This article was downloaded by:[University of Sussex]
On: 15 July 2008
Access Details: [subscription number 776502344]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

Ear differences in the recall of fricatives and vowels

C. J. Darwin [a]
[a] Department of Psychology, University of Connecticut and Haskins Laboratories,

Online Publication Date: 01 February 1971

To cite this Article: Darwin, C. J. (1971) 'Ear differences in the recall of fricatives and vowels', The Quarterly Journal of Experimental Psychology, 23:1, 46 — 62

To link to this article: DOI: 10.1080/00335557143000059
URL: http://dx.doi.org/10.1080/00335557143000059

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Quarterly Journal of Experimental Psychology (1971) **23**, 46–62

EAR DIFFERENCES IN THE RECALL OF FRICATIVES AND VOWELS

C. J. DARWIN

Department of Psychology, University of Connecticut and Haskins Laboratories

Two experiments on the free recall of dichotically presented synthetic speech sounds are reported. The first shows that the right ear advantage for initial fricative consonants is not simply a function of the recognition response class, but that it is also a function of the particular acoustic cues used to achieve that response. This is true both for the whole response, and for the constituent phonetic features. The second experiment shows that when both the response class and the particular stimuli presented on certain trials are held constant, the right ear advantage for the constant stimuli can be influenced by the range of other stimuli occurring in the experiment. Vowels show a right ear advantage when, within the experiment, there is uncertainty as to vocal tract size, but they show no ear advantage when all the vowels in the experiment are from the same vocal tract. These results are interpreted as demonstrating that there are differences between the ears, and probably between the hemispheres, at some stage between the acoustic analysis of the signal and its identification as a phonetic category.

Introduction

Under certain conditions, sounds which enter one ear may subsequently be more efficiently recalled or recognized than similar sounds entering the other ear (Kimura, 1961*a*,*b*, 1964). Differences between the ears tend to be obtained more reliably when different sounds enter the two ears simultaneously than when only one ear is stimulated, either with one (Corsi, 1967) or with two simultaneous signals (Shankweiler, in press). Monaural stimulation *can* give significant ear differences but such experiments have required larger numbers of subjects than the usual dichotic paradigm (Bakker, 1968, 1970).

The type of stimulus material used is probably the only determinant of *which* ear gives better performance. In similar recognition paradigms the right ear does better for digit triads (Broadbent and Gregory, 1964) while the left does better for orchestrated melodies (Kimura, 1964) and simple pitch patterns (Darwin, 1969). In free recall, the right ear again does better for digit triads (Kimura, 1961*b*) while the left does better for familiar melodies (Kimura, 1967) and simple pitch sweeps, whether carried on a word or on a non-verbal timbre (Darwin, 1969).

Since patients with vocal speech impaired when their *right* hemispheres are anaesthetized show an advantage for the *left* ear in free recall of digit triads (Kimura, 1961*a*), some link between the ear difference effect and cerebral dominance must be assumed. Authors differ on the nature of this link. Some attribute it to perception (Kimura, 1961*b*), some to short-term memory (Inglis, 1962), others to attention (Treisman and Geffen, 1968). Some authors have implicitly denied the stimulus specificity of the direction of the effect, and claim that there is a general tendency to report material entering the right ear before that entering the left (Oxbury, Oxbury and Gardiner, 1967).

One important limitation of free recall experiments was pointed out by Inglis (1962). Serial-order effects (see e.g. Broadbent, 1958) could account for ear differences in a free recall paradigm if there were some tendency to report certain types of material from a particular ear first. Bryden (1963) controlled for serial order effects and found a smaller, though still significant, residual advantage for the right ear with digit sequences. Thus, while serial order effects account for some of the ear difference in a free recall paradigm, they neither explain why the sounds from one ear are recalled first, nor why there is a residual difference. The tendency to report one ear first could derive from whatever causes this residual ear difference.

This residual effect may be due to differences in the efficiency with which material is either perceived or remembered. Making the distinction between perception and memory in terms of the first and second ear reported, Bryden (1967) summarizes the available data, and shows that there is no evidence that the ear difference effect is any smaller on the first than on the second reported ear. Darwin (1969) also failed to find any such evidence both for material recalled better from the right and from the left ears.

Treisman and Geffen (1968) suggested that the ear difference effect arises because of an unequal distribution of attention, the left hemisphere finding it easier to attend to the right ear than the left ear. If this were so we would expect sounds which are more easily separated by selective attention to show a greater ear difference than those which are more difficult to separate. Kirstein and Shankweiler (1969), however, find that when a subject is asked to report the sounds from a particular ear, he makes fewer errors of attention for vowels than for consonants, but consonants show a greater right ear advantage than vowels. Selective attention may interact with the mechanisms responsible for the ear difference effect but it is not a basic cause.

Kimura's (1961*b*) explanation of the ear difference effect as reflecting differences in the efficiency with which material is perceived (in the sense used above) in the two hemispheres can account for all the available data that has been obtained with adequate experimental procedures, provided that we make the assumption that the experimental differences demonstrated between the ears can be attributed to differences between the two hemispheres. No alternative explanation can do so well. What then is the nature of this "perceptual" difference? At what stage in the varied processes of perception do differences between the ears and between the hemispheres appear?

The right ear advantage does not depend on the material being meaningful. Significantly greater scores for the right ear than for the left have been detected in free recall paradigms for initial and final stop consonants (Shankweiler and Studdert-Kennedy, 1967*a,b*) and for laterals and semi-vowels (Haggard, 1969) in a simple nonsense syllable context. The right ear advantage for stops remains when order of report is controlled by a suitable method of scoring (Darwin, 1969), or by pre-instructing order of report (Kirstein and Shankweiler, 1969). However, these experiments do not tell us whether the difference between the ears occurs before or after the sound has been categorized as a particular phoneme. The failure of vowels to give a right ear advantage in free recall (Shankweiler and Studdert-Kennedy, 1967*a*; Darwin, 1969) is not relevant here since vowels differ from consonants in both their acoustic structure and their phonological class. Vowels and consonants could have different ear asymmetries at some level *after* they have been classified as phonemes. This paper examines whether there are differences between the ears in some perceptual process which occurs *before* classification of a sound as a phoneme.

Analytically, the sounds of speech form a sub-set of the sounds of the environment since they are subject to phonetic constraints deriving from the anatomy and physiology of the vocal tract and to phonological and allophonic constraints imposed by particular languages. Maximum efficiency in perception will only be obtained if these constraints are utilized. However, to preserve the efficient perception of sounds not subject to these constraints, some functional division is required in the perceptual system so that one part may deal with the special problems of speech whilst the other remains free to deal with the remaining sounds.

The phonetic constraints are of two main types, both of which lead to a complex relationship between the perceived phoneme and the acoustic signal. In one case a complex relation arises because the articulatory specifications for some phonemes are incomplete (for example, only a general movement of the lips and jaw is specified for bilabial stops); the articulators which are not specified can then assume a wide variety of positions with a correspondingly wide variety of acoustic sequelae. In the second case the complex relation arises from the variation in size and shape of the vocal tracts producing the sound. The first set of relations has been extensively studied, and the word "encoded" has been used (Liberman *et al.*, 1967) to describe this particular lack of acoustic invariance. The second type of variability has received relatively little study. However, the relationship is not likely to be a simple one since, for example, women's vocal tracts are not only smaller than men's, but have different relative proportions (Chiba and Kajiyama, 1941). So when a vowel is spoken by two different individuals with the same articulatory gestures, the formant frequencies for one cannot, in general, be obtained by multiplying each formant frequency of the other's by a constant multiple. This multiple varies between speakers, between vowels and between individual formants (Mattingly, 1966; Fant, 1966). The perceptual system at least partially compensates for these perturbations, since it can accommodate some independent variation in the range of the first two formants (Ladefoged and Broadbent, 1957).

These are by no means the only problems for the speech recognition system, but as they are specific to speech they offer the opportunity of separating speech and non-speech perceptual mechanisms, and asking whether they are equally the prerogative of the two ears and of the two hemispheres. The first experiment asks whether the ear advantage is the same for sounds perceived as the same phoneme, but requiring to different extents Liberman's "decoder", while the second experiment asks the same question of vowel sounds from different sized vocal tracts.

Experiment I. Fricatives

Fricatives are well suited to the purpose of this experiment since there are two main cues which contribute to their perception. The first, and perceptually most significant is the spectral peak of the friction itself (Harris, 1958; Heinz and Stevens, 1961); this peak shows relatively little variation with vowel context. A secondary cue is the formant transitions to adjacent vowels. These show much more contextual variation with vowel context as they depend on the shape of the whole vocal tract. In both voiced and unvoiced fricatives they only assume a major role in distinguishing /f, v/ from θ, ð/, although they do contribute to the intelligibility of the other distinctions. Fricatives synthesized with appropriate formant transitions are generally more intelligible than those synthesized without, although the latter are still highly intelligible provided that the /f,v/-/θ, ð/ distinction is not required.

Liberman *et al.* (1967) hypothesized that only those aspects of speech which show appreciable contextual variation give a right ear advantage. This predicts that fricatives containing the appropriate formant transitions will show a right ear advantage, while those without such transitions will not.

Method

The experimental tape was prepared on the Haskins parellel formant synthesizer. The six fricatives /f, s, ʃ, v, z, ʒ/ were used in the syllabic frame /-ɛp/. /θ/ and /ð/ were not used because they are highly confusable with /f/ and /v/, respectively. There were four stimulus conditions:

(1) with appropriate friction and appropriate formant transitions;
(2) as (1) but with an instantaneous transition into the vowel, which was extended to occupy the time previously allocated to the transition;
(3) as (2) but with the vowel deleted, leaving only the steady-state friction;
(4) as (1) but without the friction, leaving formant transitions and vowel. This condition sounded like plosives rather than fricatives.

The steady-state friction lasted 45 msec, the transitions 30 msec, and the final syllable 120 msec.

The sounds were assembled into a dichotic tape, using a computer program (Mattingly, 1968) which first laid down marker pulses on the recording tape, and then synthesized utterances in a predetermined sequence as the marker pulses were detected. This method allows individual dichotic pairs to be aligned almost perfectly in time, while the use of synthetic speech allows accurate control of the amplitudes and duration of the sounds.

Each sound was paired twice with every *other* sound in its own stimulus condition to give a basic experimental tape of 240 trials, the second half of which was the same as the first, but with the trial order reversed. This whole experimental tape was taken by each subject twice. Prior to the main experiment the subjects were practiced in identifying the sounds with the following letters: f, s, sh, v, z, j, p, b, d. A pilot experiment showed that the letters p, b, d were most readily assigned to the quasiplosives which constituted condition (4). This condition was not basic to the purpose of the experiment, but was included in case none of the fricative conditions gave a significant ear advantage. When the subjects were scoring above 75% on these single sounds they were given 10 practice trials with dichotic pairs. They were told to write down the two sounds they heard putting their more confident choice first. They could write down the same response twice if they wished. They were asked to try to maintain a neutral attention before each trial, rather than listen for only one ear. After the ten practice trials, if they had no questions and had not obviously flouted the instructions, they went on to the main test trials, which came in

16 blocks of 30 trials. Half the subjects started with the headphones reversed, and all subjects reversed their headphones after every 4 blocks.

The experiment was taken by 1 left-handed and 13 right-handed undergraduate and graduate subjects. No subject had any hearing defect to the best of his knowledge, and none had a difference of more than 5 dB between the ears for the threshold at 1500 Hz measured by the method of limits.

Results

Statistical tests are taken from Siegel (1956) and are all two-tailed. Unless otherwise stated, the test used is a Wilcoxon *T*-test for matched pairs. The overall percentages correct for the first and second responses together are given in Table I.

TABLE I

Percentages correct for total scores on both responses by stimulus condition

	Stimulus condition			
	1	2	3	4
Ear	Friction transition vowel	Friction vowel	Friction	Transition vowel
Left	47·4	44·4	45·4	58·6
Right	53·2	46·6	46·6	63·6
Right − left	5·8	2·2	1·2	5·0
Right + left	50·3	45·5	46·0	61·1

A Friedman analysis of variance on total right minus left ear scores between the four stimulus conditions is significant ($P < 0·01$). The total score on the right ear is significantly higher than that on the left for condition (1) ($P < 0·01$), and condition (4) ($P < 0·05$), but not for either condition (2) or (3) ($P > 0·1$). This picture holds both with and without the left-handed subject. Condition (1) gives a significantly greater right ear advantage than either condition (2) ($P < 0·02$) or condition (3) ($P < 0·1$). Adding formant transitions thus increases the score more on the right ear than on the left. The left-handed subject shows a large effect in the opposite direction with conditions (2) and (3) showing a greater right ear advantage than condition (1). He is omitted from all remaining statistics.

The total scores show a very significant tendency for the right ear to score higher on condition (1) than on condition (2) ($P < 0·001$), but only a slight tendency for the left ear to do so ($0·1 > P > 0·05$). A similar pattern prevailed between conditions (1) and (3) but not between conditions (2) and (3). Performance on the right ear is significantly better when formant transitions are added, whilst that on the left ear is not. Thus only the right ear can utilize the additional information present in the formant transitions effectively.

Since the preceding analysis has been made in terms of simple percentage correct scores the differences found between the various stimulus conditions may be due partly to changes in preferred order of report, although it is difficult to think of any interesting reason why this should be so. To counter this objection, however, a scoring system was devised which compensated for order of report effects. These "D-scores" are described in the Appendix. D_1 scores reflect the first channel reported and D_2 the second. A positive D-score indicates a right ear advantage.

TABLE II

Mean D *scores for fricatives by stimulus condition*

	Stimulus condition		
	1 Friction transition vowel	2 Friction vowel	3 Friction
D_1	−0·253	−0·050	0·109
D_2	0·161	0·072	0·022
$D_2 - D_1$	0·092	0·122	0·087

Positive D score indicates right ear advantage, subscript denotes order of report.

A Friedman analysis of variance on the D_1 scores is almost significant ($0{\cdot}1 > P > 0{\cdot}05$), but fails significance on the D_2 scores ($P > 0{\cdot}1$). The significance level of individual Wilcoxon T-tests on these scores is therefore not reliable. The following significance levels are given, however, as an indication of the pattern of the results. The important differences, those between condition (1) and conditions (2) and (3), respectively, appear large and show apparent significance levels of less than 0·025 for the D_1 scores. As in the percentage correct analysis there is a large difference between conditions (1) and (2) for the right ear scores ($P < 0{\cdot}002$), but a small one for the left ear scores ($P > 0{\cdot}1$).

Although the D scores are too variable to allow these significance levels to be accepted, the overall pattern of results is almost identical to that of the percentage correct scores. Since the D scores compensate for order of report effects, it is unlikely that the significant patterns seen in the percentage correct scores are attributable to a change in order of report preferences. It seems more probable that the D scores are inherently more variable than the simple percentage correct from which they are derived.

In summary, a similar pattern of results is obtained with both simple percentage correct scores and a more complicated score which makes some compensation for the order in which the two ears are reported and the overall level of performance. The right ear advantage is greater when appropriate formant transitions are present than when they are absent. The presence of a succeeding vowel in the absence of formant transitions, however, does not appear to influence the ear advantage. The ear difference effect is thus not simply a function of the recognition response class, but is also influenced by the particular cues used to achieve a given response. Moreover, the results are as predicted by Liberman *et al.*'s encoding hypothesis in that only those sounds with formant transitions show a right ear advantage.

So far in this analysis we have taken as correct a response which has both the appropriate voicing and place of articulation. It is of some interest to see whether there are ear advantages for these two dimensions independently. There is convincing psychological evidence that the traditional phonetic feature system is implicated in processes of perception (Miller and Nicely, 1955) and short-term memory (Wickelgren, 1966). If the ear difference indeed reflects differences in the perceptual efficacy of the two ears, these differences may be present not only for the perception of the phoneme as a whole, but also for the perception of its constituent features.

In a dichotic listening experiment using stop consonants, Halwes (1969) found that a large proportion of errors arose from a failure to combine features correctly rather than from a failure to extract them. Many "incorrect" responses in Halwes' experiment consisted of a feature from one ear combined with a feature from the other ear. Perhaps when, as in this fricatives experiment, a correct, response is scored only when both voicing and place of articulation are correct, the ear difference is due to a difference in the efficiency with which the two features are combined into a response rather than to any differences in the efficiency with which they are actually extracted. If this were entirely the case we would expect there to be no residual ear difference if the ear effects for the two dimensions are assessed separately. On the other hand, it is possible that there are differences between the ears in the efficiency with which the features are actually extracted, in which case we *would* expect ear differences when we analyse the features separately.

The results of the fricatives experiment were accordingly scored to provide separate analyses of the voicing and place of articulation dimensions. The dimension not under consideration was made irrelevant both in the stimulus and in the response. This procedure is necessary if the analyses of the two dimensions are to be truly independent.

Analysis of place of articulation was carried out in terms of overall percentage correct, making voicing irrelevant in both the stimulus and the response. A Friedman analysis of variance gave a significant overall variation over stimulus conditions for right minus left ear percentage correct scores ($X_r^2 = 7{\cdot}55$, $df = 2$, $P < 0{\cdot}05$). As in the main analysis the only condition to show a significant right ear advantage was the first, that which had friction and formant transitions ($T = 4\frac{1}{2}$, $n = 13$, $P < 0{\cdot}005$). Neither group 2 nor group 3 showed a significant right ear advantage ($P > 0{\cdot}1$). There was a significant difference between the first group and the average of the other two in this respect ($T = 14\frac{1}{2}$, $n = 13$, $P < 0{\cdot}05$). Analysis in terms of D scores was not made because of the large variance with only three response alternatives.

For the voicing dimension the only trials which contribute differentially to the ear difference are those on which the two stimuli have different voicing, but the two responses have the same voicing. Only one of the stimuli has then been incorporated into the response. A Friedman analysis of variance on the difference between right and left ear incorporations of voicing for the three fricative conditions is significant ($X_r^2 = 7{\cdot}0$, $df = 2$, $P < 0{\cdot}05$). Individual T-tests show that voicing is incorporated more often from the right ear than from the left in both the first ($T = 13\frac{1}{2}$, $n = 12$, $P < 0{\cdot}05$) and the second ($T = 11$, $n = 12$, $P < 0{\cdot}05$) stimulus conditions (the two with the succeeding vowel). There is no significant right ear advantage for the third condition with the isolated friction ($T = 20\frac{1}{2}$, $n = 11$, $P > 0{\cdot}1$). There is a significant difference between groups 2 and 3 in this respect ($T = 12$, $n = 13$, $P < 0{\cdot}02$) but not between any of the others. Combining the first two groups gives a highly significant advantage for the right ear ($T = 1\frac{1}{2}$, $n = 12$, $P < 0{\cdot}002$) and a significant difference between their mean and the third group ($T = 10$, $n = 11$, $P < 0{\cdot}05$). Thus the voicing dimension is reported more accurately from the right than from the left ear only when there is a succeeding vowel.

For fricatives there is thus a dissociation between the stimulus conditions necessary to give a right ear preference for place of articulation and those necessary to give one for voicing. Formant transitions are necessary for the former but a succeeding vowel suffices for the latter. However, these conclusions must be qualified by their possible contamination with changes in order of report preferences since they are based on an analysis of percent correct scores.

Discussion

The main result of this experiment is that the right ear advantage is not determined solely by the recognition response, but is also influenced by the particular sound used to achieve that response. This appears to be true both for the phonetic response as a whole and for the individual articulatory features which constitute that response. Moreover, the particular acoustic signals which must be present for voicing or for place of articulation to show a right ear advantage are different. For place of articulation appropriate formant transitions must be present, whilst for voicing a succeeding vowel suffices. This dissociation suggests that the difference between the ears is occurring before or during the classification of the sound into features, and that it is not simply a consequence of an overall ear difference for the phonemic response. In particular the presence of a right ear advantage for voicing under condition 2, when there is no overall right ear advantage for the entire phoneme argues that the ear difference for the individual features is not a consequence of the ear advantage for the entire response, but rather that the ear advantage for particular features logically precedes that for the entire response.

If differences between the *ears* are not simply a function of response class, can the same be said of differences between the *hemispheres*? Unfortunately, no. An important assumption in the interpretation of ear differences is that there is a functional decussation of the auditory pathways. Although there is electrophysiological evidence which shows a statistical decussation in sub-human species both for evoked potentials (Tunturi, 1946; Rozensweig, 1951) and for single unit recording (Hall and Goldstein, 1968), the main evidence we have that this decussation is both present in man and sufficient to reveal inter-hemispheric differences is the results of dichotic listening experiments. The most convincing demonstration occurs in patients with a section of the corpus callosum. These patients can report verbal material equally well from either ear when only one ear is stimulated at a time, but can report practically nothing from the left ear when similar verbal material is played simultaneously into both ears (Milner, Taylor and Sperry, 1968). Moreover, this weakening of the left ear response is dependent on the nature of the sounds in the other ear. As the sounds in the right ear are progressively distorted, performance on the left ear improves (Sparks and Geschwind, 1968).

Normal subjects show much smaller ear differences than the commissurectomized patients when undistorted digit sequences are played in both ears (Milner *et al.*, 1968; Kimura, 1961*b*). Normal subjects also show an ear difference effect which is dependent on the nature of the competing stimulus. Initial and final plosive consonants give a reliable right ear advantage when they are opposed by another such consonant (Shankweiler and Studdert-Kennedy, 1967*b*); however, plosive consonants embedded in a nonsense word and opposed by white noise give no ear difference (Corsi, 1967). An unpublished experiment by the present author showed no ear difference between the ears using initial plosives rather than embedded ones, with noise on the other ear. Thus the ear difference effect *is* influenced by the nature of the competing stimulus.

The simplest explanation of these effects is that in normal subjects considerable information about the sounds on the left ear can be transmitted across the commissures to the left hemisphere. The commissurectomized patients, being deprived of this path, must rely entirely on the direct ipsilateral path. The efficiency of this latter path is critically dependent on the nature of the sounds on the two ears. With no sound on one ear, it can function well, but as progressively less distorted speech is introduced on the other ear, it becomes less and less efficient.

A significant difference between scores from the two ears can be interpreted as showing that there is some difference between the hemispheres, and that the sounds on each ear have gone predominantly to their opposite hemispheres. However, if there is no significant difference between the ears, we cannot attribute this failure with any confidence to either an equivalence of the two hemispheres or to a failure of the relevant pathways to decussate sufficiently to reveal an inter-hemispheric difference. The differences in ear advantage between the various stimulus groups reported in this experiment could then be due either to a difference in the degree to which the two hemispheres are implicated in their processing, or to a difference in their abilities to produce a functional decussation of the relevant pathways. We can only conclude that the former is true, and thus that the hemispheres differ in their ability to classify phonemes if we have independent evidence that those sounds which did not give a right ear advantage were *in principle* capable of revealing any inter-hemispheric difference that there might have been.

All the sounds which failed to give an ear advantage for a particular feature in this experiment had a steady state along the physical dimension relevant to that phonetic feature. Thus place of articulation only shows an ear advantage when it is cued by a moving pattern of formant transitions, while the voicing feature only shows an ear advantage when it is cued by a sound which may be only partially voiced. Perhaps no steady-state discrimination can give an ear difference. The absence of any ear difference for steady-state vowels, whether in CVC context or in isolation (Shankweiler and Studdert-Kennedy, 1967*a*,*b*), and of very brief duration (Darwin, 1969) supports this idea. Furthermore, Darwin (1969) found only tenuous evidence for a left ear advantage for recall of steady-state non-verbal timbres similar to those whose discrimination was more impaired after right, than left temporal lobectomy (Milner, 1962). If an ear advantage can be demonstrated for steady-state sounds, we will have more justification for assuming that the steady-state sounds used in this fricatives experiment were in principle capable of showing ear differences.

We must now face the logical difficulty that without further assumptions we cannot tell whether any change made in the stimulus conditions which produces an ear advantage is having its effect through changing the conditions necessary to reveal differences between the hemispheres, or through changing the nature of the task in such a way as to implicate mechanisms for which the hemispheres do in fact differ.

One reasonable assumption is that the functional decussation of the auditory pathways is determined only by the particular sounds which are presented on any one trial, and is not influenced by the range of sounds which may occur in the experiment. In other words, if we know from the fact that they give an ear advantage that there is good decussation for a particular dichotic pair of sounds in one experiment, we can remove some of the other dichotic pairs from the experiment without changing the functional decussation for that particular pair. In contrast, the number of different dichotic pairs used in an experiment will generally alter the complexity of the task, and so perhaps alter the relative contribution of either hemisphere. If, then, we can show that greater ear advantages can be obtained for some sounds when the number of different stimuli used in the experiment is changed, we might assume we are measuring a change in inter-hemispheric ability rather than a change in the functional decussation of the auditory pathway.

If, then, the steady-state sounds used in this, and other experiments, have failed to show any ear difference solely because of inadequate functional auditory decussation, we should not expect such sounds to show an advantage when only the complexity of the perceptual discrimination is changed. The next experiment attempts to demonstrate that the ear advantage *is* influenced by the complexity of the perceptual discrimination by changing the range of vocal tract sizes that a set of vowels can come from.

Experiment II. Vowels from Different Sized Vocal Tracts

There is a rough correlation between voice pitch and formant frequencies, since women and children have higher voices and smaller vocal tracts than men. This correlation is utilized in estimating vocal tract size (Fujisaki and Kawashima, 1969). A recent experiment by Haggard (1971a) shows that when vowel perception depends on the fundamental frequency of the vowel, there is a right ear advantage under free recall conditions. Steady-state sounds are here showing a right ear advantage, when there is a difference in pitch between the two ears. Unfortunately for the present argument, this difference in pitch is a reasonable candidate for a factor which changes the conditions necessary to reveal the ear difference effect, as well as one which alters the perceptual complexity of the task. Can we show a right ear advantage for steady-state vowels which have the same pitch on either ear? The most direct way to answer this question is to use sets of vowels from two different sized vocal tracts.

Method

The five vowels /i, ɛ, æ, a, ʌ/ in the context /ən-t/ were synthesized on the Haskins parallel formant synthesizer using only the first two formants. Two sets of these five words were made, the formant frequencies for one set being 25% higher than those for the other set. The formant values are given in Table III.

Two different experimental tapes were then constructed. On one tape, each sound was paired with every other sound except itself and its phonemic homologue from the other vocal tract. On the other tape, only the sounds from the smaller vocal tract were used, and each sound was paired with every other sound except itself. The first tape had 160 trials and the second 40. The order of the trials on the second tape was exactly the same as the order of those trials on the first tape in which both sounds came from the smaller vocal tract.

The first tape was taken twice by one group of 18 subjects, and the second tape was taken twice by a second group of 18 subjects. All subjects were right-handed, native speakers of American English, who to the best of their knowledge had no hearing defects. The instructions and training they received were similar to those used in the fricatives experiment. The words used to identify the sounds were "a nit, a net, a gnat, a knot, a nut", and both groups of subjects used the five letters "i, e, a, o, u" as their responses. Those who took the first tape had training in identifying the sounds from both vocal tracts, whereas the second group of subjects were only introduced to the sounds from the smaller vocal tract. The usual counter-balancing procedures were observed.

TABLE III

Formant frequencies for vowels in Experiment II

Vowel	Large vocal tract F1	F2	Small vocal tract F1	F2
/i/	386	2078	489	2540
/ɛ/	537	1845	666	2307
/æ/	666	1695	844	2156
/a/	718	1075	894	1312
/ʌ/	640	1232	794	1541

Results

Five stimulus conditions are distinguished in the results. Four come from the first group of subjects and correspond to whether the dichotic pair had sounds from (1) the larger vocal tract only; (2) the smaller vocal tract only; (3) the larger on the left ear and the smaller on the right; (4) the smaller on the left and the larger on the right. The fifth condition corresponds to the second group of subjects who had the smaller vocal tract on both ears all the time. The overall percentage correct and the *D* scores are given in Tables IV and V, respectively.

The overall superiority for the right ear for the first group of subjects (summing over the first four stimulus conditions) is significant both on percentage correct ($P < 0{\cdot}001$) and on D_1 scores ($P < 0{\cdot}01$). For the second group of subjects there is no significant right ear advantage on either score ($P > 0{\cdot}1$).

A Friedman analysis of variance over the first four stimulus conditions is significant for differences in percentage correct (P <0·01) and D_1 scores (P <0·02). The variation in overall level of performance, however, is barely significant (P <0·1). Individual Wilcoxon T-tests show that ear differences are significant for the first and second stimulus conditions separately on overall percentages correct (P <0·01) and on D_1 (P <0·05 and <0·06, respectively).

For conditions (3) and (4) combined, when the two-ears had different vocal tracts, the right ear did significantly better than the left (P <0·002) but there was also a significant tendency for the vowels from the small vocal tract to be recalled better than those from the larger (P <0·05). This difference is not present when the two ears receive vowels from the same vocal tract as in conditions (1) and (2). It is not then due to markedly poorer intelligiblity for the smaller vocal tract.

TABLE IV

Overall percentages correct in Experiment II by dichotic pair composition

Vocal tract size on Left ear	 Right ear		Overall percentage correct on Left ear	 Right ear	 Right-left	 $P(L=R)$
Large	Large	(1)	46·7	50·9	4·2	<0·01
Small	Small	(2)	45·8	50·4	4·5	<0·01
Large	Small	(3)	45·7	56·2	10·6	<0·002
Small	Large	(4)	54·0	51·2	−2·8	
Total			48·1	52·2	4·1	<0·001
Small	Small	(5)	54·0	53·4	−0·6	>0·1

TABLE V

D *scores for Experiment II by dichotic pair composition*

Vocal tract size on Left ear	 Right ear		D_1	D_2	$p(D_1 = 0)$	
Large	Large	(1)	0·083	0·078	<0·06	<0·002
Small	Small	(2)	0·103	0·067	<0·05	
Large	Small	(3)	0·226	0·206		<0·002
Small	Large	(4)	−0·062	−0·089		
Small	Small	(5)	−0·060	−0·013	>0·10	

There is a significantly greater right ear advantage for the vowels in condition (2) than in condition (5) both for percentages correct (P <0·05) and for D_1 scores (P <0·02) on Mann-Whitney U-tests. But there is no difference between the averages of conditions (1) and (2) vs. conditions (3) and (4) (P >0·1). In other words, the right ear advantage for vowels in this experiment depends on the nature of the discrimination within the framework of the whole experiment rather than within the individual trial.

A reliable right ear advantage for steady-state vowels, therefore, can be obtained when there is uncertainty within the experiment as to what size vocal tract has produced them. But this right ear advantage is not influenced by whether on a particular trial the two alternative sizes of vocal tract are in fact present.

Discussion

This experiment confirms Haggard's finding that vowels can give a right ear advantage. Whether the advantage appears or not in this experiment depends on the complexity of the perceptual discrimination, rather than on the particular sounds used on any one trial. On the assumption that the sounds used for the second group of subjects were in principle capable of showing a right ear advantage, we can conclude that the hemispheres do differ in their ability to classify vowels from different sized vocal tracts. This assumption seems reasonable, since identical sounds did give a right ear advantage when played to the first group of subjects, as part of a larger experiment.

The assumption that was necessary in interpreting the results of the fricatives experiment in terms of differences between the two hemispheres has received some justification, since the vowels used here are cued mainly by a steady-state. More direct confirmation of this could perhaps be obtained by using steady-state friction from different sized vocal tracts.

Can we draw any conclusions about the stage or stages in perception at which ear or hemisphere differences become apparent? The ear difference effect is not solely a function either of the stimulus, or of the response, but rather of the processes which must mediate between the two. The fricatives experiment showed that it did not depend on the response category alone, since whether or not it appeared either for the entire phonetic response or for one of the constituent dimensions of voicing and place of articulation depended on the presence of particular acoustic cues. The vowel experiment described here shows that the effect does not depend solely on either the stimuli presented on a particular trial, or on the response category, since the same stimuli do or do not show a right ear advantage depending on the complexity of the relationship between the stimuli and the responses.

A similar conclusion has been reached by Studdert-Kennedy and Shankweiler (1970) on the basis of a feature analysis of a dichotic experiment with stop consonants. They, with Halwes (1969), find that a large proportion of errors arise from inappropriate combination of correctly extracted features. They suggest that this arises because acoustic features can be extracted correctly in either hemisphere, but that they can only be related to phonemic features and assembled into a phonemic response in the left hemisphere.

More direct evidence that particular acoustic features themselves are not entirely responsible for the ear difference effect comes from an experiment by Haggard (1971b). Haggard shows that when the voicing dimension is cued only by a change in pitch (Haggard, Ambler and Callow, 1970) in a dichotic listening paradigm, the recall of this feature shows a right ear advantage. Since Darwin (1969) has shown that simple pitch sweeps give a *left* ear advantage when carried on a word, but do not cue a phonemic distinction, it seems likely that the pitch sweeps which cued voicing in Haggard's experiment would show a left ear advantage in a suitable non-speech context. Here, then, it is not the extraction of the acoustic cue which is important, but its phonetic relevance.

The existence of some stage which mediates between an acoustic representation of the input stimulus and the phonetic output has been suggested by Hiki *et al.* (1968) on the basis of experiments on a short-term contrast effect in vowel perception (Fry *et al.*, 1962). They suggest that there is some transform which maps acoustic space into a multi-dimensional phonetic space from which decisions are made about the appropriate phonetic category. The nature of this transform is determined both by the short-term effects that they investigated and by the longer term normalization effects demonstrated by Ladefoged and Broadbent (1957).

The arguments put forward here have concentrated on identifying the earliest stage at which differences between the ears become apparent. This is not necessarily the only stage, or that at which the greatest differences may be obtained. Work on temporal lobectomized patients has shown large differences between the two temporal lobes for verbal memory in excess of the short-term memory span (Milner, 1958), but there has been considerably less evidence that verbal perceptual deficits depend on which hemisphere is damaged. Luria (1966) presents some evidence that patients with damage to the left temporal lobe are impaired in their ability to repeat simple nonsense syllables. But this is the only evidence of its kind. The work on commissurectomized patients has given no evidence that there are any perceptual differences between the two hemispheres (Milner, Taylor and Sperry, 1968), although, of course, recall is largely restricted to only one hemisphere. Perceptual differences may in fact exist at the level of phonemic analysis, and these differences may not yet have been revealed because few tests have put strain specifically on the phonetic aspects of speech perception. That no effects, other than those reported by Luria, have yet appeared does suggest that the lateralization of speech perception is considerably less than that of speech production and verbal memory. This does not necessarily mean that these latter processes are influencing the results of the experiments reported here. It may well be that the dichotic listening technique is particularly sensitive to processes which occur early in the sequence of perception and memory, if only because stimuli are more likely to be differentiated according to ear of arrival immediately after input than at some later time. We must, however, acknowledge the possibility that memory processes may show differential ear effects, although there is yet little evidence that they do.

Appendix: *D*-Scores

In a free recall dichotic listening experiment, the simple percentage correct score is inadequate for two reasons. First, it takes no account of the relative number of times one ear is reported first and the other second, so that errors arising from serial order effects are confounded with those from other sources. Second, differences in percentage correct are not strictly comparable between subjects because of varying overall levels of performance; a given difference in detectability gives rise to a wide range of differences in percentage correct at different performance levels. The two *D* scores described here give estimates of the difference in recall between the two ears on the first and second reported channels, respectively. These estimates take into account both the relative number of times each ear is reported first, and the absolute probability of being correct on each of these channels.

First and second channel here refer simply to the order of report rather than to any property of the input. The following letter combinations denote the number of trials on which each subject made the corresponding pattern of correct responses.

LR = left ear correct on first channel, right ear correct on second channel.
RL = right ear correct on first channel, left ear correct on second channel.
LZ = left ear correct on first channel, neither ear correct on second channel.
RZ = right ear correct on first channel, neither ear correct on second channel.
ZL = neither ear correct on first channel, left ear correct on second channel.
ZR = neither ear correct on first channel, right ear correct on second channel.
ZZ = neither ear correct on first channel, neither ear correct on second channel.

Then let:

$$\begin{aligned} p(\mathrm{L}_1) &= (\mathrm{LR} + \mathrm{LZ})/(\mathrm{LR} + \mathrm{LZ} + \mathrm{ZR}) \\ p(\mathrm{R}_1) &= (\mathrm{RL} + \mathrm{RZ})/(\mathrm{RL} + \mathrm{RZ} + \mathrm{ZL}) \\ p(\mathrm{L}_2) &= (\mathrm{RL} + \mathrm{ZL})/(\mathrm{RL} + \mathrm{RZ} + \mathrm{ZL}) \\ p(\mathrm{R}_2) &= (\mathrm{LR} + \mathrm{ZR})/(\mathrm{LR} + \mathrm{LZ} + \mathrm{ZR}) \end{aligned}$$

Denoting a normal transformation with a prime we now define

$$\begin{aligned} D_1 &= p'(\mathrm{R}_1) - p'(\mathrm{L}_1) \\ D_2 &= p'(\mathrm{R}_2) - p'(\mathrm{L}_2) \end{aligned}$$

This scoring method ignores trials on which neither ear was correct (ZZ), and assumes that making a normal transformation is an adequate compensation for variations in overall performance level (Green and Birdsall, 1964).

The first experiment reported here was performed at the Psychological Laboratory, Cambridge, under the supervision of Dr D. E. Broadbent, while I held a Medical Research Council Studentship. I was also supported by the Commonwealth Fund and by a grant from the National Institute of Child Health and Development to Haskins Laboratories.

This paper would not have been possible without the help of Drs M. P. Haggard, A. M. Liberman, I. G. Mattingly, D. P. Shankweiler, M. Studdert-Kennedy and T. G. Halwes. Professor O. L. Zangwill and Dr F. S. Cooper generously provided research facilities.

References

Bakker, D. J. (1968). Ear-asymmetry with monaural stimulation. *Psychonomic Science*, **12**, 62.

Bakker, D. J. (1970). Ear-asymmetry with monaural stimulation: relations to lateral dominance and lateral awareness. *Neuropsychologia*, **8**, 103–17.

Broadbent, D. E. (1958). *Perception and Communication.* London: Pergamon.

Broadbent, D. E. and Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology*, **16**, 359–60.

Bryden, M. P. (1963). Ear preference in auditory perception. *Journal of Experimental Psychology*, **65**, 103–5.

Bryden, M. P. (1967). An evaluation of some models of laterality effects in dichotic listening. *Acta oto-laryngologica*, **63**, 595–604.

Chiba, T. and Kajiyama, M. (1941). *The Vowel—Its Nature and Structure.* Tokyo.

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

Darwin, C. J. (1969). Auditory perception and cerebral dominance. Unpublished Ph.D. thesis, University of Cambridge.

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scalings. *STL-QPSR*, **4**, 22–30.

Fry, D. B., Abramson, A. S., Eimas, P. D. and Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, **4**, 171–89.

Fujisaki, J. and Kawashima, T. (1968). The roles of pitch and higher formants in the perception of vowels. *Institute of Electrical and Electronic Engineers Transactions*, AU-16, 73–7.

Green, D. M. and Birdsall, T. M. (1964). The effect of vocabulary size on association score. In Swets, J. A. (Ed.) *Signal Detection and Recognition by Human Observers*, pp. 609–19. New York: Wiley.

Haggard, M. P. (1971a). Encoding and the REA for speech signals. *Quarterly Journal of Experimental Psychology*, **23**, 34–45.

Haggard, M. P. (1971b). Stimulus and task factors as determinants of ear advantages. *Quarterly Journal of Experimental Psychology*, **23** (in press).

Haggard, M. P., Ambler, S. and Callow, M. (1969). Pitch as a voicing cue. *Journal of the Acoustical Society of America*, **47**, 613–7.

Hall, J. L. and Goldstein, M. H. (1968). Representation of binaural stimuli by single units in primary auditory cortex of unanaesthetized cats. *Journal of the Acoustical Society of America*, **43**, 456–61.

Halwes, T. G. (1969). Effects of dichotic fusion on the perception of speech. *Supplement to Status Report on Speech Research*, Sept. 1969, Haskins Labs.

Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech*, **1**, 1–17.

Heinz, J. M. and Stevens, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America*, **33**, 589–96.

Hiki, S., Sato, H. and Oizumi, J. (1968). Dynamic model of vowel perception. Paper presented to 6th International Congress of Acoustics, Tokyo, Japan, August 1968.

Hyde, S. R. (1968). *Automatic Speech Recognition.* Post Office Research Dept. Report No. 45.

Inglis, J. (1962). Dichotic stimulation, temporal lobe damage, and the perception and storage of auditory stimuli—a note on Kimura's findings. *Canadian Journal of Psychology*, **16**, 11–6.

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

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

Kimura, D. (1964). Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, **14**, 355–8.

Kimura, D. (1967). Functional asymmetries of the brain in dichotic listening, *Cortex*, **3**, 163–78.

Kirstein, E. and Shankweiler, D. P. (1969). Selective listening for dichotically presented consonants and vowels. Paper read before 40th Annual Meeting of Eastern Psychological Association, Philadelphia.

Ladefoged, P. and Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, **29**, 98–104.

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

Luria, A. R. (1966). *Human Brain and Psychological Processes.* Trans. B. Haigh. New York and London: Harper and Row.

Mattingly, I. G. (1966). Speaker variation and vocal-tract size. *Journal of the Acoustical Society of America*, **39**, 1219 (A).

Mattingly, I. G. (1968). Experimental methods for speech synthesis by rule. *Institute of Electrical and Electronic Engineers Transactions*, AU–**16**, 198–202.

Miller, G. A. and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, **27**, 338–52.

Milner, B. (1958). Psychological deficits produced by temporal lobe excision. *Publications of the Association for Research in Nervous Mental Diseases*, **36**, 244–57.

Milner, B. (1962). Laterality effects in audition. In Mountcastle, V. B. (Ed.), *Interhemispheric Relations and Cerebral Dominance.* Baltimore: Johns Hopkins.

Milner, B., Taylor, L. B. and Sperry, R. W. (1968). Lateralized suppression of dichotically-presented digits after commissural section in man. *Science, New York*, **161**, 184–5.

Oxbury, S., Oxbury, J. and Gardiner, J. (1967). Laterality effects in dichotic listening. *Nature, London*, **214**, 742–3.

Rosenzweig, M. R. (1951). Representation of the two ears at the auditory cortex. *American Journal of Physiology*, **167**, 147–58.

Shankweiler, D. P. (in press). An analysis of laterality effects in speech perception. In Kjeldergaard, P. (Ed.), *The Perception of Language.* Pittsburgh, Pa.: University of Pittsburgh Press.

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

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

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

Sparks, R. W. and Geschwind, N. (1968). Dichotic listening in Man after section of neocortical commissures. *Cortex*, **4**, 3–16.

Studdert-Kennedy, M. and Shankweiler, D. P. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, **48**, 579–94.

Treisman, A. M. and Geffen, G. (1968). Selective attention and cerebral dominance in perceiving and responding to speech messages. *Quarterly Journal of Experimental Psychology*, **20**, 139–50.

Tunturi, A. R. (1946). A study of the pathway from the medial geniculate body to the acoustic cortex in the dog. *American Journal of Physiology*, **147**, 311–9.

Wickelgren, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America*, **39**, 388–98.

Received 28 *September* 1970