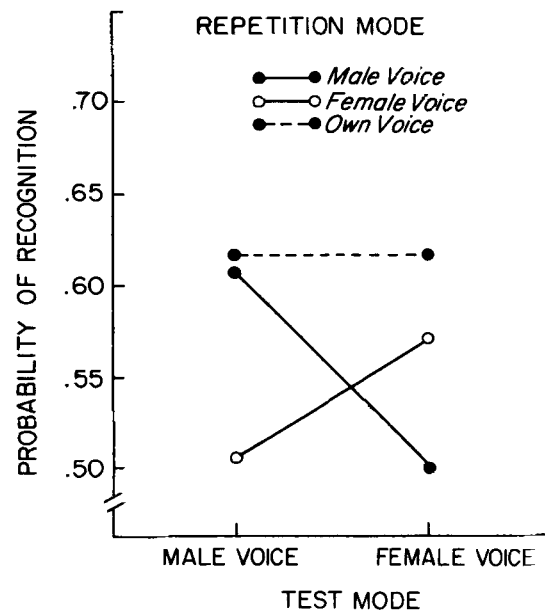


Figure 2. Probability of word recognition as a function of repetition mode and test mode in Experiment 2. The intrusion rate for new words was .17.

by the male being .16 and .17 for distractors spoken by the female.

The analysis of variance conducted on the questionnaire data indicated that the main effect of repetition mode was again significant [$F(2,48) = 8.40$, $MSe = .41$, $p < .001$]. As in Experiment 1, the subjects found it easier to repeat the words in their own voice ($p < .01$). All three mean ratings were near or above the "fairly well" category.

Discussion

The results of Experiment 2 suggest that repeating an item in terms of another person's voice influences the memory code in both a general manner (male or female) and in a manner that is specific to the particular voice used. On the one hand, the same-sex voice effect was observed even though the voices used at test did not match the voices used by the subjects to auditorily image the words at presentation. This indicates that the influence of a speaker's voice on the memory code is somewhat general in that greater similarity between the repetition mode and the test mode produces a greater probability of recognizing the item itself. On the other hand, the overall probability of recognizing an item that was repeated in another person's voice was lower in Experiment 2 than the probability of recognizing an item that was repeated in the subject's own voice. Since this was not the case in Experiment 1 and the test voice never actually matched the repetition voice in Experiment 2, the different pattern of results indicates that the influence of a speaker's voice on the memory code is somewhat specific to the particular voice used.

These results can be interpreted in terms of either the Carterette and Barnebey (1975) hypothesis or the Geiselman and Bellezza (1976, 1977) hypothesis. Specific values of a small number of attributes of an unfamiliar voice may be encoded in parallel in propositional form and/or these values may form a specific connotation that influences the meaning of the code. In either case, it is not necessary to imply that subjects have literal copies of auditory events in memory such that the events are more likely to be recognized later if the test stimulus physically matches the original stimulus (Craik & Kirsner, 1974). The subjects in both of the present experiments constructed the original "auditory" events and hence could not possibly have had auditory literal copies of the items in memory at test. The role of visual imagery in the constructive process is not clear, but several subjects reported that forming an auditory image seemed to automatically elicit a visual image of the male or female speaker, even though they had not seen the speakers. Vocal "impressionists" intentionally form such visual images.

Looking at the data from both experiments from a different perspective, the words were more likely to be recognized if they were encoded in another person's voice and were tested in that voice than if the words were encoded in the subject's own voice. However, if the

words were tested in a different person's voice than that used at encoding, the relation was reversed. Using the "own-voice" curves as a baseline, the magnitude of the deleterious effect due to a voice mismatch appeared to be inversely related to the similarity between the test voice and the voice used at encoding. This pattern suggests that the influence of a speaker's voice on the memory code is not peripheral in nature, whereby the original voice attributes can be "reinstated" at test to facilitate familiarity judgments (Light & Berger, 1976). If encoded, speaker's voice can either facilitate or attenuate the later recognition of an item depending on the voice in which the item is tested. This implies that speaker's voice may form or influence an integral part of the memory code as suggested by Geiselman and Bellezza (1977). Since the probability of recognizing an item that was repeated in the subject's own voice was not affected by the sex of the speaker at test, regardless of the sex of the subject, it appears that one's "inner voice" is neutral with respect to any influence on the code for an item in memory.

The general effects of initial context on later recognition could explain why Woodward, Bjork, and Jongeward (1973) found improved recognition of words with a longer period of rote cyclic rehearsal, even though recall performance remained unchanged. With a greater amount of rote cyclic rehearsal, the context in which an item is presented may be more likely to be stored in memory as an integral part of the item itself. Recognition judgments would be expected to be more accurate, provided that the context at presentation was preserved at test, but the storage of such contextual information would not necessarily affect the retrieval of the items. Experiments are currently being conducted to test this notion using speaker's voice as contextual information.

Another effect of imagining a speaker's voice on the processing of material presented visually has been observed by Kosslyn and Matt (1977) when subjects were given implicit voice-imagery instructions. In their experiment, subjects were introduced to two speakers by name using a tape recording where one speaker spoke at a slow rate and the other speaker spoke at a fast rate. Then the subjects were asked to read two unrelated passages, after being told that one passage was written by the slow speaker and the other passage was written by the fast speaker. It was found that approximately 75% of the subjects reported "hearing" the voice of the author while reading each passage and the passage that was said to be written by the slow speaker was read at a slower pace. Therefore, reading rate appears to be another variable that is influenced by imagining a speaker's voice for material presented visually.

In conclusion, it may be premature to assume that the same-voice effect on word recognition is the result of a physical matching process between the test stimulus and a "literal episodic record" of the original stimulus. It should also be noted that, even though imaging the words in another person's voice seemed to be effortful for the subjects, the question of whether the storage of

speaker's voice requires extra processing time is quite independent of the storage-format issue (Geiselman & Bellezza, 1976; Light & Berger, 1976). With auditory presentation, speaker's voice may automatically influence the representation in memory of what is said (Geiselman & Bellezza, 1976) and/or the storage of speaker's voice may constitute an alternative way of forming part of the memory code (Geiselman & Bellezza, 1977; Underwood, 1969) such that no extra time is required.

REFERENCE NOTES

1. Geiselman, R. E., & Bellezza, F. S. *Incidental retention of speaker's voice*. Manuscript submitted for publication, 1977.
2. Massaro, D. W. Personal communication, April 19, 1977.

REFERENCES

- CARTERETTE, E. C., & BARNEBEY, A. Recognition memory for voices. In A. Cohen & S. Nootboom (Eds.), *Structure and process in speech perception*. Heidelberg & New York: Springer Verlag, 1975.
- CICCHETTI, D. V. Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin*, 1972, 77, 405-408.
- CRAIK, F. I. M., & KIRSNER, K. The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, 1974, 26, 274-284.
- CROWDER, R. G. Language and memory. In J. P. Kavanagh & W. Strange (Eds.), *Speech and language in the laboratory, school, and clinic*. Cambridge: MIT Press, in press.
- CROWDER, R. G., & MORTON, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, 5, 365-373.
- CUTTING, J. E. There may be nothing peculiar to perceiving in a speech mode. In J. Requin (Ed.), *Attention and performance VII*. New York: Academic Press, in press.
- GEISELMAN, R. E., & BELLEZZA, F. S. Long-term memory for speaker's voice and source location. *Memory & Cognition*, 1976, 4, 483-489.
- HALLE, M., & STEVENS, J. A. Speech recognition: A model and a program for research. *IRE transactions on information theory*, 1962, IT-8, 155-159.
- HINTZMAN, D. L., BLOCK, R. A., & INSKEEP, N. R. Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 741-749.
- KIRK, R. E. *Experimental design: Procedures for the behavioral sciences*. Belmont, Calif: Wadsworth, 1968.
- KOSSLYN, S. M., & MATT, A. M. C. If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society*, 1977, 9, 250-252.
- LIGHT, L. L., & BERGER, D. E. Are there long-term "literal copies" of visually presented words? *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 654-662.
- MASSARO, D. W. *Understanding language*. New York: Academic Press, 1975.
- MORTON, J. A functional model of memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. Pp. 203-254.
- NORMAN, D. A. *Memory and attention*. New York: Wiley, 1976.
- PYLYSHYN, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 1973, 80, 1-24.
- WOODWARD, A. E., BJORK, R. A., & JONGEWARD, R. H. Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 608-617.
- UNDERWOOD, B. J. Attributes of memory. *Psychological Review*, 1969, 76, 559-573.

NOTE

1. From the early formalization of precategorical acoustic storage (PAS) (Crowder & Morton, 1969), it would be predicted that nonlinguistic aspects of auditory stimuli, such as fundamental frequency, are displaced from PAS by subsequent auditory events or that they decay within 2 sec (Morton, 1970) unless they are recoded linguistically. Even though there is currently some support for the notion that some nonlinguistic sounds are also perceived categorically (Cutting, in press), such a categorization process would yield a description of an auditory event in memory, not a literal copy. Similarly, in Massaro's (1975) model of speech processing, the utilization of synthesized auditory memory to retain speaker's voice cannot be in terms of direct literal copies but must consist of flexible heuristic routines utilized in auditory information processing (Massaro, Note 2).

(Received for publication April 19, 1977;
revision accepted June 10, 1977.)