

Mechanisms of aural encoding: VII. Differences in consonant and vowel recall in a Peterson and Peterson short-term memory paradigm*

RONALD A. COLE

University of Waterloo, Waterloo, Ontario, Canada

BRUCE D. SALES

Northwestern University, Evanston, Illinois 60201

and

RALPH N. HABER

University of Rochester, Rochester, New York 14627

Ss either saw or heard lists of three syllables which differed by their initial consonant phoneme or their final vowel phoneme. After 5 or 15 sec of mental arithmetic, Ss were required to recall the syllables. Following auditory presentation, vowels were recalled more accurately than consonants in all serial positions and at both delays. In addition, spoken consonants and vowels showed primacy and recency effects. Following visual presentation, consonants and vowels were recalled with equal accuracy at both delays, and no recency effects were observed. These data suggest that superior recall of vowels over consonants results from differential decay of these stimuli in an acoustic storage. These data are consistent with previous experiments showing that, during serial recall, the final vowels in a sequence are recalled more accurately than the final consonants.

When Ss are presented with a series of CV or VC syllables for immediate recall, they are more likely to recall the vowels than the consonants (Cole, 1973; Crowder, 1971a). Superior recall for vowels has been observed only in the final syllables of a sequence, and only for auditorily presented sequences.

Differences in recall of consonant and vowel sounds may be viewed in terms of a general theory of auditory memory proposed by Crowder and Morton (1969) and elaborated elsewhere by Crowder (1970, 1971b). This theory postulates that *spoken* items are stored in two distinct ways in memory: in a sensory, precategorical acoustic storage (PAS) and in a verbal short-term memory where information is coded in a linguistic form independently of its input modality.

It is assumed that items in PAS decay with time, and that they are replaced by more recently heard stimuli. Thus, when a S is presented with a series of auditory stimuli, only the most recently heard stimuli are preserved in a physical form in PAS. Retrieval of information from PAS is assumed to account for recency effects which are observed for auditorily, but not visually, presented stimuli.

There is much recent evidence that speech sounds are stored in a prelinguistic form in a sensory or acoustic storage (see Massaro, 1972, for a review of this literature). In fact, as Neisser (1967) points out, a

sensory storage is a necessary condition for speech perception, since many phonemic discriminations require comparisons of the acoustic input at successive points in time.

One interpretation of the difference in recall for spoken consonants and vowels is that vowels are preserved in acoustic storage more efficiently than are consonants. Thus, when S tries to recall the final items in a series of syllables, he is able to retrieve information from PAS about the most recently heard vowels in the sequence, but not about the most recently heard consonants, since they have decayed from acoustic storage. If consonant information decays from acoustic storage before S is able to report the final items in a sequence, then vowels would show a greater recency effect than consonants in the final serial positions of a sequence.

It is quite possible that an acoustic storage does preserve the spectral properties of vowels, but not consonants. Consonants and vowels differ markedly in their acoustic properties. Consonants are characterized by bursts of noise and transient waveforms, while vowels are characterized by steady-state concentrations of resonant energy.

The present experiment was designed to further test the hypothesis that the recency effect previously observed for vowels—but not consonants—results from differential decay of vowels and consonants in an acoustic storage. This hypothesis was tested by engaging S in an intervening task designed to minimize rehearsal prior to recall of a series of syllables. To the extent that an intervening task prevents S from rehearsing stimuli in short-term memory (Peterson & Peterson, 1959), any

*This research was supported in part by a grant from the National Research Council of Canada awarded to the first author. We thank Brenda McPhee for her help in running Ss and Sue Dotzenroth for her helpful comments.

†Now at the University of Nebraska at Lincoln.

differences in recall of consonants and vowels should reflect the differential decay of these stimuli in acoustic storage. If this storage truly depends upon *hearing* the stimuli, then vowels should be recalled more accurately than consonants only when syllables are *spoken*. No differences in recall of consonants and vowels should be observed when syllables are presented visually.

METHOD

Subjects

Thirty students from the University of Waterloo served as Ss. Each S was tested individually in a small room with E, in a session lasting approximately 1 h. Ss were paid for their participation.

Stimuli

The syllables /ba, ta, ga, ja, sa/ and /di, de, da, do, du/ were used as stimuli. Stimuli were presented either auditorily or visually, depending on the experimental group. In the auditory condition, syllables were spoken in a male voice at the rate of 1/sec. In the visual condition, stimuli were shown individually at the same rate. Syllables differing by their vowel phoneme were presented visually as "dee," "deh," "dah," "doe," "doo."

Procedure

The experiment consisted of 120 three-syllable trials. On 60 trials, S heard syllables differing by their initial consonant phoneme (e.g., /ba, ga, sa/), while the other 60 trials consisted of syllables differing by their terminal vowel phoneme (e.g., /di, da, du/).

On a given trial, S heard the word "ready," followed 2 sec later by the three syllables. Immediately after presentation of the third syllable, S saw a three-digit number from which he was to count silently backwards by sevens. Following a delay of either 5 or 15 sec, S heard a "click" which served as a signal to first write the number he had arrived at by counting backwards, and then report verbally the three syllables in their proper order.

In order to insure that Ss were actually engaged in the intervening task, an incentive was provided for good performance on the mental arithmetic. Each S was told that he was competing against five other Ss for a \$5 prize, and that he would win the \$5 if the total of the numbers he arrived at after mental subtraction was lower than the other Ss in his group.

Thus, all Ss were presented with a series of three syllables, which they were required to recall after a (filled) interval of 5 or 15 sec. Both delay and syllable type were varied randomly from trial to trial. However, each syllable appeared in each serial position an equal number of times at each delay.

Ss were assigned randomly to either the auditory or the visual condition. In the auditory condition, instructions and trials were presented via a Sony Model TC630 tape recorder connected to a Dynaco loudspeaker. After hearing the final syllable on each trial, S viewed the three-digit number by moving a piece of plastic one space down his answer sheet, causing a number to be exposed. In the visual condition, all stimuli were presented on a series of slides via a Kodak Carousel projector.

RESULTS

Figure 1 displays the mean number of vowels and consonants correctly recalled at each serial position at each delay and for both presentation modalities. For auditory presentation, vowels were recalled more

accurately than consonants at all serial positions after 5 or 15 sec. For visual presentation, recall of consonants and vowels was identical. Analysis of variance revealed significant main effects of presentation modality, syllable type (C vs V), delay, and serial position ($p < .01$ in all cases).

All two-way interactions were significant ($p < .02$ in all cases). Thus, vowels were recalled more frequently than consonants for auditory presentation, but not for visual. Input modality interacted with recall at the different serial positions in an interesting manner. For auditory presentation, both consonants and vowels show a U-shaped function (except for vowels at 5 sec, where virtually no forgetting occurs), while visually presented syllables show a primary effect, but no recency effect. Finally, input modality interacted with delay; decay of spoken consonants and vowels was greater between 5 and 15 sec than it was for visually presented syllables.

Separate within-Ss analyses of variance were computed for the auditory and visual conditions. For auditory presentation, all main effects (syllable type, delay, and serial position) and all interactions were significant ($p < .01$). For visual presentation, main effects were observed for delay and serial position ($p < .01$), but not for syllable type. The only significant interaction occurred between delay and serial position ($p < .02$).

DISCUSSION

The main result of this experiment is that spoken vowels are recalled more accurately than spoken consonants following 5 or 15 sec of mental arithmetic prior to recall. These data suggest that previously observed differences in recall of consonants and vowels result from decay of consonants, but not vowels, in an acoustic storage.

Did the task completely eliminate rehearsal? It is likely that Ss originally coded syllables in verbal short-term memory in a form appropriate for rehearsal. Moreover, some Ss may have rehearsed syllables silently during their presentation. However, a rehearsal mechanism does not provide a satisfactory explanation for the observed differences in consonant and vowel recall. For example, there is no reason to suspect that rehearsal would occur differentially for consonant and vowel phonemes. In fact, Cole, Haber, and Sales (1973) have shown that identical distinctive feature coding strategies are used to remember consonants and vowels in short-term memory when Ss are allowed to rehearse the material.

One explanation for our results is that spoken consonants and vowels are both originally maintained in an acoustic storage, but that consonants decay very quickly, say within 1 sec, whereas vowels are maintained for many seconds.

Unfortunately, a memory experiment of this type

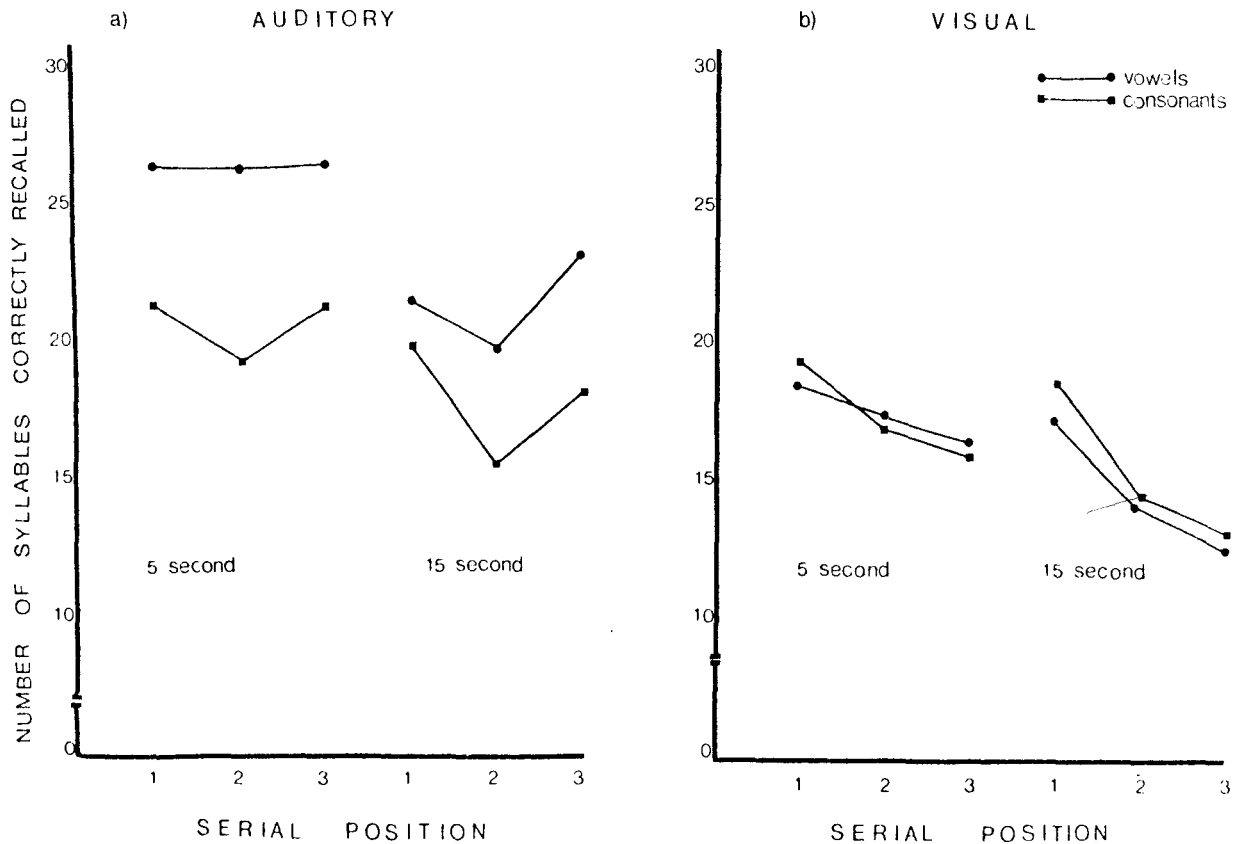


Fig. 1. Mean number of consonant and vowel phonemes recalled at each delay as a function of serial position for the two presentation modalities.

cannot yield information about the rate of decay of stimuli in acoustic storage. One difficulty is that forgetting of three items does not occur in the absence of an interpolated task, so that it is difficult to measure decay of information immediately following presentation. Thus, the relative rate of decay of consonants and vowels as shown in Fig. 1 may not represent the actual decay of items in acoustic storage. In fact, Fig. 1 shows that the drop in recall accuracy between 5 and 15 sec is greater for vowels than for consonants for spoken items, and one could infer that vowels decay *faster* than consonants. However, if consonants decay from acoustic storage very quickly, so that decay is virtually complete before 5 sec, while vowels decay gradually, then vowels would appear to decay faster than consonants when sampled at 5 and 15 sec. In order to examine decay of information in acoustic storage, one must turn to reaction time procedures or partial report procedures analogous to those employed by Sperling (1960) or Posner (1967) for visually presented letters.

One strategy that Ss may use to remember a series of stimuli is to rapidly rehearse them silently before actually recalling them.¹ In the present experiment, we may assume that Ss engaged in this rapid silent rehearsal after hearing the third syllable on each trial before beginning the interpolated mental activity. This rehearsal

would serve the purpose of establishing a memory trace for each stimulus in verbal short-term memory. Thus, Ss could retrieve information about *both* consonants and vowels during a silent rehearsal before they decayed from acoustic storage.

The overall advantage of vowels over consonants would then be accounted for by assuming that vowels decay more slowly than consonants from acoustic storage, so that Ss are able to retrieve vowels from acoustic storage *at the time of recall*. In fact, although Ss were not asked about their recall strategies, Ss mentioned that they could still "hear" the vowels at the time of recall.

An alternative explanation to the decay hypothesis is that differences in consonant and vowel recall are caused by differences in perception of the two types of stimuli. There are two major differences in perception of consonants and vowels. First, perception of consonants is categorical, while perception of vowels is not (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Second, when pairs of consonants or pairs of vowels are presented to both ears simultaneously, consonants show a right-ear (left-hemisphere) superiority, while vowels do not (Shankweiler & Studdert-Kennedy, 1967). These findings suggest that consonants and vowels are perceived quite differently. In view of these perceptual differences, it is reasonable to

suspect that errors in the recall of consonants and vowels may be caused by differences in the initial processing of the two types of sounds, rather than differences in their rate of decay from a common storage. For example, we might suppose that consonants are recognized directly from the speech wave by their distinctive features, whereas vowels are recognized only after they are analyzed in an acoustic storage.

However, recency effects were observed for both consonants and vowels following auditory presentation, while no such recency effects were observed for visually presented syllables. The recency effect for spoken consonants is most easily explained by assuming that some information about the final consonant phoneme is retained in a precategorical acoustic storage and is rescued from this storage by a fast silent rehearsal before it can decay. Until further evidence is forthcoming, it is tentatively concluded that *both* consonants and vowels are originally maintained in an acoustic storage, and that differences in their recall is due to faster decay of consonants from this storage.

REFERENCES

- Cole, R. A. Different memory functions for consonants and vowels. *Cognitive Psychology*, 1973, 4, 39-54.
- 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 & Psychophysics*, 1973, 13, 87-92.

- Crowder, R. G. The role of one's own voice in immediate memory. *Cognitive Psychology*, 1970, 1, 157-178.
- Crowder, R. G. The sounds of consonants and vowels in immediate memory. *Journal of Verbal Learning & Verbal Behavior*, 1971a, 10, 587-596.
- Crowder, R. G. Waiting for the stimulus suffix: Decay, delay, rhythm, and readout in immediate memory. *Quarterly Journal of Experimental Psychology*, 1971b, 23, 324-340.
- Crowder, R. G., & Morton, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, 5, 365-373.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-460.
- Massaro, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, 79, 124-145.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, 58, 193-198.
- Posner, M. I. Chronometric analysis of classification. *Psychological Review*, 1967, 74, 392-409.
- Shankweiler, D., & Studdert-Kennedy, M. Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, 1967, 19, 59-63.
- Sperling, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, 74(Whole No. 498).

NOTE

1. We thank Robert Crowder for this interpretation.

(Received for publication July 16, 1973;
accepted August 11, 1973.)