7

AN ANALYSIS OF LATERALITY EFFECTS IN SPEECH PERCEPTION[1]

DONALD SHANKWEILER[2]
Haskins Laboratories

It has been recognized since the time of Broca (1861) that one hemisphere of man's brain, usually the left, is specialized for speech functions. Fundamental though this fact is, it has remained somewhat isolated from the main body of knowledge about the brain, and only recently has it been subjected to experimental analysis. Many recent investigations have shown that the right hemisphere, the so-called minor hemisphere, has functions of its own which are not fully shared by the left hemisphere. Some recent studies of auditory perception, which I am about to review, point to the conclusion that the two brain hemispheres do have different functions not only in the perception of speech but also music and certain other complex sounds.

---

[2]Also at the University of Connecticut.

THE PHENOMENON

Through an application of Broadbent's dichotic listening experiment (1954), Doreen Kimura has opened a new approach to the investigation of the relative contribution of the two hemispheres to the perception of auditory patterns and speech (Kimura, 1961a, b; Milner, 1962). It was discovered by Dr. Kimura, while working in Brenda Milner's Laboratory at the Montreal Neurological Institute, that when groups of spoken digits are presented through earphones so that different digits arrive at approximately the same time at each ear, those presented to the right ear are more accurately reported than those presented to the left. On a typical test trial, a subject received three pairs of recorded digits, spaced one-half second apart. The task was to repeat all the numbers heard without regard to which ear they were delivered. Most subjects made fewer errors for the right ear than for the left. Although the ear difference was small, it was reasonably consistent from subject to subject — about 80 percent of the right-handed subjects were a little better with the right ear. This result seemed to suggest that the right ear has the better path to the speech processing areas of the left hemisphere. This suggestion was confirmed many times in subsequent experiments.

A number of animal studies had already indicated that each ear has greater representation in the opposite cerebral hemisphere. Rosenzweig (1951) had shown that for the cat the auditory projections have a stronger crossed component than uncrossed. His results were recently confirmed in the microelectrode studies of Hall and Goldstein (1968). Indications that this is true for man as well had been given by Bocca, Calearo, Cassinari and Migliavacca (1955), who discovered that the perception of speech presented to one ear alone is impaired when directed to the ear opposite a temporal-lobe lesion. Kimura's findings in perception of dichotically presented digits confirmed these earlier indications of the superiority of the crossed pathways and demonstrated, for the first time, an effect of cerebral dominance in a perceptual task performed by listeners with intact nervous systems. It was an important gain, because it meant we no longer had to rely solely on brain-damaged individuals as the only source of knowledge about the cerebral organization of auditory and speech functions.

Dichotic presentation is apparently a necessary condition for the occurrence of these lateral differences in perception. If the stimuli

are presented for recognition monaurally, neither ear has an advantage, however unfavorable the signal-to-noise ratio (given ears of equal sensitivity and free of pathology). We know that each ear is represented by a somewhat larger population of cells in the opposite cerebral hemisphere. We do not know how great the inequality of representation is. Competition of inputs brought about by dichotic presentation may well result in a decided advantage to the input channel which has even slightly stronger representation in the hemisphere specialized for processing the type of material presented. Kimura (1964) has suggested a similar possibility based on Rosenzweig's (1951) demonstration in the cat that some cortical units serve both ears. The overlap may be such that the binaural interaction which results when both ears are stimulated simultaneously could enhance contralateral transmission in the system at the expense of ipsilateral transmission. Thus, some units which respond to both ears when each ear is stimulated alone may be driven selectively by the contralateral ear when both ears are stimulated in tandem.

We must still consider the possibility that the right ear can cope more effectively with competing speech than the left ear. A necessary control is the monaural presentation of the electrically-mixed output of two channels each containing different words. Experiments conducted at McGill University by Philip Corsi (1967) have shown decisively that competition does not bring about the ear effect unless the competition is interaural; that is, dichotically produced. This finding has been confirmed by Terry Halwes and myself in our Laboratories. Therefore, we may conclude that the ear effects in dichotic listening depend jointly on differences in function of the two cerebral hemispheres and a system arranged to favor contralateral transmission during simultaneous stimulation of the two ears.

DISSOCIATED LATERALITY EFFECTS FOR SPEECH AND NONSPEECH

The findings which link the ear advantage with cerebral lateralization of speech are remarkably clear-cut and encouraged further work with the dichotic technique. One question which seemed worth exploring was whether or not some non-verbal stimuli would give a left-ear advantage when presented dichotically, thereby reflecting specialization of structures in the right hemisphere for carrying out non-verbal perceptual processes. This possibility seemed likely

in view of Milner's earlier finding (1962) that certain complex auditory discriminations are affected by right temporal-lobectomy but not by left. Working with the Seashore Measures of Musical Talents, Milner has shown that tonal pattern perception and timbre discrimination are selectively impaired by right temporal lobectomy.

Kimura (1964) devised a dichotic melodies test in which pairs of different melody fragments are presented to the ears simultaneously. After each test trial consisting of a dichotically presented pair, a set of four melodies is presented, one at a time, posing an auditory multiple-choice problem. The subject had to indicate which two melodies had been presented dichotically. The task was one of recognition and did not involve naming. Kimura found that a significantly greater number of correct choices were made for melodies presented to the left ear. The same subjects were given the dichotic digits test and were more accurate on the right ear. The ear differences for digits and melodies presented dichotically go in opposite ways as she had predicted. It was recently reported by Curry (1967) that common environmental noises are also better recognized from the left ear. Kimura's melodies and digits tasks have been presented to patients with localized lesions of the left and right temporal lobes with the outcome of a significant double dissociation of deficits on the two tasks (Shankweiler, 1966; Kimura, 1967). That is to say, patients with left temporal-lobe lesion were impaired on the digits task but not on melodies, whereas patients with right temporal-lobe lesion were impaired on the melodies task but not on digits. Thus, it was the verbal or non-verbal nature of the task that seemed to be the deciding factor in determining whether a deficit occurred in cases of left or right temporal-lobe lesion. Both sets of data, those for patients and for normal subjects, agree with other indications that the minor hemisphere plays a greater role in processing some nonspeech sounds.

CEREBRAL ASYMMETRIES IN SPEECH PERCEPTION IN RELATION TO LEVELS OF LANGUAGE

These findings are clear in their indications of hemispheric differences in perception of verbal and nonverbal materials, but they are ambiguous in one respect. You will recall that Kimura originally used spoken digits as stimuli; that is to say, meaningful words were used. We did not know whether the meanings had to be perceived

in order for the laterality effect to occur, or whether simply being perceived as speech was sufficient. The question has not often been asked whether all processes in the perception of language are lateralized. The answer has important implications. It bears on the general question of whether or not speech perception rests on mechanisms that are different from those that underlie the perception of other sound patterns. We know that the facts of hemispheric specialization point directly to the existence of specialized neural mechanisms that serve language in a broad sense. But let us consider speech in its narrowest sense, in the sense of the sounds of language and their perception as strings of phonemes. Speech this narrowly conceived is often not considered to be an integral part of the species-specific activity we call language. There are many reasons for believing, however, that the sounds of speech belong to the structure of language and are highly interconnected with other parts of the language hierarchy. If this is so, we would expect these interconnections to be evident in the neural foundations of language.

Much of the research with artificial sound alphabets has shown that not just any distinctive set of sounds can serve as an efficient vehicle for the transmission of language (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). My colleagues at Haskins Laboratories were challenged to find out why the sounds we use naturally are so well suited to the task. We have turned to analysis of articulation in order to clarify certain unique properties of speech sounds. The motor activity of speech consists largely of gestures which overlap one another in time. As a consequence of this temporal overlap, the relationship between the acoustic characteristics of speech and the perceptual units (phonemes and their features) is highly complex. The distinctive features of the phonemes are emitted in parallel fashion, not in linear phoneme-sized bundles, so that at almost any instant during the articulation of a syllable, the acoustic cues convey information about more than one phoneme. Speech, then, becomes encoded or restructured in time in the process of transmission from the speaker to the hearer. The perceptual mechanism which recovers phonemes from the sound stream functions in some sense as a decoder.

We see then that special requirements are placed on the perceptual apparatus due to the encoded nature of the speaker's message. Because of this one might expect the perception of speech to differ from that of other sounds. It seemed that the type of experiment I have been describing offered a rather direct way to study some characteristics of the speech decoder.

When Dr. Michael Studdert-Kennedy and I began our experiments at Haskins Laboratories, the first question we explored was whether the specialized mechanisms of the dominant hemisphere are engaged in perceiving speech in the narrowest sense, or whether the same processors that operate in the recognition of other auditory patterns such as environmental sounds and music are used. The second question relates to the fact that in Kimura's task the digits were presented in groups. Since our interest was in perceptual mechanisms we wanted to avoid serial presentation, if possible. First, we had to determine whether lateral differences could be obtained with single pairs of stimuli. Since the laterality effect presumably depends upon some optimal level of competition, we hoped that by careful synchronization of the stimuli, we could limit a trial to a single pair of syllables.

We presented speech in the most elementary form, directing to each ear dichotically single consonant-vowel syllables, only one pair on each trial. In our first experiments (Shankweiler & Studdert-Kennedy, 1967b) we used synthetic speech generated by the Haskins Pattern Playback (Cooper, Liberman, & Borst, 1951), a device for converting simplified hand-painted speech spectrograms into sound. The stimulus pairs were made up of all combinations of the six stop consonants (/ b, d, g, p, t, k /) followed by the same vowel / a /. Each of the 15 possible syllable pairs such as / ka-da / was presented equally often, with / ka / going to the left ear and / da / to the right half the time and then in the reverse order.

Because we were working with single pairs of syllables which contrasted only in the beginning consonant portion lasting less than 50 msec., we had to solve the problem of synchronizing the onsets more precisely than had been done before. We eventually arrived at a satisfactory method for achieving the alignment. Copies of each syllable were made using a dual-channel tape recorder; some being dubbed on track 1, some on track 2. These stimuli were then spliced into tape loops, each loop containing a pair of syllables, such as / ba-da /. The next step was aligning the stimuli, by synchronizing the onsets of each syllable and equalizing peak intensity. This was done by playing back each loop on a tape deck which had been especially modified to permit the length of tape passing between two playback heads to be varied. This was adjusted until the two utterances coincided. The final step in the alignment procedure involved making an oscillographic record of the syllables using a Visicorder with a time marker as illustrated in Figure 1. The two upper tracings in the figure show synchronized syllables,
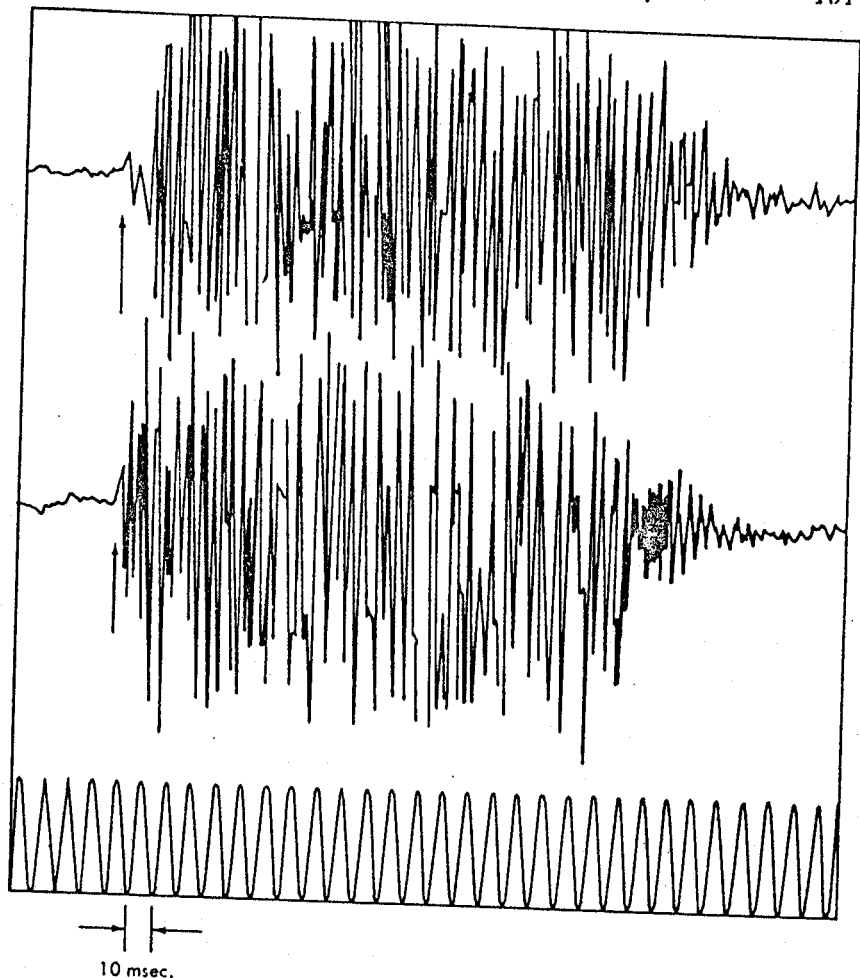
FIGURE 1. Oscillographic record displaying two syllables in temporal alignment.

The syllables shown here are real speech. The procedure was the same for aligning synthetically-produced syllables.

and the lower tracing is a time marker produced by a continuous 100 Hz. tone. We were thus able to align the stimuli within a few milliseconds.

The subjects were told that different syllables would always be presented to the two ears and that their task was to write down both

in order of confidence. They were required to make two responses even if one or both had to be guesses. The set of syllables from which they were to make their choices was displayed at the top of the answer sheet. Each listener heard each test twice, and the earphones were reversed on the second run, so that no differences in the components of the playback system could contribute to the ear differences.

We found a substantial and highly significant advantage in the accuracy of identification of syllables presented to the right ear, with 14 out of 15 subjects receiving better scores on the right, and a mean right ear advantage of 16 percent. Since our stimuli consisted of nonsense syllables that contrasted in only one phone, we concluded that the specific language processing system of the dominant cerebral hemisphere is engaged at the level of the sound structure of the language; that speech perception in the narrow sense depends on the specialized machinery of the left hemisphere. Moreover, since the effect occurred with presentation of single pairs of syllables, we concluded, in agreement with Kimura (1962, 1967), Bryden (1967) and Satz (1968), that the effect has a genuinely perceptual basis, and is not due to memory.

Dr. Studdert-Kennedy and I made up a new test containing contrasting pairs of vowels alone. These were prolonged, monotone or steady-state vowels which in speech are formed with the vocal tract in a fixed, open-tube configuration. Therefore, we would expect them to be unencoded, or at least less encoded than the stops. It would follow that these vowels would not necessarily engage the processor specialized for speech. We found for this set of vowels a much smaller right ear advantage for the same group of subjects who had shown a decisive right ear advantage for stop consonants (Figure 2). This result suggested that these vowel stimuli were processed about equally well by either hemisphere of the brain. The vowel data did show a small net trend in favor of the right ear, however; and although this was not statistically significant, we wondered whether the difference might have been larger if the vowels had not been monotone or if they had been presented in syllabic context.

These experiments raised a number of further questions. In particular, the indication of a difference in the mode of processing consonants and vowels needed re-examination. We had used synthetic speech in these experiments because of the advantages it offers in achieving uniform and specifiable acoustic structure. But because of the particular synthesizer we used, these advantages
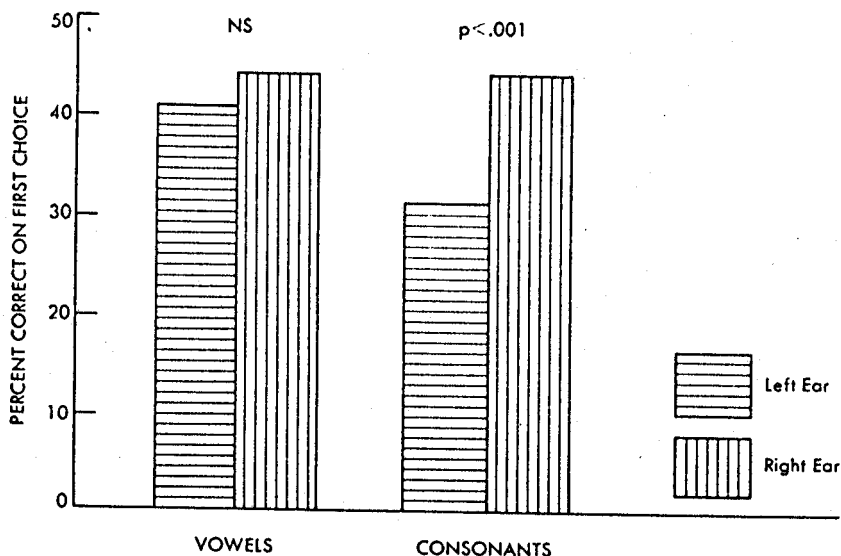
FIGURE 2. Ear differences for dichotically presented synthetic steady-state vowels and consonant-vowel syllables. Ten subjects.

were bought at the expense of natural quality. We wondered whether our results, particularly those for contrasting isolated vowels, might reflect characteristics of these particular synthetic stimuli.

IDENTIFICATION OF CONSONANTS AND VOWELS IN SYLLABLES

In order to clarify this result we performed a new experiment using real speech. In this experiment vowels as well as consonants occurred in syllabic context. The syllables were spoken by a phonetician experienced in making speech recordings. A spectrogram was made of each recorded syllable and the durations were measured. These ranged from 300 to 500 msec. For the consonant comparison, the stimuli consisted of consonant-vowel-consonant syllables formed by pairing the six stop consonants with each of six vowels (/ i, ɛ, æ, ɑ, ɔ, u /). All syllables were closed by the same consonant / p /. For every consonant, therefore, we obtained six separate estimates of the laterality effect. This required a much longer test, containing 360 trials as opposed to 120 trials for the previous experiments. Any

given syllable pair differed only in the initial consonant as before, but in the test as a whole each consonant was followed equally often by each of six vowels. This test was appropriate for our purpose of obtaining a reliable estimate of the magnitude of the ear effect with different phonetic contrasts. The alignment procedure was the same as before as illustrated in Figure 1.

A second test consisted of pairs which contrasted in the vowel alone. This test was made up of different combinations of the same set of spoken consonant-vowel-consonant nonsense syllables, yielding a test of the same length as the consonant test.

We have given some thought to the problem of how best to estimate the magnitude of the ear advantage. This is an important problem since we would like to be able to make meaningful statements about the relative size of the ear difference with various classes of stimuli, not simply to say that a significant difference is present or not present. Our initial interest was in the first response the subject wrote. Analysis of our earlier data for Playback speech indicated that second responses contained little information, while the first response seemed to be a sensitive indicator of the laterality effect. We have abandoned this measure in describing the data of the present experiment because some subjects were more accurate on the second response than the first. Simple percent correct for each ear is also an unsatisfactory measure, since the size of the ear advantage may vary with the difficulty of the task. Indeed *a priori* limits are placed on the possible size of the ear advantage when the overall level of performance is high or low. There is an optimal range of difficulty levels over which the ear difference is not severely constrained.[3] An overall level of 50 percent correct is about ideal. At this level of performance the ear difference could in principal vary from 0 to 100 percent.

One way to achieve an optimal level of difficulty is to add noise to each channel. Noise, however, has the disadvantage of masking some acoustic features more than others. We have recently adopted the policy of scoring only those stimulus pairs on which one correct response and one error was made. Trials on which both responses are correct and both are wrong can contribute nothing to the ear difference. Discarding these trials also gives us the desirable 50 percent level of difficulty. We then express the score for each ear as a percentage of the total, $(L/L+R)100$ and $(R/L+R)100$, and the

---

[3] I am indebted to Terry Halwes of the University of Minnesota for pointing this out to me.
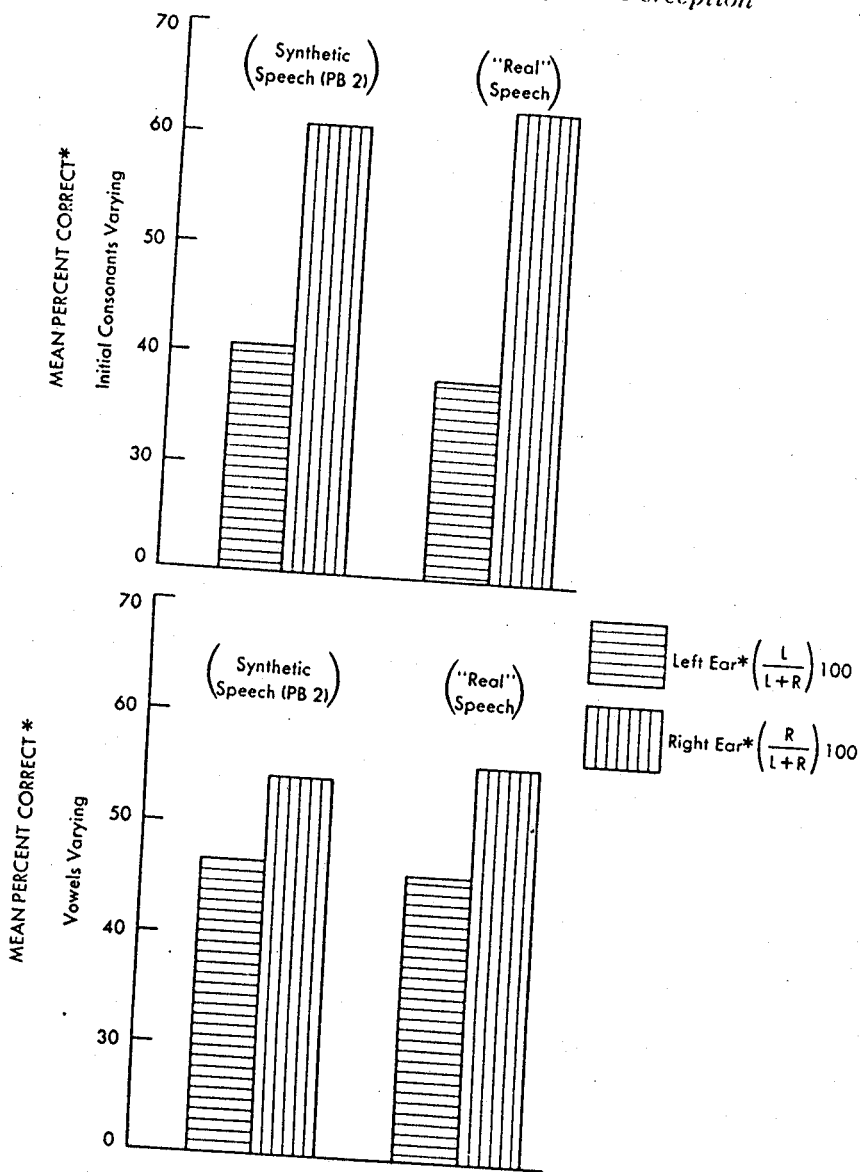
FIGURE 3. $\dfrac{R-L}{R+L}$ (100) as a measure of the ear difference for dichotically presented syllables, synthetic and real. This tabulation includes all trials on which one error occurred.

ear difference by $(L-R/L+R)100$, where L and R represent the frequencies of correct responses for presentations to the left ear and right ear.

Figure 3 shows this measure for the data from these experiments (Shankweiler and Studdert-Kennedy, 1967a). All 12 subjects were better on the right ear in the consonant test. The left portion of the Figure shows histograms for our earlier experiment with synthetic speech. The right half of the figure compares the results of the real speech consonant and vowel tests for the 11 subjects who took both tests. The two sets of data agree remarkably well although the subjects and the stimuli were different. The mean right ear advantage is greater by a factor of three for consonants than vowels in both experiments.

THE RIGHT EAR ADVANTAGE FOR INDIVIDUAL STOPS AND VOWELS

The indication that the left cerebral hemisphere participates more in the perception of stop consonants than vowels has important implications for understanding the nature of speech perception. Therefore, we would like to be sure that these differences are not the result of some artifact. We can immediately rule out the possibility that the differences in the size of the laterality effect for consonants and vowels are due to the greater difficulty of the consonants, because we have equated the difficulty of the tests by discarding trials on which both responses were correct and both were wrong. We can also rule out the possibility that the ear effects are attributable to a particular choice of stop-vowel combinations, since every stimulus was paired equally often with every other, and the whole test was presented a second time with channels and ears reversed.

We have noted that the ear effect for consonants was reliable from subject to subject and unreliable for vowels. We now examine for consistency the laterality effect for individual consonants and vowels. Figure 4 shows the ear advantages computed for each consonant and each vowel separately. The indices are arranged from left to right in order of decreasing magnitude. Consonants and vowels are perfectly segregated by this arrangement. All the consonant effects are significant at $p<.001$. One vowel / i / was significant with $p<.01$. A still more stringent test of consistency was
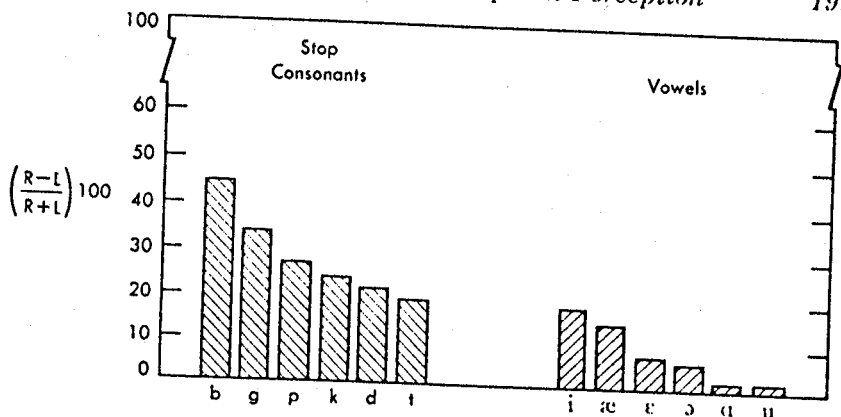
FIGURE 4. The right ear advantage for individual stop consonants and vowels.

The tabulation includes all trials on which one error occurred.

made by examining the ear effect for each of the 30 consonant pairs presented in the test. (Each subject received 12 presentations of each pair). A net right ear advantage was obtained for all but one of the 30 stimulus pairs. The vowel data are much more variable. Examination of all vowel pairs individually showed no ear difference or a *left* ear advantage on 13 out of 30 pairs. These analyses have demonstrated that the right ear advantage is present for the whole class of stop consonants, when joined in syllables with any of six vowels.[4]

When as in the earlier experiment, vowels were presented steady-state and in isolation, we were not surprised to find little advantage for the right ear because that set of vowels was relatively unlike speech. We find it remarkable that vowels embedded in CVC syllables should also fail to give a significant right-ear advantage. The result strongly supports our earlier suggestion that perception of stop consonants is highly dependent upon the dominant hemisphere function whereas vowels may or may not be processed in the dominant hemisphere. In view of the many indications that the perception of melodies and tonal patterns of complex structure de-

[4]A forthcoming paper by Studdert-Kennedy and Shankweiler will describe this experiment in detail and will present an analysis of the consonant data by feature.

pends more upon right hemisphere function than left (Milner, 1962; Kimura, 1964, 1967; Shankweiler, 1966; Curry, 1967) it appears that the vowels are processed neither like music and other nonspeech auditory patterns nor like what is most characteristic of speech. It is conceivable that the right ear advantage occurs mainly for speech cues of brief duration, such as the bursts and rapidly changing formant frequencies that characterize stop consonants but not vowels. We are doing experiments to examine this possibility, but we are doubtful that the laterality effect in perception of consonants can be tied to any particular acoustic characteristic. I think, rather, that these vowel-consonant differences demonstrate that the speech processing mechanism is particularly attuned to those phonetic properties of speech that are a consequence of encoding.

The degree of encoding should be a relative, not absolute difference between stops and vowels. Important changes in the articulation of vowels occur as the speaking rate varies. Vowels that occur in very rapid speech are not only much shorter than the vowel portion of syllables produced in isolated syllables, but they may fail to reach their "target" formant values altogether. These vowels are presumably encoded to a greater degree than the vowels contained in syllables produced in citation by paced reading from a list. We will be interested to learn whether or not these vowels behave more like stop consonants when presented dichotically than the vowels we have used.

I shall conclude by summarizing the questions examined in these experiments. The right ear advantage for contrasting consonants was relatively large and consistent when single pairs of syllables were presented, indicating that the ear advantage pertains to the recognition of the stimuli and not only to their retention. The left cerebral hemisphere has long been known to contain structures specialized for language functions. Our findings indicate that speech, in the narrowest sense, engages the specialized machinery of the left hemisphere. Perception, even of nonsense syllables, involves processes different than those involved in the perception of music and other nonspeech sound patterns. Finally, the ear effects are differentially sensitive to differences among speech sounds themselves. Syllables contrasting only in the vowel portion were identified almost equally well by either ear. We think this reflects differences between stop consonants and vowels in the degree of encoding — differences in the degree to which phonemes are restructured in time in the process of articulation.

REFERENCES

Bocca, E., Calearo, C., Cassinari, V., & Migliavacca, F. Testing "cortical" hearing in temporal lobe tumors. *Acta Oto-Laryngologica*, 1955, 45, 289-304.

Broadbent, D. E. The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 1954, 47, 191-96.

Broca, P. Remarques sur le siège de la faculté du langage articulé suivies d'une observation d'aphémie. *Bulletin de la Societe Anatomique de Paris*, 1861, 36, 330-57.

Bryden, M. P. An evaluation of some models of laterality effects in dichotic listening. *Acta Oto-Laryngologica*, 1967, 63, 595-604.

Cooper, F. S., Liberman, A. M., & Borst, J. M. The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Sciences*, 1951, 37, 318-28.

Corsi, P. M. The effects of contralateral noise upon the perception and immediate recall of monaurally presented verbal material. Unpublished masters thesis, McGill University, Montreal, 1967.

Curry, F. K. W. A comparison of left-handed and right-handed subjects on verbal and nonverbal dichotic listening tasks. *Cortex*, 1967, 3, 343-52.

Hall, J. L., & Goldstein, M. H. Representation of binaural stimuli by single units in primary auditory cortex of unanesthetized cats. *Journal of the Acoustical Society of America*, 1968, 3, 456-61.

Kimura, D. Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 1961, 15, 166-71. (a)

Kimura, D. Some effects of temporal-lobe damage on auditory perception. *Canadian Journal of Psychology*, 1961, 15, 156-65. (b)

Kimura D. Perceptual and memory functions of the left temporal lobe: A reply to Dr. Inglis. *Canadian Journal of Psychology*, 1962, 16, 18-22.

Kimura, D. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 1964, 14, 355-58.

Kimura, D. Functional asymmetry of the brain in dichotic listening. *Cortex*, 1967, 3, 163-78.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-61.

Milner, B. Laterality effects in audition. In V. B. Mountcastle (ed.), *Interhemispheric relations and cerebral dominance*. Baltimore: Johns Hopkins Univ. Press, 1962. Pp. 177-95.

Rosenzweig, M. R. Representations of the two ears at the auditory cortex. *American Journal of Physiology*, 1951, 67, 147-58.

Satz, P. Laterality effects in dichotic listening. *Nature*, 1968, 218, 277-78.

Shankweiler, D. Effects of temporal-lobe damage on perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology*, 1966, 62, 115-19.

Shankweiler, D. & Studdert-Kennedy, M. An analysis of perceptual confusions in identification of dichotically presented CVC syllables. *Journal of the Acoustical Society of America*, 1967, 41, 1581 (A). (a)

Shankweiler D., & Studdert-Kennedy, M. Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, 1967, 19, 59-63. (b)