HEMISPHERIC CONTRIBUTIONS TO THE PERCEPTUAL REPRESENTATION OF SPEECH SOUNDS

Henri Cohen

A Thesis
in
The Department
of
Psychology

Presented in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

August 1981

ABSTRACT

HEMISPHERIC CONTRIBUTIONS TO THE PERCEPTUAL REPRESENTATIONS OF SPEECH SOUNDS

Henri Cohen, Ph. D.
Concordia University, 1981

Cerebral hemispheric differences in the perception of speech sounds have been reported by several authors using the dichotic listening paradigm suggesting greater left hemisphere involvement in the processing of speech information. Four experiments, in which a variant of the dichotic listening technique was used, are reported that examined the participation of each cerebral hemisphere in the perception of speech. These experiments investigated the relationship between properties of speech sounds and their psychological representation as given in a multidimensional scaling analysis. Perception of consonants was investigated in the first two experiments and perception of vowels in the third. The results suggested that it is the case, both for consonant and vowel perception, that hemispheres preferentially managed defining features of speech sounds with no particular hemisphere dominating perception of a whole class of speech sounds. For consonants, the right hemisphere was found to be more involved than the left in processing voicing while the left hemisphere was more involved in the treatment of place information. For vowels, canonical correlation analysis results suggested greater right hemisphere involvement in F2 perception and greater left hemisphere involvement in F1 perception. The fourth experiment attempted to generalize the findings of the previous experiments and investigated each cerebral hemisphere's contribution to the acquisition of Chinese phonetic categories differentiated on voicing and aspiration. The results showed faster and better differentiation of categories contrasted on voicing and a slower but significant acquisition of the aspiration contrast for the left hemisphere only. These findings suggest active participation on the part of both hemispheres in the treatment of speech information. More importantly, the results showed that perception occurred at the level of the defining features of the stimulus rather than at the level of the stimulus complex.

Acknowledgements

I would like to sincerely express my appreciation to Dr Norman Segalowitz for his encouragement, support and advice throughout the research and composition of this thesis.

I would also like to extend my gratitude to Dr Jane Stewart and to Dr Charles White whose suggestions and criticisms were invaluable.

Thanks are also due to Dr Gilbert Taggart who kindly contributed the English speech stimuli and to Thomas Kao for the Mandarin Chinese speech stimuli and story.

But most of all it is to my wife Suzanne that I am indebted. Without her constant support and encouragement this thesis would not have been possible.

Table of Contents

Page

List of Figures

List of Tables

INTRODUCTION

Major advances have been made in the past two decades in our understanding of speech perception and of speech processing at the phonetic level. One of the fascinating problems that has emerged is the role or roles of the two cerebral hemispheres in the perception of speech. There is uncertainty about many aspects of this issue. For example, one major issue to be resolved is whether the lateralization of certain language functions is the endpoint of a developmental process originating in an initial bisymmetric cerebral representation of early verbal behavior ("progressive lateralization"; Brown & Jaffee, 1975; Krashen, 1973; Lenneberg, 1967) or whether the ultimate lateralization characteristics exist in the same form from the start ("invariant lateralization"; Kinsbourne & Hiscock, 1977). There is also argument about the exact nature of hemispheric differences and about the factors that may influence laterality patterns in both normal individuals and patient populations (Levy & Reid, 1978; Moscovitch, 1979).

Our interest in this thesis is the possibly differential involvement of the left and right hemispheres in processing phonological information. Below is a review of results and methodologies of several of the more recent laterality studies, whose contribution has been to direct attention to the importance of processing strategies and task characteristics in the evaluation of hemispheric differences. Following this review four experiments are reported that examined the participation of each cerebral hemisphere in the perception of speech sounds. In these experiments a variant of the dichotic listening paradigm was used to assess hemispheric involvement in the perception of English and foreign speech sounds. The first two experiments investigated the nature of ear differences in the perception of English consonants while ear differences in vowel perception are evaluated in the third. The fourth experiment addressed the question of the development of each hemisphere's contribution to the perception of new phonetic categories, that is, not in the perceiver's phonological repertoire.

Hemispheric Specialization and Processing Strategy

Traditionally, the lateralization of functions of the brain has been characterized in terms of simple dichotomies. For example, early behavioral research had sought to define hemispheric asymmetries of function in terms of the cortical loci for the processing of fundamentally different kinds of stimuli, such as verbal versus nonverbal material. Thus, when it was shown that right-handed subjects were better able to identify letters flashed in the right visual field, and hence to the left hemisphere, the result was interpreted as reflecting the outcome of an interaction of a verbal processor with a verbal stimulus (Bryden, 1965; Kimura, 1967). Stimuli revealing a right hemisphere advantage proved to be more elusive: a left visual field superiority was reported for facial recognition (Rizzolatti, Umilta, & Berlucchi, 1971) and a left ear advantage for recognition of dichotically presented melodies (Kimura, 1964; but see Bradshaw & Nettleton, 1981). These results were also interpreted as reflections of a match between processor and stimulus. Most functional differences between the hemispheres, however, go beyond specialization for different kinds of stimuli and may include basic differences in the ways processing is carried out at each hemisphere.

Specialization according to stimulus properties

It is becoming increasingly apparent that tasks that might superficially be labeled linguistic or visuospatial are comprised of processing stages that individually engage one or the other hemisphere. Bryden and Allard (1976) asked subjects to orally identify individual letters flashed tachitoscopically to either the right or left side of a fixation point. What varied was the typeface in which the letters were printed. They found that most typefaces were reported more accurately when presented to the right visual field (RVF), the expected effect for a verbal naming task; some typefaces, however, were reported more accurately in the left visual field (LVF), indicating that the right hemisphere was more successful at processing them. The typeface that involved more right hemisphere processing seemed to be more script-like (cursive) and in one case to be subjectively three-dimensional. It seems that the linguistic nature of the letter-naming task was insufficient to insure a left hemisphere dominance. Bryden and Allard argued that the preprocessing stage of visual identification is likely to preferentially involve the RH and that in the case of certain typefaces, this stage was sufficiently difficult that the RH-processing significantly affected the accuracy measures. It can also be concluded that processing specificity is an important aspect of dominance.

Specialization according to task demands

Other investigators have stressed the type of task required of the subject and emphasize less the nature of the stimulus. In a study by Klatzky and Atkinson (1971), pictures of common objects were the stimuli and the subjects' task was to determine if the initial letter of the name of the object belonged to a set of previously presented letters. A right visual field advantage was observed, indicating that the nature of the task attending to the name of the object pictured (a linguistic activity believed to preferentially involve the left hemisphere), was a critical determinant of the type of visual field advantage to be found.

An equivalent reversal of the hemispheric effects that might be predicted from the nature of the stimulus was obtained by Gibson, Dimond and Gazzaniga (1972) who found a left visual field advantage in a word-matching task when subjects could respond simply on the basis of the physical characteristics of the stimuli. In this study, subjects were presented with a single word displayed for 3 secs across both visual fields, followed by a second word flashed for 40 msec on the left or right of fixation. The duration of the stimulus was too brief to permit identification and subjects presumably performed the task of deciding whether or not the two letter strings were identical on the basis of gross configurational properties of the words rather than on the basis of their semantic content.

This effect of task demand is also clearly illustrated in a study by Veroff (1976) who compared the performance of patients with severe RH lesions to those with severe LH lesions on a task of arranging three pictures (place set and category set) in their correct sequence. The groups showed a complete dissociation on these tasks: the LH-lesioned group was both slower and less accurate on the change-of-category sets, while the RH-lesioned group was both slower and less accurate on the change of place sets. What would appear to be a similar task, ordering three pictures in a correct sequence, seems to involve different kinds of cognitive processing specificity which would engage both hemispheres to different degrees.

Results such as these have given rise to the suggestion that subjects can process various stimuli in different ways depending on their "psychological set" (Segalowitz & Gruber, 1977). For example, although a stimulus may be non-linguistic (music) it can be processed by some listeners (musicians) in the left hemisphere while other subjects may process it in the right hemisphere (Gates & Bradshaw, 1977; Bever & Chiarello, 1974). Thus a person's processing strategy may dictate to a certain extent how information is dealt with in the brain.

Speech Analysis and the Cerebral Hemispheres

Although processing strategies seem to contribute greatly to the effects observed in laterality research, it should be noted that there are examples where certain properties of the stimulus seem to be crucial. In a dichotic listening experiment using synthetic fricative-vowel syllables, Darwin (1971) observed that fricatives synthesized with rapid formant transitions showed a right ear advantage, while those synthesized without such transitions failed to reveal any asymmetry. Since the subjects in this study were asked to identify the dichotically presented syllables on each trial, task differences cannot alone explain the results.

One possible interpretation of these results is that the left hemisphere is specialized for the extraction of information in rapid frequency changes, and that stimuli lacking such transitions may be processed in either hemisphere. In this conception the phonetic properties of the stimuli would be more important than th acoustic properties. One would predict from this model a right-ear advantage even for phonetically impossible stimuli as long as they contain formant transitions. Cutting (1974) has reported precisely that. Complications arise, however, since speech stimuli without formant transitions often show a right ear advantage. This has led Cutting (1974) to suggest that there may be two left-hemisphere mechanisms operating on incoming auditory information, one primarily acoustic and the other primarily phonetic. In the first, the analysis of the speech signal results in a corresponding set of nonlinguistic parameters such as the frequency of the signal, its amplitude, and changes in these parameters over time. In the second, the listener presumably focuses on certain aspects of these parameters, or combination of cues, and processes them further in order to reach an organized linguistic description.

There is not unanimous agreement on the need to postulate a specialized left hemisphere processor to deal with the processing of the speech signal. According to motor theory (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), speech perception is mediated by mechanisms of speech production. Whether operating by mechanisms involving feature analysis (for detailed discussion and critique of feature detectors, see Diehl, 1981; Remez, 1979) or motor synthesis, the left hemisphere has traditionally been thought to be endowed with specialized segmentation processors that decipher the continuous speech stream into phonetic elements, track the rapid formant transitions (important cues in the identification of complex speech sounds such as consonants; for a more detailed discussion of the left hemisphere involvement in the encoded aspects of language, see Liberman, 1974) and adjust for differences in vocal tract size (Darwin, 1971; Liberman, et al., 1967). One of the functions of this specialized mechanism within the "language" hemisphere appears to be the conversion of the acoustic signal into a phonetic signal. There is, however, evidence mounting that contradicts such a notion.

Kimura and Folb (1968) claim that backward speech, which does not conserve the characteristics of speech sounds, generates a right ear advantage (REA). Allard and Scott (1975) have shown that larger asymmetries favoring the right ear are obtained for speech stimuli from which formant transitions, the very cues necessary for the identification of consonants, have been removed. On the basis of such findings, Allard and Scott have suggested that speech analysis by the left hemisphere may simply be a part of a more general system for segmenting sequential acoustic input. Other researchers have also rejected the construct of a specialized speech mode of perception. Some have suggested that feature detectors are not speech specific (Remez, 1979) and neither are categorical perception (Studdert-Kennedy, 1980) or the perception of the critical voice onset time (VOT) (Molfese & Hess, 1978). Furthermore, it has recently been shown that listeners perceive linguistic significance in acoustic patterns with properties differing widely from those traditionally held to underlie speech perception (Remez, Rubin, Pisoni, & Carrell, 1981). This phenomenon of "speech perception without traditional speech cues" led Remez and his colleagues (1981) to conclude that "speech perception makes use of time-varying acoustic properties that are more abstract than the spectra and speech cues typically studied in speech research (p. 947)".

At a more general level, another serious argument against postulating a specialized speech processor emerges from studies demonstrating that phonetic-like processing can occur for natural events, outside of speech, throughout the auditory domain. Rise time, for example, is a musical dimension that is categorically perceived (Cutting, 1977; Cutting & Rosner, 1974, 1976). Notes produced from stringed musical instruments are identified as either plucked (rapid rising times -0 through 30 msec) or bowed (more slowly rising times -50 through 80 msec). The perception of the plucked-bowed distinction meets all the criteria for categorical perception suggested by Studdert-Kennedy and his colleagues (Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). Furthermore, the plucked-bowed distinction has been shown to be categorically perceived by two month old infants (Jusczyk, Rosner, Cutting, Foard, &smith, 1977). This parallels the infant's categorical perception of consonants (Eimas, Siqueland, Jusczyk, & Vigorito, 1971).

Recent results from electrophysiological correlates of speech perception suggest an alternative viewpoint to the traditional dichotomy of function between the hemispheres in the treatment of speech information. In a series of studies focusing on certain aspects of the speech signal such as voice onset time (VOT) and consonant identification, Molfese (1980a, 1978a, 1978b) recorded auditory evoked responses to the presentation of speech and nonspeech stimuli varying in voice onset time. One of the findings of importance that has emerged from his research is that perception of the VOT cue, in both speech and nonspeech contexts, is controlled by several cortical processes, some of which restricted to the right hemisphere. These data suggest that researchers who have assigned lateralized effects in the discrimination between speech and nonspeech to a left hemisphere speech processor may have failed to apply appropriate stimulus controls (Molfese, 1980a).

Findings such as these raise questions about the role played by phonetic processing in speech perception in a broader sense than it has been considered so far. Some recent data obtained from split-brain subjects suggest that phonetic analysis -whether the output of a special speech processor or not- does not necessarily precede the extraction of meaning from speech. Springer and Gazzaniga (1975) and Zaidel (1978) have studied the capacity of the right hemisphere in commissurotomized adults to process syllables composed of a stop consonant and vowel. In both studies pairs of syllables were presented dichotically in an experimental task designed to optimize processing and output of stimuli presented to the left ear. Left ear performance was; however, at chance level compared to nearly perfect performance of stimuli directed to the right ear. The same subjects who displayed an inability to process the stop-vowel syllables in the right hemisphere were able to respond to the presentation of spoken words and sentences when response choices were restricted to the right hemisphere (Zaidel, 1978; Gazzaniga & Sperry, 1967). Thus the right hemisphere in these patients proved unable to make certin phonetic distinctions, although it did demonstrate the capacity to deal with words and sentences meaningfully. Levy (1974) argues that the inability of the right hemisphere to carry out phonetic analysis is due to its lack of a mechanism "for going from word phonologies in long term memory to articulations". She continues, "whatever phonologies are possessed by the minor hemisphere, they serve for the function of allowing some simple speech to be decoded with respect to meaning and have no interconnection among themselves. The right hemisphere may know... that 'cat' means a furry, small pet with claws, but it does not know that 'cat' rhymes with 'rat'" (p. 149).

It is possible that phonetic analysis may not be necessary to the process of speech understanding in some instances. The right hemisphere may demonstrate a mode of speech perception in which holistic processing plays an important role. The superiority of the right hemisphere for holistic processing in the visual modality has been well documented (e.g. see Levy, 1974), and it is reasonable to suppose that a similar processing strategy might apply when that hemisphere is required to process speech.

It should now be evident that differences between the hemispheres are not limited to specialization for different types of stimuli. Basic differences in the way each hemisphere processes information must also be taken into account. The dichotic listening technique, which is probably the best known and most accepted task reflecting brain lateralization, has been the most frequently employed method in the evalution of these hemispheric functional asymmtries. Most formulations on information processing by the brain derive from results obtained via the dichotic listening method. It seems then that we need further comparisons of the different tasks and response modes in order to better understand differences in hemispheric processing. Thus, in the present research, a variant of the dichotic listening procedure will be used. Before describing the technique of presentation used in this research, a critical examination of the dichotic listening technique is warranted.

Dichotic Listening

Each ear projects both contralaterally and ipsilatrally. This projection, from the level of the superior olivary nucleus upward (Berlin, 1977; Moyer, 1980) prevents lateralization to one hemisphere only in the case of monaural presentation of stimuli.

In dichotic listening, the subject wears headphones and is presented with one message to one ear and an unrelated message to the other. Thus both ears simultaneously receive different messages. The amount of information is so adjusted that the subject has difficulty in correctly reporting all the material. The subject's relative efficiency in reporting input to each ear is computed and if s/he shows an ear advantage, namely better retrieval of messages to one ear than the other, then (in the absence of differential hearing loss as an explanation) language lateralization is attributed to the side contralateral to the ear that has this advantage.

Broadbent (1954) reported that when competing speech messages are presented simultaneously, more messages are processed accurately by the right ear channel than by the left one. Kimura (1961) ascribed this right ear advantage (REA) to the presumed dominance of the left hemisphere for speech functions. These results have been interpreted as being due to suppression of the ipsilateral pathways under conditions of competition (Heilman & Watson, 1977), leaving only the contralateral pathway functional. Further evidence for this interpretation comes also from the work with commissurotomy patients. These patients experience no difficulty under monaural presentation, but are unable to report most of the left ear inputs, presumably because these inputs project only to the "mute right hemisphere" (Milner, Taylor & Sperry, 1968).

Variations of the dichotic listening technique

Researchers have sometimes found it necessary to alter the basic dichotic listening paradigm. This was usually done with the purpose of answering questions that the technique, in its original formulation, could not have apparently done or of achieving a better understanding of dichotic listening itself (e.g., Berlin, 1976). In one technique, for example, the subject is asked to report first from one ear and then from the other, the order of report being counterbalanced within the session. In another, the subject monitors both ears for one or more target stimuli. In a third, the subject knows which target stimuli to expect but is asked to lateralize them (Kinsbourne & Hicks, 1977). Researchers find it possible, in another technique, to establish asymmetries by monaural input, comparing the efficiency of retrieval from each of the two ears over a large number of trials (Bever, 1971; Morais & Darwin, 1974).

Reaction time has also been used in the study of hemispheric asymmetry of function. Its application is illustrated in a study of ear asymmetry where CV syllables are opposed by contralateral noise (Springer, 1973). Previous studies using percent correct identification as a dependent measure failed to reveal any ear asymmetries in this task (Darwin, 1971). In Springer's studies, no differences in the correct percent identification as a function of ear were found, but the reaction time measure revealed a 14-msec advantage in favor of the right ear. Reaction time has been extensively used in tachistoscopic studies (Geffen, Bradshaw, & Wallace, 1971; Gross, 1972; Moscovitch, 1972; Segalowitz & Hansson, 1979) but in only a limited way in dichotic research (Springer, 1973; Kallman & Corballis, 1975). The use of reaction time as a dependent measure in dichotic listening may thus reveal patterns of asymmetries previously undetected with other paradigms (e.g., percent correct recognition).

It appears also possible that in studies in which hemispheric asymmetry is examined with tasks involving motor production (oral reading, writing, verbal report) there is more left hemispheric activation due to motor programming. The implication is that studies using a "recognition" paradigm where subjects are required to verbally report what they have heard in a dichotic task or what they have seen in a tachistoscopic task may detect more left-hemispheric activity due to motor programming. It may well be that subjects could develop a psychological set for motor programming that reduces or increases hemispheric asymmetry. Thus, there may be in some studies an interaction between stimulus and task variables that could affect subjects' processing mode.

Technique of Stimulus Presentation

While dichotic listening studies have contributed much to our current ideas about the nature of hemispheric asymmetries in normal subjects, the meaning of results obtained from these studies is not always clear. One concern is that repeated testing of the same subjects does not always produce the same results. Some studies have found the reliability of the dichotic and tachistoscopic tests to be lower than one might expect (Blumstein, Goodglass, & Tartter, 1975). Moreover, Teng (1981) has argued that dichotic ear difference reflects an "input asymmetry" between the contralateral and ipsilateral ear-to-hemisphere projection in addition to an "ability asymmetry" between the hemispheres. She suggests that between-individual variations in ear differences may be attributable more to variations in input asymmetry than to those in hemispheric asymmetry. Dichotic listening thus appears on logical considerations to be a poor index for hemispheric specialization.

In spite of these shortcomings, dichotic listening techniques have helped raise interesting theoretical issues. One of these is whether hemispheric differences are absolute or relative. Does a difference in performance between ears reflect the fact that only one hemisphere is capable of performing the task? Or does it simply mean that one hemisphere is better at the task than the other? The typical study does not allow us to tease these alternatives out because performance in the dominant hemisphere may result either from more efficient processing by the specialized hemisphere or from processing by the specialized hemisphere after transfer of information across the cerebral commissures. The results would be presumably the same in either case: a difference in performance between the two sides. To examine this issue and to assess the relative contribution of each hemisphere to the perception of speech sounds a variant of the dichotic listening technique was used in the experiments described later.

In this variant of the dichotic listening technique subjects were presented with pairs of speech sounds given in random order to either the left, the right, or both ears. White noise in the opposite ear accompanied the occurrences of the speech stimuli only when they were presented to either the left or right ear. Thus, each subject judged each stimulus pair under three different conditions: with white noise in the left ear, with white noise in the right ear, and without noise. White noise was chosen because its broad spectral band was believed to cover the frequency range of the stimuli used in all experiments, and to serve, at a low signal-to-noise ratio, as an inhibitor of ipsilateral and cross-callosal transfer of speech information (e.g., Springer, 1973). Subjects were asked to judge the degree of dissimilarity between the stimuli within each pair. The judgment task allowed the subjects to develop their own strategies and imposed little, if any, constraint on the manner in which the stimuli were to be judged.

Presumably differences in the perceptual configurations derived from multidimensional scaling analyses would reflect differences in perception and information processing. A structural analysis of each configuration can thus yield both qualitative and quantitative information about hemispheric functional asymmetries with respect to speech perception. The manner in which these structural analyses of the configurations help us probe the processes involved in speech perception is taken up below.

Perceptual Representation

Much of the work on the representation of complex stimuli, auditory or visual, supports the notion that perception is based on an analysis of stimulus patterns along a number of psychological dimensions or features. Information-processing models often refer to this processing as "feature extraction", which is thought to reflect a selective reduction of information whereby perceptually salient features are extracted from the pattern while other information is not retained.

One may conceive of these dimensions as forming a multidimensional perceptual space in which each stimulus is represented as a point. This space, of course, is not directly observable and both the set of dimensions comprising the space and the loci of the stimuli within the space must be inferred or derived by indirect methods. Multidimensional scaling (MDS) is an important method for deriving a representation of the perceptual space (Ramsay, 1978; Romney, Shepard & Nerlove, 1972). Multidimensional scaling techniques are designed to decompose a matrix of pairwise similarity or dissimilarity judgments on a set of complex stimuli into a metric space of some (investigator-specified) number of orthogonal dimensions. Each stimulus is defined as a point in the space such that, ideally, the distances between pairs of stimuli in the space are monotonically related to the degrees of judged dissimilarity of the pairs. MDS methods are thus designed to find the dimensions given the dissimilarities. The typical input data consists of an n x n (or (n x (n-1))/2) matrix whose cell values indicate the dissimilarity of pairs of the n stimuli. Dissimilarities are generally assumed to measure the psychological distance between the stimuli (Ramsay, 1978).

The central assumption underlying MDS techniques for perception is that stimuli are "coded" internally in terms of continuously varying parameters or dimensions. The aim of MDS is to discover the number of dimensions relevant to perception of the stimuli under consideration and to determine the stimulus coordinates on each dimension, i.e. to determine the dimensionality and configuration of stimuli in multidimensional space.

The set of abstracted dimensions and the relative loci of the stimuli within the space may be interpreted to reflect the structure of psychological space. Having thus obtained an abstracted multidimensional solution, an investigator may attempt to relate the derived psychological dimensions to the known physical structure of the stimuli. Success in identifying the psychophysical functions relating psychological to physical dimensions is typically measured by a high correlation between values on a psychological dimensions and values on the candidate physical measure across stimuli. The problem of interpretation, then, is to identify physical or other correlates of these psychological dimensions, or to specify their psychological meaning in some way. It should be noted, however, that interpretation is sometimes configurational rather than dimensional; i.e., it entails description of meaningful clusters, symmetries or other patterns in the multidimensional configuration. Furthermore, dimensions need not vary continuously in the sense that all possible values will be evidenced, but may be more discreetly valued, as would be the case if strong clustering occurred (cf. Torgerson, 1965, for a more detailed discussion).

MDS methods have been used successfully to identify psychological dimensions underlying the perception of speech sounds (eg., Cohen, 1978; Klein, Plomp & Pols, 1970; Shepard, 1972; Singh & Woods, 1971). In this context, success has usually meant that the derived configuration accounts for a large proportion of the variability in the dissimilarity judgments and that the revealed meaning of the psychological dimensions is intuitively reasonable.

Shepard (1972) applied multidimensional scaling techniques to some data from Miller and Nicely's (1955) study of confusions among English consonants. The subjects in Miller and Nicely's experiment listened to female speakers read one-syllable stimuli, such as /pa/, /ta/, /ka/, etc., from randomized lists, and write down the consonants they heard after each syllable was spoken. There were 17 experimental sessions in each of which speech transmission was degraded in a different way. Shepard (1972) subjected the pooled symmetrized matrix from the noise conditions to MDS analysis. A two-dimensional (rotated) space was obtained for these data (Figure 1). Shepard assigned the interpretation "voicing" to the horizontal dimension, since it distinguishes the voiced consonants (those, which, when spoken, produce vocal cord vibration) from their voiceless cognates. The interpretation "nasality" given to the vertical dimension reflects the fact that the two nasals, /m/, and /n/, are separated from the other consonants.

Even though only two distinctive features (voicing and nasality) are explicitly used in the labels for dimensions, Shepard does point out that some information regarding affrication and place of articulation is preserved in the two-dimensional configuration where, for example, the stops

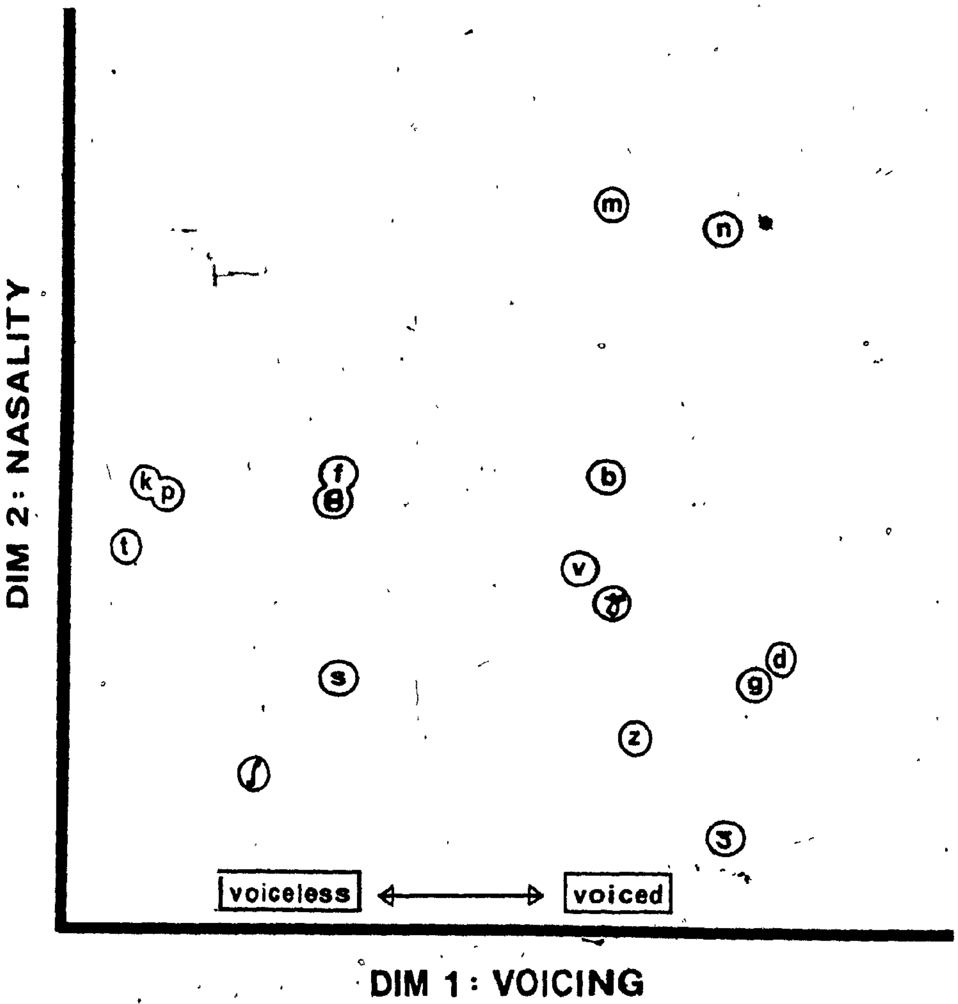

Figure 1. Spatial and hierarchical representation of the perceptual similarity between consonants derived from Miller and Nicely's (1955) data. From Shepard (1972).

are generally at opposite extremes of Dimension 1 with the fricatives in between. Figure 2 shows some of the possible places of articulation as indicated by the arrows going from one of the lower articulators to one of the upper articulators.

Research conducted on Dutch vowels has shown that superimposing the derived perceptual space onto the physical space (reduced after a principal components analysis and a simple geometric transformation) yielded practically the same loci for most vowels in both spaces (Plomp, Pols & van der Kamp, 1973). Cohen (1978), using MDS procedures, obtained similar results with English synthetic vowels presented in three different phonetic contexts. In these studies, excellent correlations between the physical and psychological configurations were found, indicating that information similar to that contained in the spectral representation may be used directly in the perceptual judgments of speech sounds.

These studies suggest that construction by subjects of a perceptual space of speech sounds definitely takes into account some important physical characteristics which can be related, in the case of consonants, to manner and place of articulation and, in the case of vowels, to tongue height and advancement which refer to the first and second formant frequencies respectively. All this suggests that any change in the physical parameters of the speech signals would

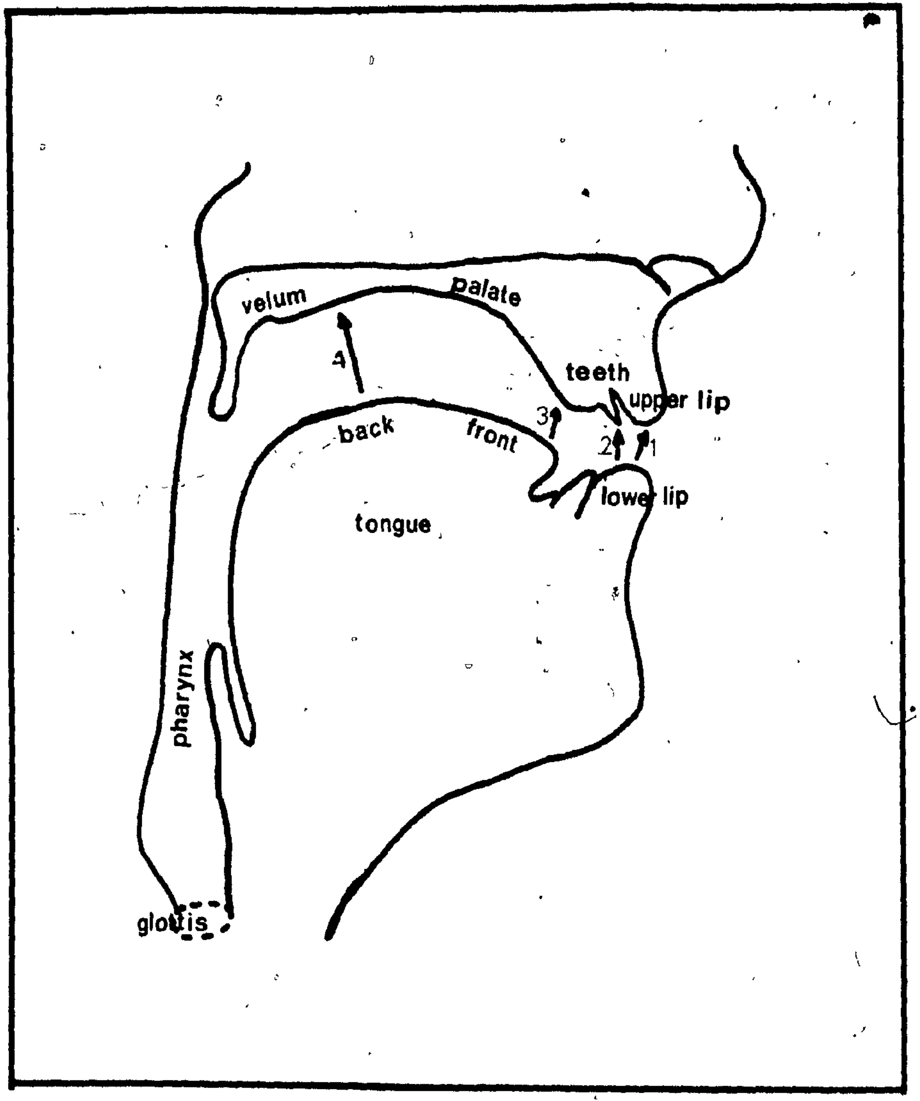

Figure 2. Diagram showing vocal organs and a few possible places of articulation. (1) bilabial; (2) labiodental; (3) alveolar; (4) velar.

translate into changes in the perceptual representation of the stimuli involved. Conversely, changes in the perceptual configurations may reflect differences in perception.

The Experiments

The purpose of the experiments presented in this thesis is to investigate the contribution of each cerebral hemisphere to the perception of speech sounds. More specifically, the issue was to try to determine whether the right hemisphere was involved at all in the process of speech perception, and given that it was, to determine the extent of its participation and the nature of this participation.

Four experiments were carried out in an effort to answer these questions.The first experiment addressed the problem of hemispheric involvement in the perception of English consonants. Stimuli similar to those used by Miller and Nicely (1955) were used in this experiment to provide reference data and to compare the derived perceptual representations to that obtained by Shepard (1972) in his reanalysis of the Miller and Nicely's data.

The second experiment explored the extent of each hemisphere's contribution to the perception of two important defining features of consonants, voicing and place of articulation. Stop consonants similar to those used by Studdert-Kennedy and Shankweiler (1970) were used as stimuli.

The third experiment investigated vowel perception in an effort to determine whether traditional classification systems of vowel sounds, in terms of tongue height and advancement could account for the perceptual representations derived from presentations to each ear.

The last experiment concerned the involvement of each hemisphere in the acquisition of new phonetic categories (Mandarin Chinese retroflexes).

EXPERIMENT 1

This first experiment addressed the question of hemispheric participation in the perception of English consonants. Several investigators have already reported differences in performance between the two ears in the identification of consonants. In a paper published in 1967, Shankweiler and Studdert-Kennedy reported the finding that under dichotic presentation, the right ear was superior to the left for recall of consonant-vowel-consonant (CVC) nonsense syllables differing only in the initial stop consonant. However, these authors were unable to detect any significant difference between the ears for the recall of similar syllables differing only in the vowel. Similarly, if two different sounds are led one to each ear but with a temporal offset of around 60 msec, the second sound is recalled more accurately than the first from the right ear, for stops but not for vowels (Studdert-Kennedy, Shankweiler & Schulman, 1970). Other classes of speech sounds have given results in the laterality paradigm intermediate between stops and vowels. For example, Darwin (1971) found that place of articulation for fricatives was only reported better from the right ear if formant transitions were present. Place of articulation from fricatives is cued mainly by the spectrum of the friction, but intelligibility is increased if appropriate formant transitions are added (Harris, 1958). To determine which aspects of the perceptual process depend upon the specific language processing machinery of the dominant hemisphere, Studdert-Kennedy and Shankweiler (1970) tested components of speech signals by presenting spoken CVC syllables in dichotic pairs that contrasted on only one phone. Significant REA's were found for initial and final stop consonants, and for voicing and place of production of stop consonants. Presumably because the right hemisphere was performing above chance level, this led them to conclude 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. They therefore suggested that the specialization of the dominant hemisphere in speech perception is due to its possession of a linguistic device, and not to specialized capacities for auditory analysis.

Although these studies report laterality differences, little remains known about the actual participation of each hemisphere. In this experiment a different approach to the study of brain involvement in speech perception was taken, in contrast to previous studies in which some form of identification or recognition was required. Subjects in this study were asked to rate the extent to which pairs of CV syllables differed. Multidimensional scaling analyses of the ratings yielded perceptual representations assumed to reflect each hemisphere's involvement in the analysis of speech sounds.

Method

Subjects

Four male and four female right-handed univesity students with no reported hearing deficits participated in this experiment. Handedness was assessed using the Montreal Neurological Institute's "Stated hand preference questionnaire". All subjects were native speakers of English and were paid $5.00 for their participation.

Stimulus materials

The stimuli used in this experiment consisted of 16 CV syllables (the same as those used in the Miller and Nicely's (1955) study): /ba, da, fa, ga, ka, ma, na, pa, sa, ta, va, za, θa, ða, ʃa, ʒa/, spoken by a phonetician. A tape was made consisting of three randomized sets of the 136 possible random pairings of the stimuli and excluding permutations.

Procedure

Subjects were tested singly or two at a time. They listened through headphones to presentation of the stimulus pairs. Pairs were presented in blocks of 17, each block randomly assigned to the left, right, or both ears with no two consecutive blocks to the same ear. White noise simultaneously accompanied presentation of the speech stimuli in the right and left ear presentation conditions only. The signal to noise ratio was -6db, measured with a Bruel and Kjoer Impulse Precision Sound Level Meter (Type 2204). Sound level was the same for all subjects. The subjects' task was to indicate by marking a stroke on a 100mm continuous line the extent to which they judged the dissimilarity of the two stimuli in each pair. The scale was labeled 'very similar' at one end and 'very dissimilar' at the other. Subjects were encouraged to use as much of the scale as possible but were not required to do so. They were not given any description of the stimuli and were left free to form their own schemes of reference in the judgment task. The session lasted about one hour.

Results and Discussion

Each subject contributed 136 observations per condition. The diagonal entries (data from trials when a stimulus was paired with itself) in the dissimilarity matrix were not input as part of the data, thus only 120 observations per subject for each condition were analyzed. Inspection of these diagonal entries revealed that all subjects assigned zero or near zero values to all 16 pairs of identical stimuli. All subjects in the experiment produced acceptable data: within-subject statistics provide no unacceptably high standard error estimates (>1.3) or unacceptably low exponents (<.3) that would suggest subjects provided ratings which were at either end of the scale or exactly in the middle (Ramsay, 1978).

Separate MDS analyses, using model M4 of the Multiscale program (Ramsay, 1978), for each subject for each of the left, right, and both channels conditions yielded configurations in a Euclidian space that reflected the underlying structure of the stimuli.

The data were scaled in two and in three dimensions in order to achieve the optimal configuration by the method of maximum log likelihood (Ramsay, 1977, 1978). Applying Ramsay's stopping rules, a three-dimensional solution was justified (corrected $X^2 > 47.1$, $p < .05$) in several cases while in the rest a two-dimensional solution was accepted.

The dissimilarity ratings produced by each of the eight subjects for the left, right and both ears conditions were further subjected to a group analysis, using model M3 of the Multiscale program, in an effort to determine the structure underlying the group's perceptual representations of the consonant stimuli. Model M3 refers to the specification of a more complex type of provision for individual variation where it is assumed that subjects share a common perspective on the stimulus objects, or "group" perceptual structure, but that they vary with respect to the proportions in which they combine the dimensions of this spatial representation (for a more detailed presentation and discussion of the modified Euclidian distance formula used in Program M3, see Ramsay, 1978).

The stopping rule of dimensionality adopted to establish the number of dimensions underlying the group's perceptual space for each condition was twice the critical chi square value $X^2=33.92$ with df=22 and p=.05 (Ramsay, 1978). Applying this stopping rule, four dimensions provide a significantly better fit of the data for consonants presented to the left ear, to the right ear and to both ears ($p< .05$, in each case).

Comparison of group perceptual representations.

The evaluation of the correspondence between the perceptual representations and the articulatory and defining charactristics of the consonant stimuli is approached next. The problem here is to determine whether there exist any differences in the structure of the perceptual representations derived from presentations of the speech stimuli to the left, right and both channels. Figures 3, 4 and 5 show the representation of consonants plotted in the space defined by the two most important dimensions, Dimension 1 and Dimension 2, for LEFT EAR, RIGHT EAR and BOTH EARS

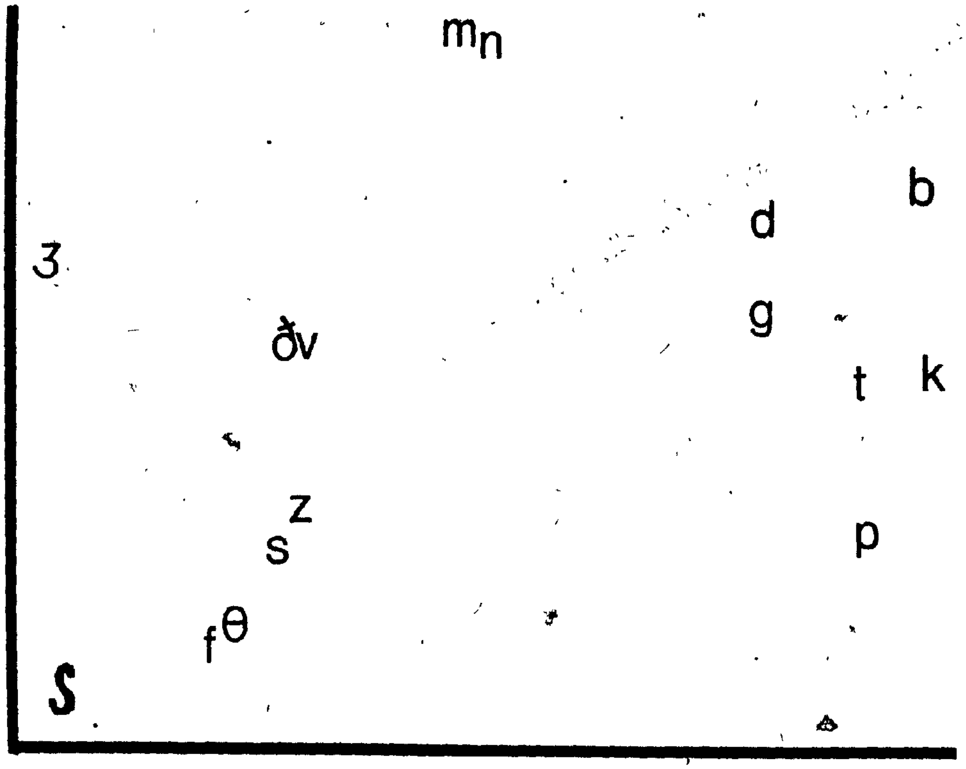

Figure 3. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of consonants presented to the left ear.

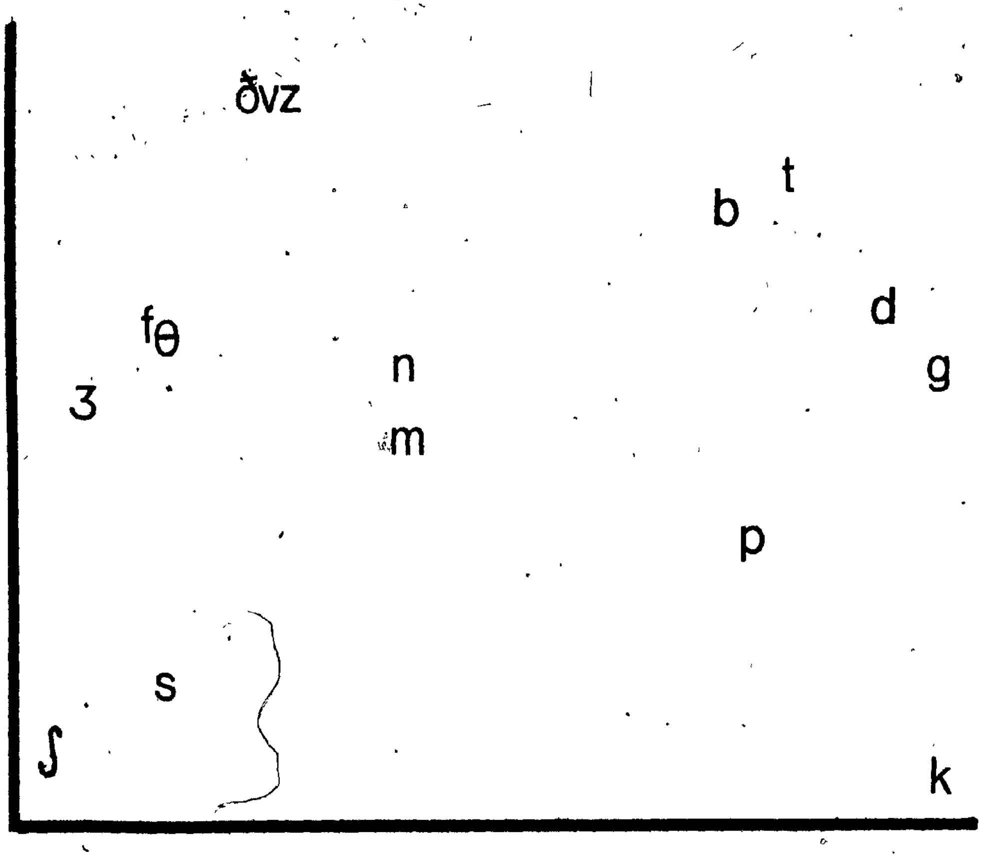

Figure 4. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of consonants presented to the right ear.

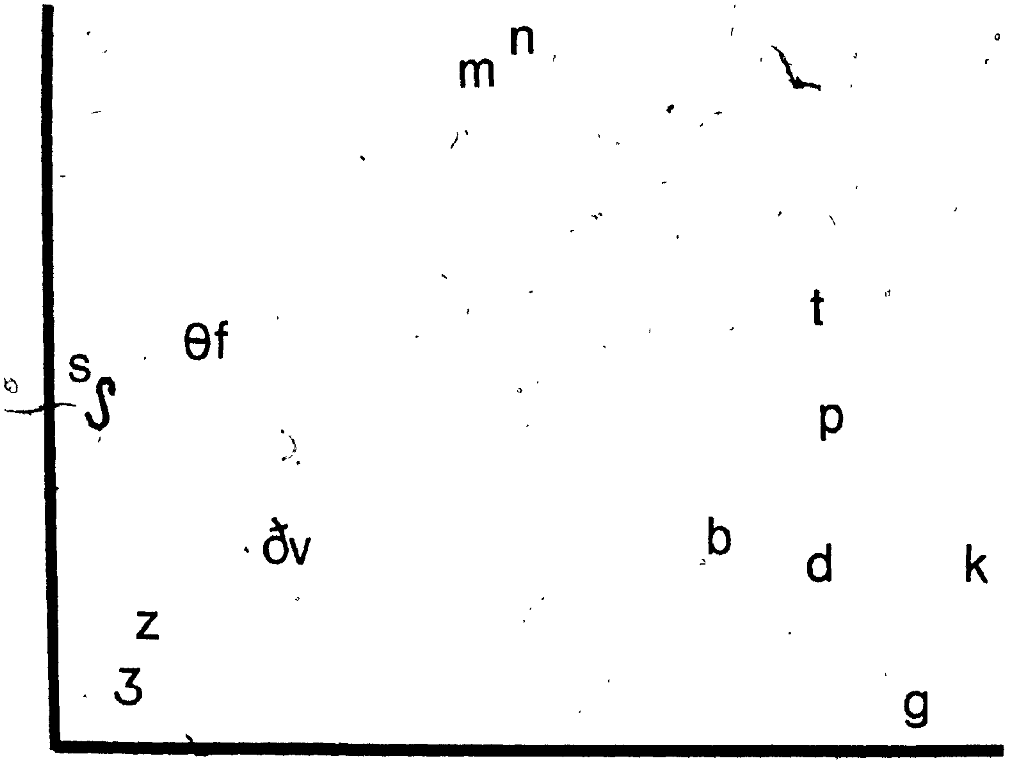

Figure 5. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of consonants presented to both ears.

conditions.

Inspection of the MDS plots of the stimuli used in this experiment revealed that the points in the space defined by Dimension 1 and Dimension 2 are clustered into three categories. These categories appear stable and invariant across conditions. They are the fricatives /f, θ, s, ʃ, v, ð, z, ʒ/, the nasal stops (nasals) /m, n/, and the oral stops (stops) /b, d, g, p, t, k/.

The first dimension extracted in the MDS analysis, Dimension 1, clearly refers in all three configurations to manner of articulation. The location of the consonant stimuli along Dimension 1, however, reveals more than manner of articulation information. In the BOTH EARS condition, for example, fricatives are further divided in groups on a purely auditory basis, the sibilant and nonsibilant sounds. The fricatives /s, z, ʃ, ʒ / have more acoustic energy at a higher pitch than the other fricatives. These sibilants constitute a class of speech sounds in which there is a high pitched, turbulent noise, as in 'sip' and 'ship'.

Another level of information that is preserved in Dimension 1, in BOTH EARS, refers to place of articulation. Among the stops, consonants are separated into velar /g, k/, alveolar /d, t/, and bilabial /b, p/. The nasal stops /m, n/ are in a region of Dimension 1 (and of the spatial representation) of their own. Among the fricatives, the labiodentals /f, v/ and the dentals /θ, ð/ are distinguished from the alveolars /s, z/ and the palato-alveolars /ʃ,ʒ/. This distinction contrasts fricatives produced at frontal places of articulation from those produced at back places of articulation.

Aside from manner of articulation the information contained in Dimension 1 is not invariant across conditions. Differences exist which have bearing on theoretical implications concerning cerebral hemispheric involvement in the perception of speech sounds. These differences will be attended to now and a more detailed discussion of their meaning will follow in a later section.

In LEFT EAR (Figure 3), there does not appear to be any meaningful interpretation in addition to manner of articulation to the way sounds are arrayed along Dimension 1. In RIGHT EAR, however, place of articulation is clearly evident among stop consonants arranged from bilabial to velar. In fricatives, except for the placement of /z/, there is, as in BOTH EARS, a contrast between those sounds produced with frontal articulators and those produced with articulators further back in the articulatory apparatus, i.e., labiodental and dental vs. alveolar and palato-alveolar. These preliminary results seem to indicate that the left hemisphere is involved in the processing of place of articulation information.

The meaning of Dimension 2 appears to be the same across the three conditions and can be interpreted as a dimension of voicing. In the LEFT EAR condition the voicing contrast is as marked as in the BOTH EARS condition with apparently a sharper distinction among fricatives than oral stops. In addition to voicing, Dimension 2 can also be interpreted to refer to distinction between sibilants and non sibilants in RIGHT EAR. This would suggest that the right hemisphere deals mainly with the processing of voicing information. The left hemisphere, however, appears to deal with both voicing and manner of articulation in Dimension 2. It is of interest to note that the two first dimensions extracted for all perceptual representations can be explained in terms of traditional consonant descriptive features, articulation and voicing distinctions.

The other derived dimensions may also be interpreted in terms of these two defining features, although to a lesser extent because the patterns are not as strong across subjects. Dimension 3 seems to contrast in BOTH EARS between voiced and voiceless fricatives as well as between place of articulation, i.e., differentiating on the one hand the palato-alveolar sounds /ʃ, ʒ/ from the other fricatives and on the other, bilabial stops /b, p/ from the other stops.

In RIGHT EAR, Dimension 3 can be interpreted as referring mainly to nasality, i.e., distinguishing nasals from other sounds. There appears, however, to be voicing information for both fricatives and oral stops preserved in Dimension 3. In LEFT EAR the meaning of Dimension 3 could not be derived.

The interpretation of Dimension 4 varies for each condition. In BOTH EARS it appears to refer to voicing contrasts between stops only; in RIGHT EAR to voicing contrasts between fricatives only. Interestingly enough, this last extracted dimension may be interpreted as reflecting a differentiation between the palato-alveolars /ʃ,ʒ/ and the velars /g, k/ -the four sounds in this experiment articulated furthest back from all other consonants in the LEFT EAR condition.

The results of Experiment 1 demonstrate that the features of manner and place of articulation and of voicing explain the perceptual representations obtained. Manner of articulation was found to be in all the ear conditions the most salient dimension, accounting for approximately half of the total explained variance. Differences between ears were observed with respect to voicing and place of articulation. In LEFT EAR and in BOTH EARS Dimension 2 was interpreted as reflecting voicing contrasts, while in RIGHT EAR, Dimension 2 and Dimension 3 were interpreted as reflecting both voicing and manner of articulation.

These findings suggest a contribution on the part of both hemispheres to the perception of English consonants. There is apparently no report in the literature of any study investigating hemispheric differences with a large corpus of consonant sounds comprising stops and fricatives. Results obtained with the dichotic listening technique have usually shown a REA for stop consonants (Studdert-Kennedy and Shankweiler, 1970) as well as for fricatives (Darwin, 1971). A different pattern of results, however, is suggested in this experiment. Ear advantages appear present in the sense that they only refer to specific features underlying the structure of the derived perceptual spaces. In other words, there appears to be an interaction between feature and ear of presentation: a greater participation in voicing revealed in LEFT EAR and a greater involvement in place of articulation in RIGHT EAR.

These results tentatively suggest that participation in the process of speech perception may not be the province of the left hemisphere only and that the right hemisphere may play an active role in the analysis of the speech signal. Each hemisphere, it appeared, processes preferentially a different property of the stimulus. The left hemisphere contributes mainly to the processing of place information while the right hemisphere processes mainly voicing information. The next experiment was therefore conducted to further investigate the preferential involvement of each hemisphere to these particular aspects of speech perception.

EXPERIMENT 2

The results of Experiment 1 suggested that although each hemisphere may process properties of speech stimuli such as voicing and place of articulation, they nonetheless appeared to be preferentially engaged in the treatment of one particular feature over the other. The left hemisphere seemed to be more involved in the treatment of place of articulation while the right hemisphere seemed to be more involved in voicing. To investigate each hemisphere's role in the perception and processing of voicing and place of articulation, native speakers of English performed a dissimilarity rating task on six consonant categories. These six consonants were contrasted on voicing (voiced /b, d, g/ versus voiceless /p, t, k/) and on place of articulation (bilabial /b, p/, alveolar /d, t/ and velar /g, k/).

Figure 6 shows a schematic representation of the six consonant categories along the voicing and place of articulation dimensions. Stimuli were embedded in three vocalic contexts to provide coarticulation variation. Thus each consonant category consisted of three exemplars. Membership of each exemplar in a particular category depends upon the extent to which it shares properties with other members of the category. Thus, consonant categories

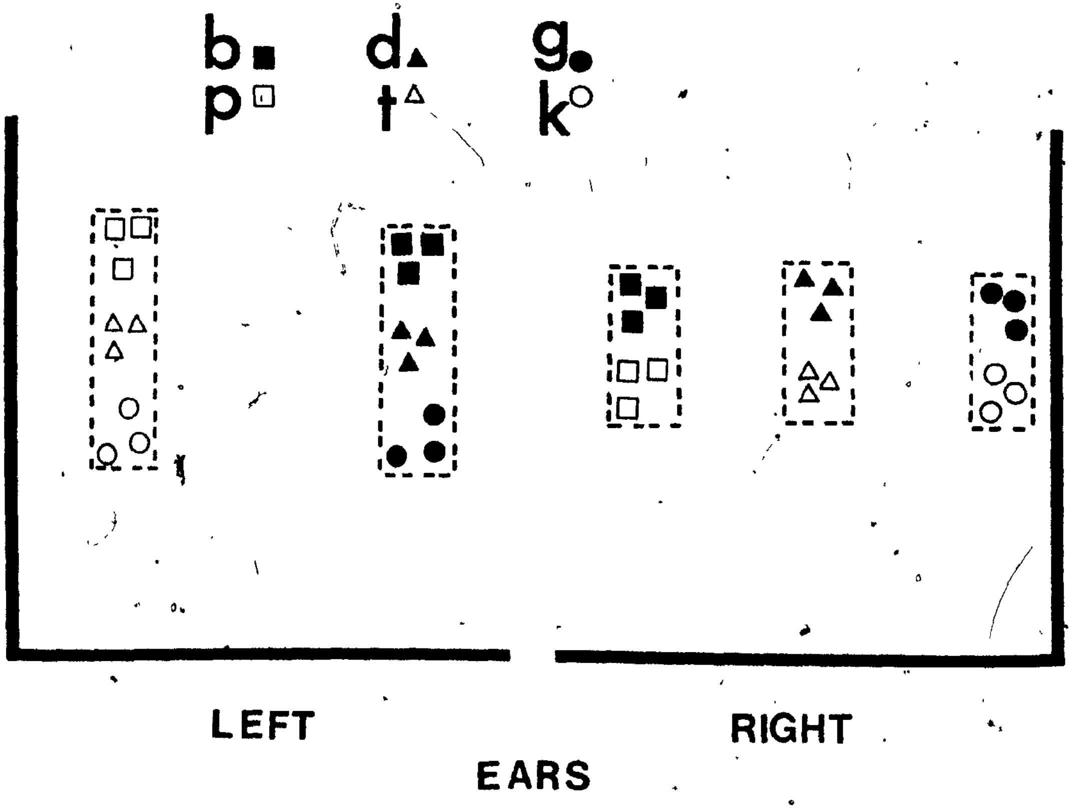

Figure 6. Schematized perceptual representation of consonant categories presented to the left and right ears. Dimension 1 (horizontal axis) is the first extracted dimension and accounts for the larger proportion of explained variance.

can further be grouped into two larger categories, each comprising all nine exemplars sharing the same voicing value. Similarly, the exemplars can also be assigned to categories on the basis of their place of production thus forming three categories, each made up of six exemplars.

The distribution of categories in the perceptual space, as depicted in Figure 6, is important. The first extracted dimension (Dimension 1) is shown as accounting for a greater proportion of the explained variance than Dimension 2. This pattern would result if subjects find the stimulus feature corresponding to this dimension more perceptually relevant.

Voice onset time (VOT) is one of several cues that is utilized in our perception and identification of various consonant sounds and reflects the temporal relationship between laryngeal pulsing and consonant release. Since voicing is known to be an important characteristic feature of speech, and speech processes have generally been thought to be controlled by the left hemisphere, most researchers have therefore concluded that voicing is controlled by mechanisms within the left hemisphere.

In an important recent study, Molfese (1980a) recorded auditory evoked responses (AER) over the right and left hemispheres while adult participants attended to repetitive series of two-tone stimuli varying in tone onset time (TOT). Portions of the AERs were found to vary systematically over the two hemispheres in a manner identical to that previously reported for VOT stimuli (Molfese, 1978a). The lateralized effects reported by Molfese (1980a) appeared to be due to the localization of a temporal processing mechanism within the right hemisphere. Furthermore, the right hemisphere responded to the TOT in a categorical fashion along boundaries previously reported for VOT: the 0- and 20-msec stimuli were discriminated from the 40- and 60-msec stimuli. The AER's, however, showed that the left hemisphere could not discriminate between the two TOT categories although it could make within categories distinctions.

In earlier studies investigating electrophysiological correlates of speech perception, Molfese (1978b, 1980b; Molfese & Molfese, 1979) noted mechanisms within the left hemisphere sensitive to second formant (F2) transitions, cues which are important for discriminating consonant place of articulation information (see also Liberman, 1974).

Taken together, these data suggest that there may be differential distribution of specific acoustic processing mechanisms across the two hemispheres, and not necessarily a left hemisphere specialized speech processor per se, that is responsible for the observed hemispheric differences in speech perception (Molfese, 1980a).

The consonants were embedded in three different vowel environments to yield 18 different stimuli. The dissimilarity ratings obtained for these exemplars were subjected to MDS analysis to obtain a perceptual representation for each condition (left, right and both ears) for each subject.

Perceptual differentiation of these consonant categories was assessed in a manner proposed by Homa, Rhoads and Chambliss (1979) for the study of visual concept formation. A perceptual configuration for the different samples of speech sounds was derived using multidimensional scaling procedures. The interpoint distances between the exemplars in this perceptual space were calculated to obtain a measure of category differentiation. The degree to which categories are differentiated from each other is reflected in both how small the interpoint distances are for items within the same category, since small distances in perceptual space imply similarity, and in how large the interpoint distances are between items taken from different categories. The ratio of these within-category to between-category distances provides an index of the degree to which the perceptual space is structured along category lines. The smaller the structure ratio, the more structured the perceptual space.

If the results obtained by Molfese accurately reflect functional hemispheric differences in the perception of voicing and place of articulation we should then obtain, in the present experiment, lower structure ratios for voicing categories presented to the left ear (right hemisphere) than for categories presented to the right ear (left hemisphere). Moreover, structure ratios for voicing categories should be lower than for place categories presented to the left ear (right hemisphere). Finally, structure ratios for place categories presented to the right ear (left hemisphere) should be lower than for place categories preseneted to the left ear (Right hemisphere).

Method

Subjects.

Twelve native speakers of English, right handed and with no reported hearing deficits were recruited for this experiment. These subjects participated only in this experimnt and were paid $6.00 for their participation.

Stimulus materials.

Three exemplars of six consonants contrasted on voicing and place of articulation, spoken by a phonetician, were used as stimuli. The voiced consonants /b, d, g/ and their voiceless cognates /p, t, k/, contrasted also on place of articulation: /b, p/ -bilabial; /d, t/ -alveolar; and /g, k/ -velar. These sounds were embedded in three different vowel environments: prevocalic /_i/, intervocalic /i_u/, and postvocalic /a_/ to yield 18 different monosyllabic and disyllabic exemplars. A tape was made consisting of three sets of all possible combinations of pairs of stimuli excepting permutations. Each set thus contained 170 pairs.

Stimulus pairs were randomly presented in blocks of 16, plus one buffer pair at the beginning of each block, with blocks randomly assigned to the left, right or both channels.

Procedure.

The procedure was exactly the same as in the first experiment.

Results and Discussion

Each subject contributed 170 observations per condition. Identity pairings and pairs in the buffer trials were not analyzed, thus only 154 observations per condition per subject were analyzed. All subjects in the experiment produced acceptable data (standard error estimate <1.3; exponents $>.3$). Applying Ramsay's stopping rules, the dimensionality retained for the final configurations derived for each condition was three in some cases (corrected X^2 $=50.4$) and two in others. The projections for each of the 18 stimuli on the axes defining the final dimensional space accepted for each subject were further analyzed to compute the structure ratios described earlier. These ratios, it will be recalled, reflect category structure of the perceptual space, that is, the degree to which the space is differentiated along category lines. Structure ratios were computed for each category, for each condition, and for each subject in the following manner.

To obtain an evaluation of the voicing contrast, structure ratios were first obtained for categories differentiated on place of production but sharing the same voicing value. Thus, structure ratios of exemplars of categories /b, d/ and /g/, for example, were computed by considering their degree of structure in the space defined by the voiced and voiceless exemplars (this permitted an assessment of how well differentiated the voiced exemplars were from the voiceless exemplars). The same operation was then performed on the voiceless categories. Structure ratios were thus obtained for voiced versus voiceless consonants as well as for voiceless versus voiced consonants, across all three places of articulation, then summed and finally averaged over the number of comparisons (two in this case) to yield an index of perceptual structure for voicing.

To evaluate the place of articulation contrast, structure ratios were computed in order to reveal the extent to which all elements sharing the same place of articulation were differentiated from all other elements in the space. Structure ratios were therefore computed for differentiating the exemplars of two consonant categories (one voiced and one voiceless (e.g., /b/ and /p/) sharing the same place of articulation (here, bilabial) from all other exemplars of the other consonant categories in the perceptual space. This was repeated for each pair of categories sharing the same place of articulation. These structure ratios were also added and then averaged by the number of comparisons to yield an index of perceptual structure for place of articulation.

These data were then subjected to a two-way analysis of variance with two levels of FEATURE (voicing, place of articulation) and three levels of EAR of presentation (left, right, both ears) with repeated measures on both factors. The data were analyzed in an effort to determine whether voicing contrasts varied as a function of ear of presentation and whether differences, if any, favoring voicing in the LEFT EAR condition would be greater than differences observed between voicing and place of articulation in the RIGHT EAR condition.

Figure 7 shows that place of articulation is more strongly differentiated overall than voicing ($p<.01$). Furthermore, structure ratios reflecting voicing contrasts were, as predicted, smaller for stimuli presented to the left ear than for those presented to the right. Consonants presented to both ears yielded smaller structure ratios than those presented to either ear alone, although statistically different from the right ear only in voicing. A two way interaction between FEATURE and EAR of presentation was statistically significant, $F(2, 22)=4.718$, $MSe=.0790$, $p<.01$. With respect to differences between structure ratios for voicing and structure ratios for place between the left and the right channels, post hoc Tukey test revealed a

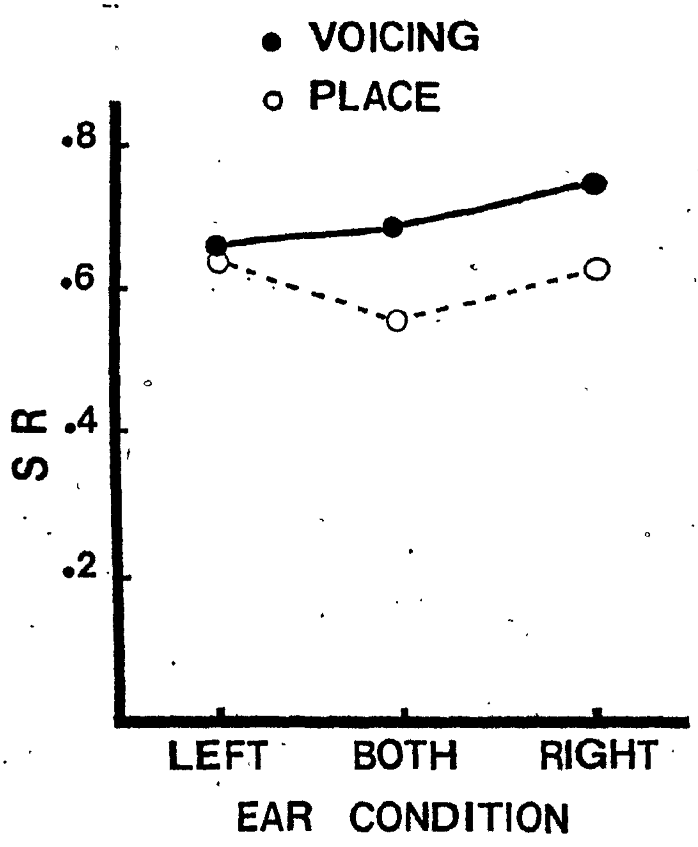

Figure 7. Ear differences (mean structure ratios) in the perception of the features of place and voicing.

statistically significant difference, p<.05, suggesting a more important involvement in place of articulation than in voicing in the RIGHT EAR condition. Post hoc Tukey tests also revealed significantly small[illegible] structure ratios for place of articulation in both ears than in either ear alone. (p<.05).

The general picture suggested by this first analysis revealed better overall performance for place of articulation than for voicing. In addition, there was a LEA in voicing. Furthermore, the contrasts between voicing and place of articulation were significantly more pronounced for consonants presented to the right ear than for consonants presented to the left, suggesting a greater involvement in place of articulation over voicing on the part of the left hemisphere. These results are in accord with those obtained by Molfese (1980a) who showed, using electrophysiological evidence, the involvement of the right hemisphere over the left hemisphere in contrasting stimuli along the voicing dimension, and of the left hemisphere being more involved in contrasting place of articulation than voicing.

In order to get a more detailed picture of the nature of the differences between the right and left ear conditions, the data were further analyzed in an effort to determine if differences would be observed when examining the manner in which stimulus categories were treated in each condition. An attempt was made to determine whether voicing influenced the extent to which members of a particular category would be differentiated along place of articulation (place contrasted along voicing) or whether place of articulation of a given category would influence the extent to which subsets of that category would be differentiated from each other (voicing contrasted along place).

To evaluate the extent to which voicing influenced differentiation of categories along place of articulation, structure ratios were therefore computed in the following way for voiced categories: /b/ exemplars differentiated from /d/ and /g/ exemplars; /d/ exemplars from /b/ and /g/ exemplars and /g/ exemplars from /b/ and /d/ exemplars. The same process was repeated for computations of structure ratios involving voiceless categories (/p/, /t/, /k/).

A three way analysis of variance was performed on the structure ratios with two levels of VOICING (voiced, voiceless), three levels of EAR of presentation (left ear, both ears, right ear), and three levels of PLACE (labial, alveolar, velar) with repeated measures on all factors. There was a significant three-way interaction ($F(4,44)=5.361$, $MSe=.038$, $p=.0013$). Post hoc tests revealed that performance in the LEFT EAR condition was better than in the RIGHT EAR condition for contrasting bilabial place of production for voiced stops (/b/), $p<.05$; performance in BOTH EARS was better than in either RIGHT EAR or LEFT EAR for contrasting voiceless velar place of production (/k/), $p<.01$; that, overall, both LEFT EAR and RIGHT EAR performed better for contrasting voiced velar place of articulation (/g/) as opposed to voiceless velars (/k/), p<.01. Figure 8 shows the results for each place of articulation category contrasted on voicing for the left, both, and right channels.

To evaluate the voicing contrast for a given category of place of articulation, structure ratios that had been computed for the voicing contrasts in the first analysis were averaged in the following manner: ((/b/ exemplars differentiated from /p/ exemplars) + (/p/ exemplars differentiated from /b/ exemplars))/2, and so on for each pair of categories on each place of articulation. These data were then subjected to a two way analysis of variance with three levels of EAR of presentation (LEFT, BOTH, RIGHT ears) and three levels of PLACE of articulation (bilabial, alveolar, velar) with repeated measures on both factors.

The analysis yielded a main effect of EAR of presentation (F(2, 22)=4.587, MSe=.045, p<.02) as well as a main effect of PLACE of articulation (F(2, 22)=9.832, MSe=.022, p=.0009). These results showed that performance in the BOTH EARS condition was better than in the separate ear conditions and that performance in the LEFT EAR condition was better than in the RIGHT EAR condition. Furthermore, the voicing contrast for alveolars (/d, t/) was better than for either bilabials or velars. However, the voicing contrast

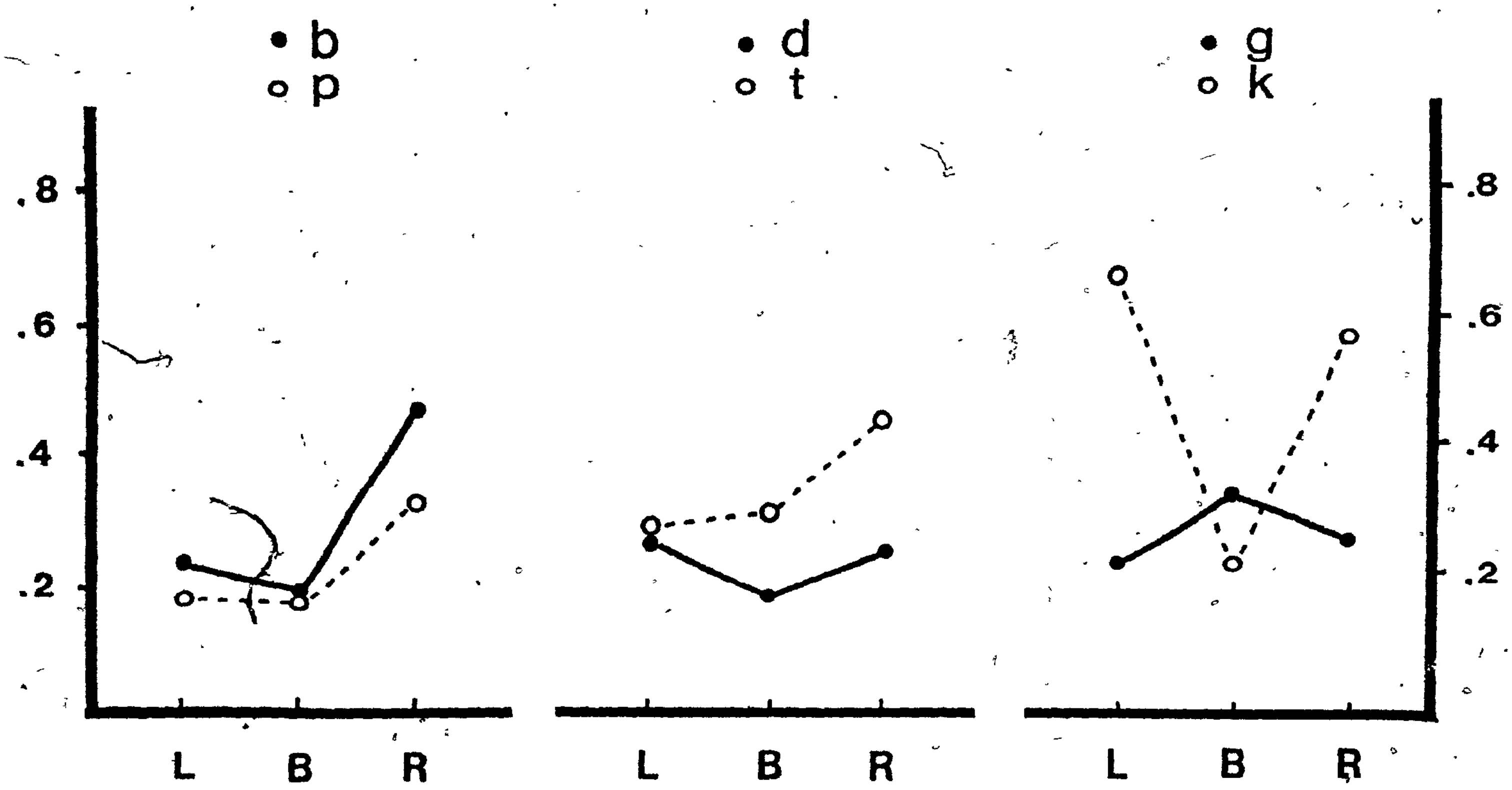

Figure 8. Differentiation of the voicing contrast at each place of articulation.

for alveolars was significantly better than that for velars only ($p<.05$).

The results of this second experiment present a complex picture of each ear's participation in the perception of stop consonants. In order to try to make sense of the information revealed by the various analyses the following interpretation of the results is offered.

The analyses performed on category structure ratios demonstrated first that place of articulation is more salient than voicing or, in other words, that overall performance is superior for place of articulation than for voicing. This is in contrast with results obtained by Studdert-Kennedy and Shankweiler (1970) who found that place of articulation was adversely affected by dichotic competition. While these authors postulate, on the basis of the number of errors made by each ear, a REA for stop consonants, the present study revealed voicing differences favoring the left ear. These results are congruent with those of Molfese (1980) using electrophysiological techniques that did not require any identification or report on the part of his subjects. Since subjects in the present study were not required to either identify or write (as in Studdert-Kennedy & Shankweiler's (1970) study) the stimuli they were listening to when doing the rating task, it may be assumed that subjects were not compelled to make a decision along categorical boundaries (e.g., yes-no, "or is it /b, p, d, t, g/ or /k/ that you just heard?") and allowed them to form and build their own schemes of reference with regard to the set of stimuli.

While a LEA was found in this study for the voicing feature, a closer examination of the data, suggested that this effect was not found for all three values of place: perceptual differentiation of the voicing contrast for bilabials was more pronouced than that for alveolars or velars. The mean values for each contrast being .263 for bilabial, .292 for alveolar, and .382 for velar, suggesting a gradual decrease of differentiation ($p<.01$) of the voicing contrast as consonants are pronounced further back in the vocal tract. It is of interest to note that the first consonants uttered by young infants are generally bilabials (/b , m, p/) (Jakobson, 1968; Salus & Salus, 1974) and require less articulatory effort in their production (Jakobson, 1962). This bias towards bilabials may have a perceptual basis since these sounds are plosives uttered at the forward extremity of the vocal tract and may carry more acoustic energy than other plosives uttered inside the vocal tract.

We may also ask whether the advantage of place of articulation over voicing shown in the first analysis, in which place of production vs voicing were contrasted, is of the same magnitude for the two ears. It was found that there was a larger difference between voicing and place contrasts in the right ear than in the left, a difference responsible in part for the interaction observed. Figures 9 and 10 show the perceptual spaces derived from the group's data for the left and right ear presentations. There are both similarities and differences in the way the speech stimuli are organized in the region of the perceptual space defined by the first two dimensions generated by Program M3 of MULTISCALE.

These representations are similar in the sense that both perceptual spaces can be explained in terms of the two defining features under study, place of articulation and voicing. They are different in the sense that each representation emphasizes a different dimension or feature. The first dimension extracted by the MDS analysis accounts for the larger portion of the explained variance in both representations. It is clear in Figure 9 that stimuli are discriminated according to their voicing value on Dimension 1, the voiced consonants at one extreme, the voiceless consonants at the other. It is also fairly evident that subjects took into consideration place of articulation when rating stimuli presented to the left ear. This can be seen from the organization of stimulus categories on Dimension 2. Figure 10 therefore suggests that subjects emphasized voicing to a larger extent than place of articulation when the speech sounds were presented to the left ear.

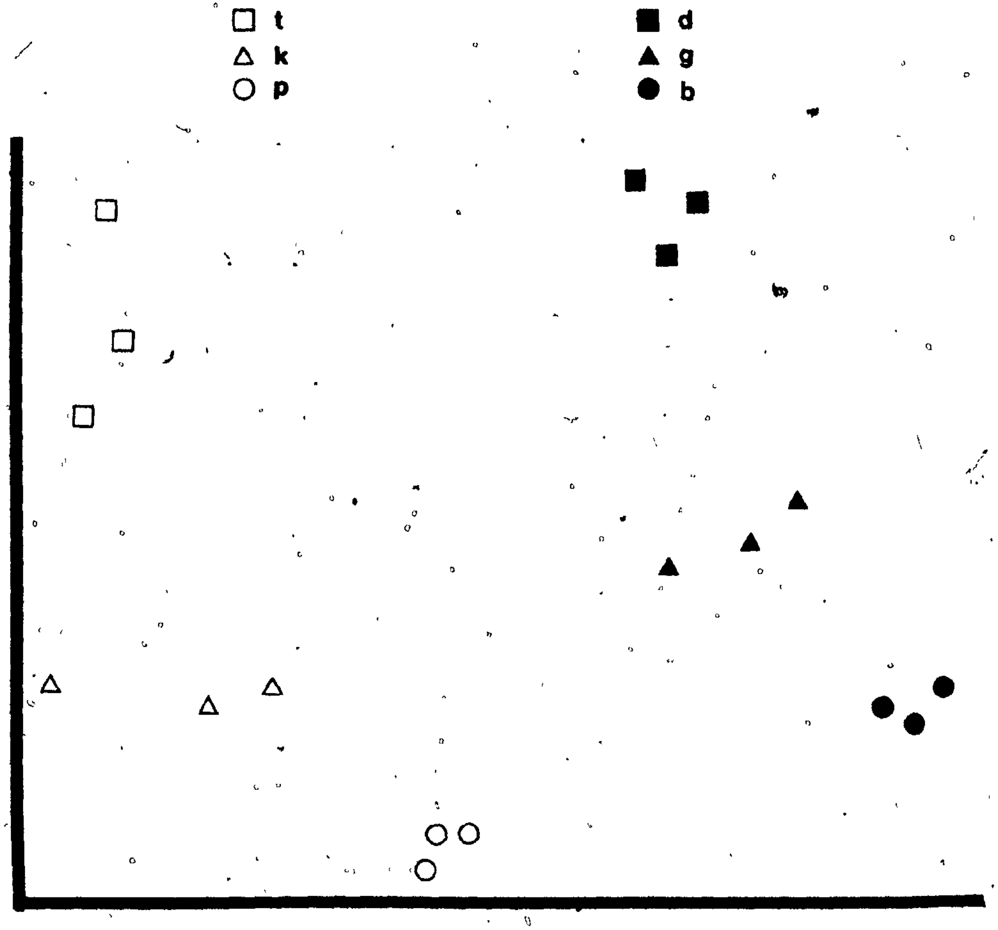

Figure 9. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of consonant categories presented to the left ear.

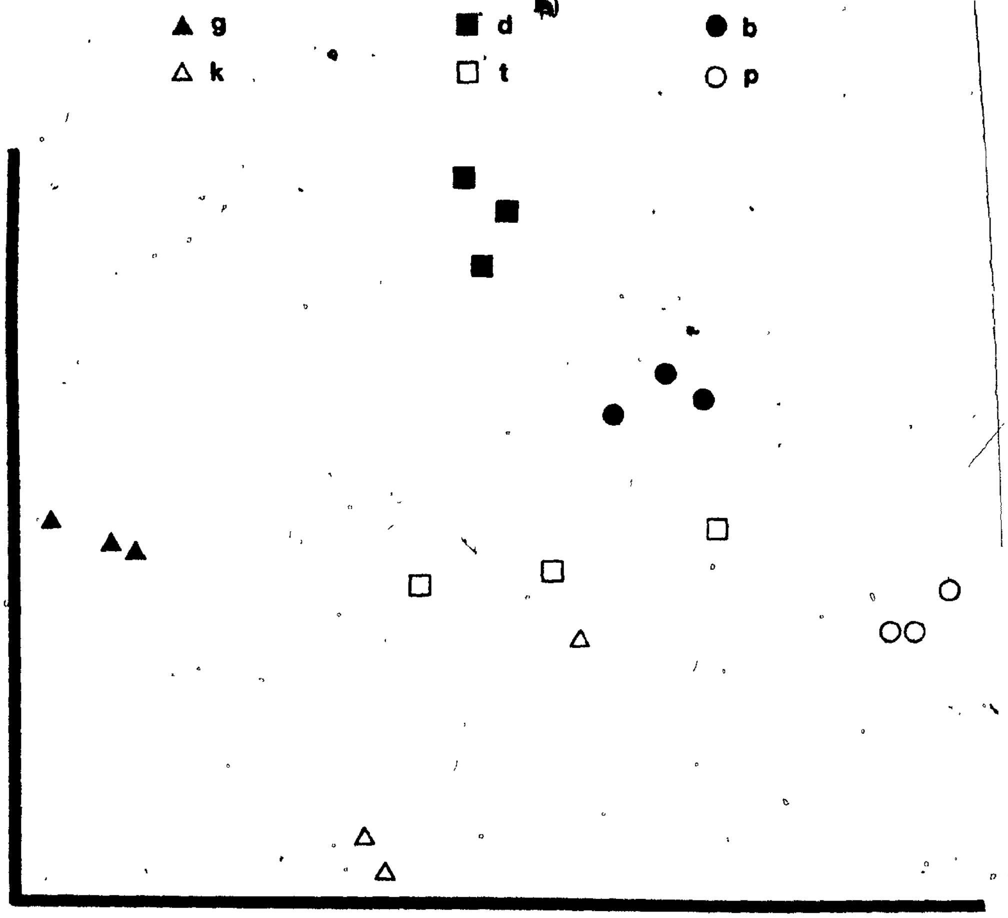

Figure 10. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of consonant categories presented to the right ear.

Figure 10, on the other hand, indicates that subjects emphasized place of articulation (Dimenson 1) over voicing (Dimension 2). Furthermore, the contrast between voicing and place is not as marked as in Figure 9. The proportion of the explained variance accounted by Dimension 1 and Dimension 2 for stimuli presented to the left ear is respectively .38 and .29; for stimuli presented to the right, the proportion is respectively .46 and .43 for Dimension 1 and Dimension 2.

There exists a further distinction between the two configurations described with respect to place of articulation. In RIGHT EAR, the organization of consonant categories in the perceptual space is isomorphic to the order of their place of production in the mouth. Such a relationship between perceptual organization and physical organization is absent in LEFT EAR and may suggest greater left hemisphere participation to place perception.

The interpretation accorded to the dimensions defining the group's psychological representations appear intuitively reasonable given the clustering and ordering of the speech sounds along those dimensions. To the extent that changes or differences in the perceptual spaces reflect changes or differences in perception, it can be concluded that the left ear is superior to the right ear for the feature voicing. Furthermore, it would also appear that for this group of subjects' perceptual representation, the place feature was superior to the voicing feature in RIGHT EAR. The results of a study by Miceli, Caltagirone, Gainotti, & Payer-Rigo (1978) add support to the proposition that the LH may be dominant for the treatment of place of articulation. These authors observed that discrimination of pairs of phonemes varying on place of articulation was more difficult than discrimination of phonemes varying on voicing for patients showing signs of aphasic impairment. No such effects were observed with three control groups (normal controls, right brain damaged patients and non-aphasic left brain damaged patients). Perecman and Kellar (1981) also obtained similar results with similar populations. These results support the findings of the first experiment where a LEA was found for voicing and a REA for place of articulation.

The results of these first two experiments, investigating hemispheric contribution to the perception of consonants, have shown ear advantages for particular components of consonant sounds. The conclusion suggested in Experiment 1, namely that the right hemisphere may be more involved in voicing and the left in place of articulation was supported by the results of the second experiment. These results finally suggest that ear differences seem to be relative, quantitative, and a matter of degree rather than absolute, qualitative, and a difference in kind. Either ear appears to process the same information as the other, but apparently at different levels of competence, and probably in different preferred directions.

In order to obtain a more comprehensive picture of hemispheric involvement to the process of speech perception, the third experiment addressed the question of whether ear differences in vowel perception can be interpreted in terms of traditional vowel classification systems.

EXPERIMENT 3

Differences between the perception of consonant and vowel sounds have led investigators to propose a "special" speech perception mode to characterize the way these latter phonetic segments are heard (Liberman, 1970). One of the findings that has been cited as evidence for a special speech perception mode is the difference in perception between synthetic stop consonants and steady-state vowels. Stop consonants have been found to be perceived in a categorical mode, unlike other auditory stimuli. Discrimination is limited by absolute identification. Listeners are able to discriminate stimuli drawn from 'different' phonetic categories but cannot discriminate stimuli drawn from the 'same' phonetic category, even though the acoustic difference between stimuli is comparable. On the other hand, steady-state vowels have been found to be perceived continuously much like nonspeech sounds (e.g, Pisoni, 1973).

Results from dichotic listening experiments suggest also a special mode of perception for vowel sounds. No ear differences are usually obtained under normal hearing conditions. However, a right ear advantage (i.e., in terms of the number of vowels correctly identified; e.g., Studdert-Kennedy & Shankweiler, 1970) may be observed if the vowel signal is degraded in dichotic presentation (Weiss & House, 1973).

This experiment investigates the particularities of each hemisphere's contribution to the perception of vowel sounds. Vowel sounds are usually defined in terms of acoustic parameters such as F1, F2, F3, the formant frequencies, and duration. Multidimensional scaling studies of vowel perception have consistently observed that F1 and F2 constitute the most important characteristics taken into account by subjects in the perception of these sounds (e.g., Cohen, 1978; Klein et al., 1970; Shepard, 1972). Third formant frequency (Shepard, 1972) and duration (Cohen et al., 1967) have also been found to be of perceptual significance, although to a lesser extent than F1 or F2. In this experiment, therefore, an attempt was made to examine the nature of differences, if any, between the hemispheres with respect to to the perception of vowels varying in these features. Native speakers of English were asked to rate the extent to which CVC stimulus pairs differing in the vowel were perceived as dissimilar.

Method

Subjects.

The same subjects who took part in Experiment 1 participated in this study three or four days later.

Stimulus materials.

Twelve vowels and three diphthongs embedded in /h-d/, spoken by a phonetician, were used as stimuli. The vowel sounds were the same as those used by Cohen (1978): /i, ɪ , e, ɛ , ae, a, ɔ , o, U, u, ʌ , ə , ai, aU, ɔɪ / as in beet, bit, gate, get, fat, father, lawn, lone, full, fool, about, but, hide, how, and toy respectively. A tape was made consisting of three sets of 119 random combinations of stimulus pairs.

Procedure.

The procedure was exactly the same as in Experiment 1.

Results and Discussion

The data were analyzed in an effort to determine whether the dimensions underlying the group's perceptual representations remained the same across the conditions and could be explained in terms of a traditional classification system of vowels. Differences, if any, between the physical and perceptual representations as well as between ear conditions were also evaluated.

Each subject contributed 119 observations per condition. Ratings for identity pairs and buffer trial pairs were excluded from the analysis. Thus 105 observations per condition for each subject were analyzed. Two subjects produced unacceptable data (exponents <.3) suggesting inconsistency in the use of the scale and had to be eliminated. The following results are therefore based on data collected from six subjects. The MDS analyses for all subjects yielded a perceptual space for the vowels that can be viewed as an approximation to an auditory-perceptual structure and that can provide information about the properties of that structure. The assumption is made, as in the other experiments, that distances of judged similarity relate meaningfully to psychological distances of similarity.

Applying Ramsay's stopping rule for model M3, four dimensions provided a significantly better fit to the data than three dimensions for vowels presented to both ears and for vowels presented to the left ear ($p<.05$). Three dimensions were accepted for stimuli presented to the right ear ($p<.05$).

Physical representation of vowel sounds.

Spectrograms of the vowel sounds were made using a spectrum analyzer (Vibralyzer 7030A, from Kay Electronics) in order to determine the physical space of the speech stimuli. The first three lower formant center frequencies (F1, F2, F3) were recorded. In terms of a geometrical model, we can say that analysis of the sound spectrograms of the 15 speech stimuli yielded a set of 15 points in a three dimensional space defined by dimensions corresponding to F1, F2 and F3. Figure 11 represents the distribution of the twelve vowel sounds in the plane of the acoustic space defined by F1 and F2.

Perceptual and physical representations compared.

To obtain a measure of the extent to which acoustic features of the stimuli correspond to their perceptual representations, canonical correlations were computed, using the SPSS Cancorr program, between the physical parameters defining the vowel sounds used (values of F1, F2, F3, and duration are given in Appendix 1) and the projections of the points on the axes defining the final group configurations.

Canonical correlation analysis produces two important types of information, the canonical variates and the canonical correlations between them. The canonical variates come in two sets and are composed of coefficients reflecting the importance of the original variables in the subset in forming the variates. The important point about canonical correlation is that the first canonical variate from the first set of variables and the first canonical variate from the second set of variables are extracted so as to have the highest intercorrelation possible given the variables involved, and successively for all pairs of canonical variates. The canonical correlation is the correlation between each corresponding pair of canonical variates, and

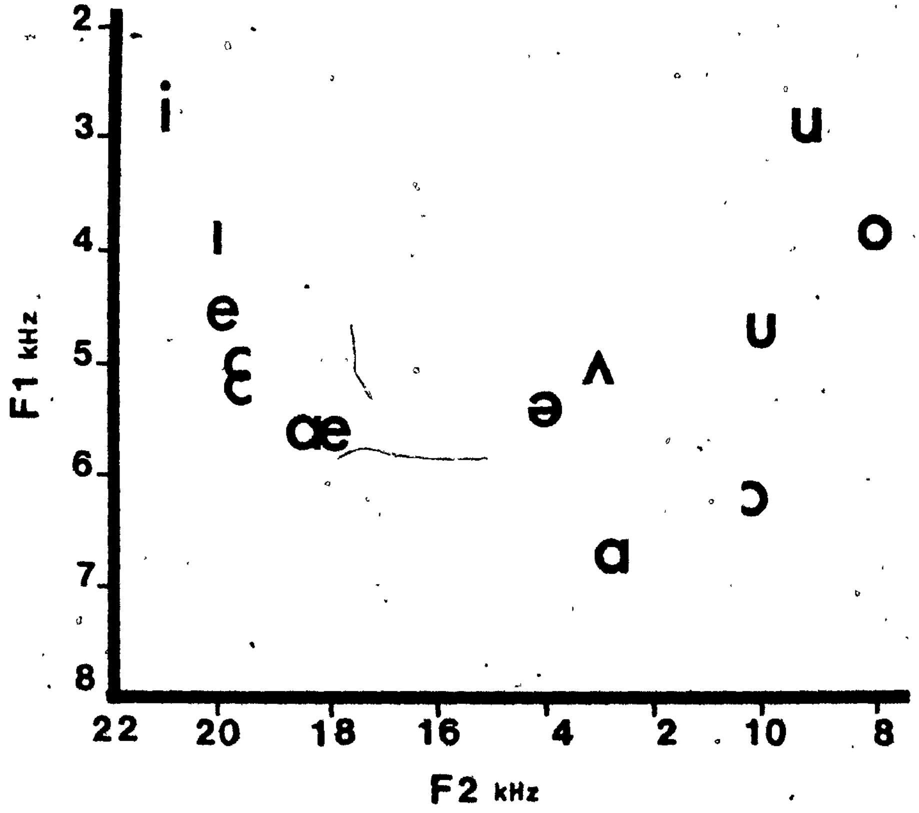

Figure 11. Physical representation of vowel sounds in the plane defined by F2 and F1.

its square, the eigenvalue, represents the amount of variance in one canonical variate accounted for by the other canonical variate. In other words, the method systematically extracts the first and largest source of variance, and the canonical correlation is an index of the relation between the two sets of variables based on this source of variance. Then the next large source of variance is extracted and analyzed independent of the first source. The second canonical correlation coefficient, which is smaller than the first, is an index of the relation between the two sets of variables due to this second source of variance, and so on for all successive pairs of canonical variates.

The canonical analyses yielded for each condition two statistically significant canonical correlation coefficients indicating that the relation between the first two pairs of canonical variates accounted for most of the variance. Table 1 is a summary of the canonical correlation source tables and identifies for each canonical variate the variable contributing most importantly to that variate.

Table 1

Summary Table of Canonical Correlation

Analysis for each Ear of Presentation Condition

	Left Channel		Both Channels		Right Channel	
	CANVAR 1	CANVAR 2	CANVAR 1	CANVAR 2	CANVAR 1	CANVAR 2
1st set	.975 (F2)	.759 (F1)	-1.087 (F2)	.976 (F3)	.884 (F1)	.868 (F2)
2nd set	.789 (D1)	-.794 (D2)	- .94 (D1)	1.106 (D2)	-.68 (D3)	.802 (D1)
CANCOR	.993	.979	.998	.915	.989	.969
X^2	56.739	26.707	59.326	21.2	52.18	23.35
d.f.	16	9	16	9	12	6
p	.0001	.002	.0001	.012	.0001	.001

The first important observation is that both F2 in the acoustic space (Figure 11) and Dimension 1 of the perceptual space make a larger contribution in forming the canonical variates than any other variable in their respective sets. The intercorrelation matrix between variables of both sets shows an excellent correspondence between F2 (a physical parameter) and Dimension 1 (a psychological dimension) for all three conditions. The correlations (Pearson r) are .944, .95, .956, p<001, df=11, for the LEFT, BOTH, and RIGHT EAR conditions respectively.

Dimension 1, which thus appears to correspond to F2, accounts in all three conditions for the larger part of the variance compared to other dimensions. The reason F2 is important is probably due to the fact that it varies more than the other formants. Frequency values of F2 in the present study ranged from 720 to 2100 Hz with only two vowels /I, e/ having the same F2 value. Frequency values of F1, on the other hand, ranged from 300 to 680 Hz with several vowels having similar F1 frequencies.

The second important observation of interest pertains to the nature of the variables with the greatest weights in the canonical variates. Second formant and Dimension 1 contribute most importantly to the first pair of canonical variates extracted in the BOTH EARS and LEFT EAR conditions, and to the second pair of canonical variates in the RIGHT EAR condition. First formant and Dimension 2 are most involved in forming the second pair of canonical variates in LEFT EAR and the first pair of canonical variates in RIGHT EAR. In BOTH EARS, F3 and Dimension 2 are the main contributors to the second pair of canonical variates. This is interpreted as reflecting a greater contribution of F2 perception in the perceptual spaces for BOTH EARS and LEFT EAR than for RIGHT EAR and a greater contribution of F1 perception in the perceptual space for the RIGHT EAR condition than for the BOTH EARS or LEFT EAR condition.

The Pearson product moment correlation coefficients indicate a strong relation between F1 and Dimension 2 in the LEFT EAR condition ($r=.821$, $p<.01$, $df=11$) as well as between F1 and Dimension 3 in BOTH EARS ($r=.691$, $p<.05$, $df=11$) and between F1 and Dimenson 2 in RIGHT EAR ($r=.710$, $p<.05$, $df=11$). These results are in accord with those of other investigators who found the first two formant frequencies, F1 and F2, to be the most important factors in vowel recognition and perception (Cohen, 1978; Cohen et al., 1967; Peterson & Barney, 1952; Shepard, 1972). The relationship between the physically measured properties (F1 and F2) of these speech sounds and the psychological structures that were derived independently of those physical measurements argues for the notion that each ear seems to be involved in the perception of these two acoustic parameters.

Differences, however, appear present in the extent to which each ear deals with these two parameters. These results suggest greater involvement in F2 perception in the LEFT EAR and BOTH EARS conditions and greater involvement in F1 in the RIGHT EAR condition.

To the extent that the derived perceptual representation for BOTH EARS reflects more accurately, than for either ear condition alone, the terminal output in the perception of vowels, we can postulate a left ear advantage for the perception of F2 and a right ear advantage for F1. Figure 12 shows the group's perceptual structure recovered from presentation of vowels to the left ear.

Let us consider first the results obtained for both ears via the canonical analysis. The first variables that load most importantly into the first two pairs of statistically significant canonical variates are, respectively, F2 with Dimension 1 and F3 with Dimension 2. Pearson correlation coefficient (from the intercorrelation matrix) shows a good correspondence between F3 and F2, r=.654, p<.01, df=11. The relationship between F2 and F3 needs, at this point, to be briefly elaborated upon.

In a well-known study by Peterson and Barney (1952) listeners attempted to identify, after aural presentation, which of 10 monosyllabic words, differing only in the vowel had been spoken. The task was rendered difficult by the fact that different speakers varying both in age and sex pronounced the same words on different presentations.

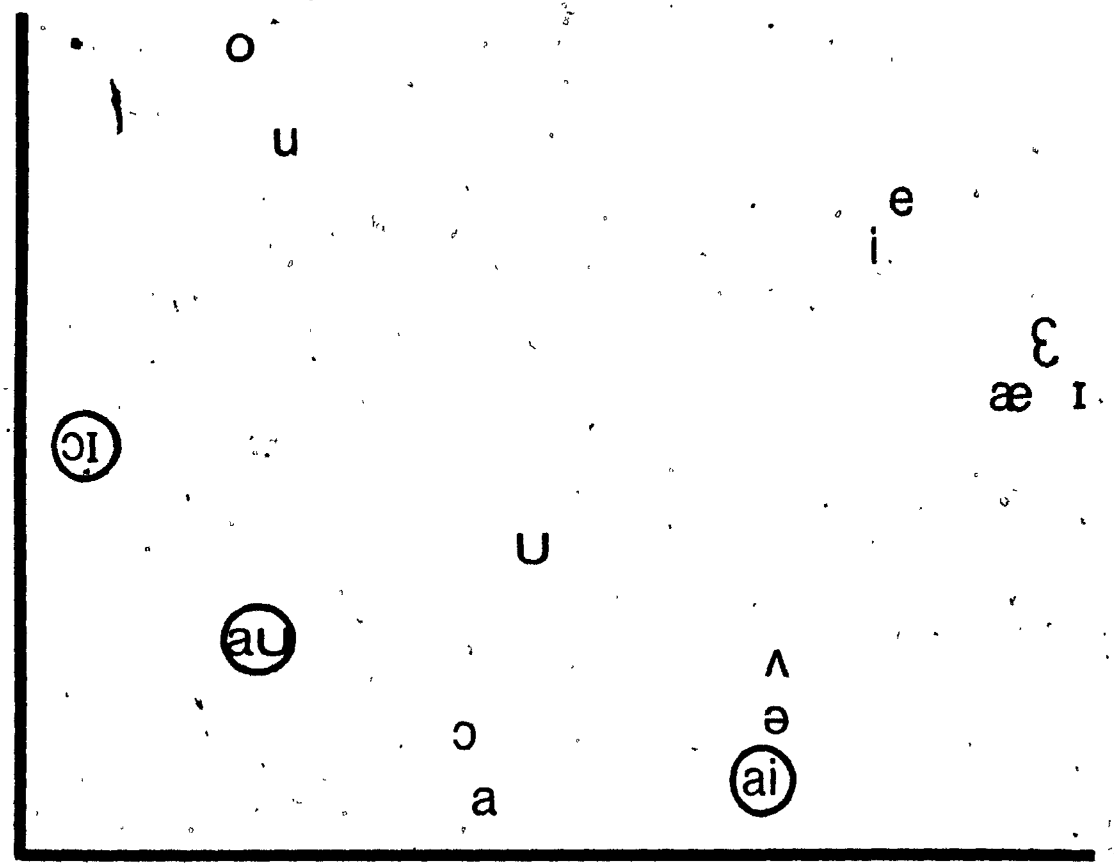

Figure 12. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of vowels presented to the left ear. Circled symbols refer to diphthongs.

Shepard (1972) computed a 10 by 10 symmetric matrix of proximity measures from the resulting confusion matrix obtained by Peterson and Barney and subjected it to an exponential analysis of proximities. Shepard fitted three rotated axes through the derived perceptual space such that each axis would best fit the average center frequency of the vowel formants as measured by Peterson and Barney. He found that the angles between the three axes corresponding to the first three formants were 99° between F1 and F2, 104° between F1 and F3, and 49° between F2 and F3. That the psychological effect or contribution of F1 is relatively independent of the other formants is suggested by the near orthogonality of the first axis to the other two. Psychological effects of F2 and F3, on the other hand, appear to be somewhat interdependent. This may be attributed to the restricted range in which F2 and F3 can vary due to constraints of the human vocal tract: the lower F2, the more determined is F3.

All this would suggest that F2 (and its correlated parameter F3) plays a major role in vowel perception. This is not surprising since F2, as was pointed out above, exhibits wide variation in the frequency range to which the ear is most sensitive (from 500 to 5000 Hz; Plomp, et al., 1967) Furthermore, F2 contains additional important information necessary for the identification of encoded speech, namely, F2 transitions, important cues in the identification of consonants (Liberman, et al., 1967).

As was shown earlier, F2 and Dimension 1 have the greatest weight in the first pair of canonical variates extracted from the LEFT EAR data suggesting a LEA in the perception of F2 information. Conversely, to the extent that F1 and Dimension 2 are the largest coefficients in the first pair of canonical variates in RIGHT EAR, and F2 and Dimension 1, in the second statistically significant pair of canonical variates, it would be reasonable to postulate a REA in the perception of F1 information.

The importance of F1 in the right channel can be accounted for by the protensity features, tense and lax. Tense vowels are produced with added muscle tension while lax vowels are produced without such tension (Jakobson & Halle, 1974). In the English vowel system, some vowels are tense, others are lax, and still others are neutral for the feature tenseness. Lindau (1971) has shown, via radiographic recordings of the vocal tract the importance of tongue height (F1) in the production of tense and lax vowel sounds.

The results of this third experiment suggest that there are quantitative ear differences present in the perception of the most important acoustic parameters of vowel sounds, F2 and F1, respectively. The left hemisphere (right ear) was found, by means of canonical correlation, to be preferentially engaged in F1 perception, and the right hemisphere (left ear) in F2 perception. It is of interest to note that each ear is involved in the perception of these two physical characteristics but that they nonetheless vary with respect to the relative importance each ear gives to each of these two physical features of vowel sounds. This would suggest that the final output in vowel perception is a coordinated combination of these two sources of perception with the left ear probably playing a larger role (given the importance of F2 in speech percepton) at the input stage.

Ear advantages for vowels have traditionally been expressed in terms of number of correct responses. The dichotic listening paradigm, probably because it requires categorical responses on the part of the subjects, does not easily allow an analysis of ear differences to the perception of acoustic parameters varying on a continuous scale. Thus, finding that the right ear performs slightly better (i.e., reports slightly more accurately) than the left in vowel perception (e.g., Studdert-Kennedy & Shankweiler, 1970) does not tell us much about what it is in the vowel sound that the right ear is more sensitive to. The right ear advantage usually observed in speech perception may thus partly be the result of producing a behavioral response in a categorical mode. The interest of the present study (and of the adopted approach) is that it presents evidence of hemispheric involvement at the level of the defining features of the stimuli and not necessarily at the level of the stimulus complex.

The three experiments described have attempted to answer the question of the participation of each hemisphere to the perception of English speech sounds. The findings have consistently suggested that it is the case, both for consonant and vowel perception, that hemispheres preferentially managed defining features of speech sounds with no particular hemisphere dominating perception of a whole class of speech sounds. In the case of consonants, the derived perceptual representations suggested that the right hemisphere was more involved than the left in the treatment of voicing information and the left more involved than the right in the treatment of place of articulation information. With respect to vowel perception, canonical correlation results suggested greater involvement in F2 perception on the part of the right hemisphere and a greater participation on the part of the left hemisphere in F1 perception.

The stimuli used in these experiments represented classes of speech sounds drawn from the participants' native speech repertoire. The fourth experiment represents an attempt to generalize the findings of the previous experiments to a completely different kind of speech sounds. In this last experiment, each hemisphere's participation in the acquisition of new phonetic categories (Mandarin Chinese speech) is investigated.

EXPERIMENT 4

Perceptual processing of speech can be characterized as the transformation of a continuous acoustic signal into an ordered sequence of discrete, phonemic units. This process of transformation is, in the young child, gradually acquired as the child is exposed to the sounds of the language spoken in his environment. Presumably, it is by encountering language in its many forms and representations -as it is spoken by speakers differing in age, sex, accent and so on- that the child is eventually able to discriminate the sounds of the language. It is therefore through such exposure to examples of the fifty or so phonemes making up the English language uttered a large number of times in a large number of contexts that the young child comes to acquire the phonetic repertoire of English.

This experiment addresses the question of the participation of each hemisphere during the acquisition of new speech sounds by adults. The results of Experiments 1, 2, and 3 showed that each cerebral hemisphere may processe speech information somewhat differently than the other. It is therefore of interest to determine whether novel linguistic material consisting of speech sound categories absent in the subjects' mother tongue would engage each hemisphere's capabilities to a different extent than would be the case if previously experienced speech sounds were used as stimuli.

Multidimensional scaling analysis of dissimilarity ratings should produce, upon first exposure to these new speech categories, ill-defined categories with members distributed in a near-random manner throughout the representational space. With increased exposure to the sounds of the new language, categories would become more distinct or better differentiated, resulting in a more compact clustering between members of the categories in the representational space. Wakefield, Doughtie and Yom (1974) have shown that a 45 minute exposure to an unfamiliar language (Korean) was sufficient to permit subjects to identify structures of sentences significantly better than a control group with no exposure to that language. In their experiment, the subjects' task was to choose between two versions of a sentence -one with a pause interrupting a structural component and one with a pause separating different structural components- the version that sounded more natural.

If increased exposure to the sounds of a new languge leads to increased discrimination between those sounds, what, then, is the role played by each hemisphere in the acquisition of these new speech categories, and is the contribution of each hemisphere quantitatively/qualitatively different throughout this learning process?

To investigate these questions, the present experiment was designed as follows. English and French speaking subjects with no previous experience in Mandarin Chinese, performed a dissimilarity rating task on the first, third and fifth days of the experiment. On days two and four they listened to a thirty minute text of a recorded story spoken by a native speaker of Mandarin.

There were four different target consonants to differentiate, none of which are found in English or French. Two of these consonants were differentiated on a voiced/unvoiced dimension, and the other two on an aspirated/unaspirated dimension, which is not phonemic in either English or French.

The dissimilarity ratings obtained for these consonants were subjected to MDS analysis to obtain a perceptual configuration for each subject. The ratios of within category to between category distances wee derived from these perceptual configurations.

There were two groups of subjects in the experiment. One group, described above, received listening experience with long passages of Mandarin containing the target sounds. A second group of English and French speaking subjects performed only the rating task and were not eposed to any listening material during the three sessions. This condition controlled for the effect of performing the rating tasks.

The target sounds used were four Mandarin retroflexes. The descriptions of the target sounds are taken from Dow (1972) and the symbols used are taken from Suen (1979). They are:

ZH- voiceless unaspirated retroflex d with voiceless retroflex z

CH- the aspirated homorganic pair of ZH, voiceless aspirated retroflex t with voiceless retroflex s

SH- voiceless apical prepalatal fricative

R- the voiced homorganic pair of SH.

These retroflex sounds were embedded in four phonetic contexts to generate 16 stimuli: __U, __OU, I__ANG, and U__I.

According to Dow (1972), ZH becomes slightly more voiced, CH less aspirated and R more liquid in intervocalic positions. These phonetic contexts were selected to create coarticulation variation among the exemplars of the same target sound.

Method

Subjects.

Sixteen right-handed native speakers of English or French participated in this experiment. They were randomly assigned to an experimental or control group. Subjects were paid for their participation. Those in the experimental condition received $20, and those in the control condition received $12.

Stimuli.

The target sounds were the three Mandarin retroflexed /ZH/, /CH/, /SH/, and /R/ embedded in the four vowel environments described above. Fourth descending tone was always used with the final vowel as this made it possible to construct a meaningful story in Chinese.

The 16 stimuli were paired in all possible combinations excluding permutations to produce 120 pairs for the dissimilarity rating task. There were three such random pairings of stimuli: one presented to both ears, one presented to the left ear with accompanying white noise in the right ear, and one presented to the right ear with accompanying noise in the left one. Order ofpresentation was random, with no two consecutive blocks of 16 stimuli to the same channel. There was about one second pause between each stimulus item within a pair and about three seconds pause between successive pairs. .

The listening text consisted of a three-minute story recorded by a male native speaker of Mandarin, and contained the 16 exemplars of the target sounds with equal frequency (5 tokens each). this text was played continuously for thirty minutes.

Procedure.

In the rating phase of the experiment, subjects first heard the 16 different sounds followed by the test trials in random blocks of 16 pairs of stimuli in random order. They indicated the extent to which they judged the dissimilarity of each stimulus pair by marking a rating scale booklet containing 100mm lines labeled similar-dissimilar for this purpose, where 0 indicated maximal similarity and 100 indicated maximal dissimilarity. The subjects were encouraged to use the full scale as much as possible but were not required to do so.

All subjects performed the rating task on the first, third and fifth days of the experiment. Order of presentation was the same for all subjects across the three sessions with order of channel stimulation counterbalanced. In the listening task, subjects in the experimental condition were told to monitor the occurrences of three different "words" by making pencil strokes where appropriate on a sheet of paper. This was done in order to ensure repeated exposure to exemplars of the stimulus categories as well as to variations in the speech of one native Mandarin Chinese speaker. The three words were phonologically unrelated to the target sounds. Subjects were not informed about the identity of the language until after completing the experiment. They were tested singly or in groups of two, with each session about one hour.

Results and Discussion

The dissimilarity ratings for the 16 stimuli produced by each of the 16 subjects on the first, third and fifth days for the left, right and both channels conditions were subjected to MDS analysis using the M4 model of MULTISCALE. All data were scaled in two and in three dimensions in order to determine the optimal dimensionality of the solutions. Application of Ramsay's stopping rules indicated a three-dimensional solution was justified (corrected $X^2 > 47.1$, $p < .05$) in several cases while in the rest only a two-dimensional solution was justified.

The projections for each of the 16 stimuli on the axes defining the final multidimensional space accepted for each subject were analysed to compute the interpoint distances between each of the stimuli. These distances were then used to calculate the degree of conceptual structure exhibited by the elements in each perceptual representation. For this, the structure ratio W/B was calculated where W is the average distance between members within a particular category relative to B, the average distance to members of other categories (see Experiment 2).

To evaluate ear differences in the acquisition of new phonetic categories, structure ratios were computed for categories differentiated on voicing (SH and R) and for categories differentiated on aspiration (ZH and CH). Thus, computations were first made for SH exemplars diffrentiated from R exemplars and, conversely, for R exemplars diffrentiated from SH exemplars. These two ratios were averaged to yield a value reflecting the degree of structure for categories differentiated on voicing. The same procedure was followed for the two categories differentiated on voicing.

The structure ratios thus computed for the two groups of subjects were subjected to a 2 X 3 X 3 X 2 analysis of variance in which the factors were GROUP (text, no text), DAY (first, second, and third testing sessions), EAR (left, right, and both ears) and FEATURE (voicing, aspiration) with repeated measures over the last three factors.

The analysis revealed two significant interactions. The first, a two-way interaction of GROUP by DAY, (F(2,28)=9.485, MSe=.064, p<.001) showed that the text group demonstrated significant overall improvement between the first and third testing sessions while the no text group did not. Overall structure ratios, calculated by collapsing across all three ear conditions and both the voicing and aspiration features, are shown in Figure 13.

The second significant interaction is a DAY by CONDITION by FEATURE interaction, (F(4, 56)=4.796, MSe=.037, p=.0021). Figure 15 shows the evolution over sessions of the voicing and aspiration differentiation for each condition on

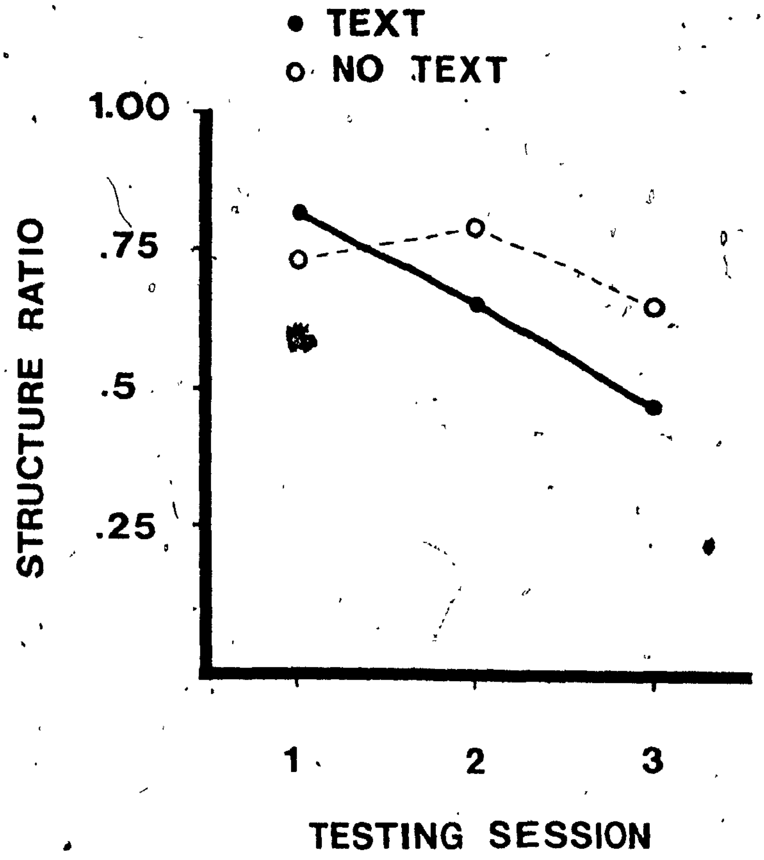

Figure 13. Evolution of the perceptual structure in the acquisition of new phonetic categories. Text group listened on two occasions to 30 min of Mandarin Chinese speech.

each testing day, for each group. It is apparent from the data displayed in Figure 14 that there are developmental differences between ear conditions, regarding the rate at which voicing and aspiration contrasted categories are acquired.

Let us first examine the results for the Text group in the LEFT EAR condition. There is marked improvement in the way categories are differentiated on voicing on each successive testing session. From Day 2 (i.e., the second day of testing), differences in the degree to which subjects discriminated between categories varying on voicing and those varying in aspiration is also apparent: the performance along the voicing dimension reflects better discrimination (post hoc Tukey test, $p<.01$). This difference becomes greater by the last testing session whereas no further improvement was evident regarding discrimination between categories differentiated on aspiration (.94 on Day 1 vs. .79 on Day 3, $p>.05$).

When we examine the results for the RIGHT EAR condition, we also find an improvement in voicing from first to second testing session of about the same magnitude as the one observed for the LEFT EAR condition ($p<.01$). There is also significant improvement in voicing from the second to the third testing session.

In contrast to the LEFT EAR condition, the RIGHT EAR shows gradual improvement by the end of the experiment for

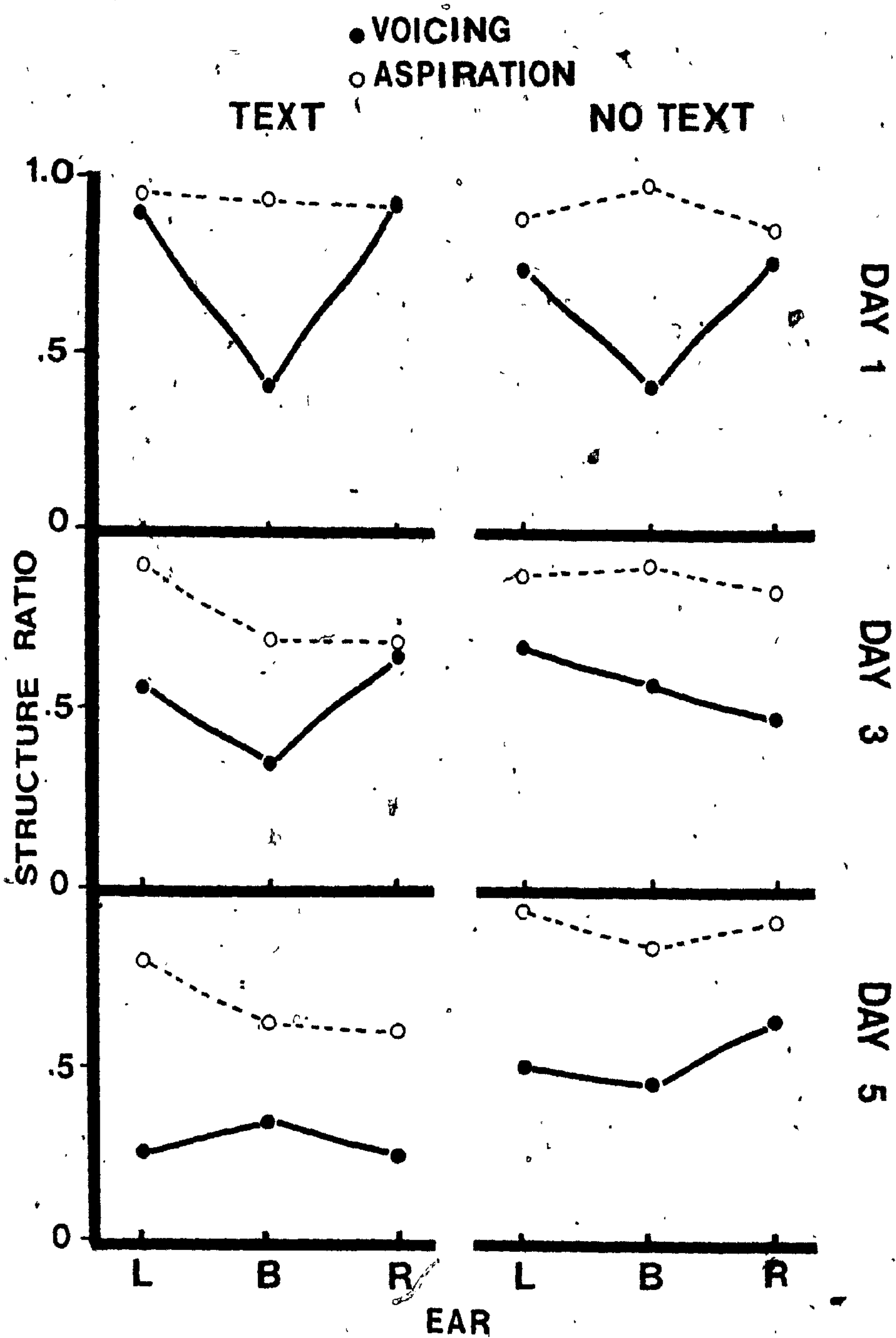

Figure 14. Rate of acquisition of aspiration and voicing contrasts.

the feature of aspiration (.90 on Day 1, and .58 on Day 3; p<.01). By Day 3, however, the voicing contrast appears much better developed than the aspiration contrast (.25 and .58 respeectively, p<.01).

Figure 14 shows that performance in the BOTH EARS condition remained stable throughout the experiment for the voicing feature. Performance was, however, better than in either RIGHT EAR or LEFT EAR on Day 1 (p<.01) as well as on Day 2, although only significantly so when compared to the right ear (p<.05) and about the same as either ear alone on Day 3. Performance in BOTH EARS for the feature of aspiration paralleled that of the right ear.

The no text group showed improvement only on the voicing contrast (.72 on Day 1 and .49 on Day 3; p<.05). Their performance on voicing in the BOTH EARS condition parallels that observed for the Text group, but they exhibit no improvement on aspiration. Both the Text and No text groups perform significantly better in discriminating voiced from voiceless categories than in differentiating categories which vary on aspiration.

Figures 15 and 16 depict representations of the perceptual space produced by the Text group in the LEFT EAR condition on Days 1 and 3 respectively. In Figure 15, the four exemplars of each category appear randomly distributed in the perceptual space, indicating poor conceptual

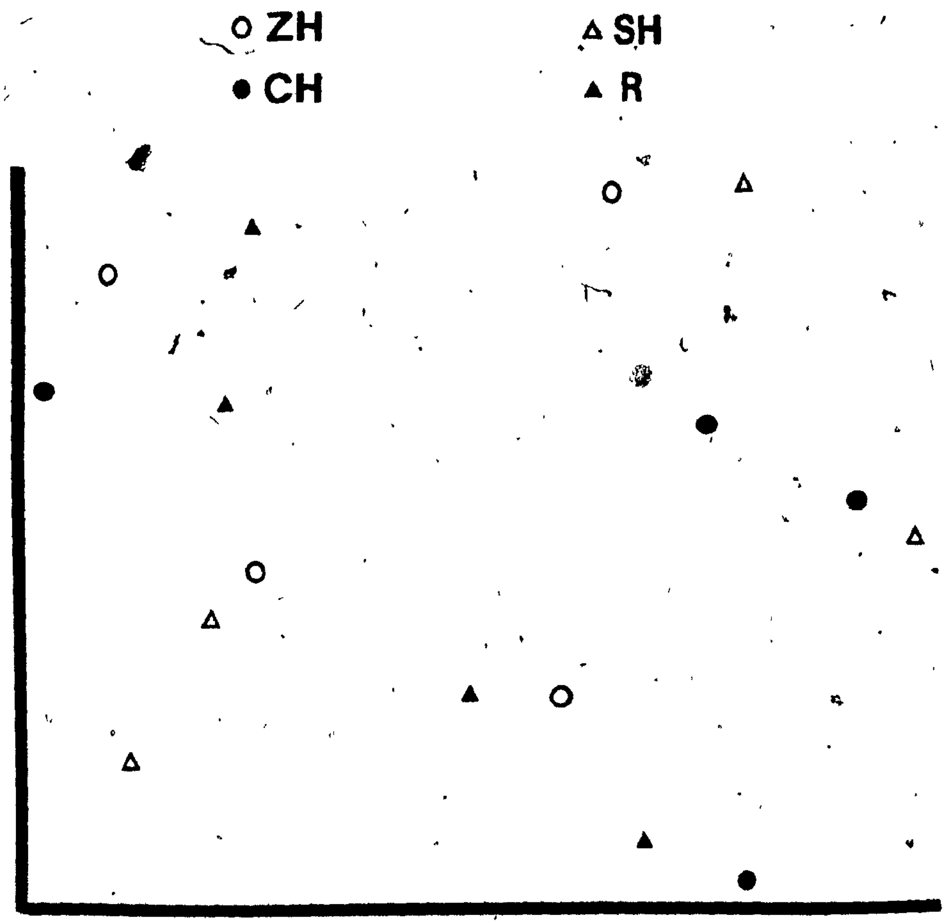

Figure 15. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of Mandarin Chinese phonetic categories presented to the left ear on Day 1.

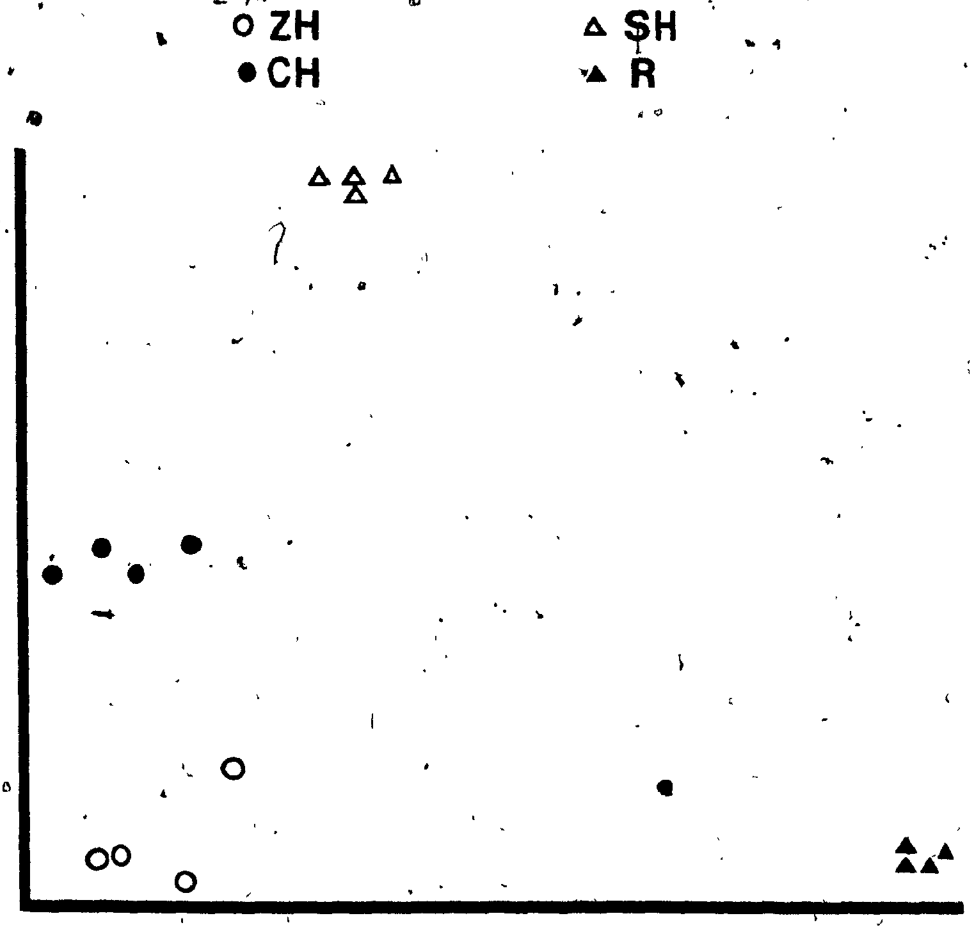

Figure 16. Plot of Dimension 2 (vertical axis) against Dimension 1 for the configuration of Mandarin Chinese phonetic categories presented to the left ear on Day 3.

structure. Figure 16 reveals that subjects clearly showed improvement over time in the structuring of their perceptual spaces. It can be seen in this figure that two categories of speech sounds -/SH/ and /R/- are very well differentiated from each other (mean structure ratio (averaged over subjects) for voicing contrast = .27). The exemplars of the aspirated category /CH/ and those of the unaspirated category /ZH/ are close to the other exemplars in their respective categories but the categories themselves are not as differentiated (mean structure ratio for aspiration contrast=.79) as the categories differing on voicing.

The aspirated-unaspirated distinction is of special interest because it does not represent a contrast found in either English or French. As can be seen from the poor structure ratios on Day 1 (Figure 14) for all conditions, the data support the expectation that these two categories would be difficult to discriminate from each other. Following our expectation, however, there was a significant improvement in differentiating the aspiration contrast although for the RIGHT EAR and BOTH EARS conditions only. Furthermore, this development was significant only for subjects in the Text condition.

A point of interest in the present study concerns the existence of ear differences in the formation of phonological perceptual categories. The Text group produced perceptual structures which showed enhanced discrimination by day 3 for categories contrasted on voicing, in both the LEFT EAR and RIGHT EAR conditions. The subjects' perceptual representations showed at the beginning ill-defined, overlapping categories which, after subsequent testing and repeated exposure to the exemplars of these categories in the listening text, resulted in relatively well-structured categories. The right ear, however, was demonstrably superior to the left by the third testing session in discriminating aspirated /CH/ exemplars from unaspirated /ZH/ exemplars.

As exposure to the unfamiliar aspirated and unaspirated phonetic categories increased, the subjects' initial perceptual representations of these categories underwent structural changes (reflecting conceptual shifts) which were presumably affected by the adoption of new criteria in the judgment of stimuli. These changes may reflect a process of feature extraction whereby stimuli are analyzed as a set of features which in turn help construct and affect the ideal, prototypical representation of these conceptual categories. These ideal representations are further refined with increased exposure to members of the concepts, suggesting gradual transformations or refinements of the prototype for these particular categories. It is not surprising that voicing contrasts were acquired faster and better than aspiration contrasts, voicing being an important feature of the English and French sound systems.

The important point, as noted above, is the REA in the extraction of the aspiration feature, evidenced by enhanced discrimination from near random representtion (mean structure ratio=.9) to significantly better structured representation (.58), a process which is believed to be analytic in nature.

Thus, there is clearly evidence of an interaction between the nature of the stimulus and ear of presentation. That this effect is not only due to the nature of the task is shown by the ability of the right hemisphere to learn to discriminate between those exemplars differentiated on voicing only. This strongly suggests the important participation of the left hemisphere in the difficult task of extracting new relevant phonetic information.

If this process, observed in adults, mirrors development of lateralization in early speech perception, it may be reasonable to assume early involvement in speech analysis on the part of each hemisphere with the left one possibly more attentive to difficult features.

In summary, the results of this experiment suggest first that listening experience leads to increased perceptual differentiation of new phonetic categories. The results also suggest faster and better differentiation of new phonetic categories varying along a previously familiar phonological contrast (voicing) and a slower but significant evolution on the part of the right ear only in the differentiation of new phonetic categories contrasted along a previously unknown dimension (aspiration) suggesting 'qualitative' hemispheric differences in the processing of new phonetic information.

Individual Differences in the Perceptual Representations of Speech Sounds

Individual differences plague neurolinguistic research just as much as in any other field involving live organisms (Segalowitz & Bryden, in press). Such differences may be due to individual variations in task performance, differences in brain morphology, as well as to differences in developmental experiences that may affect the way language comes to be organized in the brain. It can readily be accepted that no two individuals will perceive a set of stimuli in exactly the same way. Moreover, no two individuals will use the response medium in exactly the same way.

This problem is especially true with multidimensional scaling analyses. For example, MDS results obtained from grouped data may reveal a higher dimensionality than individual data because different individuals use different dimensions or properties in making the judgments. Dimensionality for grouped data should thus be interpreted partly as a consequence of intersubject variation in dimensions used as well as the dimensionality of any subject's perceptions (Ramsay, 1978).

The MDS program used in the present research (program M3 of MULTISCALE) assumes that the subjects share a common, perspective on the speech stimuli but that they nonetheless vary with respect to their relative use of the dimensions underlying the group's perceptual representations. A glance at Figures 17 and 18 reveals the extent of individual variation in each ear condition. These figures show that subjects appear to use each dimension differently (as reflected by the weight estimates, an index of the subject's consideration of a particular dimension) from other subjects and to a different extent from one condition to the next. Furthermore, there are important differences in the number of dimensions underlying the spaces derived from presentations to the two ears as well as in the salience each ear gives to each dimension. In Experiment 3, for example, four dimensions account for a better fit of the data for the left ear while only three do so for the right ear. The extraction of a larger number of dimensions may be indicative of a more complex or sophisticated perceptual skill in that more features or properties of the stimuli are taken into account.

The dimensionality derived for a particular class of stimuli presented to a given ear may also be subject to variation due to increased amounts of experience with those stimuli. In Experiment 4, only two dimensions accounted for the representation of initially unfamiliar stimuli presented to the left ear on the first day. By the fifth day, however, four dimensions now define the group's perceptual space. A similar development is observed for the right ear with three dimensions defining the group's representation of Chinese

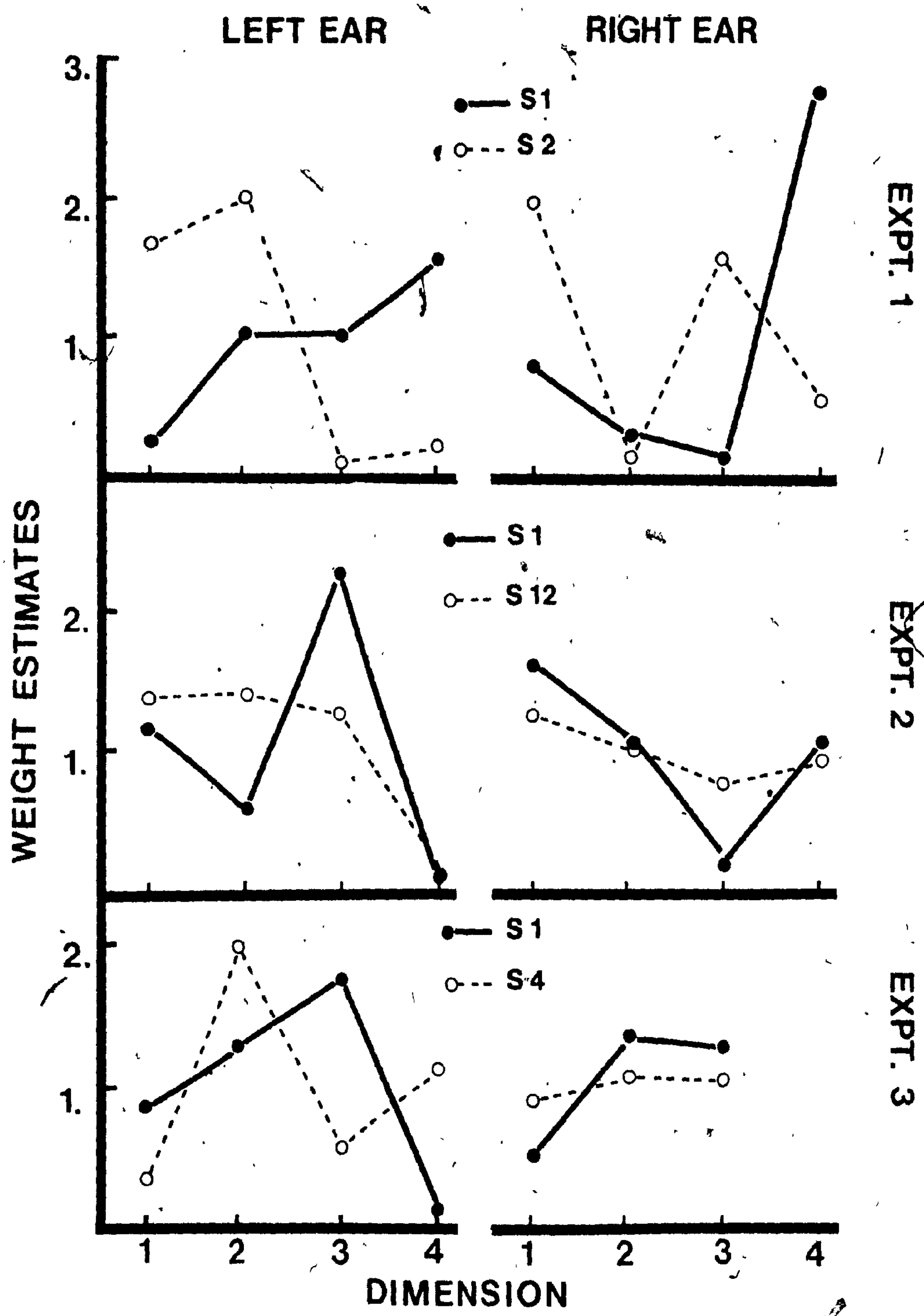

Figure 17. Final weight estimates for two subjects in Experiments 1, 2 and 3.

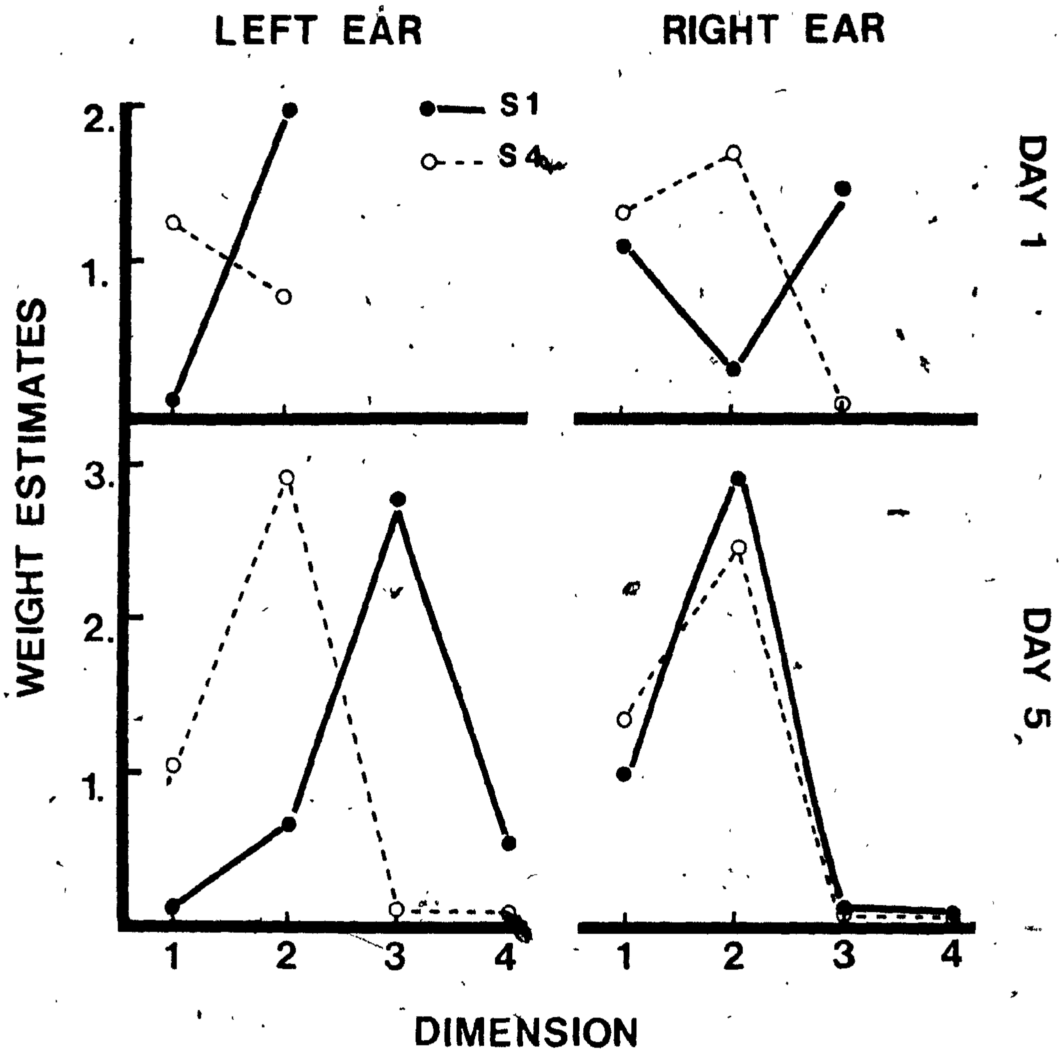

Figure 18. Final weight estimates for two subjects in Experiment 4.

sounds at the beginning of the experiment and four dimensions at the end of the experiment.

These data suggest different degrees of involvement on the part of each hemisphere in the perception of different classes of speech sounds. These results argue for the notion that the two hemispheres do not give equal consideration to the defining features or properties of a set of stimuli and that listeners vary with respect to the relative significance accorded to each dimension.

GENERAL DISCUSSION

The experiments reported in this thesis represent an attempt to advance our understanding of hemispheric specialization in speech perception. The research used a variant of the dichotic listening paradigm in which different classes of speech sounds were presented to investigate each hemisphere's contribution to the perception of consonants, vowels and new phonetic categories.

The results of the first experiment suggested participation of both hemispheres to the perception of consonants varying in manner and place of articulation as well as in voicing. The data suggested a LEA in voicing and a REA in place of articulation. Manner of articulation, it was found, constituted a parameter which appeared to be engaged in all ear conditions. These results were only suggestive in the sense that their interpretation relied principally on explaining the meaning of the dimensions defining the perceptual representations and on assuming that differences in the perceptual spaces had their origin in the differential treatment of auditory information by each ear.

The second experiment was therefore conducted to obtain quantifiable information about ear advantages in the perception of consonants. The sounds chosen varied on two important features: voicing and place of articulation. The experiment produced evidence, in the form of structure ratio differences, for different patterns of processing in the right and left hemispheres of voicing information. Specifically a LEA was found for the differentiation of the voicing contrast. Thus, the data suggested that both hemispheres may be involved in the perception and extraction of both place and voicing information, but that they do so to different degrees.

Others, of course have also found hemispheric differences in the processing of speech sounds. For example, Cutting (1974) used a modified dichotic listening procedure in which he presented subjects with synthetic CV and V syllables. The CV syllables contained an initial frequency transition which identified a specific consonant, while the V syllables did not. Cutting found that stimuli which contained an initial transition element were better discriminated when presented to the right ear and concluded that these findings reflected different processing capabilities of the two hemispheres. In a related study, however, Cutting (1974) found that the LH did not appear to discriminate between phonetic and non phonetic transitions. It should be noted that although these findings were based on small, albeit reliable, differences, Cutting overlooked the fact that performance in the left ear condition was still above chance level. Freides (1977) has argued that techniques such as dichotic listening procedures that employ competition and masking measures are heavily influenced by output factors. Cutting's procedures may therefore have measured response strategies rather than processing dominance. The failure of Molfese (e.g., Molfese, Freeman and Palermo, 1975; Molfese, 1978a) to replicate Cutting's right ear advantage effect may have been due to the sensitivity of AEP procedures to hemispheric processing of incoming activation.

The results of the present experiment go a step further towards determining hemispheric involvement in speech perception. What the data revealed was not a particular ear advantage for a certain class of stimuli (e.g., consonants, vowels) but rather a preferred (and not exclusive) involvement on the part of each hemisphere for certain components of the speech signal.

Electrophysiological investigations of linguistic events have demonstrated, at least at the level of the input stage, active responding on the part of both hemispheres and an absence of interaction with task variables indicating that the hemispheres responded to all the stimuli in a different fashion (Molfese, et al., 1975; Molfese, 1978a). More recent results suggest also a strong contribution, as revealed by AEP techniques, on the part of the RH in the perception of voicing.

The data from Experiments 1 and 2 yielded results in accord with those obtained by Molfese (1980a). The fact that the RIGHT EAR and LEFT EAR perceptual representations , derived from exactly the same information, differed on the importance, or salience, given to the features of place and voicing, adds to the suggestion that the components of the stimuli interacted in some different fashion with the processors accessed by each channel.

The same phenomenon may also be inferred from the results of the third experiment in which vowel perception was investigated. Once again, no particular ear advantage for vowel perception as such was found, but rather a differential involvement on the part of each hemisphere to different acoustic parameters important in vowel perception and recognition.

The results of the fourth experiment showed gradual involvement by each ear, as measured by increased discrimination, following listening exposure to exemplars of the foreign stimuli. That increased discrimination resulted at least in part from listening to Chinese material and not only from familiarity with the task is shown by the different developmental course followed by the experimental and control groups. That hemispheric processing as such was involved is also suggested by the differential development of acquisition of the two features investigated, aspiration and voicing. An interaction was observed showing a REA (left hemisphere participation) in the extraction of the apparently more difficult aspiration feature. These results suggest that when confronted with a difficult linguistic task, the left hemisphere may be called upon (at least initially) to a greater extent than the right in identifying the unfamiliar features of a foreign language. It would be of interest to find out, whether, given time and more experience with the exemplars, the right hemisphere could catch up with the left in the acquisition of a difficult phonetic feature. These results are, of course, only preliminary. The study, however, presents an innovative approach to the study of hemispheric functional specialization that can be used to test for the psychological reality of the processes assumed to be involved.

Many authors involved in neuropsychological research have restricted themselves to ascribing functions to either the right or left hemisphere. In such cases the hemispheres are then described as two more or less independent systems to which functions are attributed on the basis of a dichotomous principle. Suggestions that individuals may demonstrate differences in their "degree" of functional asymmetry have, however, been made by a number of investigators (e.g., Hardyck, 1977; Zangwill, 1960). Their position is that there are varying degrees of functional asymmetry between the hemispheres. Sources of evidence for the notion of laterality as a continuous phenomenon, as well as interpretive theoretical models, are increasingly abundant.

One type of model that has become popular and that may shift the focus of research away from broad characterization of left and right hemisphere capabilities (e.g., verbal-nonverbal, logical-prelogical; Bradshaw and Nettleton, 1981) to a more detailed view of laterality phenomena is the information processing model (Beaumont, 1974; Fowler, 1975; Hardyck, 1977; Moscovitch, 1979). This model views the brain as an information processing system, built up of discrete components. The actual functioning of these components, as well as their organization and representation, may vary across both subjects and tasks. So rather than providing answers to "what it is that characterizes the specific functions of the right and left hemispheres in the normal adult" these new approaches intend to explain "how the commissures act in providing information transfer, between the hemispheres, and in constraining, or modulating the activities in the parallel halves of the brain in such a way that a functional asymmetry arises and is maintained" (Teuber, 1973, p. 71).Taken together, these observations add support to the notion that it is less and less acceptable to describe left hemisphere dominance in terms associated with the traditional verbal/nonverbal dichotomy.

This study presented an innovative approach to the study of hemispheric functional specialization which can be used to test for the psychological reality of the processes assumed to be involved. This approach can be applied to the study of the specificity and plasticity of brain functions underlying speech during normal and abnormal language development, and in the dissolution and recovery of speech or language after cerebral lesions. At one level, differences in the recovered perceptual configurations, for example, would argue for the existence of specific differences between age groups in the processing of speech or language. At another level, we can advance our knowledge of the organization or reorganization of linguistic and other cognitive functions in the brain by investigating differences in perception between clinical populations (e.g., acallosal, hemispherectomized, callosotomized) and reference groups. Finally, the experiments reported here have shown, with different classes of speech sounds, the active role played by both hemispheres in speech perception. More importantly, the results also suggested that this participation to the process of speech analysis occurred along divisions defined by the features of the stimulus rather than by the class of the stimulus.

References

Allard, F., & Scott, B.L. Burst cues, transition cues and hemispheric specializations with real speech sounds. Quarterly Journal of Experimental Psychology, 1975, 80, 66-74.

Bever, T.G. The nature of cerebral dominance in speech behavior of child and adult. In R. Huxley and E. Ingram (Eds.), Models and methods. New York: Academic Press, 1971.

Bever, T.G., & Chiarello, R.G. Cerebral dominance in musicians and nonmusicians. Science, 1974, 185, 137-139.

Berlin, C.I. Asymmetry in auditory tasks. In S. Harnad, R.W. Doty, L. Goldstein, J. Jaynes, and G. Krauthamer (eds.), Lateralization in the nervous system. New York: Academic Press, 1977.

Blumstein, S., Goodglass, H., & Tartter, V. The reliability of ear advantage in dichotic listening. Brain and Language, 1975, 2, 226-235.

Bradshaw, J. L., & Nettleton, N. C. The nature of hemispheric specialization in man. The Behavioral and Brain Sciences, 1981, 4, 51-63.

Broadbent, D. E. The role of auditory localization in attenuation and memory span. Journal of Experimental Psychology, 1954, 47, 191-196.

Brown, J.W., & Jaffe, J. Hypothesis on cerebral dominance. Neuropsychologia, 1975, 13, 107-110.

Bryden, M.P. Tachistoscopic recognition, handedness, and cerebral dominance. Neuropsychologia, 1965, 3, 1-8.

Bryden, M.p., & Allard, F. Visual hemifield differences in dot localization. Perception and Psychophysics, 1976, 19, 23-28.

Cohen, A., Slis, I.H., & 't Hart, J. On tolerance and intolerance in vowel perception. Phonetica, 1967, 16, 65-70.

Cohen, H. The psychological representation of vowel sounds. Unpublished M.A. Thesis. Department of Psychology, Concordia University, 1978.

Cutting, J. E. Two left hemisphere mechanisms in speech perception. _Perception and Psychophysics_, 1974, _16_, 601-612.

Cutting, J.E. The magical number two and the natural categories of speech and music. In N.S. Sutherland (Ed.), _Tutorial essays in Psychology_. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Cutting, J.E., & Rosner, B.S. Categories and boundaries in speech and music. _Perception and Psychophysics_, 1974, _16_, 564-570.

Cutting, J.E., & Rosner, B.S. Discrimination functions predicted from categories in speech and music. _Perception and Psychophysics_, 1976, _20_, 87-88.

Darwin, C. J. Ear differences in the recall of fricatives and vowels. _Quarterly Journal of Experimental Psychology_, 1971, _23_, 46-62.

Diehl, R. L. Feature detectors for speech: A critical reappraisal. _Psychological Bulletin_, 1981, _89_, 1-18.

Dow, F. _An introduction to the pronouciation of Chinese._ Edinburgh: University of Edinburgh, 1972.

Eimas, P.D., Siqueland, E.R., Jusczyck, P.W., & Vigorito, J.M. Speech perception in infants. Science, 1971, 171, 303-306.

Fowler, C.A. A systems approach to the cerebral hemispheres. Haskins Status Report on Speech Research, SR-44, 1975.

Freides, D. Do dichotic listening procedures measure lateralization of information processing or retrieval strategy. Perception and Psychophysics, 1977, 21, 259-263.

Gates, A., & Bradshaw, J. The role of the cerebral hemispheres in music. Brain and Language, 1977, 9, 403-431.

Gazzaniga, M. S., & Sperry, R. W. Language after section of the cerebral commissures. Brain, 1967, 90, 131-148.

Gazzaniga, M. S. Cerebral dominance viewed as a decision system. In S.J. Dimond and J.G. Beaumont (Eds.), Hemisphere function in the human brain. Paul Elek: London, 1974.

Geffen, G., Bradshaw, J., & Wallace, G. interhemispheric effects on reaction time to verbal and nonverbal stimuli. Journal of Experimental Psychology, 1971, 87, 415-422.

Gibson, A., Dimond, S., & Gazzaniga, M. S. Left visual field superiority for word matching. Neuropsychologia, 1972, 10, 463-466.

Gross, M.M. Hemispheric specilization for processing of visually presented verbal and nonverbal stimuli. Perception and Psychophysics, 1972, 12, 357-363.

Hardyck, C. A model of individual differences in hemispheric functioning. In H. Whitaker and H.A. Whitaker (Eds.), Studies in neurolinguistics. New York: Academic Press, 1977.

Heilman, K.M., & Watson, R.T. The neglect syndrome -- A unilateral defect of the orienting response. In S. Harnad, R.W. Doty, L. Goldstein, J. Jaynes, and G. Krauthamer (Eds.), Lateralization in the nervous system. New York: Academic Press, 1977.

Homa, D., Rhoads, D., & Chambliss, D. The evolution of conceptual structure. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 11-23.

Jakobson, R. Why "mama" and "papa"? In selected writings: Roman Jakobson. The Hague: Mouton, 1962.

Jakobson, R. Child language, aphasia, and phonological universals. The Hague: Mouton, 1968.

Jakobson, R., & Halle, M. Phonology in relation to phonetics. In B. Malmberg (Ed.), Manual of Phonetics. Amsterdam: North-Holland/American Elsevier, 1974.

Jusczyck, P.W., Rosner, B.S., Cutting, J.E., Foard, C.F., & Smith, L. Categorical perception of nonlinguistic sounds by two month old infants. Perception and Psychophysics, 1877, 21, 50-54.

Kallman, H., & Corballis, M. Ear asymmetry in reaction time to musical sounds. Perception and Psychophysics, 1975, 17, 368-370.

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

Kimura, D. Left-right differences in the perception of melodies. Quarterly Journal of Experimental Psychology, 1964, 16, 355-358.

Kimura, D. Functional asymmetry of the brain in dichotic listening. Cortex, 1967, 3, 163-178.

Kimura, D., & Folb, S. Neural processing of backward speech sounds. Science, 1968, 161, 395-396.

Kinsbourne, M. The asymmetrical function of the brain. New York: Cambridge University Press, 1978.

Kinsbourne, M., & Hicks, R.E. Mapping cerebral functional space: Competition and collaboration in human performance. In M. Kinsbourne (Ed.), The asymmetrical function of the brain. Cambridge University Press, 1978.

Kinsbourne, M., & Hiscock, M. The development of cerebral dominance. In S. Segalowitz and F.A. Gruber (Eds.), Language: Developmental and neurological theory. New York: Academic Press, 1977.

Klatzky, R., & Atkinson, R. C. Specialization of the cerebral hemispheres for information in short term memory. Perception and Psychophysics, 1971, 10, 335-338.

Klein, W., Plomp, R., & Pols, L. C. W. Vowel spectra, vowel spaces, and vowel identification. Journal of the

Acoustical Society of America, 1970, 48, 999-1009.

Krashen, S. Lateralization, language learning, and the critical period: some new evidence. Language Learning, 1973, 23, 63-74.

Lenneberg, E. Biological foundations of language. New York: Wiley, 1967.

Levy, J. Psychobiological implications of human asymmetry. In S. J. Dimond and J. G. Beaumont (Eds.), Hemisphere function in the human brain. London: Paul Elek, 1974.

Levy, J., & Reid, M. Variations in cerebral organization as a function of handedness, hand posture in writing, and sex. Journal of Experimental Psychology, 1978, 107, 119-144.

Liberman, A.M. Some characteristics of perception in the speech mode. In D.A. Hainburg and K. Pribram (Eds.), Perception and its disorders. Baltimore: Williams and Wilkins, 1970.

Liberman, A.M. The specialization of the language hemisphere. In F.O. Schmitt and F.G. Worden (Eds.), The neurosciences: Third study program. Cambridge, Mass.: MIT

Press, 1974.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. Perception of the speech code. Psychological Review, 1967, 74, 431-461.

Liberman, A. M., & Pisoni, D. B. Evidence for a special speech perception subsystem in the human. In T. H. Bullock (Ed.), Recognition of complex acoustic signals. Berlin: Dahlem Konferenzen, 1978.

Lindau, M. Vowel features. Language, 1978, 54, 541-563.

Miceli, G., Gainotti, G., Caltagirone, C., & Payer-Rigo, P. Discrimination of voice versus place contrasts in aphasia. Brain and Language, 1978, 6, 47-51.

Miller, G.A., & Nicely, P.E. An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 1955, 27, 338-352.

Milner, B., Taylor, L., & Sperry, R. W. Lateralized suppression of dichotically presented digits after commissural section in man. Science, 1968, 161, 184-186.

Molfese, D. L. Left and right hemisphere involvement in speech perception: Electrophysiological correlates. Perception and Psychophysics, 1978, 23, 237-243 (a).

Molfese, D. L. Electrophysiological correlates of categorical perception in adults. Brain and Language, 1978, 5, 25-35 (b).

Molfese, D. L. hemispheric specialization for temporal information: Implicatins for the perception of voicing cues during speech perception. Brain and Language, 1980, 11, 285-299 (a).

Molfese, D. L. The phoneme and the engram: Electrophysiological evidence for the acoustic invariant in stop consonants. Brain and Language, 1980, 9, 372-376 (b).

Molfese, D.L., Freeman, R.B., & Palermo, D.S. The ontogeny of brain lateralization for speech and nonspeech stimuli. Brain and Language, 1975, 2, 356-368.

Molfese, D.L., & Hess, T.M. Hemispheric specialization for VOT perception in the preschool child. Journal of Experimental Child Psychology, 1978, 26, 71-84.

Molfese, D. L., & Molfese, V. J. VOT distinctions in infants: Learned or innate? In H. A. Whitaker and H. Whitaker (Eds.), Studies in neurolinguistics (vol. 4). New York: Academic Press, 1979.

Morais, J., & Darwin, C.J. Ear differences for same-different reaction times to monaurally presented speech. Brain and Language, 1974, 1, 383-390.

Moscovitch, M. Choice reaction time study assessing the verbal behavior of the minor hemisphere in normal adult humans. Journal of Comparative and Physiological Psychology, 1972, 80, 66-74.

Moscovitch, M. Information processing and the cerebral hemispheres. In M.S. Gazzaniga (Ed.), Handbook of behavioral neurobiology. New York: Plenum Press, 1979.

Moyer, K. E. Neuroanatomy. New York: Harper and Row, 1980.

Perecman,E., & Kellar, L. The effect of voice and place among aphasic, nonaphasic right-damaged, and normal subjects on a metalinguistic task. Brain and Language, 1981, 12, 213-223.

Peterson, G.E., & Barney, H.L. Control methods used in a study of the vowels. Journal of the Acoustical Society of America, 1952, 24, 175-184.

Pisoni, D.B. Auditory and phonetic memory codes in the discrimination of consonants and vowels. Perception and Psychophysics, 1973, 13, 253-260.

Plomp, R., Pols, L.C.W., & van de Geer, J.P. Dimensional analysis of vowel spectra. Journal of the Acoustical Society of America, 1967, 41, 707-712.

Pols, L.C.W., van der Kamp, L.J.Th., & Plomp, R. Perceptual and physical space of vowel sounds. Journal of the Acoustical Society of America, 1969, 46, 458-467.

Ramsay, J.O. Maximum likelihood estimation in multidimensional scaling. Psychometrika, 1977, 42, 241-266.

Ramsay, J.O. Multiscale. National Educational Resources: Chicago, 1978.

Remez, R. E. Adaptation of the category boundary between speech and nonspeech: A case against feature detectors. Cognitive Psychology, 1979, 11, 38-57.

Remez, R.E., Rubin, P.E., Pisoni, D.B., & Carrell, T.D. Speech perception without traditional cues. <u>Science</u>, 1981, <u>212</u>, 947-950.

Rizzolatti, G., Umilta, C., & Berlucchi, G. Opposite superiorities of the right and left cerebral hemispheres in discriminative reaction time to physiognomical and alphabetical material. <u>Brain</u>, 1971, <u>94</u>, 431-442.

Romney, A. K., Shepard, R. N., & Nerlove, S. B, (Eds.), <u>Multidimensional scaling: Theory and applications in the behavioral sciences</u> (vol. 2). New York: Seminar Press, 1972.

Salus, P.H., & Salus, M.W. Developmental neurophysiology and phonological acquisition order. <u>Language</u>, 1974, <u>50</u>, 151-160.

Segalowitz, N., & Hansson, P. Hemispheric functions in the processing of agent-patient information. <u>Brain and Language</u>, 1979, <u>8</u>, 51-61.

Segalowitz, S.J., & Bryden, M.P. Individual differences in hemispheric representation of language. To appear in S.J. Segalowitz (Ed.), <u>Language functions and brain organization</u>. New York: Academic Press.

Segalowitz, S., & Gruber, F. Language development and neurological theory. New York: Academic Press, 1977.

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

Shepard, R. N. Psychological representation of speech sounds. In E. David and P. Denes (Eds.), Human communication: A unified view. New York: McGraw Hill, 1972.

Singh, S., & Woods, D.R. Perceptual structure of 12 American English vowels. Journal of the Acoustical Society of America, 1972, 49, 1861-1866.

Springer, S. P. Hemispheric specialization for speech opposed by contralateral noise. Perception and Psychophysics, 1973, 13, 391-393.

Springer, S. P., & Gazzaniga, M. S. Dichotic listening in partial and complete split brain subjects. Neuropsychologia, 1975, 13, 341-346.

Studdert-Kennedy, M., & Shankweiler, D. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, 1970, *48*, 579-594.

Studdert-Kennedy, M. Speech perception. *Language and Speech*, 1980, *23*, 45-65.

Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., & Cooper, F.S. Motor theory of speech perception: A reply to Lane's critical review. *Psychological review*, 1970, *77*, 234-249.

Studdert-Kennedy, M., Shankweiler, D., & Schulman, S. Opposed effects of a delayed channel on perception of dichotically and monotically presented CV syllables. *Journal of the Acoustical Society of America*, 1970, *48*, 599-602.

Suen, C. *Computational analysis of Mandarin*. Boston: Birkhauser, 1979.

Teng, E.L. Dichotic ear difference is a poor index for the functional asymmetry between the cerebral hemispheres. *Neuropsychologia*, 1981, *19*, 235-240.

Teuber, H.-L. Why two brains? In F.O. Schmitt (Ed.), *The* neurosciences: Third study program. MIT Press: Cambridge, 1973.

Torgerson, W. S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Wakefield, J.A., Doughtie, E.B., & Yom, B.-H.L. The identification of structural components of an unknown language. Journal of Psycholinguistic Research, 1974, 3, 261-269.

Weiss, M., & House, A. S. Perception of dichotically presented vowels. Journal of the Acoustical Society of America, 1973, 53, 51-58.

Zaidel, E. Auditory language comprehension in the right hemisphere following cerebral commissurotomy and hemispherectomy: A comparison with child language and aphasia. In E. Zurif and A. Caramazza (Eds.), The acquisition and breakdown of language: Parallels and divergencies. Baltimore: Johns Hopkins University Press, 1978.

Zangwill, O.L. Cerebral dominance and the relation to cerebral function. Edinburgh: Oliver and Boyd, 1960.

Appendix 1

Physical parameters of vowels

Vowel	F1 (Hz)	F2 (Hz)	F3 (Hz)	Duration (msec)
i	300	2100	2950	660
I	400	2000	2600	300
e	450	2000	2600	470
ɛ	500	2030	2600	320
ae	530	1960	2450	450
a	680	1250	2550	470
ɔ	600	900	2630	520
o	360	720	2210	480
U	450	1020	2500	340
u	330	865	2200	420
ʌ	490	1330	2490	470
ə	510	1450	2460	470