Reprinted from THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Vol. 48, No. 2, (Part 2), 579–594, August 1970

Received 30 October 1969

9.5, 9.7

Hemispheric Specialization for Speech Perception

MICHAEL STUDDERT-KENNEDY* AND DONALD SHANKWEILER†

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510

Earlier experiments with dichotically presented nonsense syllables had suggested that perception of the sounds of speech depends upon unilateral processors located in the cerebral hemisphere dominant for language. Our aim in this study was to pull the speech signal apart to test its components in order to determine, if possible, which aspects of the perceptual process depend upon the specific language processing machinery of the dominant hemisphere. The stimuli were spoken consonant–vowel–consonant syllables presented in dichotic pairs which contrasted in only one phone (initial stop consonant, final stop consonant, or vowel). Significant right-ear advantages were found for initial and final stop consonants, nonsignificant right-ear advantages for six medial vowels, and significant right-ear advantages for the articulatory features of voicing and place of production in stop consonants. Analysis of correct responses and errors showed that consonant features are processed independently, in agreement with earlier research employing other methods. Evidence is put forward for the view that specialization of the dominant hemisphere in speech perception is due to its possession of a linguistic device, not to specialized capacities for auditory analysis. We have concluded that, while the general auditory system common to both hemispheres is equipped to extract the auditory parameters of a speech signal, the dominant hemisphere may be specialized for the extraction of linguistic features from those parameters.

INTRODUCTION

Man is a language-using animal with skeletal structure and brain mechanisms specialized for language. For more than a century, it has been known that language functions are, to a considerable extent, unilaterally represented in one or the other of the cerebral hemispheres, most commonly the left. The evidence of cerebral lateralization and localization argues powerfully for the existence of neural machinery specialized for language, but the exact nature of the language function, and characteristics of the neural mechanisms that serve it, remain to be specified. Most studies of the neural basis of language have dealt with higher-level language functions and their dissolution. An alternative approach, which may prove more fruitful, is to investigate the lower-level language functions, that is, to focus on the production and perception of speech sounds.

Study of the evolution of the vocal tract in relation to the physiological requirements for producing the sounds of speech suggests that man has evolved special structures for speech production and has not simply appropriated existing structures designed for eating and breathing (Lieberman, 1968; Lieberman, Klatt, and Wilson, 1969). We may reasonably suppose that he has also evolved matching mechanisms for speech perception. There is, in fact, much evidence that speech perception entails peculiar processes, distinct from those of nonspeech auditory perception (for a review of the evidence, see Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). There are also grounds for believing that the sounds of speech are integral to the hierarchial structure of language (Lieberman, 1967; Mattingly and Liberman, 1970). We might, therefore, expect that among the language processes lateralized in the dominant hemisphere are mechanisms for the perception of speech. Evidence of this is not easily gathered from normal subjects with intact nervous systems. But recently a plausible technique has become available and is put to work in the present study.

Kimura (1961a), using a task similar to one described by Broadbent (1954), showed that, if pairs of contrasting digits were presented simultaneously to right and left ears, those presented to the right were more accurately reported. She attributed the effect to functional prepotency of the contralateral pathway from the right ear to language-dominant left hemisphere (Kimura, 1961b). There is evidence for stronger contralateral than ipsilateral auditory pathways in dog (Tunturi, 1946), cat (Rosenzweig, 1951; Hall and Goldstein, 1968), and man (Bocca, Calearo, Cassinari, and Migliavacca, 1955), and for inhibition of the ipsilateral signal in man during dichotic presentation (Milner, Taylor, and Sperry, 1968; Sparks and Geschwind, 1968). The right-ear advantage for verbal materials has now been repeatedly confirmed, and attempts to account for it solely in terms of memory, attention, or various response factors have been found inadequate (for reviews, see Bryden, 1967, and Satz, 1968). Kimura's attribution of the effect to cerebral dominance has received support from several other pieces of evidence. She herself (1961b) showed that the effect was reversed—a left-ear advantage appeared—in subjects known to have language dominance in the right hemisphere. She and others (Kimura, 1964; Chaney and Webster, 1965; Curry, 1967) showed that the effect was also reversed for nonspeech materials (melodies, sonar signals, environmental noises). The reversal of the effect for dichotically presented nonspeech fits with other indications that perception of auditory patterns and their attributes typically depends more upon right-hemisphere mechanisms than upon left (Milner, 1962; Spreen, Benton, and Fincham, 1965; Shankweiler, 1966a,b; Vignolo, 1969).

Kimura's contention that ear advantages in dichotic listening reflect dual cerebral asymmetries of function in perception of verbal and nonverbal materials is thus supported by much evidence from a variety of sources. Dichotic listening techniques, therefore, seem to offer a new way to raise the question of the status of speech (in the narrow sense) and its relation to language. If speech is indeed integral to language, we might expect this fact to be reflected in the neural machinery for its perception. Specifically, we may ask: Are the sounds of speech processed by the dominant hemisphere, by the minor hemisphere along with music, or equally by both hemispheres? All the dichotic speech studies referred to above used meaningful words as stimuli and therefore did not speak to this question. Studies using nonsense speech have, however, been carried out in order to discover whether the right-ear advantage depends upon the stimuli being meaningful. The results show clearly that it does not (Shankweiler and Studdert-Kennedy, 1966; Curry, 1967; Curry and Rutherford, 1967; Kimura, 1967; Kimura and Folb, 1968; Darwin, 1969; Haggard, 1969). We were therefore encouraged to make further use of dichotic listening experiments as a device for probing in some detail the processes of speech perception. Our general plan was to pull the speech signal apart and to test its components (consonants, vowels, isolated formants, and so on) in order to determine, if possible, which aspects of the perceptual process depend upon lateralized mechanisms, and by looking for information contained in perceptual errors to guess at some of the characteristics of the processing machinery.

In a study employing synthetic speech (Shankweiler and Studdert-Kennedy, 1967), we compared synthetic CV syllables and steady-state vowels. Our choice of stimuli was dictated by the repeated finding at Haskins Laboratories that the identification of stop consonants and vowels engage different perceptual processes, stop consonants being "categorically," and vowels "continuously," perceived (for discussion and summary of this evidence, see Liberman *et al.*, 1967; Lane, 1965; Studdert-Kennedy, Liberman, Harris, and Cooper, 1970). In our dichotic study of these two classes of phonemes, we found a significant right-ear advantage for the stop consonants and a small, but not significant right-ear advantage for the vowels. We also found evidence implicating the articulatory features of voicing and place of production in stop consonant perception and lateralization. The present study[1] was designed to press our analysis of speech perception further by testing the lateralization of "natural" speech rather than synthetic, of final consonants as well as initials, of vowels embedded in CVC syllables rather than steady-state, and of the consonant features of voicing and place.

I. METHOD

A. Test Construction

We wished to study dichotic effects in the perception of initial and final stop consonants followed or preceded by various vowels, and of medial vowels followed or preceded by various stop consonants. We constructed four dichotic tests: two consonant and two vowel tests. The stimuli consisted of consonant–vowel–consonant (CVC) syllables formed by pairing each of the six stop consonants, /b, d, g, p, t, k/, with each of the six vowels, /i, ε, æ, ɑ, ɔ, u/. In one consonant and one vowel test, all syllables ended with the consonant /p/ [initial-consonant-varying (IC) tests], while in the other pair of tests all syllables began with the consonant /p/ [final-consonant-varying (FC) tests].

The syllables were spoken by a phonetician. He was given two randomized lists of 36 CVC syllables (six consonants × six vowels), one with initial consonants varying, one with final consonants varying. He was asked to read each list once at an even intensity (monitored on a VU meter), and to release the final stop. His utterances were recorded, a spectrogram was made of each syllable, and its duration was measured. The durations averaged around 400 msec, with a range of about 300–500 msec. Most of the variability arose from differences in the "natural" length of the vowels and from differences in the delay of the final stop release. For some few syllables, which seemed not perfectly intelligible, the phonetician was asked to make a new recording.

As an example of test construction, we will describe the procedure for the dichotic consonant test in which the initial consonant varied. The 36 recorded syllables were dubbed several times with a two-channel tape recorder: half the syllables were assigned to one track of the tape and half to the other, so that each consonant was recorded equally often on each track. The syllables were then spliced into tape loops. Each loop carried a pair of syllables contrasting only in their initial consonants (e.g., /bap/–/dap/), one on each tape track. There were 90 such loops: each consonant was paired once with every consonant other than itself (15 combinations) followed by each of the six vowels.

The next task was to synchronize the onsets of the two syllables on a loop. This was accomplished by playing the loop on a special two-channel tape deck, modified to permit the length of leader tape passing between two playback heads to be varied, until the onsets of the two syllables coincided. Onset was defined on a permanent oscillographic record, obtained from a Honeywell 1508 Visicorder, as the first excursion above noise level that was sustained and followed by clear periodicity. Synchronization of onsets was determined from a three-channel Visicorder record, with two channels displaying the speech waves and the third a 100-Hz sine wave. Figure 1 reproduces the Visicorder record of two syllables with synchronous onsets.

Once the playback of two syllables on a loop had been synchronized, the pair was dubbed on parallel tracks using an Ampex PR-10 recorder. The input channels were matched for peak intensity on the VU meter, and the pair was recorded four times, each syllable going twice to channel 1 and twice to channel 2. In view of the arduous process of construction, this master tape of synchronized, contrasting syllables, distributed evenly over channels, was preserved uncut, as a source of stimuli in possible future experiments. From it, each syllable pair was recorded twice, once in each of its two channel orientations, on an Ampex PR-10. Thus 90 loops, made from dubbings of 36 parent recordings, yielded 180 third-generation stimuli in which each consonant was paired with every consonant other than itself followed by each of the six vowels, once on each tape track.

These stimuli were then spliced into a random order with the restriction that each consonant pair should appear once with each vowel in the first half and once with each vowel in the second half of the test. There was a 6-sec interval between stimuli, a 10-sec interval after every 10th stimulus, and a 30-sec interval after the 90th.

The IC vowel test was constructed from the original 36 recordings in exactly the same way as the IC consonant test, with the single difference that the tape loops were formed from pairs of syllables contrasting only in their vowels.

The FC consonant and vowel tests were constructed in a similar manner. Here the difference was in the alignment procedure: these syllables were synchronized at their final releases. Selecting the exact point of release on an oscillographic record proved a singularly difficult task. Many arbitrary decisions had to be made and the resulting alignments were almost certainly less precise than those of the corresponding IC pairs.

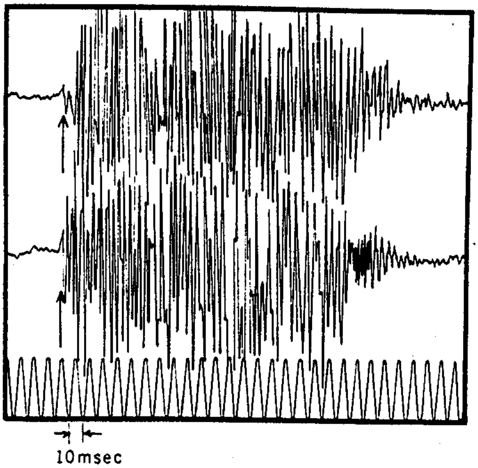

Fig. 1. Temporal alignment of syllables for dichotic presentation.

B. Subjects

There were 12 subjects: seven women and five men, aged between 18 and 26 years. Audiograms were taken separately on left and right ears. All subjects had normal hearing, considered themselves right-handed, and had no left-handed members of their immediate families. They served for four sessions of 45–50 min each and were paid for their work.

C. Procedure

Subjects took the tests individually in a quiet room, listening, over matched PDR-8 earphones, to the output of an Ampex PR-10 two-channel tape recorder.

The order in which the tests were given was counterbalanced. All subjects took a vowel test in their first and fourth sessions: half took the IC, half the FC on each occasion. All subjects took a consonant test in their second and third sessions: half of those who had taken the IC vowel test in their first session took the FC consonant test in their second and the IC consonant test in their third. The orders for the other subgroups of subjects were appropriately reversed. One subject (BZ) did not come for his final session and so gave no data on the IC vowel test.

The experimenter began a session by playing a steady-state calibrating tone (1000 Hz), spliced to the beginning of each test, on both recorder channels and adjusting the outputs to the voltage equivalent of approximately 70 dB SPL. The subject was then given the following, or analogous, instructions to read:

> This is an experiment in speech perception. You are going to listen over earphones to a series of monosyllables—consonant–vowel–consonant monosyllables, such as "pet," "bap," "doop," "pawg," and so on. They

> will be presented in *simultaneous pairs*, one to the left ear, one to the right. In any pair, the two syllables will have the same consonants, but different vowels. *The two vowels will always be different*, and will be drawn from the set of six given below.
>
> Your task is to *identify both vowels*. Opposite the appropriate trial number on your answer sheet you should write *two* of the following:
>
> ee (as in beet)
> eh (as in bet)
> ae (as in bat)
> ah (as in father)
> aw (as in bought)
> oo (as in boot)
>
> You should always write *two* vowels, *even if you have to guess*. Write them in order of confidence. That is to say, *write the one you are more sure of first, the one your are less sure of second*. There are 180 trials in the first test. You will have a short rest after 90, a longer rest after the 180. Then you will do a second test of the same length.
>
> Each batch of 90 trials takes about 10 min, and the task may not be easy. But you are asked to *give it your fullest possible attention*. Don't worry if you think you are missing a lot. Just make careful guesses, and then get ready for the next trial. There are about 6 sec between trials.
>
> Any questions? If not, put the earphones on and adjust them so that they fit comfortably on your head.

For the consonant test, the specified responses were: b, d, g, p, t, k. Appropriate changes in instructions were made for the FC tests.

Subjects wrote their responses on two 90-item response sheets, at the top of which the set of letters from which responses were to be selected was displayed. Upon completion of the 180-item test, subjects took a short rest, reversed the orientation of the earphones and took the test again. For each of the four dichotic tests, half the subjects heard channel 1 in their right ear first and half heard it in their left ear first. Channels were switched across ears by phone reversal rather than electrically so that bias due to channel and phone characteristics or phone position on the head would not be confounded with ear performance.

D. Summary

The elaborate procedure of test construction and presentation described above yielded 360 dichotic trials for each subject on each test, that is, 24 judgments on each of the 15 contrasting phoneme combinations or 60 judgments on each phoneme by each ear. Any bias due to neighboring vowel (or consonant), imprecise synchronization of onsets or offsets, recorder channels, earphone characteristics, or position of earphones on the head or sequence of testing was distributed equally over the ears of the entire group of subjects.

II. RESULTS

A. Over-All Performance

Table I summarizes the raw data and provides percentage bases for subsequent tables. Over-all performance on both ears was considerably higher for the IC vowels (82%) than for the IC consonants (68%); FC consonant performance (74%) falls midway.[2] For reasons that will become apparent (see below: an index of the laterality effect), we distinguished between trials on which both syllables were correctly identified and trials on which only one syllable was correctly identified. The distribution of total correct into the two categories is shown in the two right-hand columns of Table I. The difficulty of the IC consonant test as compared with the vowel is again shown by its lower percentage of both-correct trials (43% for consonants, 69% for vowels) and its higher percentages of one-correct trials (25% for consonants, 14% for vowels).

B. Ear Advantage

Table II presents percentage correct on the three tests, by preference and by ear, for individual subjects and for the group.

On the initial-consonant test, every subject shows a total right-ear advantage of between 4% (SB, JH) and 22% (AL). The mean total right-ear advantage of 12%

TABLE I. Over-all performance: initial consonants, medial vowels, and final consonants.

Test	No. of syllable combinations	No. of syllable presentations per ear per subject	No. of subjects	No. of syllable presentations per ear for group	No. of syllable presentations for group (both ears)	Total correct	No. correct on trials with both correct	No. correct on trials with only one correct
Initial consonants	15	360	12	4320	8640	5858 (68%)[a]	3702 (43%)	2156[b] (25%)
Medial vowels (IC)	15	360	11	3960	7920	6516 (82%)	5442 (69%)	1074[b] (14%)
Final consonants	15	360	12	4320	8640	6394 (74%)	4505 (52%)	1889 (22%)

[a] All percentages in this table are based on number of syllable presentations for group (both ears).
[b] Group percentage bases for trials on which only one syllable was correctly identified.

TABLE II. Percentage correct by preference and by ear for individual subjects.

Subject	Initial consonants						Medial vowels						Final consonants					
	1st Pref.		2nd Pref.		Total		1st Pref.		2nd Pref.		Total		1st Pref.		2nd Pref.		Total	
ear	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R
SB	42	37	25	35	68	72	42	44	35	33	77	77	40	44	27	26	67	70
JH	40	42	23	25	63	67	45	47	40	36	85	83	36	49	23	23	59	72
MJ	27	46	18	16	45	62	34	48	29	24	63	72	41	45	19	14	60	59
NK	43	45	26	29	69	74	44	44	34	33	78	77	37	44	22	25	59	69
AL	33	21	18	52	51	73	47	10	10	54	57	64	53	14	13	62	66	76
BL	23	59	43	24	67	84	64	34	28	62	92	96	37	57	46	32	83	89
LN	21	65	45	17	65	82	40	59	58	40	98	99	32	62	42	27	74	89
HW	34	44	24	23	58	67	50	46	40	47	90	93	55	38	18	32	73	70
JW	30	50	23	18	53	68	32	37	24	23	56	60	37	54	28	20	65	74
BZ	25	57	42	26	67	83	—	—	—	—	—	—	41	52	43	36	84	88
SZ	33	52	38	28	71	81	58	42	40	57	98	99	37	55	43	30	80	85
JWn	28	53	34	21	62	74	45	54	52	42	97	96	39	52	41	31	81	83
Mean	32	48	30	26	62	74	46	42	35	41	81	83	40	47	31	30	71	77
$\bar{R}-\bar{L}$	16		−4		12		−4		6		2		7		−1		6	

is significant on a two-tailed matched pairs t-test ($t=7.19$, $p<0.001$).

For the final consonants, right-ear advantages are smaller and more variable. Ten subjects show a total right-ear advantage of between 2% (JWn) and 15% (LN). Two subjects (MJ, HW) show left-ear advantages of 1% and 3%, respectively. The mean total right-ear advantage of 6% is significant on a two-tailed matched pairs t-test ($t=3.84$, $p<0.01$).

The vowel results are again variable. Seven subjects show right-ear advantages, three (JH, NK, JWn) show small left-ear advantages and one (SB) shows no advantage. The mean total right-ear advantage of 2% falls short of significance on a two-tailed test at the 0.05 level ($t=2.16$, $p<0.06$).

Over-all performance is higher on first preferences than on second for all three tests and, for both initial and final consonants, the total right-ear advantage is derived from first preferences (although some subjects—SB and AL on initials, HW and AL on finals—show their larger ear advantage on second preferences). That the right-ear advantage on consonants does not arise from a general tendency to report the right ear first, while the left-ear signal decays in storage, is shown by the fact that the ear advantage on first preferences for the vowels is to the left. Furthermore, the higher over-all performance on first preferences is due almost entirely to the right ear on initial consonants, to the left ear on vowels.[3] The tendency to attach greater confidence to correct responses combined with the relatively large number of trials on which both responses were correct leads to nonsignificant reversals of the consonant-ear advantages on second preferences.

C. An Index of the Laterality Effect

The laterality effect has been shown to be a function, under certain circumstances, of task difficulty (Satz, Achenbach, Pattishall, and Fennell, 1965; Bartz, Satz, and Fennell, 1967; Satz, 1968), and a ceiling is necessarily imposed upon it by very high or very low over-all performance (Halwes, 1969). Since the vowels evidently set the listeners an easier task than the consonants, we sought a method of data analysis by which the two levels of difficulty might be equated. We found this in trials on which only one of the syllables was correctly identified. All such trials are presumably, in some sense, of equal difficulty, and over-all performance on the subset is necessarily equal (50%) for consonants and vowels. No ear advantage can, in any event, be detected on trials for which the syllables are either both correct,[4] or both incorrect, so that restriction of a laterality measure to the trials on which only one syllable was correctly identified (see Table I, last column) confines attention to the only occasions on which the effect has an opportunity to appear. Our null hypothesis for these one-correct trials is then that the single correct syllables are identified equally often by right and left ears. Deviation from this 50–50 distribution may be expressed as a percentage: $(R—L)/(R+L)\ 100$, where R (or L) is the number of trials on which the correctly identified syllable was delivered to the right (or left) ear. The index will range from 0 (50–50 distribution) to ±100 (0–100 distribution), with negative values indicating a left-ear advantage, positive values a right-ear advantage. Its significance may be tested on the null hypothesis that $R/(R+L)=L/(R+L)=0.50$, using the normal curve as an approximation to the binomial.

Table III presents values of this index, based on one-correct-only trials, for individual subjects, on initial consonants, final consonants, and vowels. For initial consonants, the mean-percentage laterality effect is 26. Each subject contributes between 150 and 208 trials. For nine subjects, the index is positive and significant; for three subjects (SB, JH, NK), the index is positive, but not significant.

For final consonants, the mean percentage laterality effect is 17. Each subject contributes between 89 and 237 trials. For seven subjects, the index is positive and

TABLE III. Individual percentage ear advantages for initial stop consonants, final stop consonants, and medial vowels based on trials containing only one correct response.

	Initial consonants				Medial vowels				Final consonants			
Subject	R−L[a]	R+L	$\frac{R-L}{R+L}100$	P	R−L	R+L	$\frac{R-L}{R+L}100$	P	R−L	R+L	$\frac{R-L}{R+L}100$	P
SB	15	171	9	NS[b]	2	134	1	NS	10	182	5	NS
JH	18	200	9	NS	−5	99	−5	NS	54	196	28	<0.0001
MJ	62	208	30	<0.0001	33	191	17	<0.02	−3	237	−1	NS
NK	20	178	11	NS	−1	143	−1	NS	37	205	18	<0.01
AL	94	204	46	<0.0001	26	192	14	<0.06	37	177	21	<0.01
BL	62	150	41	<0.0001	4	32	12	NS	21	89	24	<0.05
LN	58	178	33	<0.0001	4	8	50	NS	52	122	43	<0.0001
HW	32	186	17	<0.05	8	50	16	NS	−10	184	−5	NS
JW	55	207	27	<0.0001	13	191	7	NS	32	188	17	<0.05
BZ	55	153	36	<0.0001	—	—	—	—	15	95	16	NS
SZ	36	156	23	<0.01	1	9	0	NS	20	106	19	<0.06
JWn	43	165	26	<0.001	−1	25	−4	NS	12	108	11	NS
Total	550	2156			84	1074			277	1889		
			Mean 26				Mean 10				Mean 17	

[a] R = Number of trials on which only the right-ear stimulus was correctly identified. L = Number of trials on which only the left-ear stimulus was correctly identified.

[b] NS = Not significant at 0.10 level.

significant; for three subjects (SB, BZ, JWn), the index is positive but not significant; for two subjects (MJ, HW), the index is negative and not significant.

For the vowels, the mean-percentage laterality effect is 10, but the reliability of this is low. Subjects vary widely in their indices and in their numbers of one-correct trials. Subject LN, for example, has an index of 50, based on only 8 trials, subject NK an index of −1 based on 143 trials, subject MJ an index of 17 based on 191 trials. For only two subjects (MJ, AL) is the index significant.

D. Laterality Effect for Individual Stop Consonants and Vowels

Up to this point, we have treated stop consonants and vowels as undifferentiated classes. But do all members of these classes show a laterality effect of the same degree? To answer this question, the group data were broken down by phonemes, and the laterality index was computed for each consonant and vowel. Figure 2 presents the results. The indices are arranged from left to right in order of decreasing magnitude. Consonants and vowels are perfectly segregated by this arrangement. /b/ and /g/ have the highest indices, and the voiced consonant at a given place value is always higher than its unvoiced counterpart. But the right-ear advantage is present for the whole class of initial stop consonants, and all indices are significant with $p<0.0001$: lateralization is strong and consistent. For the vowels, on the other hand, lateralization is weak and inconsistent: all indices are positive, but only one (for /i/) is significant with $p<0.01$, and one (for /æ/) with $p<0.10$.

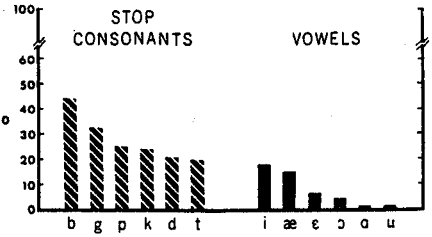

FIG. 2. The right-ear advantage for individual stop consonants and vowels on single-error trials. For explanation of the index plotted against the ordinate, see text.

E. Laterality Effect and Item Difficulty

We eliminated task difficulty as a variable affecting the apparent lateralization of consonants and vowels by analyzing one-correct trials only. But it would still be possible for differences in the lateralization of individual phonemes on these trials to be linked to item difficulty. Consonants were therefore ranked according to difficulty, measured by total number of errors (order: /k, b, t, g, p, d/) and the value of their indices (order: /b, g, p, k, d, t/). Kendall's tau (Siegel, 1956) was computed and gave a nonsignificant value of 0.20. Vowels ranked according to their levels of difficulty (/æ, ɔ, ɑ, u, ɛ, i/) and indices (/i, æ, ɛ, ɔ, ɑ, u/) yielded a nonsignificant tau of −0.13. There is, therefore, no evidence here for a relation between the observed laterality effect and item difficulty.

F. Identification of Consonant Feature Values

Having found that each of the six stop consonants is significantly lateralized, we may now ask whether the same is true of the articulatory features of which they are composed. Logically prior to this, however, is the question of whether these features are even perceived. Their psychological validity is, in fact, attested by the results of scaling the perceived distances among the stop consonants, /b, d, g, p, t, k/ (Greenberg and Jenkins, 1964), and analyses of errors in perception and short-term memory have suggested that the features are separately extracted and stored (Miller and Nicely, 1955; Singh, 1966, 1969; Wickelgren, 1966; Klatt, 1968). Experiments with dichotic listening offer a new approach to study of the perceptual process.

Each of the six stop consonants may be specified in terms of two articulatory features: voicing and place of production. In English, place of production has three values (labial, alveolar, velar), while voicing has only two (voiced, voiceless), so that we can specify each of the stops uniquely within a 2×3 matrix. The dichotic pairs may then contrast in voicing (/b, p/, /d, t/, /g, k/), in place (/b, d/, /b, g/, /d, g/, /p, t/, /p, k/, /t, k), or in voicing and place (/b, t/, /b, k/, /d, p/, /d, k/, /g, p/, /g, t/). In each of these three blocks of trials, each consonant occurs equally often at each ear. If consonants are perceptually irreducible wholes and their component features no more than useful descriptive devices, we would expect performance to display only chance variation across blocks of trials for which articulatory features were the basis of classification. But, in fact, we find, as in our earlier experiment (Shankweiler and Studdert-Kennedy, 1967b), that performance does vary significantly. Table IV shows that when a feature value is common to both ears (that is, when the dichotic pair contrasts in only one feature), an error is less likely to be made and both responses are more likely to be correct than when no feature value is common (that is, when the dichotic pair provides a double contrast, a contrast in both voicing and place). Furthermore, performance varies according to which feature is shared: more advantage accrues from shared place than from shared voicing.[5] Or, in opposite terms, the feature more adversely affected by conditions of dichotic competition is place: even when voicing is shared, the contrast in place depresses performance. The outcome confirms the perceptual reality of the features: voicing and place values are indeed separately extracted.

The same conclusion is suggested by an analysis of errors. Even if a consonant is wrongly identified, one of its feature values may be correctly identified and appropriate analysis will permit inferences about the perceptual process. The analysis is confined to trials on which a single error was made, since it is only for these that we can assign an error to its ear and stimulus. To ensure that no differential advantage accrues through a shared feature value, the analysis is also confined to trials on which each ear receives a different value of both voicing and place, that is, to double contrast trials. For these trials we may then determine the frequency with which each feature was correctly identified on erroneous responses and we may compare this frequency with that expected by chance. To make the procedure clear, suppose that the stimulus pair is /b, t/ and that the subject correctly identifies /b/, so that we know his error is on /t/. His erroneous response may then be correct on voicing (/p/ or /k/), correct on place (/d/), or correct on neither feature (/g/). Correct guesses, if made on the perceptually unanalyzed phonemes without regard to their component features, would then be distributed in the proportions 2:1:1 for voicing, place, and neither feature correct. Table V shows that, in fact, voicing alone is correctly identified an overwhelmingly large proportion of times. Chi-square for this table equals 200.34, which, with 2 degrees of freedom (df), is highly significant ($p<0.001$).

TABLE IV. Percentage of different trial outcomes as a function of feature composition of dichotic pairs.

Feature having a value shared by the dichotic pair	Trial outcomes (percent) Both correct	One correct	Neither correct
Place	61	37	2
Voice	43	52	5
Neither	33	55	12

TABLE V. Number and percentage of features correct on single-error responses in double-contrast trials.

Feature correct	Number	Percent
Voice alone	678	72
Place alone	184	19
Neither	83	9
Total	945	100

We may be confident, then, that the features are separately processed, and that voicing values are more accurately identified than place. But some advantage may yet accrue to the identification of one feature from the correct identification of the other. In other words, the two perceptual processes may be at least partially dependent. The degree of their independence may be estimated by combining correct responses and errors into a single confusion matrix and carrying out an information analysis (Miller and Nicely, 1955; Attneave, 1959). The procedure has the additional advantage of providing a comparison between voicing and place identification in which the unequal guessing probabilities for the two features may be discounted by expressing, for each feature, the information transmitted as a percentage of the maximum possible transmitted information.

Three confusion matrices were therefore constructed: a 2×2 voicing matrix in which stimuli and responses were grouped into voiced and voiceless; a 3×3 place matrix in which stimuli and responses were grouped into labial, alveolar, and velar; and a 6×6 matrix for the six individual consonants. Entries into these tables could use only those trials on which at least one phoneme was correctly perceived, since, when neither phoneme is correct, the erroneous responses cannot be assigned to their appropriate stimuli. This has two consequences for the analysis. First, since all double errors are excluded, it leads to an overestimate of the

TABLE VI. Information in bits, and percentage of maximum possible information transmitted for each feature separately, and for the features combined in individual consonants.

	Absolute amount of information transmitted in bits				Percentage of maximum possible information transmitted			
	Voice	Place	(V+P)	Combined	Voice	Place	$\left\{\frac{V+P}{2}\right\}$	Combined
	0.38	0.41	(0.79)	0.86	38	26	(32)	33
Maximum possible	1	1.58		2.58				

transmitted information for the experiment as a whole. But, since the purpose of the analysis is to compare the features and to estimate their degree of independence rather than to make a reliable estimate of information transmission, this need not concern us. A second consequence is that not all phonemes, or classes of phonemes, are equally represented in the trials to be analyzed, so that the presented information (and hence the possible transmitted information) is reduced from the value that it would have if the sample were representative of the whole set of stimuli. However, the reduction in presented information proved to be only a few thousandths of a bit for each matrix, so that maximum possible transmitted information remained effectively 1 bit on voicing, 1.58 bits on place, and 2.58 bits on the individual consonants.

The actual information transmitted was computed for each matrix and the results are displayed on the left side of Table VI. If the features of voicing and place were independently identified, the sum of the information transmitted for voicing and place separately would equal the information transmitted for the individual consonants in which the two features are combined (McGill, 1954; Miller and Nicely, 1955). Table VI shows that the required additivity holds to a close approximation. The independent perception of these features, demonstrated by previous investigators (Miller and Nicely, 1955; Singh, 1966) is again confirmed.

Table VI (right side) also expresses information transmitted as a percentage of maximum possible information transmitted on the two features, thus correcting for their unequal guessing probabilities. We again see the superiority of voicing over place identification: 12% more of the available voicing information is transmitted than of the available place information.

TABLE VII. Percentage correct responses on each feature value for trials with at least one correct response.

Feature	Value	Percent correct
Place	Labial	64
	Alveolar	82
	Velar	63
Voicing	Voiced	85
	Voiceless	83

The general superiority of voicing over place identification, shown by the three data analyses described above, may not, of course, hold for all feature values. As a rough test for the homogeneity of the effect, we can compute the percentage correct on each feature value for all trials having at least one correct response (double-error trials again being excluded since responses on these trials cannot be assigned to their stimuli). Table VII shows the results of these computations. There is little difference between performance on the labial and velar place values: both are some 20% lower than performances on either of the two voicing values. The joker in the set is the alveolar performance of 82%, suggesting that perception of this place value is no more affected by dichotic stress than is perception of voicing. However, the results must be viewed with caution, since the data reveal a heavy bias toward alveolar responses: 42% of all place responses on these trials were alveolar, as compared with 29% each for labial and velar responses. A similar, though much smaller, bias appears in the data of Miller and Nicely (1955, Table XVIII) for the set of six stop consonants.

The bias probably does not reflect listeners' expectations based on their experience with the language. Even though Denes (1963) estimates alveolar stop consonants to be roughly three times as frequent in English as either labial or velar stops, he also estimates voiceless stops to be very nearly twice as frequent as voiced, and no corresponding bias appears in our data (if anything, the reverse: 53% of listeners' responses on these trials were voiced, 47% voiceless). Furthermore, analysis of errors shows that most alveolar responses are made on trials in which at least one of the stimuli carries the alveolar place value. The "bias" therefore arises when one member of a dichotic pair is alveolar and the other is not: the alveolar value then "dominates" the contrasting labial or velar value. In other words, our first inference seems to be correct: the "bias" has a perceptual basis, and the alveolar stops in this experiment were less susceptible to dichotic stress than labial or velar stops.

G. Lateralization of Feature Perception

We may now ask whether the independence of the two features and the advantage of voicing over place shown in the combined data, holds equally for the two

TABLE VIII. Percentage correct responses for the two ears as a function of feature composition of dichotic pairs.

Feature having a value shared by the dichotic pair	Percent correct Left ear	Percent correct Right ear
Place	74	86
Voice	63	75
Neither	54	67

TABLE IX. Conditional percentages of feature errors for the two ears on single-error responses in double-contrast trials.

Feature in error	Other feature	Percent Left	Percent Right
Place	Voicing correct	86	93
Place	Voicing incorrect	14	7
Voicing	Place correct	67	73
Voicing	Place incorrect	33	27

TABLE X. Percentage correct responses on each feature value for each ear on trials with at least one correct response.

Feature	Value	Percent correct Left ear	Percent correct Right ear
Place	Labial	59	71
	Alveolar	79	84
	Velar	58	68
Voicing	Voiced	82	89
	Voiceless	80	87

ears. To answer these questions, the data were reanalyzed separately for each ear. We begin with a reanalysis of Table IV. The results are now given in terms of percentage of correct responses for each ear rather than in terms of trial outcomes, since no difference between the ears can appear on trials for which the responses were either both correct or both incorrect. Table VIII shows the outcome of the reanalysis. For both ears the ranking is exactly as in Table IV: performance is highest when place is shared, second highest when voicing is shared, and lowest when neither feature is shared.

We may notice, furthermore, that the right ear has approximately the same advantage over the left ear (about 12%) for each type of dichotic pair. This suggests that the right-ear advantage is the same for both voicing and place—that one feature is not more heavily lateralized than the other. The same conclusion is suggested by an error analysis along the lines of Table V. Again we make use only of double-contrast trials, and, to avoid any bias due to possible interaction between the features (despite their evident independence), we compute for each ear conditional percentages: that is, we compute the percentage correct on voicing, given that place was missed, and the percentage correct on place, given that voicing was missed. Table IX gives the results of these computations: the right-ear advantage is 7% on voicing, 6% on place.

However, equal lateralization of the two features is not evident in every analysis. Table X shows the breakdown of Table VII by ear. The expected right-ear advantage appears for every value of both features, but is somewhat greater for labial and velar place values than for voicing, suggesting stronger lateralization of these place values. [Both ears, incidentally, show a gain in alveolar performance: for the left ear the gain is approximately 20% as against 13%–16% for the right ear, perhaps reflecting a somewhat stronger alveolar preference on the left ear (44% of all left-ear responses, as against 39% of all right ear responses, were alveolar)].

Finally, Table XI displays the results of the information analysis. Both ears transmit a greater percentage of their voicing than of their place information. And for both ears the expected additivity, or independence, of feature information holds quite closely. However, the right-ear advantage is here greater on voicing (18%) than on place (10%). The difference cannot be tested for significance, but the disagreements between Tables VIII and IX (features equivalent in lateralization), Table X (right-ear advantage greater on two place values), and Table XI (right-ear advantage greater on voicing) are obvious.

There is also disagreement between one particular analysis in this and in our earlier study. In that study, we found differing degrees of laterality effect according to which features were shared (or contrasted) between the ears in a dichotic pair. We took this to indicate some

TABLE XI. Information in bits and percentage of maximum possible information transmitted for each feature separately and for the features combined in individual consonants, for right and left ears.

	Absolute amount of information transmitted in bits				Percentage of maximum possible information transmitted			
	Voice	Place	(V+P)	Combined	Voice	Place	$\left\{\frac{V+P}{2}\right\}$	Combined
Right ear	0.49	0.50	(0.99)	1.06	49	32	40	41
Left ear	0.31	0.35	(0.66)	0.70	31	22	26	27
Maximum possible	1	1.58	2.58					

difference in the degrees of lateralization of the two features. But in the corresponding analysis of the present study (Table VIII) we found no differences in laterality effect.

We therefore conclude that, while both features are clearly and independently lateralized, reliable estimates of their relative degrees of lateralization have eluded us.

III. DISCUSSION

The results are in general agreement with those of our previous study and of several other investigators (Curry, 1967; Curry and Rutherford, 1967; Kimura, 1967; Darwin, 1969a,b; Haggard, 1969; Halwes, 1969), in demonstrating a laterality effect for the perception of dichotic signals that differ only in their phonetic structure. They show further that the laterality effect extends to the perception of subphonemic features. Before discussing some of the problems that the results present, we briefly consider a possible mechanism of speech lateralization.

A. Mechanism for the Laterality Effect in Speech Perception

As Kimura (1961b, 1964) first suggested, the laterality effect may be accounted for by the assumptions of cerebral dominance and functional prepotency of the contralateral over the ipsilateral auditory pathways. Contralateral prepotency rests upon the greater number of these neurons and upon inhibition of ipsilateral neurons during dichotic stimulation. Strong corroboration of Kimura's argument has come from the work of Milner, Taylor, and Sperry (1968). (See also Sparks and Geschwind, 1968.) They studied right-handed patients (presumably left-brained for language) for whom the main commissures linking the cerebral hemispheres had been sectioned to relieve epilepsy. Under dichotic stimulation, these subjects were able to report verbal stimuli presented to the right ear, but not those presented to the left; under monaurual stimulation, they performed equally well with the two ears. Milner *et al.* attribute their results to suppression of the ipsilateral pathway from left ear to left (language) hemisphere during dichotic stimulation and, of course, to sectioning of the callosal pathway that should have carried the left-ear input from right hemisphere to left. Their data justify the inference that, when under dichotic stimulation normal left-brained subjects correctly perceive a left-ear verbal input, the signal has been suppressed ipsilaterally, has traveled the contralateral path to the right hemisphere, and has been transferred across the lateral commissures to the left hemisphere for processing. Inputs to both ears therefore converge on the dominant hemisphere, that from the right ear by the direct contralateral path and that from the left ear by an indirect path, crossing first to the right hemisphere, then laterally to the left. The right-ear advantage in dichotic studies of speech must then arise because the left-ear input, traveling an indirect path to the left cerebral hemisphere suffers, on certain trials, a disadvantage or "loss" to which the right-ear input, traveling a direct path, is less susceptible.

The locus of this loss can be broadly specified. We first assume that the two contralateral pathways are equivalent, so that the two signals reach their respective hemispheres in equivalent states; there is, of course, ample opportunity for the signals to interact at subcortical levels, but presumably whatever loss such interaction may induce is induced equally on both signals. If we further assume that the two signals upon arrival in the dominant hemisphere are served by the same set of processors (as evidence, discussed below, suggests), loss in the left-ear signal must occur immediately before, during, or after transfer to the dominant hemisphere.

The nature and source of the left-ear loss are matters of great interest to which we return briefly in a later section of the discussion. Here we merely remark that a preliminary attack on the problem might be made through careful comparison of error patterns for right- and left-ear inputs. As we have seen, in the limited data of the present study the general pattern of errors is rather similar for the two ears. This suggests that the left-ear input is subject to stress that differs in degree, but not in kind, from that exerted on the right-ear input. The notion of a generalized auditory stress common to both ears, whatever its source, is encouraged by the fact that the error pattern in this experiment is remarkably similar to that found in other studies. The superiority of voicing identification over place, for example, was observed by Miller and Nicely (1955) and by Singh (1966) in studies of speech perception through masking noise.

B. Nature of Cerebral Dominance in Speech Perception

To speak of cerebral dominance in speech perception is to imply that at least some portion of the perceptual function is performed more efficiently, or even exclusively, by the dominant hemisphere. The problem is to define that portion. That dichotic inputs must, at some point in their time course, converge on a final common path is evident from the fact that the two inputs ultimately activate a single articulatory response mechanism. But how early the inputs converge is the matter of interest. We would like to know, for example, whether convergence occurs before any linguistic analysis of the signal whatever (as would be true if both ears were served by a single set of specialized speech processors in the speech-dominant hemisphere), after partial linguistic analysis (as would be true if, for example, features were separately extracted in the two hemispheres, but were recombined in the dominant hemisphere), or after complete linguistic analysis and immediately before response (as would be true if the two hemispheres were equivalent in their capacities to

TABLE XII. Number and percentage of errors on double-contrast trials that arose by blending or not blending features from opposite ears. Trials affording two errors and trials affording one error are distinguished.

Trial outcome	Number of "blend" errors	Number of "nonblend" errors	Total number of errors	Percent "blend" errors
Double error	263	147	410	64
Single error	673	272	945	71
Total	936	419	1355	69

analyze the signal, but were served by a single set of specialized output mechanisms in the speech-dominant hemisphere). More generally, is the signal from the nondominant hemisphere transferred to the dominant hemisphere in a linguistic or in an auditory code? Some leverage on this question may be gained from a further analysis of errors in the present study.

Independent processing of subphonemic features requires that, at some point between input and output, a syllable be broken into its component features and that, at some later point, these features be recombined into a unitary response. If convergence of the two inputs occurs before features are recombined, a feature value has an opportunity to lose its local sign, that is, to lose information about its ear of origin. A correctly perceived feature from one ear might then be incorrectly combined with a correctly perceived feature from the opposite ear. The resulting response would be a "blend" of features from opposite ears. However, if convergence of the two inputs occurs *after* features are recombined, local sign could only be lost for the entire syllable, not for its component features. Blend responses would then occur only by chance. Evidence for greater than chance occurrence of blends is therefore evidence for loss of local sign on features and, by inference, for convergence of the inputs before the features are recombined.

Blends cannot be detected on single-contrast trials: even if the error occurs in combining the features, any resulting response will be correct, since one of the crossed feature values is presented to both ears. But on double-contrast trials, blending errors may be detected. For example, if the stimulus pair is /b, t/, the erroneous responses /p/ or /d/ are blends (drawing place values from one ear and voicing values from the other), while the erroneous responses /g/ and /k/ are not blends. Both classes of error would occur equally often, if there were no tendency for errors of local sign to occur on the features and if subjects were distributing their errors at random. In fact, blending errors occur with high frequency. Table XII shows that, of 410 errors on double-error double-contrast trials, 263 (64%) were blends; of 945 errors on single-error double-contrast trials, 673 (71%) were blends. The over-all percentage of blends (69%) is far in excess of chance expectation (50%). For each row of the table, $p<0.0001$ on a test of the chance hypothesis by the normal approximation to the binomial.

Errors of local sign on the features do then occur in these data, as in those of Kirstein and Shankweiler (1969), with very high frequency. The result is additional evidence for the independent processing of the features. More importantly, it suggests that inputs to left and right ears converge on a common center at some stage *before* combination of the features into a final unitary response.

We may now ask whether convergence occurs immediately before feature combination or at some earlier stage. In other words, is the signal that is transferred from right hemisphere to left coded into separate linguistic features or is it in some form of nonlinguistic auditory code? If the first were true, features of the left-ear syllable and features of the right-ear syllable would be extracted in separate hemispheres, and the feature composition of one syllable should have no effect on the probability of correctly identifying the other. If the second were true, interaction could occur between auditory parameters of the two inputs during the process of feature extraction, and this interaction should be reflected in performance. In fact, we already know from Tables IV and VIII that a response is more likely to be correct if the two inputs have a feature value in common. Furthermore, the advantage of sharing a feature value accrues more frequently if place is shared than if voicing is shared. We conclude that the inputs converge before rather than after feature extraction, and that duplication of the auditory information conveying the shared feature value gives rise to the observed advantage. In other words, we take the systematic relation between performance and the feature composition of dichotic pairs to be evidence consistent with the hypothesis of interaction during, or immediately before, the actual process of feature extraction.

Also consistent with this interpretation are the similar error patterns for left and right ears that we have already reported. As a further example, Table XIII shows the breakdown of Table XII by ear. (Only single-error trials are considered, since double errors cannot be assigned to their ears. An example of a single-error "blend" would be the response /d/ in the response pair /b, d/, given to stimulus pair /b, t/.) While the percentage of "blend" errors is greater for the right ear (75%) than for the left (69%), the difference is not significant at the 0.05 level, and both ears

TABLE XIII. Number and percentage of errors on double-contrast trials that arose by blending or not blending features from opposite ears, for right and left ears. Single-error trials only.

Ear	Number of "blend" errors	Number of "nonblend" errors	Total number of errors	Percent "blend" errors
Right	268	91	359	75
Left	405	181	586	69
Total	673	272	945	71

show a heavy preponderance of "blend" over "nonblend" errors.

We therefore tentatively conclude that convergence of the two signals in the dominant hemisphere occurs before the extraction of linguistic features, and that it is for this process of feature extraction that the dominant hemisphere is specialized. On this hypothesis, we would assign to the dominant hemisphere that portion of the perceptual process which is truly linguistic: the separation and sorting of a complex of auditory parameters into phonological features. Such a specialized "decoding" operation has been shown, on quite other grounds, to be entailed in speech perception (Liberman *et al.*, 1967).

C. Role of the General Auditory System in Speech Perception

The foregoing argument has suggested that the role of the dominant hemisphere is due to its possession of a special linguistic device rather than to superior capacities for auditory analysis. We should therefore emphasize the distinction between extraction of the auditory parameters of speech and linguistic "interpretation" of those parameters. It is for the latter that specialized processing is required and for which the dominant hemisphere seems to be equipped, while the former is the domain of the general auditory system common to both hemispheres. In other words, the peculiarity of speech may lie not so much in its acoustic structure as in the phonological information that this structure conveys. There is therefore no *a priori* reason to expect that specialization of the speech perceptual process should extend to the mechanisms by which the acoustic parameters of speech are extracted.

Consider, for example, an acoustic variable underlying the identification of place in stop consonants: the extent and direction of the second formant transition (Liberman, Delattre, Cooper, and Gerstman, 1954). Data bearing on the perception of such frequency transitions in nonspeech have been reported for resonant frequencies (Brady, House, and Stevens, 1961) and, more recently, for tone bursts (Pollack, 1968; Nabelek and Hirsh, 1969). Nabelek and Hirsh determined the optimal glide durations for the discrimination of frequency change to be, in general, between 20 and 30 msec. They remark that these values are "close to the durations that were found by Liberman, Delattre, Gerstman, and Cooper (1956) to be important for the discrimination of speech sounds" (p. 1518). They conclude that this optimum transition duration "is a general property of hearing and . . . does not only appear in connection with speech sounds" (p. 1518).

Their conclusion does not, of course, imply that there may be no functional differences between the hemispheres in auditory perception. There is, in fact, much evidence that for nonspeech the right nondominant hemisphere plays a greater role than the left in recognition of auditory patterns and in discrimination of their attributes (Milner, 1962; Kimura, 1964; Benton, 1965; Chaney and Webster, 1965; Shankweiler, 1966a,b; Curry, 1967; Vignolo, 1969). But whatever the peculiar auditory capabilities of the right hemisphere may be, there is reason to believe that each hemisphere can perform an auditory pattern analysis of the speech signal without aid from the other. The isolated left hemisphere can, in fact, go further and complete the perceptual process by interpretation of these auditory patterns as sets of linguistic features (as the data of Milner *et al.* cited in Sec. III-A, show).

Whether the right hemisphere can go so far is open to question. Sperry and Gazzaniga (1967) (see also Smith and Burkland, 1966; Gazzaniga and Sperry, 1967; Sparks and Geschwind, 1968) found that commissurectomized patients, instructed orally to select an object from a concealed tray with the left hand, were able to do so. Since left-hand stereognostic discrimination was known, from other of their tests, to be controlled only by the right hemisphere, it was evident that this hemisphere, in some sense, "perceived" the speech. However, the hemisphere was unaware of what it had "heard": the patients were unable to name the object they had selected and were holding. Similar results have been reported by Milner *et al.* (1968) for commissurectomized patients to whom instructions had been presented dichotically, thus presumably confining left-hand instructions to the right hemisphere. These authors conclude that "the minor, right hemisphere does show some rudimentary verbal comprehension" (p. 184).

Interpretation of such results is not easy, particularly since these patients had pre-existing epileptogenic lesions in addition to surgical disconnection of the hemispheres. However, it seems possible that the right hemisphere's "rudimentary comprehension" rested on auditory analysis which, by repeated association with the outcome of subsequent linguistic processing, had come to control simple discriminative responses. Certainly, a capacity for the auditory analysis of speech would seem to be the least we can attribute to the right hemisphere.

We therefore conclude that the auditory system common to both hemispheres is probably equipped to track formants, register temporal intervals, and in general extract the auditory parameters of speech. But to the dominant hemisphere may be largely reserved the tasks of linguistic interpretation: for example, selecting from a formant transition the relevant overlapping cues to consonantal place of articulation and to neighboring vowel, or selecting from the infinity of temporal intervals automatically registered in the auditory stream the one interval relevant to the perception of voicing (Lisker and Abramson, 1964; Abramson and Lisker, 1965). Completion of such tasks is presumably prerequisite to conscious perception of speech.

The interpretation of the laterality effect outlined in preceding sections has implications for future work that may best be drawn by first discussing the results for consonants and vowels in the present study.

D. Consonant-Feature Lateralization

Underlying lateralization of consonants are the independent lateralizations of their component features. Since the bulk of consonantal errors is due to the loss of a single feature (see Tables V and IX), any reduction in the laterality effect of one feature would lead to a reduction in the laterality effect of the consonants as a whole. An example of such an effect may have been provided by the final consonants of this study.

The right-ear advantage for the final consonants, though significant, was relatively small. The result is at variance with that of Darwin (1969a,b), who found a strong right-ear advantage for final consonants in dichotically presented synthetic VC syllables.[6] If we accept the difference as genuine and not due to some artifact such as poor synchronization of the final consonants in this study, an interesting explanation might be that our reduced effect arose from reduced place lateralization, and that place lateralization only occurs for cues carried by a formant transition. A formant transition was the sole source of cues in the unreleased synthetic stops used by Darwin, but not in the released "natural" speech stops of the present study, where final bursts may sometimes have provided enough information for clear place identification.

The implication, in light of our previous argument, is that a final burst, standing in relative isolation from the rest of the syllable, may be estimated as well by the minor as by the major hemisphere and that information about its parameters (intensity, duration, frequency band) is liable to relatively little loss during transfer to the dominant hemisphere for feature extraction. A formant transition, on the other hand, in which cues for both vowel and consonant are delicately implicated, even if correctly estimated auditorily by the minor hemisphere, may be subject to degradation during transfer to the dominant hemisphere. The presence of a formant transition was found by Darwin (1969a,b) in an experiment with synthetic (initial) fricatives (/f, s, ʃ, v, z, ʒ/ followed by /ɛp/) to be a necessary condition of right-ear advantage: fricatives synthetized from friction alone, without transition, were clearly identifiable, but gave no right-ear advantage. The likely importance of formant transitions in the laterality effect may also bear on the results for the vowels to which we now turn.

E. Vowel Lateralization

A main purpose of the present study was to determine whether natural vowels embedded in a consonantal frame would show a greater right-ear advantage than the synthetic, isolated, steady-state vowels of our previous study. They did not. Nonetheless, some tendency toward a right-ear advantage for the vowels is evident. In both studies, the mean advantage, though not significant, was to the right (4%, 2%). Of the 21 subjects in the two studies, 13 gave right-ear advantages (two significant), seven gave left-ear advantages (none significant), and one gave no ear advantage. For the six vowels in the present study, all ear advantages were to the right (one significant). In short, the vowels display a weak, variable, right-ear advantage, and by this are distinguished from consonants for which a stronger right-ear advantage is the rule, and also from musical or other nonspeech sounds for which a left-ear advantage is the rule (Kimura, 1964; Shankweiler, 1966a,b; Chaney and Webster, 1965; Curry, 1967).

The vowels studied up till now seem to occupy a position on the margin of speech. But we should note that the vowels of this experiment, though embedded in CVC syllables, were still of relatively long duration, each syllable lasting between 300 and 500 msec. Presumably, were they synthetic, we could push them (or isolated steady-state vowels) toward nonspeech and a left-ear advantage by systematic manipulation of their spectral composition, musicalizing them, perhaps, by reducing the bandwidths of their formants and increasing their duration. But under what conditions might the tentative right-ear advantage be magnified into a full right-ear advantage comparable with that of the consonants?

If the vowels are isolated and steady-state, merely reducing their duration from 150 to 40 msec has no effect: neither the longer nor the shorter vowels show a significant ear advantage (Darwin, 1969a,b), and reduction of duration much below 40 msec is not possible without loss of vowel quality and approach to a nonspeech click. But for vowels placed in CVC syllables the story may be different. We know that the identification of synthetic CVC vowels may be affected by the rate of articulation (Lindblom and Studdert-Kennedy, 1967). Such vowels may be said to be "encoded" (Liberman *et al.*, 1967) in the sense that cues for their identification are provided simultaneously (in parallel) with cues for the identification of their neighboring consonants. Identification of both vowels and consonants entails a judgment, in some form, of the formant transitions. From the dichotic work of Haggard (1969) we know that synthetic semivowels and laterals (/w, r, l, j/), for which important cues are carried by relatively slow formant transitions, may give a right-ear advantage of the same order as that given by stop consonants. And finally, we have the evidence of Darwin (1969a,b), cited above, on the possible importance of formant transitions in the laterality effect for fricatives. We may then reasonably hypothesize that reduced, rapidly articulated, "encoded" vowels in CVC syllables, dependent for their recognition on the perception of formant transitions, would show a significant right-ear advantage. Experiments to test this hypothesis are now being planned.

F. Cerebral Dominance and Information Loss in the Laterality Effect

In the foregoing discussion, we have suggested that differences in right-ear advantage among stops and vowels may be due to differences in the susceptibility of these signal classes to information loss during transmission. In earlier discussions (for example, Shankweiler and Studdert-Kennedy, 1967a; Shankweiler, 1970), we have taken such differences in ear advantage to reflect differences in the degree to which consonants and vowels engage the specialized perceptual mechanisms of the dominant hemisphere. We should now make explicit the reasons for this shift in interpretation and, at the same time, summarize our current understanding of the laterality effect.

There are two necessary conditions of an ear advantage in dichotic listening. First, some part of the perceptual process must depend upon unilateral neural machinery; second, the signal from the ipsilateral ear must undergo a significant loss due either to degradation of the signal during transmission to the dominant hemisphere or to its decay during the time it is held before final processing. Wherever a reliable contralateral ear advantage is observed, both these conditions must have been fulfilled. However, Darwin (1969a,b) and Halwes (1969) have independently pointed out that where an ear advantage is not observed, or is small, the outcome is ambiguous: it may indicate either no unilateral processing or no significant information loss in the ipsilateral signal. In other words, the absence of an ear advantage is not inconsistent with complete lateralization of some portion of the perceptual function, since the outcome may simply indicate that the acoustic materials being studied are not susceptible to information loss under certain experimental conditions.

This is the interpretation that the reduced effect for final consonants seems to demand, since, in the interests of parsimony, we must suppose that final consonants require the operation of specialized feature extractors in the dominant hemisphere no less than initials. For the vowels, the situation is not so clear. The "continuous" nature of vowel perception (for a recent discussion, see Studdert-Kennedy *et al.*, 1970) may perhaps be related to vowels not engaging discrete feature extractors in the dominant hemisphere. At the same time, transfer of vowel information to the dominant hemisphere for final perceptual response is unavoidable, and the most parsimonious interpretation again seems to be that the reduced or null laterality effect for vowels is also due to reduced information loss rather than to absence of cerebral dominance.

We may, finally, distinguish two broad directions that future research with dichotic materials might take. First, there is research of general auditory interest. Much remains to be learned about the experimental and acoustic conditions of ipsilateral transmission loss. Appropriate research may increase our understanding of those features in the design of the auditory system that make it possible to demonstrate laterality effects. Second, there is research directed primarily to the understanding of speech perception. Wherever a laterality effect for speech materials clearly occurs, we may exploit the effect to infer underlying perceptual processes. Here we should emphasize a point that may easily be missed: the size of the laterality effect is not a measure of its importance or of its value for research. We are not concerned in dichotic experiments to estimate the contribution of a variable to control over perception. We are, rather, exploiting the apparently trivial errors of a system under stress to uncover its functional processes.

IV. CONCLUSIONS

This study of dichotically presented "natural" speech CVC syllables showed: (1) a significant right-ear advantage for initial stop consonants; (2) a significant, though reduced, right-ear advantage for final stop consonants; (3) a nonsignificant right-ear advantage for six medial vowels; and (4) significant and independent right-ear advantages for the articulatory features of voicing and place in initial stop consonants.

We have argued, following Kimura (1961b), that the right-ear advantages are to be attributed to left cerebral dominance and functional prepotency of the contralateral pathways during dichotic stimulation. From analysis of the errors made in perception of the initial stop consonants, we have tentatively concluded that, while the general auditory system may be equipped to extract the auditory parameters of a speech signal, the dominant hemisphere is specialized for the extraction of linguistic features from those parameters. The laterality effect would then be due to a loss of auditory information arising from interhemispheric transfer of the ipsilateral signal to the dominant hemisphere for linguistic processing.

ACKNOWLEDGMENTS

We wish to thank A. M. Liberman, M. P. Haggard, C. J. Darwin, and T. Halwes for many hours of fruitful discussion during the preparation of this paper. Acknowledgment is due to the Charles E. Merrill Publishing Company for permission to reprint Figs. 1 and 2, which are also to appear in Shankweiler (1970). This work was supported in part by a grant to Haskins Laboratories from the National Institute of Child Health and Human Development.

---

[*] Also Queens College of the City Univ. of New York, Flushing, N. Y.

[†] Also Univ. of Connecticut, Storrs, Conn.

[1] Reports of some of the findings of this study were included in a paper read before the Acoustical Society of America (Shankweiler and Studdert-Kennedy, 1967a), and in a presentation by one of us (D.S.) at the ONR conference on Perception of Language, University of Pittsburgh, January 1967. (Shankweiler, 1970).

[2] Main results for the FC consonants are presented in Tables I, II, and III. All further consonant data analysis is for IC consonants only, largely due to our dissatisfaction with the FC stimuli. Accordingly, since vowel data were intended for comparison with consonant, only the IC vowel data have been fully analyzed: all reported vowel results are for this test only.

[3] Order of report effects have been shown to be present, but insufficient to account for the entire laterality effect, in many studies. For reviews, see Bryden (1967), Satz (1968), and Halwes (1969).

[4] A measure of ear advantage might be derived from both-correct trials by use of preference scores, but these trials may not all be of equal difficulty.

[5] We note here a discrepancy between this result and a finding of our earlier study. There, performance was improved by the sharing of voicing (suggesting the greater difficulty of that feature); here, performance was improved by the sharing of place. Since the inference from Table IV of greater difficulty in the perception of place than of voicing is borne out by every other relevant analysis in the present study [as also by the findings of Miller and Nicely (1955) and Singh (1966)], we have discounted the discrepancy in our subsequent discussions.

[6] Trost *et al.* (1968) report equal right-ear advantages for initial and final consonants in "natural" CVC syllables. But since their test lists included fricatives and liquids, and voiced, voiceless, and nasal stops (not all of which occurred equally often in initial and final position), their results are difficult to compare with those of this study.

REFERENCES

ABRAMSON, A. S., and LISKER, L. (1965). "Voice Onset-time in Stop Consonants: Acoustical Analysis and Synthesis," Int. Congr. Acoust., 5th, Liège, Vol. Ia, p. A51.

ATTNEAVE, F. (1959). *Applications of Information Theory to Psychology* (Holt, Rinehart, and Winston, New York).

BARTZ, W. H., SATZ, P., and FENNELL, E. (1967). "Grouping Strategies in Dichotic Listensing: The Effects of Instructions, Rate and Ear Symmetry," J. Exp. Psychol. *74*, 132–136.

BENTON, A. L. (1965). "The Problem of Cerebral Dominance," Canad. Psychologist *6*, 332–348.

BOCCA, E., CALEARO, C., CASSINARI, V., and MIGLIAVACCA, F. (1955). "Testing 'Cortical' Hearing in Temporal Lobe Tumors," Acta Oto-Laryngol. *45*, 289–304.

BRADY, P. T., HOUSE, A. S., and STEVENS, K. N. (1961). "Perception of Sounds Characterized by a Rapidly Changing Resonant Frequency," J. Acoust. Soc. Amer. *33*, 1357–1362.

BRANCH, C., MILNER, B., and RASMUSSEN, T. (1964). "Intracarotid Sodium Amytal for the Lateralization of Cerebral Speech Dominance," J. Neurosurg. *21*, 399–405.

BROADBENT, D. E. (1954). "The Role of Auditory Localization in Attention and Memory Span," J. Exp. Psychol. *47*, 191–196.

BRYDEN, M. P. (1967). "An Evaluation of Some Models of Dichotic Listening," Acta Oto-Laryngol. *63*, 595–604.

CHANEY, R. B., and WEBSTER, J. C. (1965). "Information in Certain Multidimensional Signals," U. S. Navy Electron. Lab. Rep. No. 1339, San Diego, Calif.

CURRY, F. K. W. (1967). "A Comparison of Left-handed and Right-handed Subjects on Verbal and Non-verbal Dichotic Listening Tasks," Cortex *3*, 343–352.

CURRY, F. K. W., and RUTHERFORD, D. R. (1967). "Recognition and Recall of Dichotically Presented Verbal Stimuli by Right- and Left-handed Persons," Neuropsychologia *5*, 119–126.

DARWIN, C. J. (1969a). *Auditory Perception and Cerebral Dominance*, (Unpublished PhD thesis, University of Cambridge).

DARWIN, C. J. (1969b). "Laterality Effects in the Recall of Steady-State and Transient Speech Sounds," J. Acoust. Soc. Amer. *46*, 114(A).

DENES, P. B. (1963). "On the Statistics of Spoken English," J. Acoust. Soc. Amer. *35*, 892–904.

GAZZANIGA, M. S., and SPERRY, R. W. (1967). "Language after Section of the Cerebral Commissures," Brain *90*, 131–148.

GREENBERG, J. H., and JENKINS, J. (1964). "Studies in the Psychological Correlates of the Sound System," Word *20*, 157–177.

HAGGARD, M. P. (1969). "Perception of Semi-Vowels and Laterals," J. Acoust. Soc. Amer. *46*, 115(A).

HALL, J. L., and GOLDSTEIN, M. H. (1968). "Representation of Binaural Stimuli by Single Units in Primary Auditory Cortex of Unanaesthetized Cats," J. Acount Soc. Amer. *43*, 456–461.

HALWES, T. (1969). "Effects of Dichotic Fusion in the Perception of Speech," PhD thesis, University of Minnesota.

KIMURA, D. (1961a). "Some Effects of Temporal Lobe Damage on Auditory Perception," Canad. J. Psychol. *15*, 156–165.

KIMURA, D. (1961b). "Cerebral Dominance and the Perception of Verbal Stimuli," Canad. J. Psychol. *15*, 166–171.

KIMURA, D. (1964). "Left-Right Differences in the Perception of Melodies," Quart. J. Exp. Psychol. *16*, 355–358.

KIMURA, D. (1967). "Functional Asymmetry of the Brain in Dichotic Listening," Cortex *3*, 163–178.

KIMURA, D., and FOLB, S. (1968). "Neural Processing of Backwards-Speech Sounds," Science *161*, 395–396.

KIRSTEIN, E., and SHANKWEILER, D. (1969). "Selective Listening for Dichotically Presented Consonants and Vowels," Haskins Labs. Status Rep. on Speech Res. SR 17/18, 133–141.

KLATT, D. H. (1968). "Structure of Confusions in Short-Term Memory Between English Consonants," J. Acoust. Soc. Amer. *44*, 401–407.

LANE, H. L. (1965). "The Motor Theory of Speech Perception: A Critical Review," Psychol. Rev. *72*, 275–309.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D., and STUDDERT-KENNEDY, M. (1967). "Perception of the Speech Code," Psychol. Rev. *74*, 431–461.

LIBERMAN, A. M., DELATTRE, P. C., COOPER, F. S., and GERSTMAN, L. J. (1954). "The Role of Consonant–Vowel Transitions in the Perception of the Stop and Nasal Consonants," Psychol. Monogr. *68*, No. 379.

LIBERMAN, A. M., DELATTRE, P. C., GERSTMAN, L. J., and COOPER, F. S. (1956). "Tempo of Frequency Change as a Cue for

Distinguishing Classes of Speech Sounds," J. Exp. Psychol. *52*, 127–137.

Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1958). "Some Cues for the Distinction Between Voiced and Voiceless Stops in Initial Position," Language and Speech *1*, 153–167.

Lieberman, P. (1967). *Intonation, Perception and Langauge* (MIT Press, Cambridge, Mass.).

Lieberman, P. (1968). "Primate Vocalizations and Human Linguistic Ability," J. Acoust. Soc. Amer. *44*, 1574–1584.

Lieberman, P., Klatt, D. L., and Wilson, W. A. (1969). "Vocal Tract Limitations on the Vowel Repertoires of Rhesus Monkey and Other Nonhuman Primates," Science *164*, 1185–1187.

Lindblom, B. E. F., and Studdert-Kennedy, M. (1967). "On the Role of Fomant Transitions in Vowel Recognition," J. Acoust. Soc. Amer. *42*, 830–843.

Lisker, L., and Abramson, A. S. (1964). "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements," Word *20*, 384–422.

Mattingly, I., and Liberman, A. M. (1970). "The Speech Code and the Physiology of Language," in *Information Processing in the Nervous System*, K. N. Leibovic, Ed. (Springer, New York), pp. 97–117.

McGill, W. J. (1954). "Multivariate Information Transmission," Psychometrika *19*, 97–116.

Miller, G., and Nicely, P. E. (1955). "An Analysis of Perceptual Confusions Among Some English Consonants," J. Acoust. Soc. Amer. *27*, 338–352.

Milner, B. (1962). "Laterality Effects in Audition," in *Interhemispheric Relations and Cerebral Dominance*, V. B. Mountcastle, Ed. (Johns Hopkins U. P., Baltimore), pp. 177–195.

Milner, B., Taylor, L., and Sperry, R. W. (1968). "Lateralized Suppression of Dichotically-presented Digits after Commissural Section in Man," Science *161*, 184–185.

Nabelek, I., and Hirsh, I. J. (1969). "On the Discrimination of Frequency Transitions," J. Acoust. Soc. Amer. *45*, 1510–1519.

Pollack, I. (1968). "Detection of Rate of Change of Auditory Frequency," J. Exp. Psychol. *77*, 535–541.

Rosenzweig, M. R. (1951). "Representations of the Two Ears at the Auditory Cortex," Amer. J. Physiol. *167*, 147–158.

Satz, P. (1968). "Laterality Effects in Dichotic Listening," Nature *218*, 277–278.

Satz, P., Achenbach, K., Pattishall, E., and Fennell, E. (1965). "Order of Report, Ear Asymmetry and Handedness in Dichotic Listening," Cortex *1*, 377–396.

Satz, P., Fennell, E., and Jones, M. B. (1969). "Comments on: A Model of the Inheritance of Handedness and Cerebral Dominance," Neuropsychologia *7*, 101–103.

Shankweiler, D. (1966a). "Defects in Recognition and Reproduction of Familiar Tunes after Unilateral Temporal Lobectomy," Paper read before the 37th Ann. Meeting of the Eastern Psychol. Ass., New York, N. Y.

Shankweiler, D. (1966b). "Effects of Temporal-Lobe Damage on Perception of Dichotically Presented Melodies," J. Comp. Physiol. Psychol. *62*, 115–119.

Shankweiler, D. (1970). "An Analysis of Laterality Effects in Speech Perception," in *Perception of Language*, P. Kjeldergaard, Ed. (Chas. E. Merrill, Columbus, Ohio).

Shankweiler, D., and Studdert-Kennedy, M. (1966). "Lateral Differences in Perception of Dichotically Presented Synthetic Consonant–Vowel Syllables and Steady-State Vowels," J. Acoust. Soc. Amer. *39*, 1256(A).

Shankweiler, D., and Studdert-Kennedy, M. (1967a). "An Analysis of Perceptual Confusions in Identification of Dichotically Presented CVC Syllables," J. Acoust. Soc. Amer. *41*, 1581(A).

Shankwelier, D., and Studdert-Kennedy, M. (1967b). "Identification of Consonants and Vowels Presented to Left and Right Ears," Quart. J. Exp. Psychol. *19*, 59–63.

Siegel, S. (1956). *Non-parametric Statistics.* (McGraw-Hill, New York).

Singh, S. (1966). "Crosslanguage Study of Perceptual Confusions of Plosive Phonemes in Two Conditions of Distortion," J. Acoust. Soc. Amer. *40*, 635–656.

Singh, S. (1969). "Interrelationship of English Consonants," in Proc. Intern. Congr. Phonetic Sci.,6th, Prague, pp. 542–544.

Smith, A., and Burkland, C. W. (1966). "Dominant Hemispherectomy: Preliminary Report on Neuropsychological Sequelae," Science *153*, 1280–1282.

Sparks, R., and Geschwind, N. (1968). "Dichotic Listening in Man After Section of Neocortical Commissures," Cortex *4*, 3–16.

Sperry, R. W., and Gazzaniga, M. S. (1967). "Language Following Surgical Disconnection of the Hemispheres," in *Brain Mechanisms Underlying Speech and Language*, C. H. Millikan and F. L. Darley, Eds. (Grune and Stratton, New York), pp. 108–121.

Spreen, O., Benton, A. L., and Fincham, R. W. (1965). "Auditory Agnosia Without Aphasia," Arch. Neurol. *13*, 84–92.

Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., and Cooper, F. S. (1970). "The Motor Theory of Speech Perception: A Reply to Lane's Critical Review," Psychol. Rev. *77*, 234–249.

Trost, J. E., Shewan, C. M., Nathanson, S. N., and Sant, L. V. (1968). "A Dichotic Study of Ear Superiority in Perception of Consonants," Paper read before the 44th Annual Convention of the American Speech and Hearing Association, Denver.

Tunturi, A. R. (1946). "A Study on the Pathway from the Medial Geniculate Body to the Acoustic Cortex in the Dog," Amer. J. Physiol. *147*, 311–319.

Vignolo, L. A. (1969). "Auditory Agnosia: A Review and Report of Recent Evidence," in *Contributions to Clinical Neuropsychology*, A. L. Benton, Ed. (Aldine, Chicago), pp. 172–208.

Wickelgren, W. A. (1966). "Distinctive Features and Errors in Short-Term Memory for English Consonants," J. Acoust. Soc. Amer. *39*, 388–398.