

# Same-different reaction time to the sequential visual presentation of vowels and consonants

KIMBERLY D. PETERSON, J. RICHARD SIMON, and JYH-HONE WANG  
*University of Iowa, Iowa City, Iowa*

Subjects saw either two vowels (E and I) or two consonants (D and G) presented sequentially on a screen (ISI of either 250 or 2,000 msec) and pressed a key to indicate whether the letters were the same or different. As predicted, reactions were faster to vowel pairs than to consonant pairs. Results are explained in terms of (1) coding of letters in terms of their speech properties; (2) representation of vowel sounds but not consonant sounds in the covert rehearsal loop; (3) transfer of vowel sounds to the echoic store; and (4) faster comparison of those features represented in the echoic store.

In their model of information processing, Crowder and Morton (1969) postulated a precategorical acoustic storage (PAS) that retains auditory information in a prelinguistic form for at least a few seconds. In a later study, Crowder (1971) found that, with auditory presentation, syllables with the same consonant sounds but different vowel sounds (BOO, BIH, BEE) were confused less than were syllables with different consonant sounds but the same vowel sounds (BAH, DAH, GAH). This finding suggested that the PAS may selectively retain vowel sounds and not consonant sounds.

Conrad (1964) suggested that there may be selectivity in retaining vowel sounds even with visual presentation. He reported that letters that were most easily confused in an auditory presentation were also most easily confused in a visual presentation. These confusable letters were similar to syllables used in Crowder's study; that is, they differed only in their stop consonant sounds but had the same vowel sounds (B, D, P).

Sperling (1967) developed a model for visual information processing that may help to explain these phenomena. One version of this model (Crowder, 1982) includes a covert rehearsal loop in which "letters become coded (defined) in terms of their speech properties, which, in turn, leads to confusion errors based on speech" (p. 76). The model also suggests that the covert rehearsal loop transfers sounds into the echoic memory store.

The present study derives from the Sperling model. We hypothesized that, in a same-different task involving the sequential visual presentation of single vowels and consonants, reaction times (RTs) would be faster for vowel pairs than for consonant pairs. This difference was predicted based on the following assumptions: (1) letters are coded in terms of their speech properties; (2) vowel sounds, but not consonant sounds, are represented in the

covert rehearsal loop; (3) the covert rehearsal loop transfers the vowel sounds into the echoic store; and (4) reactions are faster to features represented in the echoic store (Posner & Rogers, 1978).

Two interstimulus intervals (ISIs), 250 and 2,000 msec, were employed to explore the time frame in which the covert rehearsal loop operates. We assumed that covert rehearsal begins very shortly after stimulus presentation, and that, as a result, sounds are transferred to the echoic store almost immediately. It is generally agreed that sounds in the echoic store persist for approximately 2 sec (Crowder, 1976). At 250 msec after presentation, the echo should be fairly strong, but it should fade considerably by 2,000 msec. Posner and Rogers (1978) have shown that RTs are faster to physical (echoic) than to phonetic representations of sounds. Therefore, we predicted that the RT advantage of vowel pairs over consonant pairs would be greater at the short than at the long ISI, because at the short ISI, the comparison would be based on a strong echo, whereas at the long ISI, it would be based on either a weaker echo or a phonetic code.

## METHOD

### Subjects

Subjects were 56 undergraduates, 28 males and 28 females, enrolled in an introductory psychology course at the University of Iowa. Subjects ranged in age from 18 to 22 years.

### Apparatus

The apparatus measured same-different RT to two letters presented one after the other on a 60-cm square rear-projection screen located 76 cm directly in front of the subject. Two response keys, one on the right and one on the left, rested on a table in front of the screen. One key was labeled "Same" and the other was labeled "Different." Subjects operated the keys with the index fingers of both hands.

A DEC PDP-11/34 laboratory computer was used to control the sequence of trials, timing of stimulus presentation, and recording and storing of responses. Kodak Ektagraphic slide projectors with attached Uniblitz Model 225 LOAOX high-speed shutters were used to present the stimuli, which were the vowels E and I and consonants D and G. The letters were chosen specifically so that there was a greater similarity in visual features between the two vowels than between the two consonants (Briggs & Hocevar, 1975). Each projector contained a slide with

This research was supported by funds from the Graduate College of the University of Iowa. K. D. Peterson was supported by the Undergraduate Scholar Assistantship Program. Requests for reprints should be sent to J. R. Simon, Department of Psychology, University of Iowa, Iowa City, IA 52242.

a single letter. Opening the shutter in front of a projector resulted in the letter's appearing on the screen. The opening and closing times for the shutters were 3 and 4 msec, respectively. The shutters were connected to the computer through two custom-made D/A converters. The apparatus, then, formed a precise, highly flexible, multichannel tachistoscope. Stimulus rise time was eliminated as a complicating factor since projector bulbs remained on throughout the session. Also, since the stimulus slides did not move during the experiment, perfect stimulus alignment was maintained.

Each trial began with a ready signal, that is, a 1-sec presentation of a small white dot with a luminance of 4.5 cd/m<sup>2</sup> in the center of the screen at eye level. A sequence of two letters was then presented. The first letter appeared 1 sec after the ready signal offset for a duration of 1,250 msec. After an ISI of either 250 or 2,000 msec, the second letter appeared and remained on the screen until the subject pressed a key. If there was no response within 5 sec, the program advanced automatically to the next trial. A 2-sec interval elapsed between the subject's response and the onset of the ready signal for the next trial. The first letter was presented slightly to the left of the center of the screen, and the second letter was presented slightly to the right. All letters were uppercase, white, 3.9 cm high, with a stroke width of 0.6 cm and a luminance of 9.0 cd/m<sup>2</sup>.

### Procedure and Experimental Design

Subjects were told that a sequence of two letters would appear on the screen. They were asked to silently rehearse the first letter and to respond after the second letter by pressing the "Same" or "Different" key to indicate whether the two letters were the same or different. Subjects were instructed to respond as quickly and as accurately as possible on each trial.

Eight combinations of the two letters were used: EE, EI, II, IE, DD, DG, GG, and GD. Note that, for each combination, both letters were either consonants or vowels and were either the same or different. There were no combinations involving a consonant and a vowel. Each combination was presented at each of two ISIs, 250 and 2,000 msec, thus creating 16 different trial types. Each subject performed on 128 test trials. The sequence was randomized in such a way that each trial type appeared once in each successive block of 16 trials. The test trials were preceded by 16 unscored warm-up trials and 16 practice trials in which each trial type appeared once.

The experimental design consisted of three within-subjects factors and two between-subjects factors. The within-subjects factors were letter type (consonant or vowel), response (same or different), and ISI (250 or 2,000 msec). The between-subjects factors were sex and key position. For half of the males and half of the females, the right key was labeled "Same" and the left key was labeled "Different." For the other half of each sex group, the key labels were reversed.

## RESULTS

Median RTs were computed for each subject for each of the eight treatment conditions, that is, for "same" and "different" responses to pairs of vowels and consonants presented with 250- and 2,000-msec ISIs. Trials on which errors were made (2.1%) were excluded from the analysis. Table 1 summarizes the means of the median RTs. An analysis of variance revealed that reactions to vowel pairs were significantly faster than to consonant pairs [496 vs. 503 msec;  $F(1,55) = 5.75, p < .02$ ]. "Same" responses were significantly faster than "different" responses [485 vs. 513 msec;  $F(1,55) = 19.99, p < .001$ ]. Reactions were significantly faster for the 2,000-msec ISI than for the 250-msec ISI [492 vs. 507 msec;  $F(1,55) = 9.36, p < .01$ ]. There were no significant interactions between any of the main effects.

**Table 1**  
Same-Different Reaction Time (in Milliseconds) to Vowel and Consonant Pairs Presented at Interstimulus Intervals of 250 and 2,000 msec

Letter Pairs	Reaction Time		
	Same	Different	Mean
	ISIs Combined		
Vowels	482	510	496
Consonants	489	517	503
Mean	485	513	
	250-msec ISI		
Vowels	486	517	501
Consonants	495	530	512
Mean	490	524	507
	2,000-msec ISI		
Vowels	478	502	490
Consonants	482	505	493
Mean	480	503	492

## DISCUSSION

Results of this experiment revealed that (1) "same"-"different" reactions were faster to vowel pairs than to consonant pairs; (2) "same" reactions were faster than "different" reactions; (3) RT was faster when the ISI was long (2,000 msec) than when it was short (250 msec); and (4) the RT difference between vowel and consonant pairs did not vary significantly as a function of ISI.

We had predicted the faster reactions to vowel pairs than to consonant pairs based on the notion that vowels would be retrieved from memory more quickly than consonants due to the presence of vowel sounds but not consonant sounds in the echoic store. Since stop consonant sounds are not represented in the rehearsal loop and, hence, are not available for transfer to the echoic store, the comparison of two consonants requires the retrieval of the visual features of the initial letter from visual working memory. Accessing working memory as opposed to sensory (echoic) memory results in longer RT (Posner & Rogers, 1978).

Consider, for example, a trial in which the letters D and G are presented sequentially. When the letter D is presented, the subject covertly rehearses the letter name "dee." Presumably, the covert rehearsal loop retains only the "ee" sound rather than the complete "dee" sound. The "ee" sound is then transferred to the echoic store. When the letter G appears for comparison, the subject cannot determine from the contents of the echoic store whether the initial letter was D or G, since both letters contain the vowel sound "ee." Therefore, in order to distinguish between the consonants, the subject must retrieve the visual features of the first letter. This retrieval involves accessing the visual working memory and results in slower RTs. Contrast this with a trial in which the letters E and I are presented. The sounds of the letter names, "ee" and "aye," are phonetically different representations in the echoic store and are, therefore, more rapidly discriminated.

An alternative explanation for our finding of faster reactions to vowels than to consonants might be that the specific vowels we employed were more easily distinguishable than the specific consonants in terms of their visual features. We think this explanation is highly unlikely since, based on visual features, the letters E and I are much more similar (confusable) than are the letters D and G (Briggs & Hocevar, 1975).

Our data confirmed the well-established finding that "same" reactions are faster than "different" reactions (e.g., Proctor, 1981). We also found, somewhat surprisingly, that reactions were significantly slower at the 250- than at the 2,000-msec ISI. In some sequential processing tasks, RT is slower at very short ISIs due to incomplete processing of the first stimulus before the second stimulus is presented (e.g., Posner & Boies, 1971). In our procedure, however, the first letter was exposed for 1,250 msec, which should have permitted complete processing even on trials with a 250-msec ISI. Perhaps the longer RT at the 250-msec ISI resulted from subjects' searching both the iconic and echoic

sensory registers. At the 2,000-msec ISI, the icon is no longer present (Crowder, 1976), so the search would involve only one sensory register and would result in shorter RT.

Finally, we found that although the RT difference between vowels and consonants did not vary significantly as a function of ISI, the vowel superiority tended to be greater at the 250- than at the 2,000-msec ISI. This finding is consistent with our notion that vowel sounds in the echoic store have faded considerably in 2,000 msec and that the comparison process at this longer ISI is based on either a weak echo or a phonetic code.

#### REFERENCES

- BRIGGS, R., & HOCEVAR, D. J. (1975). A new distinctive feature theory for upper case letters. *Journal of General Psychology*, **93**, 87-93.
- CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75-84.
- CROWDER, R. G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning & Verbal Behavior*, **10**, 587-596.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- CROWDER, R. G. (1982). *The psychology of reading*. New York: Oxford University Press.
- CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, **5**, 365-373.
- POSNER, M. I., & BOIES, S. J. (1971). Components of attention. *Psychological Review*, **78**, 391-408.
- POSNER, M. I., & ROGERS, M. G. K. (1978). Chronometric analysis of abstraction and recognition. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Human information processing* (Vol. 5). Hillsdale, NJ: Erlbaum.
- PROCTOR, R. W. (1981). A unified theory for matching-task phenomena. *Psychological Review*, **88**, 291-326.
- SERLING, G. (1967). Successive approximations to a model for short term memory. *Acta Psychologica*, **27**, 285-292.

(Manuscript received for publication March 13, 1986.)