Two left-hemisphere mechanisms in speech perception*

JAMES E. CUTTING†

Yale University and Haskins Laboratories, New Haven Connecticut 06510

Right-ear advantages of different magnitudes occur systematically in dichotic listening for different phoneme classes and for certain phonemes according to their syllabic position. Such differences cannot be accounted for in terms of a single mechanism unique to the left hemisphere. Instead, at least two mechanisms are needed. One such device appears to be involved in the auditory analysis of transitions and other aspects of the speech signal. This device appears to be engaged for speech and nonspeech sounds alike. The other mechanism, the more accustomed "speech processor," appears to make all phonetic decisions in identifying the stimulus.

Early studies in dichotic listening (Broadbent, 1956; Kimura, 1961) presented different series of digits simultaneously to each ear. Results showed that this task overloaded the perceptual system, and that numerous errors occurred. The errors, however, were differentially distributed; more occurred in recalling digits presented to the left ear than for those presented to the right. The superior performance of the right ear over the left is known as the right-ear advantage and has been explained, in part, as a reflection of linguistic capabilities of the cerebral hemispheres. In the dichotic situation, it appears that linguistic signals can best travel the path from the right ear to the left hemisphere (Kimura, 1961; Studdert-Kennedy & Shankweiler, 1970). Clinical work in the latter half of the 19th century showed that the left hemisphere of the brain was primarily responsible for language functions (Boring, 1950, p. 70). Nevertheless, it has not known which specific aspects of dichotic stimuli contributed to the right-ear advantage. Paired digits differ in duration, acoustic structure, syllabic form, and many other aspects; any one of these differences might have been responsible for laterality.

Shankweiler and Studdert-Kennedy (1967) showed that the right-ear advantage was closely related to the perception of certain parts of the sound pattern of speech, but not to others. The identification of stop consonants in dichotic consonant-vowel (CV) nonsense syllables showed a large right-ear advantage. The identification of steady-state vowel syllables, on the other hand, showed a much smaller

and statistically unreliable right-ear advantage (see also Studdert-Kennedy & Shankweiler, 1970). Other classes of phonemes appeared to yield results which were intermediate between stop consonants and vowels. Liquids and semivowels (Haggard, 1971) and fricatives (Darwin, 1971) have been shown to yield right-ear advantages, but the magnitude of the advantages appears to be smaller than often found for stops. Experiment I was designed to determine if such differences could be reliably observed in stops, liquids, and vowels.

EXPERIMENT I:
EAR ADVANTAGES FOR STOPS, LIQUIDS, AND VOWELS WITHIN THE SAME SYLLABLE[1]

Method

Stimuli. Eight consonant-consonant-vowel (CCV) syllables were prepared on the Haskins Laboratories parallel-resonance synthesizer. There were three phoneme classes within each syllable: stops, liquids, and vowels. Each phoneme class was represented by two phonemes: /g/ and /k/ were the stops, /l/ and /r/ the liquids, and /ɛ/ and /ae/ the vowels. All possible combinations were used: /glɛ, klɛ, grɛ, krɛ, glae, klae, grae, krae/. The stimuli were 455 msec in duration and had the same falling pitch contour. The duration of the formant transitions in the stop + liquid clusters was 210 msec followed by 245 msec of the steady-state vowel. The stimuli /glɛ/ and /krae/ are shown in Fig. 1, and the approximate acoustic domain for each phoneme class is noted.

Subjects and Apparatus. Sixteen Yale University undergraduates participated in two tasks. The subjects were all right-handed native American English speakers with no history of hearing trouble. None had previous experience in listening to synthetic speech or to dichotic stimuli. They were tested in groups of four, with stimuli played on an Ampex AG500 dual-track tape recorder and sent through a listening station to matched Telephonics earphones (Model TDH39). The same apparatus was used for all studies in this paper. All stimuli were presented at approximately 80-dB sound pressure level.

Task 1: Diotic Identification

A brief identification test was conducted to assess the quality of the stimuli.

Procedure

The subjects listened to two tokens of each stimulus to familiarize themselves with synthetic speech. They then listened to a diotic identification test of 64 items. Each of the eight stimuli was presented singly, eight times in random sequence with a 3-sec interstimulus interval. The subjects identified each stimulus, writing their responses using the following orthography: GLEH, KLEH, GREH, KREH, GLAA, KLAA, GRAA, KRAA.

Results

The stimuli were highly identifiable. Subjects correctly identified the stimuli on more than 97% of the trials.

Task 2: Dichotic Ear Monitoring

Procedure

The same eight stimuli were used; however, this time, instead of presenting one stimulus at a time, two different stimuli were presented simultaneously, one to each ear. Dichotic test tapes were recorded using the pulse code modulation (PCM) system (Cooper & Mattingly, 1969) at Haskins. Each stimulus was paired with all other stimuli, but not with itself. There were 112 dichotic items per test: (28 possible pairs) by (2 channel arrangements) by (2 observations per pair). Two such tests were prepared with different random orders. Both had a 4-sec interval between pairs. The subjects listened to two repetitions of each test for a total of 448 trials. They attended to one ear at a time and wrote down which of the eight stimuli they heard presented to that ear. The order of ear monitoring was done such that half the subjects attended first to the left ear for a quarter of the trials, then to the right ear for half the trials, and finally back to the left ear for that last quarter (LRRL). The other half attended in the opposite order (RLLR). There was a brief rest between blocks of 112 trials. The order of the channel-to-ear assignments and the order of listening to the tapes were also counterbalanced across subjects. Six practice trials familiarized them with the task.

Results

There are two levels at which to analyze the data, the syllable level and the phoneme level.

Syllable level. Although the subjects were familiar with the eight stimuli, many errors occurred in reporting the correct syllable in the monitored ear. A syllable was scored correct when all three phonemes were correctly identified. Overall performance was 58% correct. The subjects performed significantly better when they monitored the right ear than when they monitored the left [F(1,15) = 20.96, p < .001]; they were 62% correct in reporting the syllable when they attended to the right ear and only 53% correct when they attended to the left, a net 9% ear difference.

Phoneme level. Since there is a stop, a liquid, and a vowel in each stimulus, the syllable can be parsed to observe the overall performance and ear advantage for each phoneme class. Considering each phoneme as a stimulus, there are two types of trials, *contrast* trials and *identity* trials. For example, there are those trials in which the two stimuli shared the same stop, such as GREH/GLAA and KRAA/KLAA, and those which had different stops, GREH/KLAA and KRAA/GLAA. The first type of trial is a
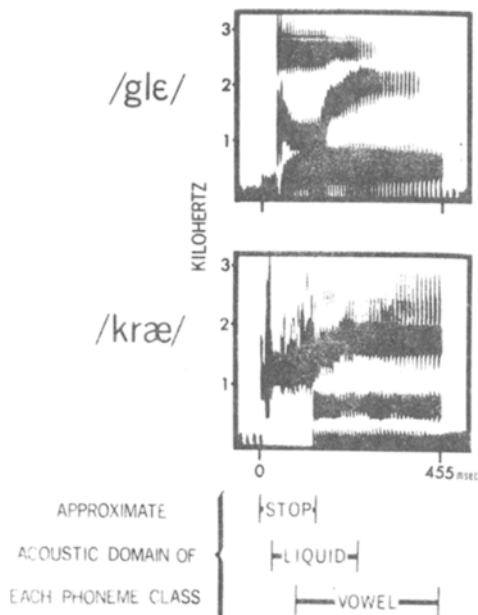
CCV Stimuli

Fig. 1. Sound spectrograms of sample consonant-consonant-vowel stimuli.

stop-identity trial, while the second is a stop-contrast trial. (Note that in considering a given phoneme class the other phoneme classes are temporarily disregarded.) Likewise, there are two types of liquid trials: for example, KLEH/KLAA as a liquid-identity trial and GLEH/KRAA, a liquid-contrast trial. Vowels are treated in the same manner; there are vowel-identity trials (for example, KRAA/GLAA) and vowel-contrast trials (KRAA/GLEH). It was on the contrast trials that 92% of the errors occurred, and it is those which are discussed first.

Contrast trials. First consider the stops. Overall, the subjects were 66% correct in reporting the stop in the monitored ear. There was a large, significant right-ear advantage [F(1,15) = 22.55, p < .001]: the subjects were 72% correct in reporting the stops while monitoring the right ear and only 60% correct while monitoring the left, a net 12% difference. Eight of the 16 Ss had statistically reliable right-ear advantages and none had a reliable left-ear advantage. These results were calculated using a phi index discussed below.

The liquids yielded a pattern similar to the stops. The subjects correctly identified the liquid in the monitored ear on 64% of the trials. Again there was a significant right-ear advantage [F(1,15) = 13.33, p < .005], but somewhat smaller than that for the stops: the subjects were 68% correct in reporting the liquid when monitoring the right ear and only 59%

correct when monitoring the left, a net 9% difference. Only four subjects had statistically reliable right-ear advantages; but, again, none had a significant left-ear advantage.

Vowels yielded a very different pattern of results. Overall performance was considerably higher: the subjects were 81% correct in identifying the vowel in the monitored ear. Furthermore, there was no ear effect for the group data. The group average, however, is somewhat misleading, since seven subjects did show statistically reliable ear effects: three had a right-ear advantage, and four had a left-ear advantage.

To pursue the idea of differential ear advantages for stops, liquids, and vowels, compensation must be made for the different performance levels for each phoneme class. A phi coefficient analysis takes this consideration into account (Kuhn, 1973). The analysis is performed on a 2 by 2 contingency table. The cell entries are the number of trials for the (a) right ear correct, (b) left ear correct, (c) right ear incorrect, and (d) left ear incorrect. Phi coefficients were computed for all subjects for all three phoneme classes. Since this calculation is a monotonic transformation of the original data, the phi indices were used in further analyses.

Figure 2 shows the ear advantages and ranges for the stops, liquids, and vowels. The right-ear advantage for stops was greater than that for liquids, which in turn was greater than that for vowels. Comparing phi indices for each phoneme class, nine subjects showed the relationship of stops greater than liquids, greater than vowels. By chance alone, this is a very unlikely outcome [z = 3.91, p < .0001]. Only one subject had ear advantages in the reverse order. Although differences between stops and vowels and between liquids and vowels were statistically reliable [T(16) = 13, p < .01, and T(16) = 28, p < .05, respectively], the difference between stops and liquids only approached significance [T(16) = 38, p < .10].

Identity trials. The remaining 8% of all errors occurred for phonemes shared by both stimuli on a particular trial. That errors occurred at all on these trials probably stems from acoustic differences between the two instances of the same phoneme: for example, a /k/ before /lae/ is slightly different from a /k/ before /rɛ/. Although identity-trial errors were relatively few, significantly more errors were made when the subjects monitored the left ear than when they monitored the right [z = 3.52, p < .0005]. There were no significant differences among stops, liquids, and vowels for these errors.

Discussion

The results of the present study support the notion that right-ear advantages of different magnitudes can occur systematically for different sets of phonemes. Certainly the overall trend shown in Fig. 2 is
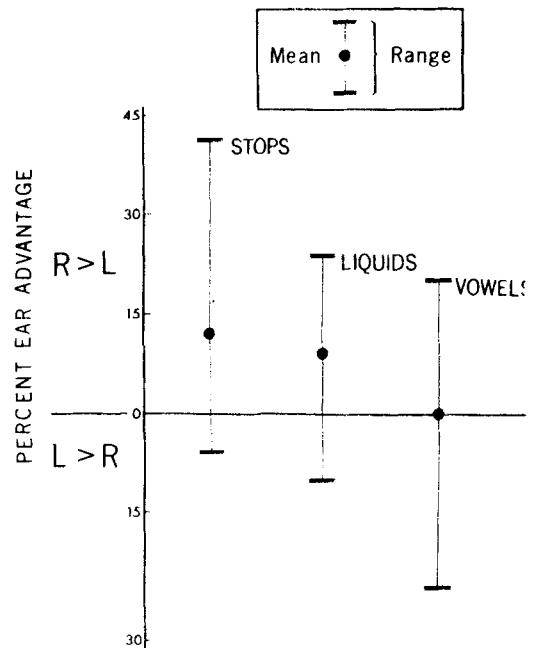

Fig. 2. Percent ear advantages for stops, liquids, and vowels within the same syllable.

compelling, and agrees remarkably well with ear advantages found by Day and Vigorito (1973) and Crystal and House (1974) for stops, liquids, and vowels in temporal-order judgment tasks. Although there was no significant difference between the stops and the liquids here, a difference is suggested. Moreover, such a difference may be related to the ear difference between consonants and vowels.

A number of explanations are possible for these systematic differences. The present paper will entertain four relatively simple explanations, and then suggest a fifth. They are not all mutually exclusive, nor are they exhaustive. First is the possibility that the results of the present study are explicable in terms of temporal order of arrival: stops were always first in the stimulus, liquids second, and vowels third. Although this hypothesis is viable for the present study, it cannot account for the temporal-order judgment data of Crystal and House (1974) and Day and Vigorito (1973). In their tasks, stops, liquids, and vowels were always in initial position in the stimulus. Nonetheless, since the present study was not a temporal-order task, the explanation should be considered. It would seem to predict that the same phoneme class would yield differential ear advantages according to its position within the syllable.

A second explanation is that the differences found in the present study stem from the durations of the particular phonemes. As shown in Fig. 1, the acoustic influences of the stops are found 120 msec into the syllable, the influence of the liquid occurs over as

much as 225 msec of the syllable, and vowel information spans as much as 400 msec. Perhaps left-hemisphere mechanisms are merely faster at extracting information from the signal. If so, the differential ear advantages might reflect the facility with which the left hemisphere can perform such tasks against the comparative sluggishness of the right hemisphere. This explanation would seem to predict that phonemes of different durations would yield different ear advantages and that phonemes of the same duration would yield similar advantages. Again, the present experiment cannot offer counterevidence.

The third explanation is concerned with perceptual difficulty. Many consonants are more difficult to perceive than vowels. The results of the present study, for example, demonstrate that in the dichotic situation vowels were identified correctly on 81% of all trials, whereas stops and liquids were identified on only 66% and 64% of the trials, respectively. Perhaps these consonants tax the perceptual system more than vowels. Furthermore, perhaps in the dichotic situation such difficulty differentially favors the left-hemisphere processors. This explanation would seem to predict that phonemes which are more difficult to perceive (as measured by some perceptual metric) would yield larger ear advantages than those which are relatively easy to perceive.

The fourth explanation concerns encodedness. Classes of phonemes differ to the extent which they must be restructured in different contexts. Stop consonants undergo massive context-conditioned variation, liquids perhaps less so, and vowels certainly less (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Perhaps the extent of restructuring (encoding) necessary to produce a speech sound is reflected in the extent to which decoding is necessary and, in turn, the extent to which left-hemisphere speech mechanisms have an advantage over their counterparts in the right hemisphere. The results of the present study are certainly congruent with this explanation. Such a hypothesis would seem to predict that the same phoneme which has undergone the same amount of contextual variation would yield the same ear advantage regardless of where it occurred within a syllable.

Experiment II is addressed to each of these four hypotheses.

EXPERIMENT II:
EAR ADVANTAGES FOR STOPS AND LIQUIDS
IN INITIAL AND FINAL POSITION[2]

Method
Stimuli. Twenty-four syllables were prepared on the Haskins parallel-resonance synthesizer: 12 were CV syllables and 12 were VC syllables. All possible combinations of the consonants /b, g, l, r/ and the vowels /i, ae, ɔ/ were used: thus, there were six stop-vowel syllables (/bi, bae, bɔ, gi, gae, gɔ/), six vowel-stop syllables (/ib, aeb, ɔb, ig, aeg, ɔg/), six liquid-vowel syllables (/li,
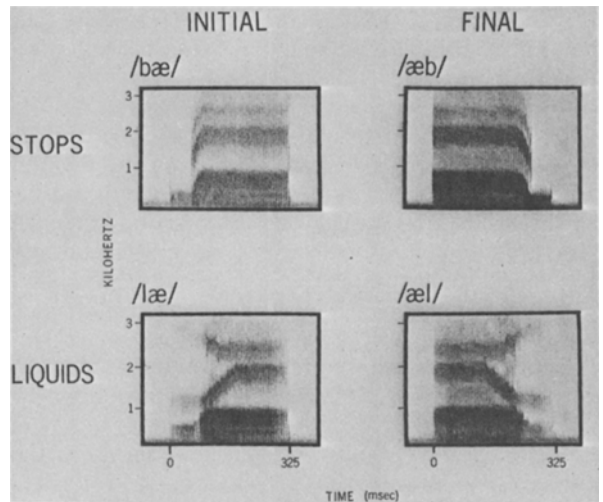

Fig. 3. Sound spectrograms of sample stimuli with initial and final stops and liquids.

lae, lɔ, ri, rae, rɔ/), and six vowel-liquid syllables (/il, ael, ɔl, ir, aer, ɔr/). The stimuli were 325 msec in duration and all had the same falling pitch contour. Except for pitch, the CV and VC syllables were exact reversals of one another. Appropriate acoustic compromises were made between the values for naturally occurring initial and final stops and liquids. As shown in Fig. 3, the acoustic structure of /bae/ was identical to that of /aeb/ except that the time axis had been reversed. Note that final stops were unreleased. The syllables /lae/ and /ael/ were also exact reversals, as were all other CV-VC pairs with the same consonant and vowel. This reversal was accomplished using the PCM system, which enables the experimenter to reverse the time ordering of a stimulus in the memory buffer without changing any other parameters. Stop consonants which preceded the same vowel differed only in the direction and extent of the second-formant (F_2) transition. Formant transitions were 50 msec in duration with an accompanying 25-msec F_1 voice bar. Liquids which preceded the same vowel differed only in the direction and extent of the F_3 transition and accompanying F_3 steady state. F_2 and F_3 transitions were 100 msec in duration with 75 msec of prerelease steady-state voicing in all formants. These same rules for stimulus construction apply to stops and liquids which followed the same vowel.

Subjects. Sixteen Yale University undergraduates participated in two tasks. They were selected according to previously discussed criteria. They listened in groups of four.

Task 1: Diotic Identification
A brief identification test was conducted to assess the quality of the stimuli.

Procedure
Subjects listened to one token of each stimulus to familiarize themselves with synthetic speech. They then listened to two diotic identification tests, each of 120 items. One test consisted of stop stimuli and the other consisted of liquids. Each of the 12 stimuli within the stop and the liquid sets was presented 10 times in random sequence with a 2-sec interstimulus interval. The subjects were asked to identify only the consonant in each of the stimuli, writing B or G for the stops and L or R for the liquids.

Results
The stimuli were highly identifiable. All were identified at a rate of 89% or better. There was no

difference in the identifiability of initial and final consonants.

Task 2: Dichotic Ear Monitoring

Procedure

Two dichotic test tapes were made for the set of stop stimuli and two for the set of liquids. Within each set, two rules governed the pairing of stimuli: (a) both stimuli in a dichotic pair were either CV syllables or VC syllables, and (b) the two stimuli shared neither the same consonant nor the same vowel. Thus, /bae/ was paired with /gi/ or /gɔ/, while /aeb/ was paired with /ig/ or /ɔg/. The same pattern was followed for the liquid stimuli.

Each test consisted of 72 dichotic pairs: (6 possible pairs within a syllable class) by (2 syllable classes, CV and VC) by (2 channel arrangements) by (3 observations per pair). Two such tests with different random orders were prepared for the stop stimuli, and two similar tests for the liquids. All had a 3-sec interval between pairs. The subjects listened to two repetitions of each 72-item test for both stops and liquids, yielding a total of 576 trials per subject.

The subjects were instructed to attend to only one ear at a time, and to write down the consonant that was presented to that ear, B or G, L or R. For each set of stimuli, the order of ear monitoring was done in the same manner as in Experiment I: half the subjects attended in the order RLLR, and half in the order LRRL. The order of headphone assignments and the order of listening to the stop and liquid tests were also counterbalanced across subjects. Six practice trials familiarized them with the task.

Results

Overall performance for all stimulus pairs was 67%. There was no difference in the overall performance for the stop stimuli and the liquid stimuli. Syllable position, however, proved to be important; performance on the initial consonants was better than on the final consonants. The subjects were 70% correct for both initial stops and initial liquids, while they were only 64% correct for final stops and final liquids. This net 6% difference was statistically significant [$F(1,15) = 15.4$, $p < .005$].

Stops. Initial and final stops yielded right-ear advantages, and both were of the same general magnitude. While monitoring stimuli with initial stops, the subjects were 74% correct for the right ear and 67% correct for the left ear, yielding a net 7% right-ear advantage. Final stops yielded a similar pattern. The subjects were 67% correct for the right ear and 61% for the left ear, yielding a net 6% right-ear advantage. Both right-ear advantages were significant [$F(1,15) = 8.5$, $p < .025$], and there was no significant difference in their magnitude, as suggested in Fig. 4.

Liquids. The liquids yielded a different pattern of results: initial liquids yielded a right-ear advantage, but final liquids did not. While monitoring initial liquids, the subjects were 73% correct for the right ear and only 66% correct for the left ear, a net 7% right-ear advantage. Final liquids, however, yielded no significant ear advantage. The subjects were 63% correct for the right ear and 65% correct for the left ear, a net 2% left-ear advantage. Thus, unlike the stops, the liquids show a pattern of results which varies according to the position of the target phoneme
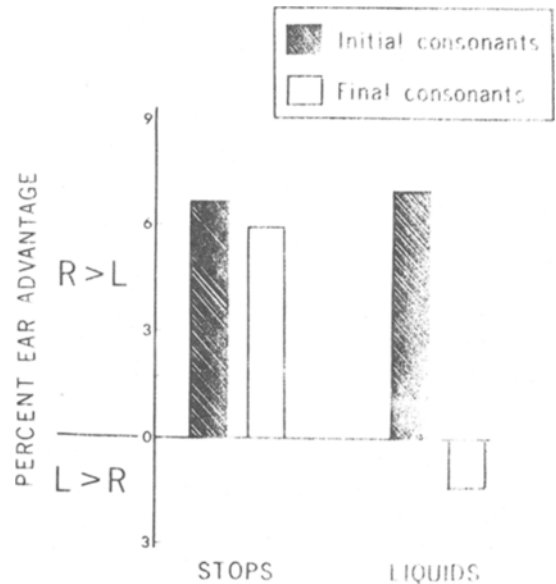

Fig. 4. Percent ear advantages for stops and liquids in initial and final syllable position.

within the syllable. The Ear by Syllable Class interaction for the liquid stimuli was significant [$F(1,15) = 6.4$, $p < .025$], as suggested in Fig. 4.

Discussion

How compatible are the results of the present study with the four previously outlined explanations for differential ear advantages? Explanation 1 accounts for the results of Experiment I in terms of the order of arrival of each phoneme; earlier arriving phonemes may have to overcome some form of perceptual inertia, may be harder to process, and may invoke the speech mechanisms of the left hemisphere to a greater degree during the early portions of the syllable. The results of the liquids in the present study support this notion in that they yielded a right-ear advantage in initial position but no advantage in final position. The results of the stop consonants, however, do not support this explanation, since they yielded nearly identical results regardless of syllable position. Moreover, Explanation 1 cannot account for the interaction of phoneme class with syllable position.

Explanation 2 contends that differential ear advantages are accountable in terms of durations of particular phonemes and phoneme classes. The results of the present study are embarrassing to this hypothesis as well. As shown in Fig. 3 and as explained previously, the duration of the acoustic manifestation of the stop consonants (including voice bar) was 75 msec. The duration of the liquids, on the other hand, was 175 msec. According to Explanation 2, liquids should yield smaller right-ear advantages than stops. As shown in Fig. 4, no such difference occurred for initial stops and liquids, and,

in fact, any tendency for such a difference is in the opposite direction. Again, this hypothesis makes no prediction of an interaction of phonemes with syllable positions.

Explanation 3 accounts for differential laterality results in terms of perceptual difficulty: phonemes which are more difficult to process yield larger right-ear advantages than those which are relatively easy to process. The results of the present study do not support this notion. Presented diotically, all stimuli were highly identifiable. Presented dichotically, the final consonants were significantly more difficult to perceive (identify) than their initial-position counterparts. This difference was found for both liquids and stops, and was 6% in both cases. Such a finding, according to Explanation 3, might lead one to expect larger right-ear advantages for the final consonants. The stop consonants do not verify this prediction, and the liquids demonstrate exactly the opposite result.

Explanation 4 predicts that phonemes of varying degrees of encodedness should yield ear advantages of different magnitudes, and that encodedness should correlate with ear advantage. The results of the present study do not support this hypothesis either. Liquid phonemes in initial and final position, preceding and following the same vowel, were constructed so that their relative amounts of context-conditioned variation were identical—they were "mirror images" of one another—and yet they yielded different ear advantages.

Each of the four explanations appears to fail because each is too simplistic. All such notions would have difficulty in accounting for the pattern of results in the present study. In most cases, it is the results of the liquids which they cannot predict. Thus, before looking to a more comprehensive explanation, it is necessary to consider the liquids in more depth.

Liquids as Consonants and Liquids as Vowels

A right-ear advantage in dichotic listening is a typical outcome for consonant stimuli, while no ear advantage is common for vowels. With this scheme in mind, the results of the present study might be interpreted in the following manner: initial liquids yield results which are typical of consonants, while final liquids yield results which are typical of vowels. This notion has a certain amount of linguistic plausibility. Liquids are maverick phonemes. In several distinctive feature systems (Halle, 1964; Jakobson, Fant, & Halle, 1951), they are considered to have both consonantal (consonant-like) and vocalic (vowel-like) features. Perhaps it is the initial liquids which are more like consonants and the final liquids which are more like vowels.

Consider differences in phoneme transcription for initial and final liquids. For example, the two /r/ sounds in the word RAIDER may be transcribed

differently. A typical transcription of RAIDER is /redɚ/. The initial /r/ is treated as a consonant, while the final /ɚ/ is treated as a vowel[3] (see Bronstein, 1960, p. 116). A similar treatment may be found for /l/. Following Bronstein, the two /l/ sounds in LADLE may be transcribed in a different manner. One such transcription is /ledł/. The initial /l/ is considered to be "light," whereas the final /ł/ is considered to be "dark" and is identified with a bar across the middle of the symbol. Bronstein considers the dark /ł/ sound to be similar to a back vowel. Perhaps the light /l/ functions more like a consonant. No phonemes other than /l/ and /r/ are transcribed differently as a function of their position within a syllable.

Perceptual differences among liquids may stem from their acoustic cues. That is, /l/s and /r/s may be cued differently in initial and final position even though they are "mirror images" of one another. Preliminary studies indicate that, given mirror-image stimuli, the cues for initial /l/ and /r/ are found primarily in the formant transitions, whereas the cues for final /l/ and /r/ are found primarily in the steady-state offglides. In some sense, then, initial liquids are cued in a manner similar to stops, and final liquids in a manner more like vowels. If the perceptual system takes advantage of where these cues are located in the sound pattern of speech, the differential ear advantages for initial and final liquids are explicable in *acoustic* terms, not *phonetic* terms. That is, some contribution to the right-ear advantage may stem from the acoustic form of the signal, not from the phonetic labeling of it.

A comprehensive explanation of differential ear advantages may need to consider the possibility that contributions are made to the right-ear advantage by auditory (acoustic) as well as phonetic aspects of the speech signal. Experiment III was designed to observe this possibility in clearer detail.

EXPERIMENT III:
PERCEPTION OF SPEECH AND NONSPEECH, WITH AND WITHOUT TRANSITIONS

The work of Pollack (1968) and Nabelek and Hirsh (1969) has revealed much about how rapidly changing frequency patterns are perceived in isolation, but the role of these patterns in the perception and lateralization of complex sounds is less clear. Speech sounds, which typically have a wealth of independently moving formants, are processed differently than nonspeech signals, which often do not have such variation (see Kimura, 1967). Moreover, speech sounds with transitions yield larger right-ear advantages than those without (Shankweiler & Studdert-Kennedy, 1967). The data of Halperin, Nachshon, and Carmon (1973) suggest that nonspeech signals with transitory information yield

results which reflect more left hemisphere processing than those without such changes. The present study, therefore, was designed to observe the effect on lateralized perception of the presence and absence of transitions in speech and nonspeech stimuli.

Method

Stimuli. Nine speech stimuli and nine nonspeech stimuli were prepared using the facilities at the Haskins Laboratories. Three-formant speech stimuli were prepared on the parallel-resonance synthesizer and included three steady-state vowel (V) syllables /i, ae, ɔ/ and six consonant-vowel (CV) syllables /bi, gi, bae, gae, bɔ, gɔ/. Each of the speech stimuli was synthesized to have the same steady-state fundamental frequency (110 Hz). They were then transferred to the PCM system for the preparation of dichotic test tapes. Nonspeech stimuli were computed numerically and stored by the PCM system as three pure tones. The frequency of the tones in each stimulus corresponded exactly to the central frequency of the formants in the speech stimuli. Thus, one nonspeech stimulus was composed of sine waves of 660, 1,720, and 2,410 Hz, the same mean values for the formants in the vowel /ae/. Two other stimuli corresponded to the vowels /i/ and /ɔ/. Six of the nonspeech stimuli consisted of steady-state sine-wave segments which followed 50 msec of frequency modulation. These modulations rose or fell to the steady-state frequency of the tones and corresponded to the transitions in the stop-vowel stimuli. All stimuli were identical in peak intensity and all were 300 msec in duration. The amplitude relationships among the component tones of the sine wave stimuli corresponded to the amplitude relationships of the formants in the speech stimuli. The sine-wave stimuli which resembled the CV stimuli are designed CVsw, and the sine-wave stimuli which resembled the V stimuli are designated Vsw. Figure 5 shows a schematic representation of four stimuli used in the present study, one from each category. Thick bars indicate formants in the speech stimuli, while narrow lines indicate sine waves in the nonspeech stimuli. The speech stimuli shown are /bae/ and /ae/, and the nonspeech stimuli are those which correspond to them. The actual bandwidths of speech formants from the Haskins synthesizer are 60, 90, and 120 Hz for F_1, F_2, and F_3, respectively. The bandwidth of sine waves is essentially zero.

Subjects. Ten Yale University undergraduates participated in Task 1, and 16 others in Task 2. They were selected by the same criteria used in Experiment I.

Task 1: Diotic Identification

A brief identification task was conducted to see if the sine-wave stimuli were identifiable as speech.

Procedure

Only CV and CVsw stimuli were selected to be identified. The subjects listened to tests of 120 items presented in random order one at a time: (2 classes of stimuli) by (6 stimuli per class) by (10 observations per stimulus). They were instructed to write down their responses as BEE, DEE, GEE, BAA, DAA, GAA, BAW, DAW, or GAW for each item. In this manner, the subjects were forced to try labeling the sine-wave stimuli as speech. Note that there were, in fact, no stimuli which began with the phoneme /d/. No practice or training was given.

Results

CV stimuli were correctly identified on 82% of all trials, while CVsw stimuli were identified as their appropriate counterparts on only 19% of all trials, only a few percentage points above chance. Responses were parsed into their component segments, as in Experiment I, and performance was measured for
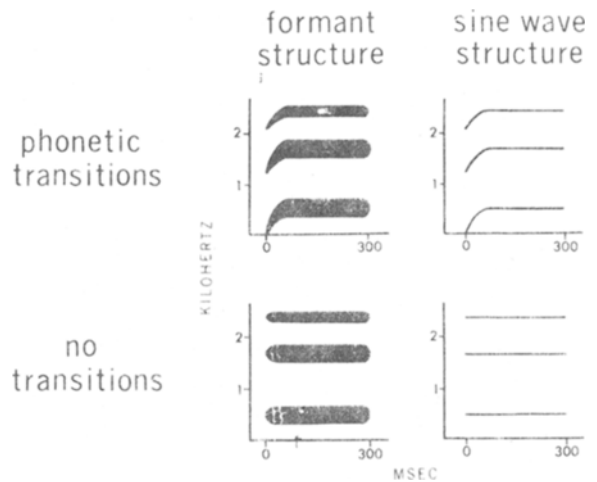

Fig. 5. Schematic spectrograms of speech and nonspeech stimuli, with and without transitions.

both consonants and vowels. For the CV stimuli, consonants were correctly identified 82% of the time and vowels 97% of the time. The corresponding scores for the CVsw stimuli were 38% and 45% "correct."

These results strongly indicate that the formant stimuli were processed as speech but that the sine-wave stimuli were not. Although the subjects gave consistently correct identifications for the CV stimuli, they scattered their identifications over the nine possible responses for the CVsw stimuli. The fact that speech stimuli were identified correctly at such a relatively low rate is almost entirely due to /d/ intrusions in the identification of /b/ and /g/ stimuli. There were few /b/-/g/ confusions. The effect of the number of stimuli perceived to be in the stimulus repertoire may need further attention.

Task 2: Dichotic Temporal-Order Judgment

A dichotic temporal-order judgment task was devised that did not require subjects to give verbal labels to the stimuli.

Procedure

Trials for the dichotic temporal-order judgment task required that subjects judge whether the leading stimulus in a given dichotic pair was identical with a subsequent probe stimulus. A trial consisted of a dichotic pair with a temporal onset asynchrony of 50 msec, followed by 1 sec of silence, followed by a diotic stimulus which was one of the members of the dichotic pair. Subjects were instructed to regard the diotic stimulus as a probe which asked the question: "Is this the stimulus which began first?" Figure 6 shows a schematic representation of two such trials. Consider Sample Trial 1, where Stimulus 1 begins before Stimulus 2 by 50 msec, and the probe stimulus is Stimulus 1. Since the probe is identical to the stimulus which began first, the correct response is *yes*. In Sample Trial 2, the dichotic pair is the same as in Trial 1, but the probe stimulus is different. Since Stimulus 2 did not begin before Stimulus 1, the correct response for Trial 2 is *no*. The 50-msec onset asynchrony was chosen because it is a particularly difficult interval at which to judge stimulus order (see Day & Cutting[4]; Day,
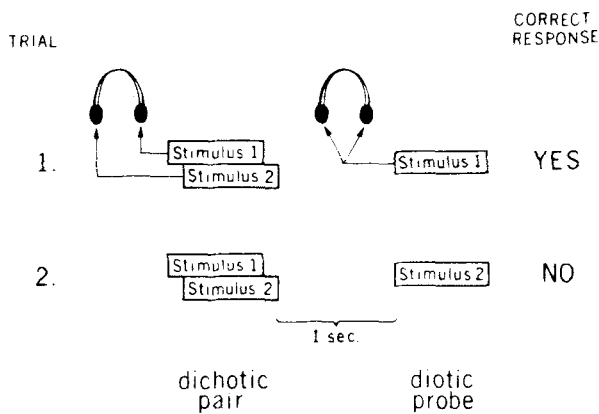
Fig. 6. A paradigmatic display to two dichotic temporal-order judgment trials with a diotic probe.

Cutting, & Copeland[5]; Day & Vigorito, 1973).

Four tests were constructed, one for each class of stimuli. Each test tape consisted of 48 trials: (6 possible pairs) by (2 channel arrangements) by (2 possible probes) by (2 observations per pair). The stimuli used in each trial were always selected from the same class of stimuli. CV trials were constructed out of CV stimuli which shared neither the same vowel nor the same consonant: thus, for example, /bi/ was paired with /gae/ or /gɔ/. In V trials, different vowel stimuli were paired: for example, /i/ was paired with /ae/ or /ɔ/. CVsw trials were constructed using the same rules applied to CV stimuli: the two stimuli in the dichotic pair neither shared the same sine-wave structure nor had pitch modulations which corresponded to the transitions of the same stop consonant. The rules for the construction of Vsw trials were the same as those for V trials. Stimuli in the dichotic pair were counterbalanced for leading and lagging position. The probe stimulus chosen for each trial and the channel assignments of the stimuli in the dichotic pair were also counterbalanced in the random sequence of trials.

Subjects listened to each test twice, reversing the earphones before the test tape was rewound and played again. The order of channel assignments was counterbalanced across subjects. Each group listened to the four tests in a different order, determined by a Latin-square design. Subjects listened to a total of 384 trials, consisting of a dichotic pair and a diotic probe, writing Y for "yes" or N for "no" for each trial. Four practice trials were given before each stimulus class in order to familiarize subjects with the stimuli.

Results

In general, the task was quite difficult; overall performance for all trials and all types of stimuli was 65% correct. Performance for each of the four types of stimuli was comparable: the average score for each was between 64% and 66%, with no significant differences among them. Because of the comparability of performance levels, no phi analyses were performed.

Before discussing the pattern of ear advantages, it is necessary to note how the results were scored. Consider again the sample trials in Fig. 6. The correct response for Sample Trial 1 is *yes*, while the correct response for Trial 2 is *no*. If, in the dichotic pair, Stimulus 1 was presented to the right ear and Stimulus 2 to the left ear, and if the S responded *yes* for the first trial and *no* for the second, he would have gotten both correct. This would be scored as two

correct responses for the right-ear leading stimulus. If the subject, on the other hand, had responded *no* and *yes* for the sample trials, respectively, both would be wrong and his score for the right-ear leading stimulus would be docked for two incorrect responses. (Of course, if the channels had been reversed with Stimulus 1 presented to the left ear and Stimulus 2 to the right, the logic would be entirely reversed).

CV trials. There was a large, significant ear difference for the consonant-vowel trials. When the leading stimulus of the dichotic pair was presented to the right ear, subjects were 72% correct in responding to the probe stimulus. When the leading stimulus was presented to the left ear, on the other hand, subjects were only 57% correct, yielding a net 15% right-ear advantage. Fourteen of the 16 subjects showed results in this direction, a ratio which is statistically reliable by a sign test [z = 2.75, p < .003].

V trials. There was a significant ear difference for the steady-state vowels as well. When the leading stimulus was presented to the right ear, subjects were 69% correct, while they were only 62% correct when the leading stimulus was presented to the left, yielding a net 7% right-ear advantage. The overall low performance for the vowels is not unusual in temporal-order judgment tasks. Twelve of 16 Ss showed results in this direction [z = 1.87, p < .06].

CVsw trials. No significant ear difference was found for the nonspeech stimuli most resembling the synthetic consonant-vowel syllables. Ear scores were 65% and 64% correct for trials in which the right- and left-ear stimuli led, yielding a net 1% right-ear advantage.

Vsw trials. A small left-ear advantage was found for the nonspeech stimuli which were most like vowels. When the steady-state tone stimulus led in the right ear, subjects were 63% correct, while they were 68% correct when it led in the left ear, yielding a net 5% left-ear advantage. The ear advantage for these trials was not significant. Ten of 16 subjects showed results in this direction (z = .75, n.s.).

The results of each of the four conditions are shown in Fig. 7. Observe the relationships of the ear-difference scores for the different classes of stimuli.

Speech vs nonspeech. The dimension of speech vs nonspeech proved to be a significant factor in the temporal-order judgment of these stimuli. Subjects showed a large right-ear advantage for the speech stimuli. Averaging the ear scores for the CV and V stimuli yields 11% right-ear advantage. This superiority of the right-ear-leading stimuli was highly significant [F(1,15) = 19.3, p < .001]. Nonspeech stimuli, however, showed a different pattern of results. Collapsing CVsw and Vsw results yields a 2% advantage to the left-ear leading stimulus. This small effect was not significant itself, but the difference in the two classes of stimuli, speech vs nonspeech, was
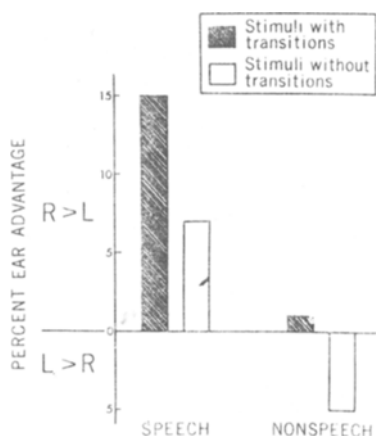
Fig. 7. Percent ear advantages for speech and nonspeech, with and without transitions.

highly significant [F(1,15) = 19.4, p < .001].

Stimuli with and without transitions. Independent of the speech vs nonspeech contrast, transitions played a definite role in temporal order judgments. Summing CV and CVsw results yields 8% advantage in favor of the right-ear-leading stimuli. Summing the other two sets of stimuli, however, yields no significant difference: V and Vsw stimuli yielded a net difference of only 1% in favor of the right-ear-leading stimulus. The difference in the results for the stimuli with transitions and those without transitions was significant [F(1.15) = 6.61, p < .05], although neither the CV and V or the CVsw and Vsw comparisons were statistically reliable.

Discussion

Differences in the perception of speech and nonspeech in dichotic listening have been well documented. In general, speech tasks, whether identification or temporal-order judgment, yield right-ear advantages (Day & Cutting[4]; Kimura, 1961). Nonspeech tasks, on the other hand, whether identification or temporal-order judgment, generally yield left-ear advantages (Chaney & Webster, 1966; Day & Cutting[4]). The results of the present study are in accord with these findings. Speech stimuli yielded a large advantage to stimuli leading in the right ear, nonspeech stimuli yielded a small advantage to stimuli leading in the left ear, and although the ear advantage for nonspeech stimuli was not significant, it was significantly different from the speech stimuli. Thus, as in previous studies, processing appears to be different for the two types of stimuli, each type of stimuli requiring different general amounts of processing in each hemisphere.

Since speech/nonspeech differences have received much attention (Kimura, 1967; Mattingly, Liberman, Syrdal, & Halwes, 1971), consider the other, and perhaps more intersting, general result. That is,

stimuli with transitions appear to be perceived differently from stimuli without transitions *regardless* of whether or not they are speech or nonspeech. It appears that a stimulus which contains rapid changes in frequency may need special processing in the left hemisphere, and that this processing is not dependent on the stimulus having been classified as speech.

The processing capabilities of the left hemisphere appear to be fundamentally different from those of the right hemisphere. Most of these differences have been thought to be related to language, and range through all levels of language from phonetics to semantics. Since speech and language processing is believed to be primarily confined to one hemisphere. It may be advantageous to have certain subsidiary systems in that hemisphere to assist in the processing of the speech signal. A subsystem which tracks rapidly changing frequencies may not be needed in the right hemisphere which is geared for less highly analytic processing (Semmes, 1968). The notion of an independent transition analyzer in the left hemisphere is congruent with the results of Darwin (1971). He found that fricatives with formant transitions yielded a right-ear advantage, while the same fricatives without transitions yielded no advantage.

In the present study, there are two variables: whether or not the stimulus is speech, and whether or not the stimulus has transitions. Speech processing is primarily a left-hemisphere task, and it may be that the analysis of transitions is also a left-hemisphere task. From the results of the present study, the two tasks appear to be independent of one another. CV stimuli, [+ speech] and [+ transition], yield a large right-ear/left-hemisphere advantage since the two variables favor left-hemisphere processing. V and CVsw stimuli, however, have only one positive value on the two dimensions and thus yield smaller ear difference scores: V stimuli are [+ speech] and [- transition] and CVsw stimuli are [- speech] and [+ transition]. Vsw stimuli are both [- speech] and [- transition], and consequently yield a left-ear advantage. Speech/nonspeech appears to be a more potent binary dimension than transition/nontransition: V stimuli yielded a larger right-ear advantage than did CVsw stimuli.

One intriguing possibility is that the speech processor may be an all-or-none device. It is possible that the difference in ear advantages for the CV and V stimuli may be entirely attributable to the auditory, but not phonetic, analysis of the stop consonants. Evidence for this stems from the results of the CVsw and Vsw stimuli, which differ to exactly the same extent as the CV and V stimuli, but which were not identified as speech and cannot be phonetically categorized. Just as there is a high degree of categorization within speech (Liberman, 1957; Pisoni, 1971), there may be categorical decisions as to whether a stimulus is speech or not (see House,

Stevens, Sandel, & Arnold, 1962).

From the results of Experiment III, it cannot be determined whether the alleged transition analyzer is a language-independent device. It is possible that some speech mechanism, perhaps a phonetic feature analyzer (Eimas, Cooper, & Corbit, 1973; Eimas & Corbit, 1973), was activated by the sine-wave stimuli. Transition values and steady-state values chosen for the sine-wave stimuli were exactly those appropriate for speech stimuli. That subjects could not identify them as speech may only reflect the cognitive outcome of the entire perceptual process. Phonetic decisions about the sine-wave stimuli may have been rejected by some higher level system for lack of evidence that the signal was speech. Experiment IV was designed to observe whether stimuli with transitions not corresponding to phonetic segments yield ear advantages equal to those stimuli which have similar transitions, but which also carry phonetic information.

EXPERIMENT IV:
PERCEPTION OF SPEECH AND NONSPEECH WITH PHONETIC AND NONPHONETIC TRANSITIONS

### Method

Stimuli. The CV and CVsw stimuli of Experiment III were used again. They were, or corresponded to, the consonant-vowel syllables /bi, gi, bae, gae, bɔ, gɔ/. Stimuli containing the same vowel, or sine waves which corresponded to the same vowel (such as /bae, gae/), were identical in all respects except for the second-formant transition. In all cases, the second-formant transition of /b/ rose for 50 msec to the resting frequency of the formant, while the transition of /g/ fell to that frequency. First- and third-formant transitions were both always upgliding. Two other sets of stimuli were synthesized. One set was similar to the CV stimuli in that its members contained formants and formant transitions typically found in speech. The particular array of formant transitions, however, could never have been produced by a human vocal tract. As in the CV stimuli, the second-formant transition could be either upgliding or downgliding in frequency. The first- and the third-formant transitions, on the other hand, were always downgliding. Since these stimuli had transitions which did not correspond to any consonant speech segment, they are designated C'V stimuli. The fourth stimulus set consisted of sine-wave analogs of the C'V stimuli, and are designated C'Vsw stimuli. Four stimuli, one from each class, are displayed in Fig. 8.

Subjects and procedure. Sixteen subjects participated in one task, dichotic temporal-order judgment. They were selected from a summer-subject pool and screened using the same criteria discussed previously. Four dichotic test tapes were prepared on the PCM system, one for each of the four classes of stimuli: CV, C'V, CVsw, and C'Vsw. Tests and trials were otherwise identical to those in Experiment III. Subjects again listened in groups of four. Test order was counterbalanced across groups of subjects according to a Latin-square design. No identification test was conducted.

Results

Overall performance for all trials and all stimulus sets · was 60%, only slightly lower than in Experiment III and not atypical for 50-msec temporal-order judgments. Again, there was no
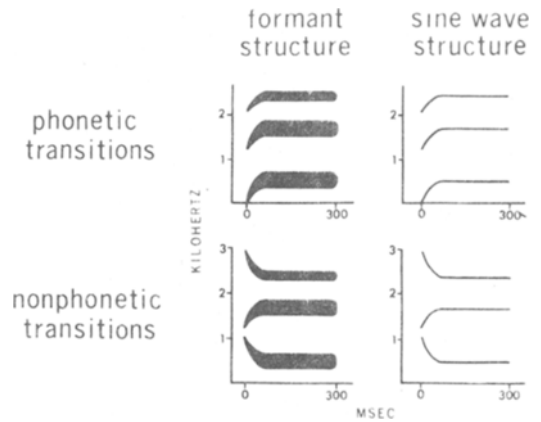

Fig. 8. Schematic spectrograms of speech and nonspeech stimuli, with phonetic and nonphonetic transitions.

significant difference in overall performance for the four types of stimuli: all were between 59% and 61%, making phi analyses unnecessary.

CV and C'V trials. There was a large significant ear difference for the formant stimuli. Subjects were 65% correct in responding to the probe stimulus when the leading stimulus was presented to the right ear, and 54% correct when presented to the left, yielding a net 11% difference [$F(1,15) = 5.38$, $p < .05$]. There was no significant difference between CV and C'V stimuli. Subjects' scores were more variable than in Experiment III, and only 11 of 16 subjects had right-ear advantages for these trials [$z = 1.25$, n.s.).

CVsw and C'Vsw trials. No significant ear advantage was found for sine-wave stimuli. Subjects were 61% and 60% correct for right-ear and left-ear leading trials, respectively—a net 1% right-ear advantage.

The result of each of the four conditions is shown in Fig. 9. The effect of phonetic vs nonphonetic
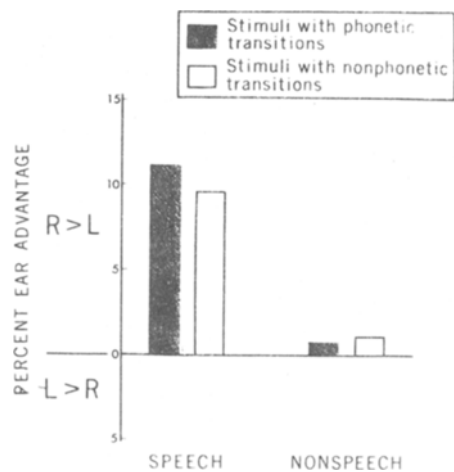
Fig. 9. Percent ear advantages for speech and nonspeech, with phonetic and nonphonetic transitions.

transitions was not an important factor in either formant or sine-wave stimulus perception.

Speech vs nonspeech. Although the C'V stimuli did not have phonetic transitions, their vowel portions were speech-like enough to make the stimuli sound like syllables with a garbled beginning, and thus may be justifiably called "speech." Sine-wave stimuli, regardless of the nature of their transitions, never sounded like speech. Accepting this redefinition of the speech/nonspeech distinction, there was a significant difference in the ear advantages for the two types of stimuli [F(1,15) = 16.41, p < .001]. The magnitude of the speech/nonspeech difference in the present study was comparable to that in Experiment III.

Discussion

The results of the present study suggest that transitions are analyzed, in part, by left-hemisphere mechanisms regardless of whether they are phonetic or nonphonetic. Stimuli with transitions not corresponding to any phoneme yield results identical to stimuli which have phonetically appropriate transitions. This result has been replicated by Cutting (in press) in a recognition task not involving judgments of temporal order. It is also congruent with that of Kimura and Folb (1968), who found similar right-ear advantages for the perception of normal speech sounds and speech played backwards. Backwards speech, like C'V stimuli, has transitions which are often inappropriate for the perception of any specific speech segments, but it is nevertheless heard and processed as speech.

Any explanation of the results of Experiments I-IV must be more complex than the four previously suggested. Each of the four appears to have a kernel of truth, but not all of it.

Explanation 1 suggests that phonemes that begin a syllable are more difficult to process than those which follow it. This may be true, but only because of the fact that CV and CCV syllables are much more common than VC or VCC syllables (for the data on English, see Trnka, 1968). Explanation 2 suggests that duration of a stimulus (phoneme) maps directly onto the magnitude of the ear advantage. In a general sense, this is also true; stops are often briefer than liquids which are often briefer than vowels. What this explanation does not consider is the relative amounts of acoustic change over the course of an utterance. Explanation 3 suggests that processing difficulty is reflected in the magnitude of the right-ear advantage. Again, in general this also appears to be true, but it fails to distinguish between the various reasons why a signal may be difficult to process. Not all of these reasons appear to contribute to the ear advantage. In Experiment II, for example, final liquids were more difficult to process than their initial counterparts (in terms of overall dichotic performance) but they did not yield a right-ear advantage. Explanation 4 also

has some value. Certainly, C'V and C'Vsw stimuli in Experiment IV may be thought to be encoded, even if the encoding is not relevant to phonetic decisions. This explanation, however, cannot account for different ear advantages for the same phoneme in the same context but in different syllable position.

A comprehensive alternative to these four explanations, and any other simple explanation that might be advanced, cannot be tightly formulated at present; only a sketch of what it must include can be suggested. Although the results of Experiment I suggest that some simple mechanism may account for differential ear advantages, the Syllable Position by Phoneme Class interaction of ear advantages in Experiment II dampens that possibility. I have interpreted the results of Experiment III as indicating that at least two separate mechanisms are involved in the perception of speech. One mechanism appears to be an auditory device whose task is to process complex acoustic aspects of the speech signal. It appears to be auditory, because the results of Experiment IV indicate that it does not discriminate between phonetic and nonphonetic transitions in making contributions to laterality results. This auditory device cannot be the only mechanism involved in speech perception, because the results of Experiments III and IV demonstrate that stimuli that are classifiable as speech yield larger right-ear advantages than sine-wave stimuli which resemble speech tokens but which are not identifiable as speech. Thus, it appears that there must be at least two mechanisms involved in the left-hemisphere specialization for speech perception, one that is auditory (at least in the sense that it is language-independent) and one that is phonetic.

REFERENCES

BORING, E. G. *A history of experimental psychology.* New York: Appleton-Century-Crofts, 1950.

BROADBENT, D. E. Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology,* 1956, **8,** 145-162.

BRONSTEIN, A. J. *The pronunciation of American English.* New York: Appleton-Century-Crofts, 1960.

CHANEY, R. B., & WEBSTER, J. C. Information in certain multi-dimensional sounds. *Journal of the Acoustical Society of America,* 1966, **40,** 447-455.

CHOMSKY, N., & HALLE, M. *The sound pattern of English.* New York: Harper and Row, 1968.

COOPER, F. S., & MATTINGLY, I. G. Computer-controlled PCM system for investigation of dichotic speech perception. *Journal of the Acoustical Society of America,* 1969, **46,** 115(A).

CRYSTAL, I., & HOUSE, A. S. The effect of local signal level on differential performances in dichotic listening. *Journal of the Acoustical Society of America,* 1974, **55,** 434(A).

CUTTING, J. E. Different speech-processing mechanisms can be reflected in the results of discrimination and dichotic listening tasks. *Brain and Language,* in press.

DARWIN, C. J. Ear differences in the recall of fricatives and vowels. *Quarterly Journal of Experimental Psychology,* 1971, **23,** 46-62.

DAY, R. S., & VIGORITO, J. M. A parallel between encodedness and the ear advantage: Evidence from a temporal-order judg-

ment task. *Journal of the Acoustical Society of America*, 1973, **53**, 358(A).

EIMAS, P. D., COOPER, W. A., & CORBIT, J. Some properties of linguistic feature detectors. *Perception & Psychophysics*, 1973, **13**, 247-252.

EIMAS, P. D., & CORBIT, J. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, **4**, 99-109.

HAGGARD, M. Encoding and the REA for speech signals. *Quarterly Journal of Experimental Psychology*, 1971, **23**, 34-45.

HALLE, M. On the bases of phonology. In J. A. Fodor and J. J. Katz (Eds.), *The structure of language*. Englewood Cliffs, N.J: Prentice-Hall, 1964.

HALPERIN, Y., NACHSHON, I., & CARMON, A. Shift in ear superiority in dichotic listening to temporal pattern nonverbal stimuli. *Journal of the Acoustical Society of America*, 1973, **53**, 46-50.

HOUSE, A. S., STEVENS, K. N., SANDEL, T. T., & ARNOLD, J. B. On the learning of speech-like vocabularies. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 133-143.

JAKOBSON, R., FANT, C. G., & HALLE, M. *Preliminaries of speech analysis*. Cambridge, Mass: M.I.T. Press, 1951.

KIMURA, D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 1961, **15**, 166-171.

KIMURA, D. Dual functional asymmetry of brain in visual perception. *Neuropsychologia*, 1967, **4**, 275-285.

KIMURA, D., & FOLB, S. Neural processing of backwards-speech sounds. *Science*, 1968, **161**, 395-396.

KUHN, G. M. The phi coefficient as an index of ear differences in dichotic listening. *Cortex*, 1973, **9**, 447-457.

LIBERMAN, A. M. Some results of research on speech perception. *Journal of the Acoustical Society of America*, 1957, **29**, 117-123.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, **74**, 431-461.

MATTINGLY, I. G., LIBERMAN, A. M., SYRDAL, A. K., & HALWES, T. Discrimination in speech and nonspeech modes. *Cognitive Psychology*, 1971, **2**, 131-;157.

NABELEK, I., & HIRSH, I. J. On the discrimination of frequency transitions. *Journal of the Acoustical Society of America*, 1969, **45**, 1510-1519.

PISONI, D. B. On the nature of categorical perception of speech sounds. PhD thesis, University of Michigan (Psycholinguistics), 1971. (Published as Supplement to Haskins Laboratories Status Report on Speech Research.)

POLLACK, I. Detection of rate of change of auditory frequency. *Journal of Experimental Psychology*, 1968, **77**, 535-541.

SEMMES, J. Hemispheric specialization: A possible clue to mechanism. *Neuropsychologia*, 1968, **5**, 11-26.

SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, 1967, **19**, 59-63.

STUDDERT-KENNEDY, M., & SHANKWEILER, D. P. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, 1970, **48**, 579-594.

TRNKA, B. *A phonological analysis of present-day American English*. University, Alabama: University of Alabama Press, 1968.

NOTES

1. Presented as a paper before the 84th Convention of the Acoustical Society of America, Miami Beach, November 1972.

2. Presented as a paper before the 85th Convention of the Acoustical Society of America, Boston, April 1973.

3. Brackets [——] are often used to signify phonetic transcriptions and slashes /————/ are often used for phonemic transcriptions (Chomsky & Halle, 1968, p. 65). I have adopted the convention of using slashes for both throughout this paper for the sake of simplicity.

4. Day, R. S., & Cutting, J. E. *What constitutes perceptual competition in dichotic listening?* Paper presented at the Eastern Psychological Association meetings, New York, April 1971.

5. Day, R. S., Cutting, J. E., & Copeland, P. M. *Perception of linguistic and nonlinguistic dimensions of dichotic stimuli*. Paper presented at the annual meetings of the Psychonomic Society, St. Louis, November 1971. (Also in Haskins Laboratories Status Report on Speech Research, SR-27, 1971, 193-197.)