

### EXPERIMENT 3. PERCEPTION OF STOP CONSONANTS AND VOWELS IN DICHOTICALLY PRESENTED CV SYLLABLES

From the results of the selective listening experiments, it seems safe to assume that the lag effect is a genuine perceptual phenomenon. However, it is still not known at what level of perceptual processing the effect originates. It has been thought that the lag effect is peculiar to the perception of encoded speech sounds and reflects the operation of special speech decoding apparatus (Studdert-Kennedy et al., 1970). On the other hand, it is conceivable that the lag effect is a more general phenomenon of auditory perception. One approach toward investigating this issue would be to see whether the lag effect can be obtained for other stimuli besides the stop consonants. Of particular interest are the unencoded speech sounds like steady-state vowels.

There is evidence from dichotic listening studies and from other sources that steady-state vowels can be processed in either the speech mode or the nonspeech mode. When vowels are presented dichotically, the results are quite different from the effects observed with competing stops. For stops, most people give highly reliable right-ear effects. For dichotically presented steady-state vowels, ear effects are typically smaller than ear effects for stops, and more people show left-ear effects for vowels than for stops. Some studies have reported statistically significant right-ear effects for vowels (Chaney and Webster, 1966; Kirstein and Shankweiler, 1969; Weiss and House, 1970), but in other studies, the effects were not significant (Shankweiler and Studdert-Kennedy, 1967; Darwin, 1969; Studdert-Kennedy and Shankweiler, 1970).

There is considerable support for the view that the inconsistent lateralization of steady-state vowels reflects the ability of vowels to be processed either in the speech mode, in the left hemisphere, or in the nonspeech mode, in the right hemisphere. In a recent paper Spellacy and Blumstein (in press) reported that dichotic syllables differing in the vowel gave a left-ear effect when embedded in a test with nonspeech filler items but that the same stimuli gave a right-ear effect when the filler items were words. No context effect was found for dichotically presented stops, which gave a right-ear effect under both conditions. An effect of test context on vowel lateralization was also reported by Darwin (1970). He compared the laterality effect for a set of vowels in a test where all the vowels were from the same vocal tract with the laterality effect obtained for the same vowel pairs embedded in a test where vocal tract size for the other pairs could vary from trial to trial. The same vowel pairs gave no ear effect in a test by themselves but gave a significant right-ear effect in the context of the larger test. Darwin (1970) attributed the shift toward a right-ear effect to the greater complexity of the perceptual task when the vocal tract size varies.

The fact that vowels usually behave in experimental tasks in a manner intermediate between speech and nonspeech sounds made the steady-state vowels seem an appropriate class of sounds to study in connection with the lag effect.

If the lag effect is a phenomenon found only with encoded speech sounds, then the lag effect should not occur or should be much reduced if the stimuli are steady-state vowels. A pilot study by Porter, Shankweiler, and Liberman (1969) gave results consistent with this prediction. They used as stimuli six isolated steady-state vowels, each 400 msec in duration, and presented these vowels for identification in a dichotic test with difference of 0 to 120 msec in the onsets of vowels at the two ears. They tested four subjects, all of whom had lag effects for stops, and found that three of the four were more accurate in identifying the leading vowels than the lagging vowels. One subject had a lag effect for vowels, but this was considerably smaller than the effect seen with stops. The authors interpreted the absence of a lag effect for vowels as evidence that the lag effect is associated with perception in the speech mode.

Up to this point we have learned that dichotically presented syllables contrasting in the initial stop consonant give lag effects, while dichotically presented isolated steady-state vowels apparently give lead effects. The main purpose of Experiment 3 was to determine whether this difference between encoded and unencoded speech sounds in the effects of interaural delay could be observed when stops and vowels were made to contrast between ears within the same syllables. The subjects in Experiment 3 listened to pairs of syllables differing between ears in both the stop and vowel, and they were asked to report either the stops or the vowels. This experiment would show whether the presence of the shared vowel is an essential condition for obtaining a lag effect for stop consonants. The experiment was also expected to show whether the same effects can be obtained for vowels in isolation and in syllabic context.

#### Method

Test tapes. Experiment 3 required the use of two different tapes. One was the tape described in Experiment 1 where syllables differed in the stop but not in the vowel. The second tape was constructed specifically for Experiment 3. For this tape the same nine syllables described in Experiment 1-- [bɛ], [dɛ], [gɛ], [ba], [da], [ga], [bɔ], [dɔ], [gɔ]--were recorded in pairs onto a two-channel tape in such a way that both the consonant and vowel differed between ears on each trial. There are thirty-six possible pairs of this sort; these are listed in Table II. Three delay intervals were used-- 0, 50, and 70 msec. Each of the thirty-six pairs of syllables appeared twice on the tape at each delay, once with channel 1 delayed and once with channel 2 delayed. The pairs and delays were arranged randomly within the 180-trial tape.

Test conditions. Each of the subjects in this experiment took three different tests.

(1) Condition "C." This condition is an exact replication of the dichotic presentation condition in Experiment 1. Subjects listened to the tape in which stops contrasted between ears and vowels were shared. The instructions to the subjects were to report both stops, guessing if necessary, and to record the clearer stop in the first column of the answer sheet. Responses were limited to "B," "D," and "G."

Table II. Syllable pairs included on the stimulus tape in Experiment 3.

<u>Syllable on</u>	
<u>Channel 1</u>	<u>Channel 2</u>

BE . . . DA
DA . . . BE
BA . . . DE
DE . . . BA
BE . . . DO
DO . . . BE
BO . . . DE
DE . . . BO
BA . . . DO
DO . . . BA
BO . . . DA
DA . . . BO

BE . . . GA
GA . . . BE
BA . . . GE
GE . . . BA
BE . . . GO
GO . . . BE
BO . . . GE
GE . . . BO
RA . . . GO
GO . . . BA
BO . . . GA
GA . . . BO

DE . . . GA
GA . . . DE
DA . . . GE
GE . . . DA
DE . . . GO
GO . . . DE
DO . . . GE
GE . . . DO
DA . . . GO
GO . . . DA
DO . . . GA
GA . . . DO

Each of these combinations occurs with simultaneous onset on the two channels, with channel 1 leading by 50 and by 70 msec and with channel 1 lagging by 50 and by 70 msec.

The letters E, A, and O represent the phonetic qualities [ɛ], [a], and [ɔ], respectively.

(2) Condition "CV-C." In this condition the subjects listened to the tape in which both stops and vowels differed between ears on any trial. The subjects were instructed to report both of the stops on each trial, recording the clearer of the stops in the first column of the answer sheets. The subjects were asked to ignore the vowels and concentrate on the consonants.

(3) Condition "CV-V." The subjects listened again to the tape in which both stops and vowels differed between ears on each trial. In this condition the subjects were instructed to attend to the vowels and to try to ignore the stops. They were to identify both vowels on each trial and to record the vowels in order of clarity on the answer sheet. The letters "E," "A," and "O" were used to indicate the vowel sounds in the words "bet," "hot," and "law" respectively.

Each condition required a one-hour testing session during which 360 trials were given. The three sessions were held on different days. The subjects were randomly assigned to one of two testing orders: (1) C, CV-C, CV-V, or (2) CV-V, CV-C, C.

Subjects. Twelve subjects took all three test conditions. The subjects were volunteers from the introductory psychology classes at the University of Connecticut and received credit toward a course requirement by participating in the experiment. All were right handed, had normal hearing, and were native speakers of English.

## Results

Accuracy of identification. Figure 22 shows the percent correct first and second responses for the three test conditions. There were no significant differences among the three conditions in first response accuracy, but significant differences among the conditions were found in accuracy of second responses (Friedman two-way analysis of variance,  $\chi^2=18.5$ , 2 d.f.  $p < .001$ ). In second responses stop consonants were more often correctly identified when vowels differed between ears than when vowels were shared ( $p < .006$ , two-tailed sign test). Within the same pair of syllables differing between ears in both the stop and vowel, vowels were more accurately identified than stops ( $p < .038$ , two-tailed sign test). Very few errors were made in vowel identification in either first or second responses.

Effect of relative onset time on overall percent correct. The experimental procedure provides two possible indicators of the lag effect: (1) more accurate identification of lagging than leading syllables and (2) judgments that lagging syllables are clearer than leading syllables. Figure 23 shows the percent correct report of lagging and leading stimuli considering both first and second responses. The lag effect is clearly present only in condition "C," where vowels are shared at the two ears. The lag effect is barely visible in the overall percent correct for conditions "CV-C" and "CV-V," where vowels contrast between ears.

Effect of relative onset time on clarity judgments. Figure 24 shows the pattern of clarity judgments for the three test conditions. In all conditions, lagging stimuli were judged as clearer than leading stimuli. In order to compare the magnitude of the lag effect in the three conditions, lag effect scores were computed for each subject for each test using the formula (Leading - Lagging)/(Leading + Lagging), where "Leading" and "Lagging" refer to the number of first responses corresponding to leading and lagging stimuli. For condition "C," where only stops varied, the lag effect scores were computed on the

Mean Percent Correct First and Second Responses as a Function of the Interval Between Stimulus Onsets for Conditions C, CV-C, and CV-V

FIRST RESPONSES SECOND RESPONSES

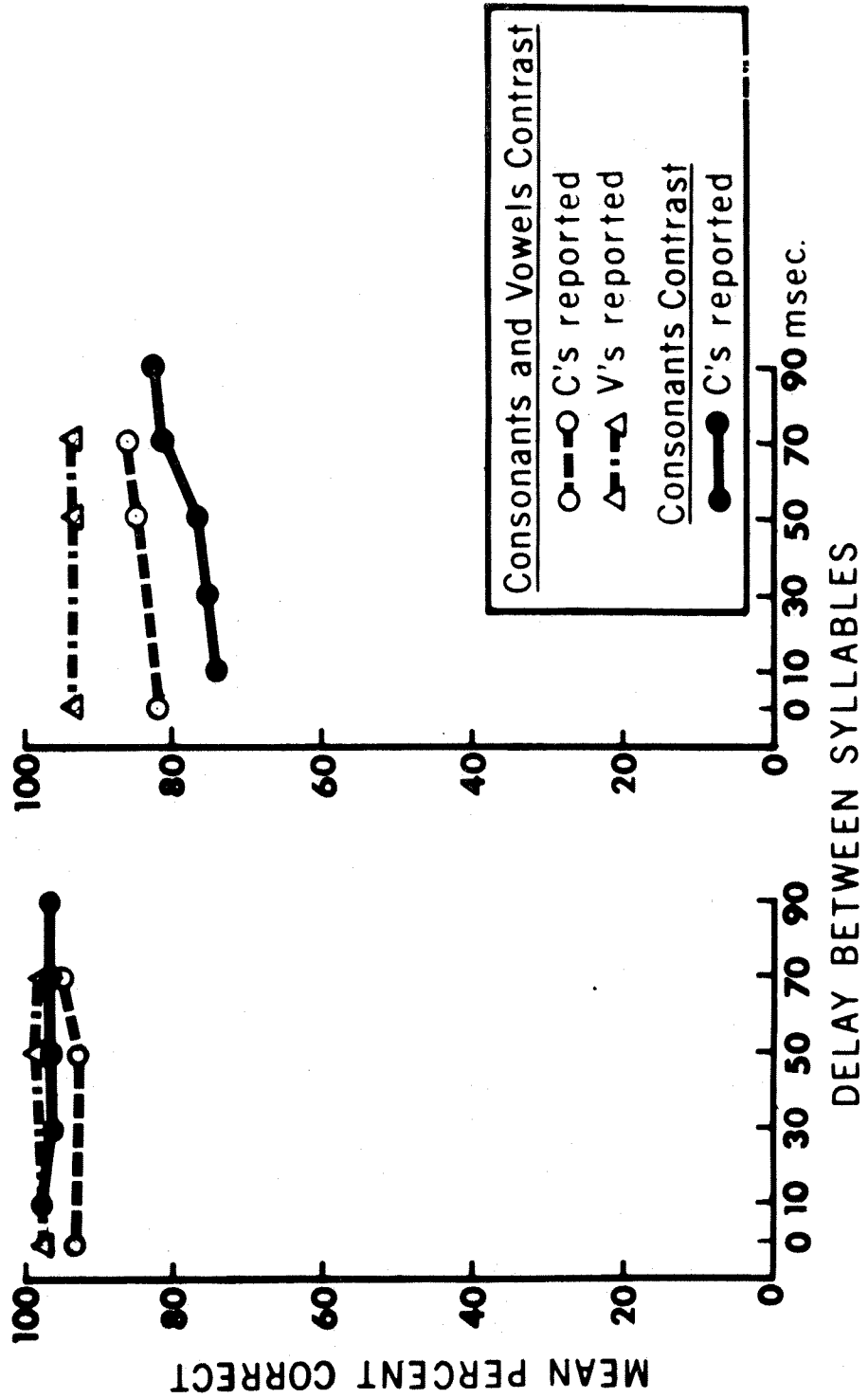


Fig. 22

Mean Percent Correct Responses for Lagging and Leading Stimuli  
for Conditions C, CV-C, and CV-V

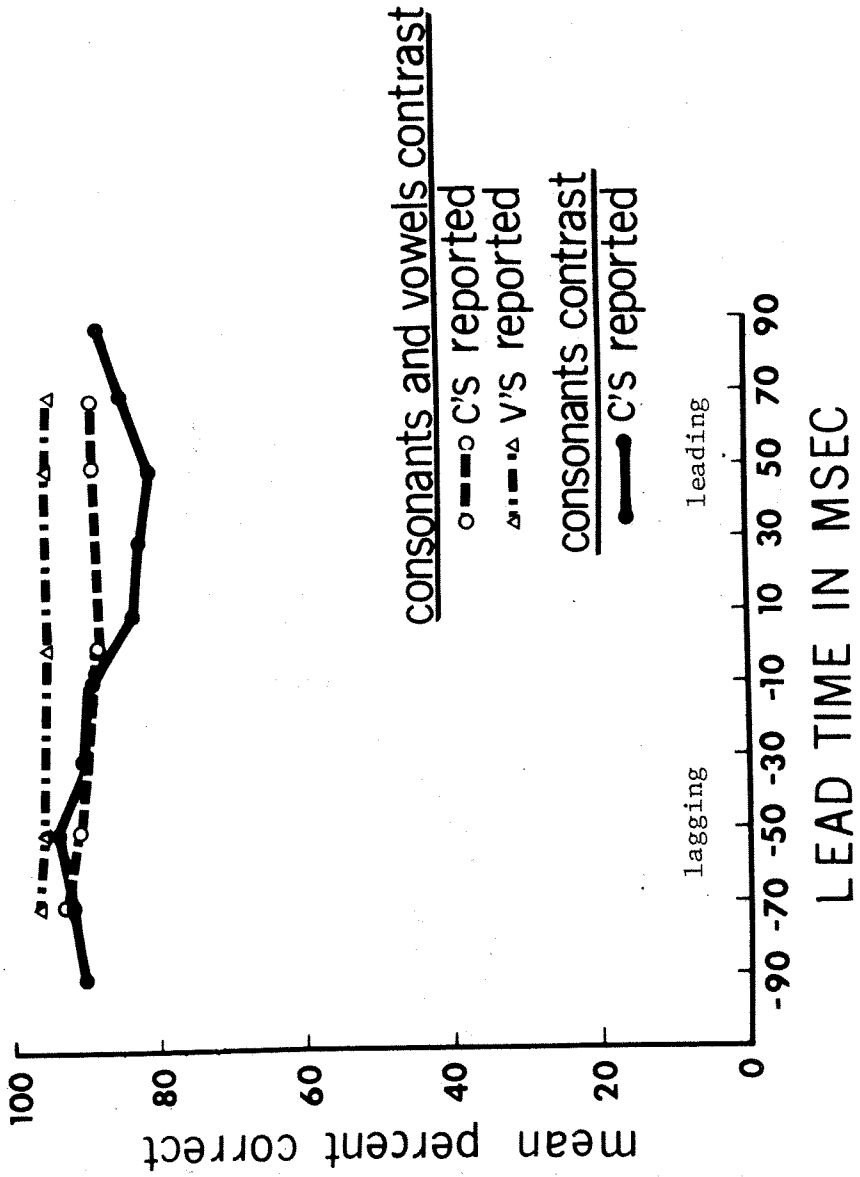


Fig. 23

Mean Percent Correct First Responses Corresponding to Lagging and Leading Stimuli for Conditions C, CV-C, and CV-V

## FIRST RESPONSES (N=12)

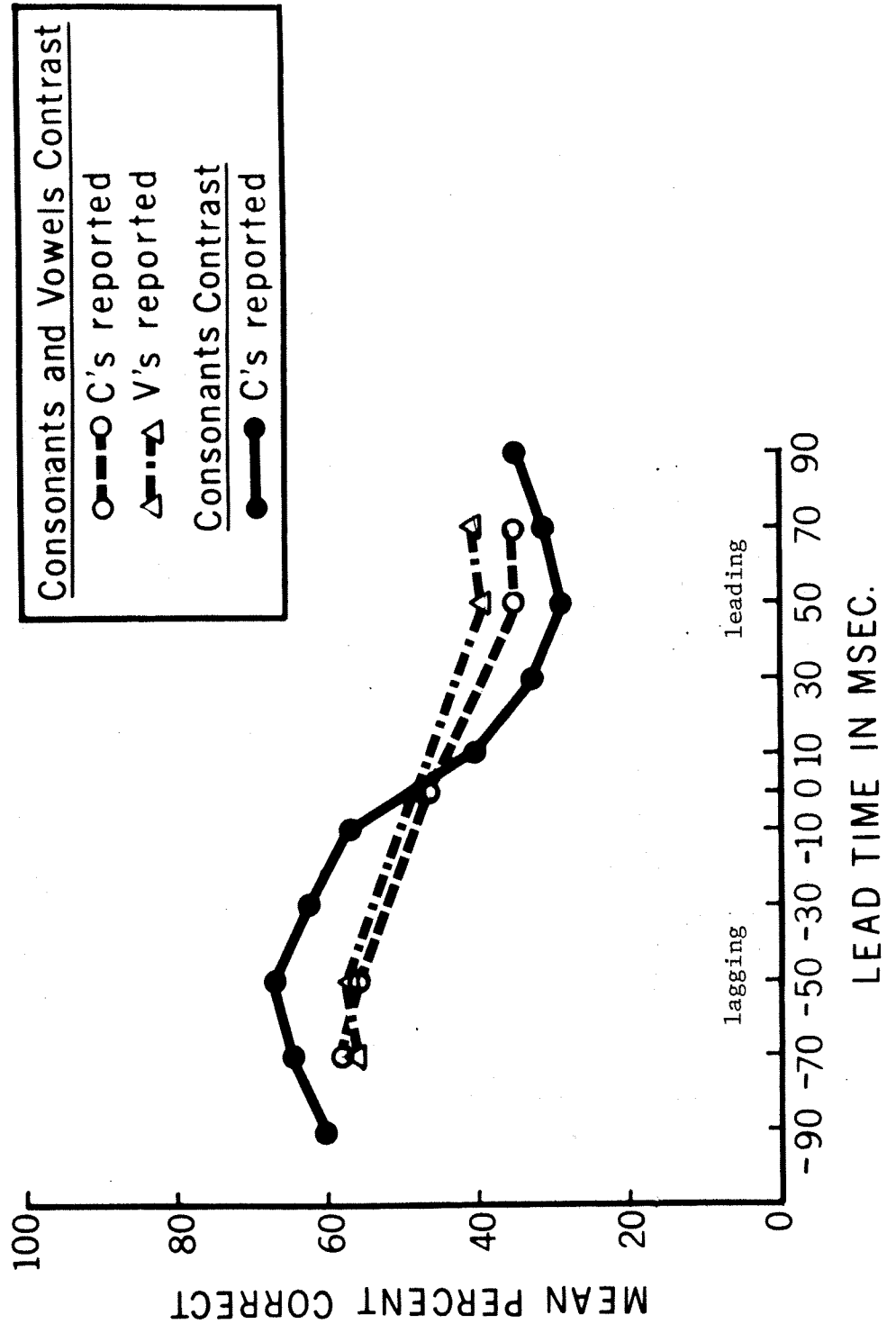


Fig. 24

basis of the 50- and 70-msec delay trials, since these were the delay intervals which the three conditions had in common. A Friedman two-way analysis of variance (Siegel, 1956) indicated that these scores differed significantly among the three conditions ( $\chi^2=9.5$ ,  $p < .01$ , 2 d.f.). The lag effect in clarity judgments for stop consonants was significantly greater when vowels were shared than when vowels contrasted between ears (C>CV-C,  $p < .006$  by a two-tailed sign test). However, there was no significant difference in the lag effect for the CV-C and CV-V conditions.

Of the twelve subjects, eleven had a lag effect in clarity judgments in condition "C" and all twelve had lag effects in conditions "CV-C" and "CV-V." Individual differences in lag effect scores were highly correlated across the three conditions. The Spearman rank correlation coefficients indicating the consistency in individual differences across conditions are shown in Table III.

Table III. Individual differences in the magnitude of the lag effect correlated across the C, CV-C, and CV-V conditions (Spearman rank correlation coefficients).

Conditions	$r_s$	
C and CV-C	.951	$p < .001$
C and CV-V	.811	$p < .01$
CV-C and CV-V	.713	$p < .01$

Laterality effects. Right-ear effects in clarity judgments were shown to ten subjects in condition "C" and eight in condition "CV-C" and "CV-V." There were no significant differences among the three conditions in the size of the ear effect. Individual differences in the size of the right-ear effect were positively correlated across the three conditions (Table IV).

Table IV. Individual differences in the magnitude of right-ear effect correlated across the C, CV-C, and CV-V conditions (Spearman rank correlation coefficients).

Conditions	$r_s$	
C and CV-C	.783	$p < .02$
C and CV-V	.537	$p < .10$
CV-C and CV-V	.678	$p < .10$

It is interesting that the two conditions involving stop consonant identification (C and CV-C) correlate more highly with each other than either of them correlates with the vowel identification test. This was true for the lag effect correlations as well as the ear effect.

## Discussion

In an earlier study, Porter et al. (1969) reported that isolated steady-state vowels did not give a lag effect. One of the purposes of Experiment 3 was to see whether such a dissociation of effects could be observed for consonant and vowel segments within the same syllable. However, when both stops and vowels were made to differ between ears, a lag effect of approximately the same extent was found whether subjects were attending to the stops or to the vowels. The lag effect observed in first responses when both stop and vowel contrasted could be described as intermediate between that obtained for stops when vowels were shared and the effect obtained for isolated vowels.

In trying to account for this result, one of the factors which seems relevant is the effect of the vowel contrast on overall accuracy of identification. Significantly more stop consonants were correctly identified when vowels differed between ears than when vowels were shared. This rise in performance may be a consequence of a reduction in fusion of stimuli from the two ears when vowels contrast. With a vowel contrast, the dichotic syllables differ acoustically for a full 350 msec, not just for the duration of the stop consonant transitions. It may, therefore, be easier to localize the syllables by ear when vowels contrast than when vowels are shared. One might speculate that this superior spatial separation of stimuli facilitates the retention of an auditory image of one syllable while the other is being decoded and thus causes an increase in the number of stimuli correctly identified.

It seems reasonable to suppose that the factors underlying the improved performance also underlie the reduction in the lag effect for stops when vowels contrast. Although in principle the extent of lag effect in clarity judgments is not constrained by the performance level, it seems likely that clarity judgments become more difficult and hence less reliable when both of the stops can be correctly identified.

In the case of the vowel identification task, the main result to be explained is why there was a lag effect rather than a lead effect. There are several reasons why vowels in syllables and vowels in isolation might give different results. When vowels are embedded in CV syllables where the stop consonants differ between ears, the listeners might find it difficult to separate their judgments about the relative clarity of the two vowels from their perception of the stops. Another possibility is that the presence of stop consonants may induce a "set" to process stimuli in the speech mode. The lag effect for vowels could reflect the fact that they are being processed in the speech perception mode, whereas the lead effect for isolated vowels might be associated with perception in the nonspeech mode. If lag effects can be obtained for vowels when they are being perceived in the speech mode, then it should be possible to obtain a lag effect for isolated vowels as well as vowels in syllables under conditions which induce a "set" toward the speech mode. Certain findings in Experiment 4 uphold this prediction.