EXPERIMENT 1. A COMPARISON OF THE EFFECTS OF MONOTIC AND DICHOTIC PRESENTATION ON THE PERCEPTION OF TEMPORALLY OVERLAPPED STOP CONSONANT-VOWEL SYLLABLES

Experiment 1 was intended to confirm the reports of Studdert-Kennedy et al. (1970) and of Lowe et al. (1970) of differences between the effects of monotic and dichotic presentation on the perception of competing temporally staggered stop consonants. In those experiments the stimuli used were the stop consonant-vowel syllables [ba], [da], [ga], [pa], [ta], and [ka] presented in pairs with delays of 5 to 120 msec between the onsets of syllables within a pair. Both studies showed that when the two syllables arrived at opposite ears, the lagging stop was more accurately reported than the leading stop, but when the syllables arrived at the same ear, the leading stop was more accurately reported than the lagging one.

Experiment 1 repeated the comparison between the dichotic and monotic conditions using a somewhat altered stimulus set. The stimulus set was expanded to nine syllables—[bɛ], [dɛ], [gɛ], [ba], [da], [ga], [bʊ], [dʊ], [gʊ]—formed by combining each of the three voiced stop consonants [b], [d], and [g] with each of three vowels, [ɛ], [a], and [ʊ]. Although the set of syllables was enlarged over that used in previous experiments, the response set was reduced to three rather than six possibilities because competing syllables always shared the same vowel. The subjects needed only to identify the stops and could ignore the vowels.

Experiments 2, 3, and 4 made use of the same stimulus set as Experiment 1. Each of these later experiments evaluated the influence of some deviation from the task described in Experiment 1. The results of Experiment 1 thus form a base line against which the later findings can be compared.

Method

Stimuli. Preparation of the stimulus syllables and of the test tape was accomplished with the aid of the computer facilities at Haskins Laboratories (Cooper et al., 1971). Nine syllables—[bɛ], [dɛ], [gɛ], [ba], [da], [ga], [bʊ], [dʊ], [gʊ]—were generated on the Haskins Laboratories computer-controlled parallel resonance speech synthesizer. Spectrograms of these syllables are shown in Figure 1. Each syllable is a 350-msec three-formant pattern. The rapidity of the changes in formant frequency (transitions) in the initial portion of each syllable is the acoustic cue which distinguishes stop consonants as a class from other classes of speech sounds. These transitions occupy only the first 45 to 70 msec of any syllable. The direction and extent of the second and third formant transitions is the primary acoustic cue for the place of articulation of the stops, while the rising first formant transition, which is constant among the nine syllables, is characteristic of voiced stops as opposed to voiceless stops or nasals. The formant transitions are followed by steady-state patterns with formant frequencies and amplitudes appropriate for each vowel. The steady-state formant frequencies for a particular vowel are independent of the place of articulation of the accompanying stop,
Spectrograms of the Nine Stimulus Syllables

Note: The spectrographic representation of an acoustic signal shows intensity by the darkness of the display, frequency (in kHz) by position on the y-axis, and time by position in the x-axis.
but the direction and extent of the formant transitions for a particular stop vary depending on which vowel follows.

The intelligibility of these syllables was assessed by asking four people with no previous exposure to synthetic speech to identify the syllables in a 180-trial randomized sequence. In the 720 trials only one error was made in identifying the vowels of the syllables. The stop consonants in the syllables [ba], [da], [ga], [bε], [gε], [dθ], and [gθ] were identified correctly 100 percent of the time. On one occasion [dθ] was heard as [gε]. The syllable [bθ] was often heard as [gθ]; this error was idiosyncratic in that some subjects consistently made the error while others always identified the sound as it was intended. The overall percent correct identification of 95 percent was considered adequate for the dichotic experiments.

Preparation of the test tape. A two-channel tape was assembled under computer control using the PCM system (Cooper and Mattingly, 1969). The PCM program allows one to specify a random order of pairs of syllables and affords precise control over the interval between the onsets of syllables on the two channels. Using this program, the syllables were first stored on a disc file and then recalled in pairs for recording onto a two-channel tape.

The tape contained eighteen different pairs of syllables— all possible pairings of the nine syllables in which stop consonants differed between channels while vowels were shared (for example, pairs like [ba]-[da] and [gθ]-[dθ]). The eighteen pairs of syllables are listed in Table I. For any pair the

<table>
<thead>
<tr>
<th>Syllable on Channel 1</th>
<th>Channel 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE       . .  DE</td>
<td></td>
</tr>
<tr>
<td>DE       . .  BE</td>
<td></td>
</tr>
<tr>
<td>BA       . .  DA</td>
<td></td>
</tr>
<tr>
<td>DA       . .  BA</td>
<td></td>
</tr>
<tr>
<td>BO       . .  DO</td>
<td></td>
</tr>
<tr>
<td>DO       . .  BO</td>
<td></td>
</tr>
<tr>
<td>BE       . .  GE</td>
<td>Each of these combinations occurs once with a channel 1 lead of 10, 30, 50, 70, and 90 msec and with a channel 2 lead of 10, 30, 50, 70, and 90 msec.</td>
</tr>
<tr>
<td>GE       . .  BE</td>
<td></td>
</tr>
<tr>
<td>BA       . .  GA</td>
<td></td>
</tr>
<tr>
<td>GA       . .  BA</td>
<td></td>
</tr>
<tr>
<td>BO       . .  GO</td>
<td></td>
</tr>
<tr>
<td>GO       . .  BO</td>
<td></td>
</tr>
<tr>
<td>DE       . .  GE</td>
<td></td>
</tr>
<tr>
<td>GE       . .  DE</td>
<td></td>
</tr>
<tr>
<td>DA       . .  GA</td>
<td></td>
</tr>
<tr>
<td>GA       . .  DA</td>
<td></td>
</tr>
<tr>
<td>DO       . .  GO</td>
<td></td>
</tr>
<tr>
<td>GO       . .  DO</td>
<td></td>
</tr>
</tbody>
</table>

The letters E, A, and O represent the phonetic qualities [ε], [a], and [θ], respectively.
syllable onset at one channel was delayed relative to the other channel by 10, 30, 50, 70, or 90 msec. Each of the eighteen pairs occurred twice at each delay interval, once with channel 1 being the delayed channel, and once with channel 2 delayed. The tape contained 180 stimuli--18 pairs x 5 delays x 2 channels--randomly ordered in a 180-trial tape. Each trial consisted of the presentation of a single pair of syllables. The intertrial interval was 6 seconds for most trials, with 10-second pauses inserted between blocks of ten trials and a 30-second pause after 90 trials.

Testing procedure. The same tape was used for both the dichotic and monotic presentation conditions. For the dichotic test each tape channel was delivered to a different ear. For the monotic test the two channels were mixed electronically and presented to the same ear. In both the dichotic and monotic tests each subject listened to the tape twice within a test session, with the headphones physically reversed on the second run. This gave a total of 360 trials for each subject, 72 trials at each of the five delay intervals. The headphone reversal is an important control on dichotic tests where it is necessary to balance stimulus conditions over the two ears. For the first 180 trials on the dichotic test the subjects always received channel 1 at the left ear and channel 2 at the right ear. For the second 180 trials, the left ear received channel 2 and the right channel 1. For the monotic test the mixture of the two channels was played to the left ear for the first 180 trials and to the right ear for the second 180 trials.

The subjects were tested in groups of one to six people in a quiet room. The tape was played on a General Radio tape deck (Type 1525) into a "listening station" designed and built by D. Zeichner of Haskins Laboratories for group dichotic experiments. The listening station is a two-channel amplifier with multiple outputs for headphones. The output level of the signal from each channel can be adjusted in 1 db calibrated steps. The listening station can be used to present stimuli monaurally to either ear, monotonically (both channels mixed electronically and presented to the same ear), and dichotically (a different channel to each ear).

The stimuli were presented at a comfortable listening level (about 75 db at each ear) over headphones (Grason-Stadler TDH-39-300Z). The intensity of the output from the two channels was equated as closely as possible by the use of calibration signals laid down with the same intensity on the two channels at the time of recording of test tape.

Instructions to subjects. As far as possible identical instructions were given for the monotic and dichotic tests. The subjects were told that they would receive two different syllables on each trial, that the syllables would always differ only in the consonant, and that only the consonants "B," "D," and "G" would be used. The task was to identify both consonants on each trial and then to decide which of the two sounded clearer. The clearer of the two sounds was to be recorded in the first column of the answer sheet (first responses) and the less clear was to be recorded in the second column (second responses). Both columns were to be filled in with different responses even if it was necessary to guess. The subjects were told that some trials would be more difficult than others, but they were not told that relative onset time was a variable in the test.
Subjects. All subjects were students in the introductory psychology classes at the University of Connecticut. Their participation in these experiments fulfilled part of their course requirement. All subjects were native speakers of English, were right handed, and had no known hearing defect.

Separate groups of subjects were run on the dichotic and monotic tests. Each group contained twelve subjects.

Results

The listeners had two tasks to perform on each trial—to identify both consonants and to indicate which was the clearer by recording the clearer consonant in the first column (first response) and the less clear consonant in the second column (second response). A response is considered correct if it corresponds to either of the stimulus consonants presented on that trial.

Accuracy of identification. Many more errors in identification were made on second responses than on first responses. On the dichotic test 94.1 percent of all first responses were correct while only 66.6 percent of all second responses were correct. The monotic test did not differ significantly from the dichotic test with respect to accuracy of performance. For the monotic test 90.4 percent of all first responses and 65.6 percent of all second responses were correct.

Figure 2 shows the accuracy of identification of stops in syllable pairs with temporal offsets of 10, 30, 50, 70, or 90 msec. First responses were nearly always correct and the accuracy of first responses was not systematically affected by the amount of time separating the onsets of the competing syllables. The ability of subjects to report both stops correctly was, however, influenced by the delay between syllables. Second responses became more accurate with longer delays for both the dichotic and monotic tests.

Effects of lag or lead on the accuracy of identification. Figure 3 shows the effect of lag or lead time on the accuracy of identification. On the monotic test leading stops were more often correctly reported than lagging stops, but with dichotic presentation lagging stops had an advantage over leading stops. The percentage of correct responses expected by chance in this task is 67 percent. When monotonically presented syllables were separated by 10 msec, the lagging stop was reported no more frequently than would be expected by chance, and identification of lagging stops did not rise above chance level until the syllables were separated by 70 msec. Leading stops, in contrast to lagging stops, were accurately reported at all delays on the monotic test. On the dichotic test the suppression of the leading syllables was not so great as the suppression of lagging syllables on the monotic test. Identification of leading stops never fell to chance level, but there was an advantage for lagging syllables at all delay intervals.

Analysis of first responses. The subjects were instructed to indicate which of the stops sounded clearer by recording the clearer stop in the first column of the answer sheet. Figure 4 shows the percentage of trials on the two tests in which lagging and leading stimuli were given as the first response.
Mean Percent Correct Responses as a Function of the Interval Between Syllable Onsets for the Monotic and Dichotic Conditions

**First Responses**

- **DELAY BETWEEN SYLLABLE ONSETS**
- **% CORRECT**

**Second Responses**

- **DELAY BETWEEN SYLLABLE ONSETS**
- **% CORRECT**

- **MONOTIC**
- **DICHTIC**
Mean Percent Correct Responses as a Function of Stimulus Lag or Lead Time for the Monotic and Dichotic Conditions

% CORRECT

LEAD TIME IN MSEC

lagging

leading

MONOTIC

DICHOTIC
Mean Percent Correct First Responses as a Function of Stimulus Lag or Lead Time for the Monotic and Dichotic Conditions
The monotic lead effect and dichotic lag effect seen in Figure 3 are enhanced in the clarity judgments. It will be seen in Experiments 3 and 4 that the pattern of clarity judgments is a much more sensitive measure of the lag effect than the overall percent correct. For this reason much of the subsequent analysis is based entirely on first responses.

**Analysis of second responses.** The subjects were instructed to record two different responses on each trial even if they had to guess. According to the instructions, all guesses should have been recorded in the second column of the answer sheet. Thus, of the responses in the second column, many were undoubtedly simply guesses, although others represent stimuli which were perceived correctly but judged as less clear than the first responses. In order to separate the guesses from the correct second responses, it is necessary to consider the pattern of first responses. The frequency with which a lagging or leading stimulus would be correctly guessed as a second response depends on the frequency of lagging and leading stimuli given as first responses. For example, on the dichotic test, lagging stimuli were given as first responses 61.1 percent of the time, so that lagging stimuli could be given as a second response on not more than 38.9 percent of all second responses. Leading stimuli were given as the first response on only 33.3 percent of the trials and thus could potentially be given as the second response on 66.7 percent of the trials. If all second responses were guesses, second responses would contain 33.3 percent leading stimuli, 19.5 percent lagging stimuli, and 47.5 percent errors. Thus, simply by chance, the lag advantage would appear to be reversed in second responses.

An analysis of second responses was performed which takes into account the fact that there are an unequal number of trials in which lagging and leading stimuli could appear as the second responses, given the first response pattern. Accuracy of identification for second responses was calculated as number of possible correct second responses. A rise above 50 percent shows that stimuli which were judged as less clear were reported with better than chance accuracy.

Figure 5 shows the percent correct second responses for the dichotic and monotic tests after the scores have been corrected in the manner just described. The first response patterns—the monotic lead effect and the dichotic lag effect—also appear in the second responses. On the monotic test, stops which were lagging by 10 or 30 msec were reported with no more than chance accuracy as second responses, but stops leading by all intervals were given as second responses much more often than would be expected if subjects were guessing. On the dichotic test it would appear that subjects more often guessed about the second response when the leading stop would have been correct than when the lagging stop would have been correct.

**Individual differences in performance on the monotic and dichotic tests.** Up to this point the comparison of the dichotic and monotic tests has considered the averaged performance of all subjects. It is also of interest to look at the distribution of lag effects and lead effects among individuals within each of the test groups. A lag effect or lead effect score was computed for each subject using as the measure of these effects the expression: (Leading-Lagging)/(Leading + Lagging) where "Lagging" and "Leading" refer to the number of first responses corresponding to lagging or leading stops summed over all delay intervals. The distributions of these scores for the dichotic
Mean Percent Correct Second Responses for the Dichotic and Monotic Tests

Note: The percent correct at each onset time has been corrected for the pattern of first responses by using the formula percent (correct/possible correct) second responses. With this correction the probability of giving a correct second response by chance is .50 at each onset time condition.
and monotic tests are shown in Figure 6. These distributions give further evidence of the reliability of the monotic lead effect and the dichotic lag effect. All twelve of the subjects on the monotic test had lead effects. On the dichotic test, eleven of the twelve subjects had lag effects; one subject had a lead effect on the dichotic test, but this was smaller than any of the monotic scores. There was also a difference in variability between the two conditions, with greater variability among subjects in the dichotic than in the monotic condition.

Distribution of dichotic laterality effects. In previous dichotic listening studies, report of stop consonants at the right ear was found to be more accurate than report of stops at the left ear. Evidence of ear asymmetry was, therefore, sought in the present experiment. The number of subjects showing a right-ear advantage was established by computing a laterality effect score for each subject. The size of the ear effect was calculated from the formula \((\text{Right} - \text{Left})/(\text{Right} + \text{Left})\) where "Right" and "Left" refer to the number of first responses corresponding to stimuli delivered to the right and left ears.\(^4\) The index ranges from \(0.0\), for no difference between ears, to \(\pm 1.00\), which would indicate that all first responses are from the right ear (positive scores) or left ear (negative scores). Figure 7 gives the frequency distribution of these laterality effect scores. Of the twelve subjects, ten had a right-ear advantage and two a left-ear advantage. For the subjects with left-ear effects, the extent of ear asymmetry was less than that seen for the typical right-ear effect.

Relationship between the ear effect and lag effect. The fact that both a right-ear effect and lag effect are reliably observed with dichotic presentation of temporally staggered stop consonants leads one to inquire into the nature of the relationship between the ear effect and lag effect. One possibility is that the two phenomena are essentially independent. On the other hand, interaction between the two effects might be observed.

One way of approaching this question was to determine whether the subjects who had large lag effects were also the subjects who had large ear effects. A Spearman rank correlation coefficient (Siegel, 1956) was computed to determine the extent of association between the lag effect and right-ear effect scores. The coefficient was \(-.26\), a nonsignificant correlation. The negative sign probably reflects a ceiling effect imposed by the task; extremely large lag effects and ear effects are mutually exclusive. Thus, the subjects with the largest ear effects had small lag effects and vice versa. However, subjects with moderate ear effects and lag effects did not show any correlation in the size of the two effects.

Effects of relative onset time on the perception of syllables at the right and left ears. Another way of looking for an interaction between the ear effect and lag effect is to see whether the consequences of interaural delay are the same for the two ears. Figure 8 shows the percentage of first responses from the right and left ear for each delay condition for all twelve

\(^4\)This measure of the right-ear effect was introduced by Studdert-Kennedy and Shankweiler (1970).
Frequency Distributions of Lag Effect Scores, \((\text{Leading-Lagging})/(\text{Leading+Lagging})\), for the Dichotic and Monotic Conditions

Note: Positive scores indicate lead effects and negative scores indicate lag effects. Scores are based on first responses.
Frequency Distribution of Ear Effect Scores, (R-L)/(R+L), Based on First Responses in the Dichotic Condition

Dichotic

\[
\frac{(r - l)}{(r + l)}
\]

Note: A positive score indicates a right-ear effect and a negative score indicates a left-ear effect.
Comparison of Left-Ear Lag and Right-Ear Lag Trials in the Dichotic Condition

Note: The figure shows the mean percentage of first responses corresponding to stimuli at the left and right ears at each delay interval.
subjects on the dichotic test. It can be observed that stimuli at the right ear were given as first responses more often than stimuli at the left ear but that a left-ear lag can counteract the right-ear effect. Thus, with a 10-msec separation between ears, stops at the right ear were judged to be clearer than those at the left, but when the left ear was delayed 30 msec behind the right, stimuli at the left ear were judged as clearer.

Figures 9 and 10 show the right and left ear curves for two of the subjects. One of the subjects (SV) performed in a manner very similar to the group performance, but the other subject (BR) was remarkable in the magnitude of the right-ear effect. BR judged the right ear to be clearer than the left even when the left ear was lagging behind the right. Nevertheless, BR showed some improvement for the left ear when it received the lagging syllable. Thus, both subjects showed a right-ear advantage and an advantage for the lagging ear, regardless of which was the delayed ear.

Figure 11 plots the data in Figure 8 in a different format which compares the frequency of report of the two ears under identical delay conditions. For any delay between ears the right ear is superior to the left, but the shape of the function relating frequency of first response to relative onset time is similar for the two ears. There is one noticeable difference between the ears—the consistent 20-msec displacement between the left- and right-ear curves. The maximum right-ear advantage occurs with a right-ear delay of 50 msec, while the maximum left-ear advantage occurs with a left-ear delay of 70 msec. Figure 12 compares the curves for the two ears after all left-ear points have been advanced 20 msec along the x-axis. Following this displacement it is observed that the curves for the two ears have the identical form.

**Discussion**

The results of Experiment 1 confirm the earlier finding that when two stop consonant-vowel syllables sharing the same vowel are overlapped in time, a lag effect is obtained with dichotic presentation and a lead effect is obtained with monotic presentation. These contrasting effects of monotic and dichotic competition were observed in the overall frequency of correct responses, in judgments of the relative clarity of the two stops, and in the accuracy of identification of the "less clear" stimulus.

Other differences between the dichotic and monotic conditions were discovered in Experiment 1. With dichotic presentation the maximum advantage for the lagging syllable occurred with delays of 50 to 70 msec between syllable onsets, whereas the monotic lead effect was maximal with shorter delays (10 to 30 msec). At these short delay intervals the lagging stop was completely obscured on the monotic test, as indicated by the fact that the lagging stimuli were reported correctly no more often than would be expected by chance. In contrast, the leading syllables on the dichotic test were reported with greater than chance accuracy at all delays. Finally, it was found that the performance of individual subjects with respect to the magnitude of the lag or lead effect was more variable with dichotic presentation than with monotic presentation.

It is not clear why there should be such a great difference between the effects of dichotic and monotic competition. Studdert-Kennedy et al. (1970)
Percent Correct First Responses by Ear on the Dichotic Test for Subject SV
Percent Correct First Responses by Ear on the Dichotic Test for Subject BR
Mean Percent Correct First Responses by Ear Comparing the Same Delay Conditions for the Two Ears
Comparison of the Form of the Lag Effect Function for the Two Ears After a Displacement of the Left-Ear Curve 20 msec Along the X-Axis

% CORRECT

-90 -70 -50 -30 -10 10 30 50 70 90 LEAD TIME IN MSEC

---

Fig 12
have attributed the monotic lead effect to peripheral masking and the dichotic lag effect to central processes. It is evident that the lag effect must arise centrally because with dichotic presentation the stimuli do not interact at the peripheral receptor. The puzzling question is why there is no evidence of a central lag effect with monotic competition as well. Perhaps the lag effect occurs only when the stimuli arrive over separate input channels, or the explanation may be that the peripheral suppression of the lagging stops is so extensive with monotic presentation as to preclude subsequent central enhancement of the lagging stimuli.

The finding of a right-ear advantage with dichotic presentation was also consistent with the results of previous research. The stops at the right ear were judged as clearer than those at the left ear for all delays when the same delay conditions were compared for the two ears. In comparing the effects of delay time for the two ears it was found that the shape of the function relating delay time to the percent correct first responses was the same for the left and right ears. From this result one might infer that the lag effect and right-ear effect are separate phenomena which combine in an additive manner to determine the relative intelligibility of the competing stop consonants. Further evidence of the independence of the two effects was the finding that individual differences in the lag effect and right-ear effect were not correlated. The only data suggesting an interaction between the lag effect and right-ear effect were the results dealing with the location of the maximum lag effect for the two ears. The maximum lag effect for the left ear was at a left-ear lag of 70 msec, while the maximum lag advantage for the right ear was with a right-ear lag of 50 msec. Certain findings in Experiment 2 lead to the conclusion that this temporal displacement is an artifact of the particular delay intervals tested and that the true peak is in the same place for the two ears.