

## Identification of Consonants and Vowels Presented to Left and Right Ears<sup>1</sup>

Donald Shankweiler and Michael Studdert-Kennedy<sup>2</sup>  
Haskins Laboratories, New York City

Several lines of evidence suggest that speech perception is characterized by a process different from that for the perception of other sounds (see, for example, House et al., 1962; Kozhevnikov and Chistovich, 1965; Liberman et al., 1966). Recent work indicating that speech and nonspeech are processed primarily in different places in the brain strengthens this hypothesis. Studies by Kimura (1961a) and others (Bryden, 1963; Broadbent and Gregory, 1964) have shown that if random sequences of digits are dichotically presented (Broadbent, 1954) so that a different digit arrives at each ear at the same time, listeners retain digits presented to the right ear more accurately than those presented to the left. If the stimuli are brief melodies (Kimura, 1964) or sonar signals (Chaney and Webster, 1965), listeners retain more accurately those presented to the left ear. These lateral differences in efficiency of handling competing stimuli seem to reflect the greater strength of the crossed auditory pathways (Rosenzweig, 1951) and the specialization of the auditory areas of each hemisphere of the brain for processing different classes of stimuli. Studies of persons with left- and right-sided lesions of the auditory cortex have provided independent evidence of functional differences between the hemispheres in auditory perception (Kimura, 1961a; 1961b; Milner, 1962; Shankweiler, 1966).

The relative contribution of each cerebral hemisphere to the

perception of different classes of sounds merits further study. A basic question is at what stage, in the processing of speech, hemispheric differences in function become evident. Do they appear at the level of phonetic structure as well as at higher levels? Data of Chaney and Webster (1965) suggest that they do. We may then ask whether all phonetic elements, or all features of the phonetic elements, are processed in the same way. The technique of dichotic presentation offers an approach to these questions in persons with intact nervous systems, and one by which we may hope to learn more explicitly how speech perception differs from the perception of other sounds.

In the present study we compared identification of dichotically presented single pairs of synthetic steady-state vowels and of consonant-vowel (CV) syllables. We also examined the effects of interaural competition for the various combinations of articulatory features contained in the consonant syllable pairs.

#### Method

Two separate tests were made up, one consisting of synthetic consonant-vowel syllables and the other of synthetic steady-state vowels in isolation.

Six consonant-vowel syllables lasting 300 msec. were prepared on the Haskins Pattern Playback (Cooper et al., 1951)<sup>3</sup> and recorded on magnetic tape. They were the voiced stop consonants /b,d,g/ and the unvoiced stop consonants /p,t,k/, each followed by the vowel /a/. The fifteen possible syllable pairs (no syllable was paired with itself) were made into loops, aligned for simultaneity of onset and re-recorded on a dual channel tape recorder (Ampex PR-10). Each member of a pair was recorded twice on channel A and twice on channel B and the resulting sixty stimulus pairs were spliced into a random order.

A similar procedure was followed with five, equal duration (300 msec.), steady-state vowels, /i,ε,æ,o,u/. The ten possible

pairs, with each member of a pair recorded twice on each channel, yielded a random order of forty pairs.

The subjects for the consonant test were 15 right-handed members of the Laboratory staff, 10 women, 5 men between the ages of 20 and 50. Ten of these listeners also served as subjects in the vowel test. None had a known hearing loss.

Subjects were tested individually in a quiet room. They listened over earphones (Permoflux, PDR-8), wired for dichotic presentation, to the output from the tape recorder. The subjects operated the recorder themselves, starting and stopping the tape before and after each trial, and listening to each trial pair only once. They wrote their judgments on an answer sheet at the top of which were displayed the six consonants (or five vowels) from which they were to make their choices. Listeners were instructed that each stimulus pair would contain two different syllables and that they were to attempt to identify both, guessing if necessary. They were asked to write the judgment of which they were more confident in the first column and their other judgment in the second column. Each listener heard each test twice, the earphones being reversed on the second run, so that channels and ears were balanced. Approximately half the listeners heard channel A in their left ear first; half heard channel B in their left ear first.

### Results and Discussion

#### Consonant-vowel comparison.

The results of the consonant test showed that syllables presented to the right ear were identified with greater accuracy than were those presented to the left by 14 of the 15 subjects. The figures given below are mean percent correct for each ear when only the response given first is counted.

Left Ear  
29 per cent

Right Ear  
45 per cent

The difference is highly significant ( $p < 0.001$  by a two-tailed test). First preferences only are recorded since it turned out that second preferences were largely guess work.

The right ear advantage shows up again when we compare the results of the consonant and vowel tests for the ten subjects who took both. Table I shows first preference figures for these two conditions. Right ear performance is identical for both consonants and vowels -- 45 percent correct. Left ear performance is slightly, but not significantly, lower for the vowels (41%), but is significantly lower for the consonants (31%) ( $p < 0.001$  by a two-tailed test). The right ear advantage occurred unreliably on the vowel test (6 of the 10 subjects were better on the right ear).

In view of Kimura's finding (1964) of a left ear advantage for musical melody recognition, as against a right ear advantage for spoken digits, the neutral status of steady-state vowels, midway, as it were, between speech and music, is perhaps not surprising. Whether this status will be maintained by vowels placed in dynamic context remains to be seen.

#### The role of articulatory features in identification of consonants.

The set of six stop consonant syllables can be classified into groups according to features of voicing (which has two values: present, absent) and place of articulation (which has three values: labial, alveolar, velar). This breakdown is shown in Table II. Of the fifteen possible syllable pairs, six contrast in two features (voicing and place), nine contrast in one feature (three in voicing, but not in place; six in place, but not in voicing.).

The percentages of correct responses on first preferences for these feature contrasts are given for each ear in Table III. Right ear performance is the same for all three feature contrast conditions; it is only for the left ear that the conditions

differ in difficulty. The resulting ear differences show a marked trend. The advantage is greatest for the double contrast (17%), next greatest for the voicing contrast (12%), least for the place contrast (6%). The overall effect is significant by analysis of variance with  $p < 0.025$ . The difference in right ear advantage for the contrast conditions of voicing alone and place alone is not significant, but the difference between the average of these two and the double contrast is significant with  $p < 0.05$ . There is nothing in auditory psychophysics that would lead one to predict that simultaneous stimuli differing along two feature dimensions should be more difficult to identify than stimuli differing on one. This outcome supports other studies, such as those reported by Kozhevnikov and Chistovich (1965) suggesting that the perception of consonant-vowel syllables may involve some process of analysis by feature.

The effect is due, as we saw, to the greater error rate for syllables presented to the left ear when two features vary, suggesting that speech signals of this kind are more readily processed in the left hemisphere when the identification task is difficult. A detailed analysis of the errors is needed to discover precisely what the left hemisphere does better than the right. This will require a confusion matrix analysis which we are currently carrying out on a larger body of data collected on dichotically presented natural speech syllables. The analysis may throw some light on the feature system used in the perception of consonants. The system may then be compared with that used in memory as exemplified in recent experiments by Conrad (1964) and Wickelgren (1966). A masking experiment such as that of Miller and Nicely (1955) does not speak to this issue because the masking stimulus has an unequal effect on different features.

#### Conclusions.

We can summarize the main findings and implications of this

experiment as follows: (1) Relatively large and stable laterality effects occur on dichotic presentation of nonsense syllables displaying phonemic contrasts. This strongly suggests that left hemisphere dominance in speech perception operates at the level of speech sound structure. (2) The effect can be demonstrated when only a single pair of syllables is presented on each trial, indicating that it pertains to the registration of the stimuli and not only to their retention. (3) The effect is significant for synthetic consonant syllables, but not for synthetic steady-state vowels. (4) The effect is greater for consonant-vowel pairs differing on two articulatory features than for pairs differing on one. This suggests that the perception of such consonant syllables may involve a process of analysis by feature.

This work was supported in part by a grant to Haskins Laboratories from the National Institute of Child Health and Human Development.

Table I

Mean Percent Correct on First Preferences: Comparison of Consonants and Vowels: Ten Subjects.

	Left Ear	Right Ear	P
Vowels	41	45	Not significant
Consonants	31	45	< 0.001

Table II

Paired Combinations of Six Stop Consonants According to Features of Voicing and Place of Articulation.

	Place of Articulation		
	Labial	Alveolar	Velar
Unvoiced	p	t	k
Voiced	b	d	g

Pairs differing in:

<u>Voicing and Place</u>	<u>Voicing Alone</u>	<u>Place Alone</u>
p-d	p-b	p-t
p-g	t-d	p-k
t-b	k-g	t-k
t-g		b-d
k-b		b-g
k-d		d-g

Table III

Synthetic CV Syllables: First Preference  
 Mean Percent Correct for Each Ear According to Feature  
 Differences. Fifteen Subjects.

Pairs differing in:

---

<u>Voicing and Place</u>		<u>Voicing Alone</u>		<u>Place Alone</u>	
Left	Right	Left	Right	Left	Right
11	28	17	29	22	28

---

References

- Broadbent, D.E. (1954). The role of auditory localization in attention and memory span. J. exp. Psychol. 47, 191-6.
- Broadbent, D.E. and Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. Quart. J. exp. Psychol. 16, 359-60.
- Bryden, M.P. (1963). Ear preference in auditory perception. J. exp. Psychol. 65, 103-5.
- Chaney, R.B. and Webster, J.C. (1965). Information in certain multidimensional acoustic signals. U.S. Navy Electronics Laboratory Reports (San Diego, Calif.) No. 1339.
- Conrad, R. (1964). Acoustic confusions in immediate memory. Brit. J. Psychol. 55, 75-84.
- Cooper, F.S., Liberman, A.M. and Borst, J.M. (1951). The interconversion of audible and visible patterns as a basis for research in the perception of speech. Proc. Nat. Acad. Sci. 37, 318-28.
- House, A.S., Stevens, K.N. Sandel, T.T. and Arnold, J.B. (1962). On the learning of speechlike vocabularies. J. verb. Learn. Behav. 1, 133-43.

- Kimura, D. (1961a). Cerebral dominance and the perception of verbal stimuli. Canad. J. Psychol. 15, 166-71.
- Kimura, D. (1961b). Some effects of temporal-lobe damage on auditory perception. Canad. J. Psychol. 15, 156-65.
- Kimura, D. (1964). Left-right differences in the perception of melodies. Quart. J. exp. Psychol. 16, 355-8.
- Kozhevnikov, V.A. and Chistovich, L.A. (1965). Rech' Artikulatsia i Vospriatie. Moscow-Leningrad. Tr. Speech: Articulation and Perception. Joint Publications Research Service, Washington, D.C., JPRS: 30, 543.
- Lieberman, A.M., Cooper, F.S., Studdert-Kennedy, M., Harris, K.S. and Shankweiler, D.P. (1966). Some observations on the efficiency of speech sounds. Proc. XVIII Intl. Cong. Psychol. Moscow, in press.
- Miller, G.A. and Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. J. Acoust. Soc. Amer. 27, 338-52.
- Milner, B. (1962). Laterality effects in audition. In Mountcastle, V.B. (Ed.), Interhemispheric relations and cerebral dominance. Baltimore: Johns Hopkins Univ. Press.
- Rosenzweig, M.R. (1951). Representations of the two ears at the auditory cortex. Amer. J. Physiol. 167, 147-58.
- Shankweiler, D. (1966). Effects of temporal-lobe damage on perception of dichotically presented melodies. J. comp. physiol. Psychol. 62, 115-119.
- Wickelgren, W.A. (1966). Distinctive features and errors in short-term memory for English consonants. J. acoust. Soc. Amer. 39, 388-98.

## Footnotes

1. An early version of this work was presented at the 71st meeting of the Acoustical Society of America, Boston, 1-4 June 1966 [J. acoust. Soc. Amer., 39, 1256 (A) 1966]. This version is to appear in Quart. J. exp. Psychol., XIX.
2. Now at Inter American University, San German, Puerto Rico.
3. The Pattern Playback is a device for producing controlled synthetic speech stimuli: Hand-painted patterns, modelled on spectrograms, are converted photoelectrically into sound.