Short-lived auditory memory for pitch

KAREN G. FOREIT

Brooklyn College of CUNY, Brooklyn, New York 11210

The effect of a stimulus suffix on immediate serial recall of lists of tones was assessed with 50- and 350-msec complex tones. Although recency effects were not found under the control conditions, the addition of a stimulus suffix significantly degraded recall at the final serial position for both stimulus durations. The implications of these results for models of auditory memory and speech perception are discussed.

According to one model of auditory memory (Crowder & Morton, 1969), each new item of an aurally presented list sets up a short-lived sensory trace and, in so doing, degrades traces of previous items. The recency effect typically found in serial recall of aurally presented lists may be attributed to utilization of sensory information for the last item, information which is unavailable for items earlier in the list. By this reasoning, adding a final, not-to-be-recalled item at the end of the list (a stimulus suffix) should selectively degrade recall for the final serial position(s) because it degrades sensory information which would otherwise be available.

In the studies that followed the original Crowder and Morton paper, the test items and stimulus suffixes were speech stimuli—words and/or syllables. It seemed logical, therefore, to test the paradigm with nonspeech stimuli, since the model did not appear necessarily to exclude that kind of information. (A suggestion had been made by Morton, 1970, that precategorical acoustic storage, PAS, was a function of the speech system, but this suggestion appears to have been dropped in light of later evidence.) The present study, then, was designed as an empirical test of the stimulus suffix effect with nonlinguistic stimuli.

In addition to simply extending the model to the nonlinguistic domain, the present study was designed to test some of the explanations for recall differences among different classes of speech stimuli first reported by Crowder (1971). Using lists of synthetic consonants and vowels, Crowder showed that while written recall of vowel lists showed both recency and stimulus suffix effects, recall of consonant lists showed neither. Analogous results were obtained by Cole (1973) using recall of real-speech vowels and consonants: spoken recall of vowel lists

showed greater recency effects than did recall of consonants. Apparently, then, auditory sensory memory shows a greater sensitivity for traces of vowels than for consonantal information.

Several possible interpretations of these data have been suggested (see Crowder, 1973b), based on the different acoustic properties and differential auditory processing of consonants and vowels. (1) The portion of the consonant-vowel syllable carrying consonantal information is rather short—approximately 50 msec in the stimuli used by Crowder—while the vocalic portion persists for a far longer time—upwards to 300 msec. Sensory storage of very short stimuli may be inferior to storage for longer stimuli. (2) Consonantal information is carried by rapidly changing formant transitions, while vowels are characterized by steady state acoustic energy (see Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Transitional acoustic information may be poorly represented in auditory memory. (3) Studies of identification and discrimination of speech sounds (Liberman et al., 1967) indicate that consonant perception is rather categorical in nature, while vowel perception is more continuous. It is more often the case that subjects can detect differences between stop consonants only when the stimuli belong to different phonetic classes, while small differences between vowels having no phonetic significance are readily discriminable. Phonetic processing of consonants may reduce their representation in a precategorical sensory store.

Of the preceding hypotheses, Crowder appears to favor the explanation of categorical perception, based on his studies of long and short vowels, and of fricatives (Crowder, 1973a, b). In a paper published after the present research was completed, Darwin and Baddeley (1974) have challenged this conclusion, reducing Crowder's findings to the dimension of auditory confusability. Using recall of lists of vowels, Darwin and Baddeley significantly reduced the size of both the recency and suffix effects by increasing interitem similarity. The vocalic portions of their stimuli were quite short—on the

This research was supported by a Bio-Medical Sciences subgrant to Brown University and by FRAP 10701 from the Research Foundation of CUNY. The author wishes to thank Dr. Franklin S. Cooper, for making available the facilities of Haskins Laboratories, and Dr. Peter D. Eimas, for his critical reading of early drafts of this paper.

Figure 1. Mean total errors by serial position for long and short tones.

order of 30 msec—thus casting doubt on the first hypothesis, that PAS requires longer stimulus durations, and their finding of a significant suffix effect with transient-energy diphthongs specifically rules out the second hypothesis, that PAS preserves steady state but not transient acoustic energy.

The study reported below, comparing written recall of lists of long and short wide-band tones, was similar in design to Crowder's (1973b) vowel study. The implications of the study are twofold: (1) it is an empirical test of the stimulus suffix effect with nonlinguistic stimuli; and (2) by comparing recall of stimuli of differing durations, it offers additional evidence concerning the hypothesis that only sufficiently long stimuli receive PAS representation.

METHOD

Stimuli

Complex steady state tones were synthesized on the parallel resonance synthesizer at the Haskins Laboratories. They were created by imposing a 100-Hz first formant on carrier waves of differing fundamental frequencies. Because they had some, but not all, of the spectral characteristics of speech, the stimuli appeared to the listener to have the resonance, but not the speech, characteristics of synthetic speech. Three easily discriminable tones with fundamental frequencies of 112, 154, and 196 Hz were used for the test items; a fourth, 70-Hz, tone served as stimulus suffix. The test tones were randomly combined to generate 50 seven-item lists, with the restrictions that such obvious patterns as repeating alternations and repeating ascending and/or descending sequences be eliminated and that across lists each tone appear equally often in each serial position. The block of 50 lists was recorded twice on an Ampex tape recorder, once with 50-msec test tones and once with 350-msec test tones. An interstimulus interval of 500 msec was used for both long and short tones, with 15 sec between trials. The same 350-msec tone served as stimulus suffix for all lists; in the control conditions, the suffix tone was replaced with a light.

Procedure

Subjects were tested in groups of two to four. Twelve subjects heard the short (50-msec) test tones, and the other

12, the long (350-msec) tones. The stimuli were played in a quiet room on a Sony tape recorder over a single external speaker. Subjects were given the usual instructions for this paradigm (see Crowder & Morton, 1969), familiarization with the test tones, and 10 practice trials in the condition they were to receive first. After this training, all subjects heard the block of 50 lists twice, once with the stimulus suffix and once under the control condition; before the second block of trials, they received another five practice trials for the new condition. Order of presentation of control and suffix conditions was counter-balanced across subjects.

Recall was written in serial order on a three-line musical "staff" marked off into seven "measures." High (196-Hz) tones were written as circles on the top line, medium (154-Hz) tones on the middle line, and low (112-Hz) tones on the bottom line. The subjects were told to guess when uncertain and to visualize the list as a spatial array rather than to verbally encode the test items as they were presented. Although many subjects attempted a verbal strategy during the first set of practice trials, none reported verbal encoding during the test lists.

RESULTS

Figure 1 presents the mean total errors by serial position for the control and suffix conditions with the 50- and 350-msec stimuli. A 2 (recall condition) by 2 (stimulus duration) by 7 (serial position) analysis of variance was performed on the raw error scores. The between-subject variable, stimulus duration, failed to show a significant effect, while all other main effects and interactions did show statistical significance. Recall steadily deteriorated with increasing serial position [$F(6,286) = 184.68$, $p < .01$] with only a nonsignificant improvement at the final position under the control conditions. Correlated t tests between the control and stimulus suffix conditions showed that, with the exception of Position 3 for the long tones only, the groups did not differ until the final two positions for the long tones [at Position 6, $t(11) = 3.93$, $p < .01$; at Position 7, $t(11) = 2.68$, $p < .05$], and at Position 7 for short tones [$t(11) = 2.70$, $p < .05$]. This suffix effect was the basis for the overall inferiority of the suffix conditions [$F(1,286) = 32.31$, $p < .01$] and the significant Recall Condition by Serial Position interaction [$F(6,286) = 3.07$, $p < .01$].

The significant interactions involving the factor of stimulus duration can be attributed to differences in the slopes of the curves and greater overall debilitation under the stimulus suffix condition for long tones. The suffix and control curves for long stimuli showed significant divergence at serial Positions 3, 6, and 7, while for the short tones the curves differed only at Position 7. However, the seemingly larger stimulus suffix effect at Position 7 for long tones than for short tones was not statistically reliable.

DISCUSSION

These results clearly indicate auditory sensory

storage of nonlinguistic material and, in so doing, extend the PAS model into the nonspeech domain. They further indicate PAS sensitivity to rather short stimuli, arguing against the hypothesis that failure to store consonantal information is a function of the brevity of the consonantal portion of the stimulus. Changing the duration of the stimulus may attenuate the duration of its sensory representation, accounting for the Stimulus Length by Recall Condition interactions found in the present study. However, it should be emphasized that reducing stimulus duration to 50 msec neither abolished the overall suffix effect nor reduced its magnitude at the final serial position. (Just the reverse might be argued, using the significant interactions with stimulus duration. But the assertion that decreasing stimulus duration reduces the size of the suffix effect ultimately rests on the difference between conditions at serial Position 6, an argument which does not seem altogether convincing.)

In addition, these results are compatible with Darwin and Baddeley's (1974) acoustic confusability hypothesis. The subjects in the present study seemed to experience little difficulty in discriminating among the stimuli used, and other researchers (see Flanagan & Saslow, 1958) have reported accurate discrimination for shorter tones closer together in fundamental frequency than used in the present study. Reducing stimulus duration might seem intuitively to make the tones less discriminable. However, this manipulation did not increase con-

fusability, in that the size of the suffix effect was not diminished by using shorter stimuli. Thus, the Darwin and Baddeley argument is still left with the problem, first noted by the authors, of defining stimulus confusability a priori.

REFERENCES

- COLE, R. A. Different memory functions for consonants and vowels. *Cognitive Psychology*, 1973, 4, 39-54.
- CROWDER, R. G. The sounds of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 587-596.
- CROWDER, R. G. Precategorical acoustic storage for vowels of long and short duration. *Perception & Psychophysics*, 1973, 13, 502-506. (a)
- CROWDER, R. G. Representation of speech sounds in precategorical acoustic storage. *Journal of Experimental Psychology*, 1973, 98, 14-24. (b)
- CROWDER, R. G., & MORTON, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, 5, 365-373.
- DARWIN, C. J., & BADDELEY, A. D. Acoustic memory and the perception of speech. *Cognitive Psychology*, 1974, 6, 41-60.
- FLANAGAN, J. L., & SASLOW, M. G. Pitch discrimination for synthetic vowels. *Journal of the Acoustical Society of America*, 1958, 30, 435-442.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- MORTON, J. A. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

(Received November 21, 1975;
revision received January 27, 1976.)