69-19,040

FRANKLIN, Barbara, 1933-
THE EFFECT OF A LOW-FREQUENCY
BAND (240-480 Hz) OF SPEECH ON
CONSONANT DISCRIMINATION.

The City University of New York, Ph.D., 1969
Speech

University Microfilms, Inc., Ann Arbor, Michigan

THE EFFECT OF A LOW-FREQUENCY BAND (240 - 480 Hz) OF SPEECH ON CONSONANT DISCRIMINATION

by

BARBARA FRANKLIN

A dissertation submitted to the
Graduate Faculty in Speech in partial
fulfillment of the requirements for the
degree of Doctor of Philosophy, The
City University of New York.

1969

This manuscript has been read and accepted for the University Committee in Speech in satisfaction of the Dissertation requirement for the degree of Doctor of Philosophy.

J Donald Harris
4/1/69 Dr. J. Donald Harris
date Chairman of Examining Committee

4/1/69 Dr. Joel Stark Joel Stark
date Executive Officer

Dr. Hayes Newby

Dr. Katherine Harris
Supervisory Committee

The City University of New York

ACKNOWLEDGMENTS

I want to thank Dr. Moe Bergman for his assistance in formulating this study. I am indebted to Dr. J. Donald Harris for exposing me to hearing science, and to Dr. Katherine Harris for introducing me to speech science. I should also like to acknowledge the influence of Dr. Beatrice Jacoby and particularly Mrs. Eleanor Ronnei on my professional career, and their inspiration which resulted in my return to graduate school in order to prepare myself to train others.

I want to thank my committee - Dr. J. Donald Harris, Dr. Hayes Newby, and Dr. Katherine Harris - for their guidance in the presentation and analysis of the data of the present investigation, and for all their suggested revisions leading to the final manuscript.

I am particularly grateful for the understanding, help, and encouragement of my husband, Dick Franklin, and the cooperation of my children, Marc and Matthew, during the events leading to the degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Figure Page

CHAPTER I

INTRODUCTION AND REVIEW OF THE LITERATURE

INTRODUCTION

The classical description of the relative contribution to intelligibility of various parts of the speech spectrum designates the high frequencies as the information-bearing and the low frequencies as the energy-bearing portion of the speech signal. Typical experiments to test communication equipment, such as those done by Fletcher (1929) and Steinberg (1929) at Bell Laboratories, confirmed this analysis, showing negligible impairment of intelligibility with high-pass filtering. Current hearing aids which have a cut-off at approximately 3000 Hz at the high- and 500 Hz at the low-frequency end of the frequency spectrum would therefore seem to be adequate. However, there is a population of profoundly deaf children who have most of their residual hearing below 500 Hz. A new type of broad-spectrum hearing aid which extends the low-frequency range down to 100 Hz was recently designed for the deaf child with a severe high-frequency drop, but with some residual hearing in the low frequencies. Ling (1965) reported in The Young Deaf Child that he did not find an improvement in discrimination scores with this type of hearing aid but believed that there was a marked improvement in rhythm and voice quality in addition to an extension of the field of audition.

The present writer's experience with young deaf children who were fitted with such broad-spectrum hearing aids is that these children showed improvement, not only in their ability to discriminate vowel sounds, but also to produce velar stops and to eliminate /b/ for /m/ substitutions. The improvement in vowel discrimination could have been predicted from the work in speech science which has located the first formant of vowels between 300-600 Hz. The improved labials could also have been expected since the first formant of consonants which contains manner-of-articulation information has been located in the low frequencies.

The improvement in velar production, however, was unexpected, since place-of-articulation information had not been located in the low frequencies. These empirical observations led to the present investigation of the effect on normal-hearing subjects of low frequencies on consonant discrimination, the low frequencies mixed either peripherally or centrally with the high frequencies.[1] It seemed to the writer that the effect of the low frequencies could best be studied by selecting a low-frequency band (LB) of speech which theoretically contains minimum consonant information and then add that low-frequency band to a high-frequency band (HB) known to be rich in consonant information.

Following is a brief summary of the review of the literature in which the writer will outline the contents of each section and indicate how each contributed both to the design of the present study and to the analysis of the experimental data.

Reports of binaural fusion and filtered speech experiments are reviewed in the first section. Tests designed to detect central hearing disorders were included, since, in addition to their intended diagnostic significance, they also provide a quantitative analysis of the information present in various band widths of the speech spectrum. The selection of the particular high-frequency band used in the present investigation was based on this information. Also included in the first section is a review of the literature which compared discrimination scores when various high- and low-pass band

---

[1] Deatherage (1966) reported ". . . that interaction 'between ears' can possibly take place at any level of the system, with several degrees of complexity, and over considerable amounts of time." (p. 233)

band combinations were (1) sent to the same ear, and (2) split, one band to each ear. This procedure was followed in the present investigation; the low-frequency band being added to the ear receiving the high-frequency band as well as to the opposite ear.

The second section of the literature review is concerned with reports which describe the effects of low frequencies, at various intensities, on the speech signal. This information enabled the writer to select a low-frequency band for the present study that would satisfy the following conditions: (1) theoretically contains negligible consonant information, (2) contributes to the speech signal when added at favorable HB/LB ratios, and (3) detracts from the speech signal at unfavorable HB/LB ratios. In the present investigation, the low-frequency band selected by the writer was added at various sensation levels relative to the high-frequency band, and a comparison was made of the effect of the LB on the HB in two conditions: (1) monotic, the two bands sent to the same ear, and (2) dichotic, the two bands sent to opposite ears.

The third section of the literature review involves reports dealing with factors which contribute to consonant discrimination. This literature provided the basis for the selection of the materials used by the writer, and also contributed to the design of the confusion matrices by means of which the data of this study are presented.

The final section of the review, dealing with spectrographic analysis of consonants, provided the background information used by the writer in the analysis of the experimental data.

REVIEW OF THE LITERATURE

Binaural Fusion and Filtered Speech Tests

Fletcher, in the 1929 edition of Speech and Hearing, reported an experiment which could be considered the precursor of all binaural fusion and distorted speech tests. The speech signal was divided into two channels, one passing all frequencies below 1000 Hz, and the other all frequencies above 1000 Hz. When only one channel was presented at a time, the speech signal was distorted and intelligibility was impaired. Fletcher found that when the two channels were sent to opposite ears, the brain was capable of combining the two signals ". . . to complete the proper picture." Black (1962) filtered words into two bands of equal intelligibility, one containing the frequencies up to 1600 Hz, the other the frequencies from 1600-15,000 Hz. When the bands were sent to opposite ears the scores were only slightly lower than when both bands were sent to the same ear.

When designing binaural fusion tests one must be aware of certain phenomena of audition. Broadbent and Ladefoged (1957), using synthetic speech, found that two different formants would fuse only when their fundamental frequencies were the same, and in unison, whether the two formants were presented to the same ear, or split, one formant to each ear. Pollack (1948) stated that one cannot consider the effects of individual bands of speech frequencies, and one must always remember "... there is an interaction among the contributions of the various bands." (p. 263) David, Guttman, and van Bergeijk (1958) found that while fusion of pulses of 1000 and 3000 readily occurred, fusion failed to occur when the pulses were far apart, such as 800 and 6000 Hz. They hypothesized that "Perhaps this indicates that while frequency overlap is not required, excitation of some common region of the basilar membrane in the two ears is required." (p. 802)

Licklider (1948) noted without further comment that "The average intelligibility scores were higher when the speech was presented to the listeners' right ears than when it was presented to their left ears." (p. 153) This matter of earedness has been investigated by two Canadian researchers, Milner (1962) and Kimura (1961^a, 1961^b, 1962), who found that subjects perceived verbal material better with the right ear than the left ear. The superiority of the right ear is most evident in a dichotic listening situation where there are competing stimuli, each ear receiving a different signal. Both authors offered as an explanation of this phenomenon that the crossed fibers leaving the cochlea have been found to be more effective than the uncrossed fibers so that each temporal lobe receives maximum stimulation from the contralateral ear. In addition, it has been found that, in the majority of cases, the left temporal lobe is the dominant lobe for speech reception. Kimura and Milner suggested therefore that the "right-ear" effect is the result of the dominant left hemisphere receiving more impulses from the crossed auditory pathway leaving the right cochlea. Kimura (1961^b) further reported that for those individuals who had speech reception in the right hemisphere there was indeed a "left-ear" superiority in the reception of verbal material. Shankweiler and Studdert-Kennedy (1967) reported a significant "right-ear" advantage of consonant-vowel syllables, but not for steady state vowels. Kimura, Milner, and Shankweiler all noted that the left temporal lobe appears to be the dominant lobe for the processing of the more difficult verbal material.

In Relation to the Diagnosis of Central Hearing Disorders

Many binaural fusion and filtered speech experiments were performed in an attempt to develop audiometric tests which could be used as diagnostic tools in determining lesions along the central auditory pathway.

Bocca, Calearo, and Cassinari (1954) stated that temporal lobe tumors are rarely characterized either by an increase of threshold for pure tones or by a lowering of speech discrimination scores. In designing a test to be used to diagnose central auditory disorders the researchers made two assumptions: first, that the speech signal would be more difficult to understand if its redundancy was reduced, and second, that the degraded speech signal would require more integrative functions which were located in the areas of the brain they wanted to investigate. Subjects with temporal lobe tumors were tested with phonetically balanced (PB) Italian bisyllabic words sent through a low-pass filter with a cut-off at 800 Hz so that their correct identification would ". . . involve psychic integration to a higher extent." (p. 220) The discrimination scores for most of the subjects were poorer in the ear contralateral to the brain lesion. Their follow-up report (1955) on these subjects, which compared the pre-operative and post-operative scores, reported that discrimination in the ear contralateral to the site of the tumor showed improvement from 20-30 days after its surgical removal.

These initial tests which examined the effects of filtered speech signals on discrimination scores were all done monaurally. Bocca (1955) reported the application of distorted speech tests to determine binaural summation and integration. Nonsense bisyllables and PB words were first presented undistorted to normal-hearing subjects, at an intensity which yielded a maximum articulation score of 30%. The same material was then sent through a low-pass filter, with a 30-35 dB/octave attenuation and a cut-off at 500 Hz, and was presented to the subject at 45 dB SL[2], yielding a maximum articulation score of 50%. The average scores for the monaural presentation of the PB words were 28% for the faint undistorted condition and 27% for the loud distorted condition. When the two bands were presented to the same ear, the score was 27%, but when the two bands were split, one to each ear, the score rose to 65%. The authors added "The louder stimulus behaves exactly as a masking stimulus when delivered to the same ear together with the weaker one." (p. 1169) Fletcher (1929) also found that the low-frequency sounds do not have the same interfering effect when the two tones are introduced into opposite ears. Licklider (1948) found that articulation scores were almost 100% when speech and noise were introduced into opposite ears.

---

[2] Sensation level (SL) refers to the number of decibels (dB) above threshold. In Bocca's experiment, the threshold for filtered speech was the reference intensity.

Calearo (1957), an associate of Bocca at the University of Milan, presented Italian PB words to subjects with diagnosed brain lesions in the temporal lobe area, at an intensity which yielded a maximum articulation score of 40%. The words sent through a low-pass filter with a cut-off at 500 Hz were presented to the subject at an intensity 40% above threshold[3] which yielded a maximum articulation score of 50%. Scores improved only when these two conditions were presented simultaneously, one to each ear, for "In this way, one ear provides quality and the other ear provides power." (p. 393) Summation failed to occur whenever the undistorted signal was sent to the ear contralateral to the lesion.

Matzker's (1962) test was based on the principle of binaural fusion. He filtered German PB words of two and three syllables into two bands: (1) a low-pass band of 500-800 Hz, and (2) a high-pass band of 1815-2500 Hz. Neither band alone contained enough information for discrimination, but when the low-pass band was sent to one ear and the high-pass band to the other, the normal brain fused the signals, and recognition of the words resulted. Matxker required only that the subject correctly identify the vowels in the words. Patients with brain lesions achieved lower scores

---

[3] The author does not clarify what he means by "40% above threshold."

when the bands were split to opposite ears than when both bands were sent to the same ear "... because of their failure of bilateral integration." (p. 210)

Hayashi, Ohta, and Morimoto (1966, 1967) designed their experiment to test Matzker's theory that audiometric tests involving the principle of binaural fusion can be used to diagnose central auditory disorders. Japanese nonsense monosyllables were filtered into two bands: (1) a low-pass band of 300-600 Hz, and (2) a high-pass band of 1200-2400 Hz. These researchers felt that Matzker's method was defective in that only one intensity level was used and no comparison was made of the monaural scores for the right and left ear. Hayashi et al. presented the material to patients with hearing losses at both 20 and 40 dB above threshold[4], and in many combinations of right and left ears, both monaurally and binaurally. The person's ability to fuse was based on the difference in his discrimination score when the two bands were presented to both ears and when the bands were split, one to each ear. The results seemed to indicate that binaural fusion is poor when the high-pass

---

[4] When both bands were presented to both ears, threshold was determined by each subject's pure-tone average for 500, 1000, and 2000 Hz. When the bands were split, one to each ear, the threshold for a 500-Hz and 1500-Hz tone was used for the low and high bands, respectively.

band is sent to the ear contralateral to the lesion. They concluded that binaural fusion takes place in cortical or subcortical areas and disagreed with Matzker's opinion that the phenomenon is a function of the brain stem.

Linden (1964) developed a "speech resynthesis test" for the purpose of discovering central disorders. Swedish spondees of fairly good intelligibility were filtered into two bands: (1) a high-pass band of 1800-2200 Hz, and (2) a low-pass band of 560-715 Hz. By inspection of the filter attenuation curves, there appears to be approximately a 50-60 dB/octave drop. At 40 dB SL[5], the discrimination score for normal-hearing subjects was 14% for the low-pass band, and 11.4% for the high-pass band. When the two bands were split, one to each ear, the discrimination score rose to 89% at 40 dB SL.

Jerger (1960) investigated the distorted speech technique in diagnosing lesions in the central auditory pathways. He sent PB words through a low-pass filter, with a cut-off at 500 Hz, and a 17 dB/octave attenuation. These words were presented on one channel 45 dB above each subject's pure-tone average for 500, 1000, and 2000 Hz.[6] The words were presented undistorted on the second

---

[5] The author does not define threshold.

[6] It seems to the writer that at this intensity the filter would produce negligible distortion.

channel at an intensity which would result in a 50% articulation score, the level varying between 5 and 15 dB above the pure-tone average for 500, 1000, and 2000 Hz. The filtered and unfiltered words were then presented alone to the right and left ear, and then binaurally, each channel sent to opposite ears. The case histories reported had lower scores when the distorted words were presented to the ear contralateral to the affected hemisphere.

Palva (1965) stated that "The literature contains little information on the understanding of binaurally presented filtered speech. . . " (p. 30), and added that ". . . no systematic studies of binaural discrimination with bands of varying widths taken from different parts of the speech area have been published." (p. 31) For this reason, before attempting to design a test to diagnose central hearing disorders, he first experimented with all possible band combinations taken from the frequency areas that contain the essential formants: (1) between 400-1000 Hz, and (2) 1200-3000 Hz. After much pre-testing, he chose the desired band widths, and proceeded to filter Finnish PB words, discarding the difficult ones, into two bands: (1) a low-pass band of 480-720 Hz, and (2) a high-pass band of 1800-2400 Hz. He found that a 30 dB/octave drop was not adequate for a "hearing synthesis test," and that a

60 dB/octave attenuation was necessary. At 50 dB SL[7], the discrimination score for normal-hearing subjects was 19.1% for the low-pass band and 17.9% for the high-pass band. The discrimination scores rose to 79% regardless of whether both bands were sent to one ear, or one band was sent to each ear. When almost any two bands were combined, one from the high-frequency and the other from the low-frequency area, the discrimination scores were improved in both the monaural and the binaural condition, in some cases more than twice their arithmetic sum.

In Relation to the Peripheral Hearing Mechanism

Information pertinent to the present investigation was also found in binaural fusion and filtered speech experiments concerned more with the peripheral than the central hearing mechanism.

Huizing (1961) felt that speech audiometry should provide, in addition to the threshold value, an analysis of the patient's hearing function under difficult hearing conditions, and so he designed a distorted speech test as a means of further assessing the function of the peripheral mechanism. Dutch PB words were filtered into two one-octave bands: (1) a low-pass band of 140-280 Hz, and (2)

---

[7] The threshold of detectability for the bands themselves was used "... to obtain a relative 0 dB value." (p. 37)

a high-pass band of 1128-2256 Hz. By inspection of the filter attenuation curves, there seems to be approximately a 40 dB/octave drop. When each of these two bands was presented separately to subjects with normal hearing, the low-pass band was completely unintelligible, whereas the high-pass band yielded a 50% articulation score at 34 dB.[8] In order to imitate a sloping audiogram, the low-pass band was presented to one ear at a high intensity of 75 dB. When the high-pass band was added to the same ear, a 50% articulation score was reached at 26 dB, but when the high-pass band was added to the opposite ear, only 12 dB was required to achieve a 50% score, so that it appeared there is less masking in the binaural as compared to the monaural condition. Huizing felt that this information could be applied to the fitting of binaural hearing aids which he felt might "... depend on the development of an interaural discriminative cooperation." (p. 151)

Huizing (1960, 1962) also reported a method of "triplet audiometry" whereby one can assess a patient's discriminative ability in particular pitch ranges. He filtered speech into three bands, each of which contributed equally to intelligibility: zona gravis which is up to 1000 Hz, zona media which covers the range

---

[8] The author does not specify a reference intensity.

from 1000-2000 Hz, and zona acuta which is 2000 Hz and up. By determining the discrimination score for each band one can then recommend selective amplification in order to restore the patient's "tonal equilibrium." Huizing added that qualitative speech audiometry of this nature has been applied for many years in the Groningen Clinic in the Netherlands.

Kryter (1960) reported that the frequency band 250-750 Hz, in addition to adding greatly to intelligibility, also gives speech a "natural" sounding quality. He filtered PB words into band widths of 500 Hz with a 70 dB/octave attenuation, and presented them to the subject at a comfortable listening level to determine which band widths would combine to yield the highest discrimination scores. He selected three bands from the richest formant regions, with center frequencies of 500, 1500, and 2500 Hz. Articulation scores obtained with a combination of these three bands closely approximated scores obtained with the unfiltered speech spectrum.

Tillman and Carhart (1965) filtered monosyllabic CNC[9] words into two bands of equal intelligibility: one a high-pass band with a cut-off at 1830 Hz and a 36 dB/octave attenuation, and the other a low-pass band with a cut-off at 1480 Hz and a 54 dB/octave

---

[9] Both the CNC word lists and the CVC word lists to be referred to later are composed of consonant-vowel-consonant words.

attenuation. The filtered words were presented to normal-hearing subjects, and it was found that the binaural and monaural functions for each band was almost the same, with approximately a 4%/dB slope. However, when the high-pass and low-pass bands were sent to opposite ears, the resulting binaural function of 5.7%/dB closely approximated the 6%/dB slope obtained with the monaural presentation of the original unfiltered words.

Speaks (1967) presented subjects with seven-word sentences which had been artificially constructed from word-triplets, that is, groups of three logically connected words strung together. A typical test item was "small-boat-with a-picture-has become," and the subject had to identify correctly such seven-word messages from among ten synthetic sentences. Speaks found it almost impossible to render these sentences unintelligible by filtering, for even when all frequencies above 125 Hz were eliminated, a 100% intelligibility score was attained at 62 dB SPL. The maximum performance for discrimination scores using PB words under similar filtering conditions was 6% at 80 dB SPL. The author felt that he would have obtained similar results even if a steeper filter attenuation than the 24 dB/octave drop had been used. Speaks reported that there is approximately an octave difference between the range necessary for sentence identification and monosyllabic identification. The crossover frequency which divides the frequency scale into two equivalent parts, each contributing equally to intelligibility, was 725 Hz for sentences and 1600-1700 Hz for monosyllables. Speaks stated that the energy in the low frequency range of the speech spectrum appears to be of prime importance for the identification of the synthetic sentences used in the reported experiment. Even in the high-pass filter conditions the cut-off frequency had to be below 300 Hz before there was no further improvement in the scores.

Egan, Carterette, and Thwing (1954) presented two messages simultaneously to the subject: one was the message to be received and the other was the interfering message. They first presented the signals to the same ear, and then to opposite ears. In the dichotic reception, one signal in each ear, the received signal can be understood at a level 32 dB lower than in the monaural reception of both signals. The authors prefer to use the term "confusion" to explain their results, rather than the term "masking" with all of its connotations. An interesting finding was that high-pass filtering of either the received or the interfering message resulted in improved intelligibility of the received message. The authors studied the interference effects of both white noise and speech and found that speech interfered more with signal reception than did noise.

The Low Frequencies as a Masker

The present investigation is concerned with the effect of low frequencies on consonant discrimination scores, the low-frequency band being presented at varying intensities relative to the intensity of the high-frequency band. This section discusses the reports which investigated the effect of low frequencies on intelligibility and the dependence of that effect on the intensity of the lows relative to the intensity of the rest of the speech signal.

Fletcher (1929) found that when complex sounds are amplified, the low frequencies gain in loudness faster than the high frequencies, and so the high tones begin to be masked as the low tones become more prominent. Palva (1965) commented that the high-intensity, low-frequency sounds mask the higher frequencies that aid discrimination. Kryter (1960) noted the possibility of "interband masking," that is, the masking of "speech by speech itself." House, Williams, Hecker, and Kryter (1965) also considered the possibility of interference with speech perception due to the "... intrasyllabic masking of consonants by vowels." (p. 165) Huizing (1961) stated that when the slope of an audiogram exceeds 10 dB/octave, the essential overtones begin to be masked by the stronger low tones, and intelligibility suffers.

Pollack (1948) stated that at low intensity levels frequencies below 350 Hz improved intelligibility and he concluded that the low frequencies contribute by increasing the audibility of the speech signal. However, at higher intensities the greater energy present in the low frequencies masks the weaker, information-bearing high frequencies. Pollack also noted that when a masking noise is added to a speech signal, it is not the absolute level of either the signal or the noise, but rather the signal-to-noise ratio that will determine intelligibility.

Hirsh and Bowman (1953) reported that the masking effect of the 20-160 Hz band was negligible because of the very small amount of speech energy it contains. They found that the 670-1000 Hz band was the most effective masker of speech. Miller (1947) also found that frequencies below 1000 Hz have the greatest disruptive effect on communication. At low noise levels, high-frequency bands of noise are the most effective maskers, but as the intensity is raised, the high frequencies, which contain the unvoiced consonants, are soon masked by the more powerful low frequencies.

French and Steinberg (1947) noted that the decrease in articulation scores at high intensity levels could be caused partly by the speech sounds in one region masking other frequency regions. The two low-frequency bands of 250-375 Hz and 375-505 Hz caused the greatest amount of interference with the higher bands. The authors added that the concept of the articulation index, which is based on the theory that the contribution of each band is independent of other bands, must therefore be modified to account for the possibility of interband masking.

Webster and Klumpp (1963) agreed with the previously published data of Pollack, Dyer, and Webster that the crossover frequency, that frequency which divides the speech signal into two bands of equal intelligibility, lowers as much as an octave as the signal-to-noise differential decreases, from the crossover frequency of 1700 Hz in quiet reported by French and Steinberg, to around 800 Hz in noise.

Factors Contributing to Consonant Discrimination

This section is concerned with factors which contribute to consonant discrimination. This information provided the basis for the selection of the word lists used in the present investigation. There is a wealth of information on the effects upon communication of noise and frequency distortion produced by various high- and low-pass filters. The results are usually expressed in terms of an articulation score reflecting only the number of errors in speech perception. Miller and Nicely (1954) performed an experiment to determine not only how many but also what kind of errors were being made. The confusion matrices used in the present study were designed after those they developed. They were concerned only with consonants which they categorized according to five relatively independent features of voicing, nasality, affrication, duration, and place-of-articulation -- the dimensions used by phoneticians to classify consonants.

The subjects in the Miller and Nicely study listened to nonsense syllables consisting of a consonant followed by the vowel / ɑ/, in signal-to-noise ratios which ranged from +12 to -18 dB.[10] Specific areas of the speech spectrum were selected by introducing various filters with a 24 dB/octave attenuation. The authors decided that low-pass filtering and noise have similar effects on speech perception, since noise penalizes the weak high frequencies more than the stronger low frequencies. They concluded that ". . . low-pass filters affect the several linguistic features differently, leaving the phonemes audible but similar in predictable ways, whereas high-pass filters remove most of the acoustic power in the consonants, leaving them inaudible and, consequently, producing quite random confusions." (p. 350)

---

[10] A positive signal-to-noise (S/N) ratio indicates that the signal is being presented at a higher intensity than the noise, whereas a negative S/N ratio indicates that the noise is at a higher intensity than the signal.

The effects of noise on the speech signal were of particular interest, for while voicing and nasality were only slightly affected at the -12 S/N ratio, place-of-articulation confusions were apparent even at the +6 S/N ratio. The authors also reported the crossover frequencies, which varied considerably for the different linguistic features: 450 Hz for nasality, 500 Hz for voicing, 750 Hz for affrication, 1900 Hz for place-of-articulation, and 2200 Hz for duration. An additional crossover point noted is of particular interest: 1250 Hz for the information measure, which is at least 300 Hz lower than the articulation crossover point.

Steinberg (1929) reported that, particularly for stop and fricative consonants, high-pass filtering with a cut-off at 1000 Hz had almost no effect on their intelligibility. When all frequencies above 1000 Hz were eliminated with a low-pass filter, articulation scores of 50-70% were still obtained for the stop and fricative consonants, a finding which he noted as being inconsistent with traditional expectations. He offered as part of the explanation for this phenomenon that ". . . a given consonant is identified as belonging to a certain group from the nature of the beginnings or endings of the vowel sounds[11], and also from the transmitted portion of the

---

[11] That part of the transmitted sound from the vowel to the consonant and from the consonant to the vowel is what will be referred to later as a "transition."

sound itself." (p. 129)

According to Fletcher's (1929) articulation curves for individual sounds, when all frequencies below 500 Hz are eliminated, discrimination scores of close to 100% are still obtained for all consonants except /f/ and /ð/. The fact that discrimination is possible without the low frequencies does not necessarily imply that there is no information below 500 Hz, since Fletcher himself stated:

> A consonant sound may sometimes be identified by the modifications produced on the following or preceding vowel even though it is below the threshold as determined by an isolated sound. It might seem logical to consider this modification of the vowel as part of the consonant. If it is so considered, then it is evident that as long as the vowel is heard there is always a chance of identifying the consonant preceding or succeeding it... (p. 277)

Fletcher found that, in general, vowels are more correctly identified than consonants except for /l/, /r/, and /ŋ/ which are among the most easily recognized speech sounds at normal intensities. He added that the sounds /ð/, /f/, and /v/ accounted for more than half of the discrimination errors, regardless of the intensity.

Lehiste and Peterson (1959) reported that spectrographic analysis has demonstrated the inter-effects of sounds that precede and follow each other, thus eliminating the possibility of phonetic balancing of articulation tests such as Egan's (1948) PB word lists. They devised their CNC lists which rely primarily on phonemic balancing, realizing that articulations are similar only when combined in identical linguistic sequences.

Fairbanks (1958) devised an open-response CVC test which he referred to as a test of "phonemic differentiation." The word lists were constructed so that "... transitions are adequately represented and that there are no undue biases." (p. 597) The words were presented at a -2 dB V/N (vowel-to-noise) ratio, and it was found that at the 50% articulation point, in the initial position, the nasals /m/ and /n/ had high and the voiceless stops had low articulation scores. In addition, it was found that the voiced form of each stop had higher discrimination scores than its voiceless form: /b/ > /p/, /d/ > /t/, and /g/ > /k/. Fairbanks concluded that both the acoustical characteristics of the consonants themselves and the vowel transitions account for the differences in consonant recognition.

House, Williams, Hecker, and Kryter (1965) based the design of their CVC closed-response set on the Fairbanks test, so that "... the materials retain a high degree of phonemic similarity from test form to test form and, in addition, contain representatives from the major classes of speech sounds." (p. 159) They found that, in the presence of various S/N ratios, in the initial position, the voiceless form of each stop tended to be discriminated better than its voiced form: /p/ $>$ /b/, /k/ $>$ /g/, and /t/ $>$ /d/, and also that the voiceless stops had considerably higher articulation scores than the nasals. These findings are the direct reverse of Fairbanks' data which conformed to classical results. House *et al.* offered as a partial explanation the fact that the uniform white noise spectrum used by Fairbanks may have masked the high frequencies more than the shaped white noise used in their experiment. The shaped noise decreased 9 dB/octave after 500 Hz, and its masking effects would, therefore, be less for the voiceless consonants.

Spectrographic Analysis of Consonants

Spectrograms of both real and synthetic speech have revealed the distinctive features that separate one phoneme from another. Haskins Laboratories has been a pioneer in the "... systematic exploration of the acoustic cues to the perception of some of the consonant sounds,"[12] and their findings were used in the analysis of the results of the present experiment.

The burst has long been known to contain manner cues for the stop consonants (plosives), but it was not known whether these bursts also contained place information. Cooper, Delattre, Liberman, Borst, and Gerstman (1952) found that, in general, high fre-

---

[12] Cooper *et al.*, 1952, p. 597.

quency bursts were all heard as /t/, regardless of the following vowel. The lower frequency bursts were perceived as /p/, except when they were level with or slightly above the second formant of the vowel, and then the same bursts were heard as /k/. An example of this burst-vowel relationship was the famous /pi, kɑ, pu/ experiment, in which a burst centered at 1400 Hz was perceived as /p/ before /i/ and /u/, and as /k/ when placed before /ɑ/.

> It is clear that for /p/ and /k/ the identification of the consonant depended, not solely on the frequency position of the burst of noise, but rather on this position in relation to the vowel. In other words, the perception of these stimuli, and also, perhaps, the perception of their spoken counterparts, requires the consonant-vowel combination as a minimal acoustic unit. (p. 597)

Delattre, Liberman, and Cooper (1955), using synthetic speech, investigated the first two formants of stops and their respective transitions. F_1 was determined by the manner of consonant production and began at the base line representing closure of the vocal tract. F_2 was determined by the place of articulation, and it was found that the second formant of each consonant seemed to begin at the same place on the spectrogram. The term "locus" was given to this apparent starting point. It was found that the plosives which were formed in the same part of the articulatory mechanism also shared the same "locus": 720 Hz for the labials /b/, /p/, and /m/; 1800 Hz for the alveolars /d/, /t/, and /n/; and 3000 Hz for the velars /g/, /k/, and /ŋ/ when preceding a front or mid vowel, and 1200 Hz when preceding a back vowel.

Liberman, Delattre, and Cooper (1958), using synthetic speech again, sought to determine the acoustic cues that separate the voiced form of each stop from its respective voiceless form, that is, what distinguished the /b/ from the /p/, the /d/ from the /t/, and the /g/ from the /k/. The researchers were careful to make the intensity of the plosive burst neutral so as to eliminate all voicing information. They did not believe that the presence or absence of the voice bar, the low-frequency region around 100-200 Hz representing vibration of the vocal cords, was alone responsible for the voiced-voiceless opposition. When the voice bar was removed the labial stops were perceived as voiceless, but the velar and alveolar stops were still heard as voiced. The purpose of the reported experiment was to investigate "... the effects of progressively eliminating the transition of the first formant" (p. 155) in order to determine the amount of voicing information present in F_1. This was accomplished by cutting back the first formant from the base line in steps ranging from 10-50 msec. It was found that in order for the sound to be perceived as voiceless, the labials required only a negligible amount of cutback, the alveolars required the greatest amount, and the velars slightly less degree of cutback.

Malécot (1956), using real speech, investigated the role of the nasal resonances in the perception of /m/, /n/, and /ŋ/. It was found that the resonances primarily provided manner cues, with the /m/ resonance giving some place information, the /n/ resonance only a little, and /ŋ/ resonance none at all. It was further noted that all nasal resonances tended to be identified as /m/. The important cues for distinguishing the nasals from each other were found to lie in the transitions, and the conclusion was reached that the nasals should be classified (as was done by Delattre *et al.*, 1955) "... as stops rather than as continuants. Although they can be pronounced in isolation without the vowels, they generally cannot be separated from them if, perceptually, they are to preserve their phonemic individuality." (p. 281)

Lisker (1957), using synthetic speech, performed one of the few experiments designed to determine the identifying features of the phones /w/, /r/, /l/, and /j/, often referred to as "Whirlys." These sounds were examined in the intervocalic position, and the steady state frequency components were labeled R_1, R_2, and R_3, which together with the transitions, provide most of the information necessary for the discrimination of these four sounds. It was found that R_1 was at or near the base line for all four sounds, depending upon the degree of closure of the vocal tract. The /w/ and /j/ could be separated on the basis of R_2, which was 600 Hz and 2600 Hz, respectively. The /r/ and /l/ shared the same R_2 of 1200 Hz; it was the very low R_3 of the /r/, around 1300 Hz, which separated the /r/ from the /l/. None of the "Whirlys" adhere too well to the "locus" theory, the /l/ least of all, for all of its three formant regions vary considerably according to the vowel environment.

Harris (1958), using real speech, investigated the identifying features which separate the fricatives. It was found that the noise or friction component of the sound was sufficient to identify the /s/ and /ʃ/, whereas the transition was the important cue for the /f/ and /θ/. The author commented that the listener seemed to sort the fricatives into two categories depending upon whether the dominant cue was the friction or the transition, and then used frequency information to distinguish between the sounds within each category. The friction cut-off is approximately 3500 Hz for the /s/ and 2400 Hz for the /ʃ/, and the "locus" for the second formant is about 800 Hz for the /f/ and 1800 Hz for the /θ/.

SUMMARY

The purpose of the present study was to investigate the effect on consonant discrimination when a low-frequency band is added at various intensities, either to the ear receiving a constant high-frequency band, or to the opposite ear. The reports reviewed in the first section of the review of the literature dealing with binaural fusion and filtered speech tests provide a quantitative analysis of the information present in various band widths of the speech spectrum, as well as a comparison of discrimination scores when various high- and low-pass band combinations were sent to the same ear, or split, one to each ear. This information contributed to the design of the present experiment by providing the basis for the selection of the particular high-frequency band used, in addition to providing guidelines for the procedure to be followed when adding the low-frequency band to the ear receiving the high-frequency band or to the opposite ear.

The reports reviewed in the second section describe the effects of low frequencies, at various intensities, on the speech signal. The information in this section enabled the writer to select a low-frequency band that would satisfy the following conditions: (1) theoretically contains negligible consonant information, (2) contributes to the speech signal when added at favorable HB/LB ratios, and (3) detracts from the speech signal at unfavorable HB/LB ratios.

The third section of the literature review presents reports dealing with factors contributing to consonant discrimination and provided the basis for the selection of the materials used in the present study, as well as contributing to the design of the confusion matrices by means of which the data of this experiment are presented. The material presented in the final section of the review indicates how spectrographic analysis of both real and synthetic speech has revealed the distinctive features that separate one phoneme from another. This material provided the background information used by the writer in the analysis of the experimental data.

CHAPTER II

PROCEDURES

Problem

The purpose of this experiment was to investigate the effect on consonant discrimination when a low-frequency band is added at various intensities and under two modes to a constant high-frequency band. More specifically, the purpose was to compare the consonant discrimination scores obtained with a high-frequency band alone of 1020-2040 Hz with those obtained when that band is accompanied by a low-frequency band of 240-480 Hz at three intensity levels (a) in the same ear, and (b) in the opposite ear.

Materials

For this experiment, the Fairbanks Rhyme test[1] was chosen because it attempts to have all transitions adequately represented, and so that qualitative as well as quantitative information on consonant

---

[1] The Fairbanks Rhyme test is concerned only with consonants in the initial position. The word lists used in this study can be found in Table 10 in the Appendix.

discrimination could be obtained. An analysis of errors could then be made, and confusion matrices be plotted along the dimensions of voicing, manner, and place-of-articulation.

Recording Technique

The words were spoken with equal effort[2] at approximately five-second intervals by a trained male into two Electrovoice 654A microphones, equidistant from the mouth, words peaking from -3 to -7, averaging -5 on the VU meter. The material was recorded[3] on two channels on Scotch recording tape, using an Ampex AG500 tape recorder. The tapes were then played through two Allison 2B filters in series, providing 60-70 dB/octave attenuation. Channel one was set for 1020-2040 Hz, chosen because of its high information content, and channel two for 240-480 Hz, chosen because of its high energy and low information content when used in isolation.[4]

---

[2] Fairbanks (1958) recorded the Rhyme test with average vocal effort.

Egan (1948) stated "... no attempt should be made to compensate for the typical differences in the speech power used in pronouncing the different sounds in the test items." (p. 977)

[3] The writer would like to thank Dr. Tom Gay for his help in preparing the taped material.

[4] Palva (1965) found the maximum discrimination score of 54% for filtered words occurred in the octave band 1080-2160 Hz, and the minimum discrimination score of 1% occurred in the octave band 240-480 Hz, at 50 dB SL.

Calibration

Two calibration tones were put on the tape, each matched in intensity to the SPL of the average peak values of the filtered words. For the 1020-2040 Hz band, a 1500-Hz pure tone was used, and for the 240-480 Hz band, a 360-Hz pure tone was used. The tones lasted for 30 sec. and were used to set the VU meter on the audiometer to zero. The tones were then pulsed for one minute, with a rise-fall time of 25 msec. and a duration of 330 msec. The 1500 Hz pulsed tones were 7 dB lower, and the 360 pulsed tones were 3 dB lower than their respective steady calibration tones. These pulsed tones were used to establish pure-tone thresholds for both ears for these center-band frequencies. The audiometer dial readings for each subject at his threshold for the 360 and 1500-Hz pulsed tones were converted to 0 dB SL and then used as the reference intensity for the presentation of the corresponding bands.

Testing Equipment

The tape was played on a Sony stereophonic tape deck,[5] Model 250, through the phonograph circuit of a Maico two-channel audiometer, Model MA-8, into TDH-39 earphones in MX 41 AR

---

[5] The writer would like to thank Dr. Larry Deutsch for his help in adjusting the testing equipment.

cushions. The Rudmose Sound Analyzer, set at 80 dB in slow meter position, indicated that the SPL of the calibration tones at the earphone was higher (by 15 dB for the 1500 and 17 dB for the 360-Hz tone) than the audiometer dial reading.[6]

Subjects and Test Presentation

The subjects were thirty-six normal-hearing college students,[7] ten male and twenty-six female, ages seventeen to twenty-four. After thresholds for the 360 and 1500-Hz pulsed tones were established, the subjects were given the following instructions:

> You are about to hear words which have been distorted. They are all real words of one syllable. You are to fill in the blanks with a single letter. Write the sound

---

[6] Subjects 1-16 were tested using the original tape deck. A second tape deck, identical to the original, was used to test subjects 17-36. Calibration of the second Sony tape deck showed that the SPL of the calibration tones at the earphone was higher (by 14 dB for the 1500 and 18 dB for the 360-Hz tone) than the audiometer dial reading. Converting from SPL to hearing threshold level, when the audiometer dial was set at zero, the pulsed tones at the earphone were 7-8 dB and 14-15 dB for the 1500 and 360-Hz pulsed tones, respectively. These values closely approximate the new 1964 ISO reference threshold values.

[7] Two subjects were tested whose results were not included in this study because they were unable to adjust to the filtered speech situation. Details as to age, sex, and hearing thresholds for the other 36 subjects can be found in Table 11 in the Appendix.

> you think you hear. If there are some[8] words which you do not recognize at all, you may leave a blank space. Start at the top of Column A and fill in the spaces down the column. When you finish the last word in Column A, repeat the procedure for Columns B, C, D, and E. This first test is a trial, and then there will be nine more tests.

Each subject received a trial on List III, consisting of five columns (A-E) of ten words (_ot, _ay, _op, etc.), in order to adjust to listening to filtered speech. Only the high-frequency band was presented to one ear, Column A at 30 dB above the threshold for the 1500-Hz pulsed tone, Column B at 20 dB above, Column C at 10 dB above, and Columns D and E at threshold. Lists I, II, and IV were then presented to each subject in randomized sequence and repeated twice for a total of nine list presentations:

a. High-frequency band alone: Lists I, II, IV.

b. High-frequency and low-frequency bands to the same ear: Lists I, II, IV.

c. High-frequency and low-frequency bands to opposite ears: Lists I, II, IV.

Since the purpose of this experiment was to explore the interaction of the bands, meaningful results could be obtained only

---

[8] House (1965) stated "The instructions emphasized the importance of responding to every stimulus even if guessing was necessary." (p. 160)

if the same list was used for each set of comparisons.[9] The high-frequency band for all three lists was presented to each subject at the same intensity as his threshold for the 1500-Hz pulsed tone, whereas the low-frequency band was presented at a different intensity for each list:

a. the threshold for the 360-Hz pulsed tone (0 dB SL)

b. the threshold plus 20 dB (20 dB SL)

c. the threshold plus 40 dB (40 dB SL)

The testing procedure involved three variables:

1. Lists

    List I - A_1

    List II - A_2

    List IV - A_3

---

[9] The advantage to using filtered speech is that "... the test can be repeated immediately and frequently, if required, using the same word material without any deterioration in the results." (Palva, 1965, p. 75)

"Not only the same list of words, but the identical recording was used for each set of comparable conditions on the two ears and for the two combined conditions. This is believed to be a critical point, since the interlist equivalence previously demonstrated for the PB-50 words need not, and, in our opinion, does not, hold under conditions of gross distortion involved in the above-described procedures. The possible learning effect involved in the use of the same list on both ears was felt to be slight in comparison with the problem of interlist equivalence." (Jerger, 1960, p. 799)

2. LB intensities

   0 dB SL - B_1

   20 dB SL - B_2

   40 dB SL - B_3

3. Earedness

   Right ear - C_1

   Left ear - C_2

A summary of the test presentations for each subject can be found in Table 12 in the Appendix.[10] The following describes how the variables were combined so as to eliminate an order effect:

C_1

#1.	$A_1A_2A_3$		
#2.	$A_1A_3A_2$		$B_1B_2B_3$
#3.	$A_2A_1A_3$	X	$B_2B_3B_1$
#4.	$A_2A_3A_1$		$B_3B_1B_2$
#5.	$A_3A_1A_2$		
#6.	$A_3A_2A_1$		

C_2

#1.	$A_1A_2A_3$		
#2.	$A_1A_3A_2$		$B_1B_2B_3$
#3.	$A_2A_1A_3$	X	$B_2B_3B_1$
#4.	$A_2A_3A_1$		$B_3B_1B_2$
#5.	$A_3A_1A_2$		
#6.	$A_3A_2A_1$		

---

[10] e. g. Subject #1 received the HB alone in the R ear for Lists I, II, and IV in that order, the LB was added to the HB at 0, 20, and 40 dB SL for Lists I, II, and IV in that order to the opposite ear, and the LB was then added to the HB at 0, 20, and 40 dB SL for Lists I, II, and IV in that order to the same ear.

#1, 3, and 5: the LB was added to the opposite ear first, the same ear second.

#2, 4, and 6: the LB was added to the same ear first, the opposite ear second.

CHAPTER III

RESULTS AND DISCUSSION

Summary of Articulation Scores

The high-frequency band was presented at threshold throughout the experiment, "threshold" referring to the acuity in SL for a pulsed tone (1500 Hz) at the approximate center of the filtered band. In order to study the effect of adding the low-frequency band to the high-frequency band, it was desirable to minimize the information content of the HB by presenting it at a low sensation level. Pre-testing indicated that when the HB was presented at threshold the articulation scores for normal-hearing subjects averaged 40%. Even at 5 dB SL, the articulation scores began to increase sharply, so 0 dB SL was selected for the presentation of the high-frequency band to each subject for all seven conditions.

The low-frequency band was presented at three intensities: threshold (0 dB), 20 dB, and 40 dB SL, "threshold" referring to the acuity in SL for a pulsed tone (360 Hz) at the center of the filtered band. Pre-testing indicated that the LB contained a negligible amount of consonant information when presented alone, at each of the three sensation levels. Pre-testing also indicated that when the LB was added below threshold to the HB, there was no change in the HB-alone score, but when added at threshold, a slight increase over the HB-alone score was noted. When the LB was added at 20 dB SL, the equivalent of PB-max[1] resulted, and when added at 40 dB SL, the articulation curve began to drop. When the LB was added at 60 dB SL, the pre-test articulation scores showed a further decrease, but the writer did not test at this intensity in the present investigation because of the insufficient interaural attenuation for the low-frequency band at this sensation level.[2] When the LB was added, either to the opposite or the same ear at 80 dB SL, the LB caused discomfort as well as further deterioration of the pre-test articulation scores.

Figure 1 presents the mean articulation scores for each condition: (1) the solid line represents the mean scores for all

---

[1] The point at which the articulation score for PB words is at its maximum, followed either by a plateau or a decrease in the articulation curve, is referred to as PB-max.

[2] The writer considered using insert phones which would increase the interaural attenuation and allow the LB to be presented at higher sensation levels. However, since the lack of research using insert phones would have introduced a further variable in the present investigation, the writer chose to use standard earphones and limit the presentation of the LB to 40 dB SL.

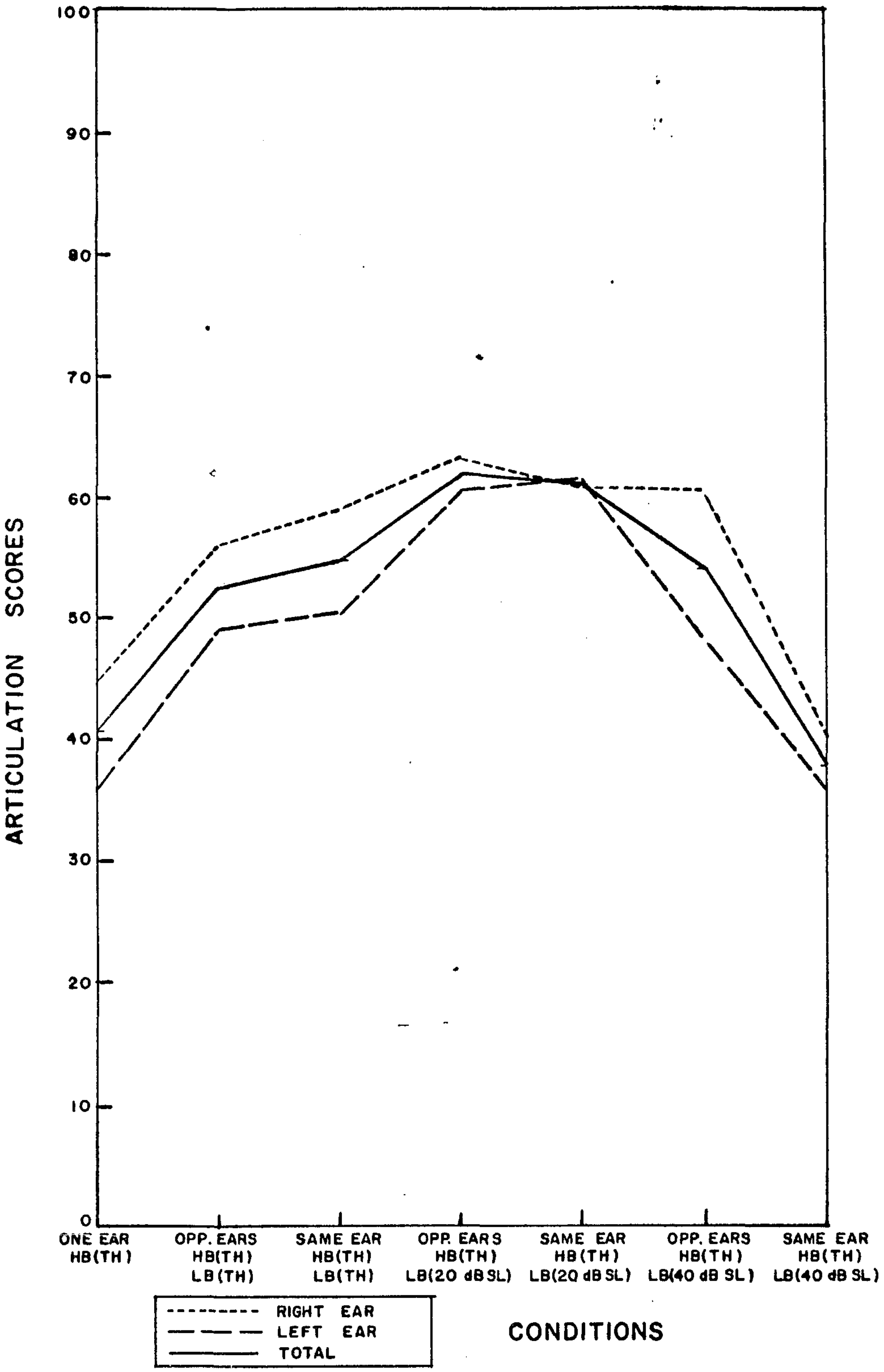

Figure 1. Mean articulation scores for each condition.

thirty-six subjects,[3] (2) the dotted line the mean scores for the eighteen subjects who received the HB in the right ear, and (3) the broken line the mean scores for the other eighteen subjects who received the HB in the left ear.

In the review of the literature, the writer included the work of Kimura, Milner, and Shankweiler which showed a "right-ear" advantage for verbal material. These researchers found the "right-ear" superiority most evident for difficult verbal material, a condition met in the present investigation wherein the speech signal was degraded by filtering. A comparison of the scores for the right and left ear presented in Figure 1 shows that in six out of seven conditions the scores were higher when the HB is presented to the right ear. The difference was: 9% points when the HB alone is presented, 7 and 9% points when the LB is added at threshold to the opposite and the same ear, respectively, 3% points when the LB is added at 20 dB SL to the opposite ear, and 13 and 4% points when the LB is added at 40 dB SL to the opposite and the same ear, respectively. The lack of earedness when the LB is added at 20 dB SL, either to the same or opposite ear, could possibly be a result of the relative ease of these two listening conditions which produced the equivalent of PB-max scores

---

[3] The scores for each subject can be found in Tables 13 and 14 in the Appendix.

in the present experiment.

The solid line in Figure 1 represents the combined results for the eighteen subjects receiving the HB in the right ear and the eighteen subjects receiving the HB in the left ear. There is a negligible difference between the scores (2% points) when the LB is added at threshold to the same and to the opposite ear, but a gain of 15 and 13% points, respectively, over the HB-alone score. Adding the LB at 20 dB SL resulted in the maximum increase in articulation scores - 21 and 22% points, respectively, for the same and opposite ear over the score obtained for the HB alone - and again there is almost no difference between the scores (1% point). When the LB is added at 40 dB SL to the same ear, there is a decrease of 2% points, but when the LB is added to the opposite ear there is an increase of 14% points over the HB-alone score, making a net difference of 16% points between the two scores.

When the LB is added at 40 dB SL to the same ear receiving the HB at 0 dB SL, the decrease in the articulation score from the maximum articulation score (PB-max equivalent) can be explained on the basis of the low frequencies "masking" the high frequencies. The mean decrease of 6-7% points from the maximum articulation score which occurred when the LB is added at 40 dB SL to the opposite ear, can possibly be due to the phenomenon of "central masking," described by Zwislocki (1953) as "... physiological interaction between the ears." (p. 752)

Confusion Matrices

The first column in Table 1 shows the consonants that were used in this study. The second column lists the frequency of occurrence of each consonant in condition #1 where the HB alone is presented, and the third column lists their frequency of occurrence in each of the other conditions, #2-7, where the LB is added to the HB. Tables 2-8 present confusion matrices, one for each of the seven conditions, expressed as rounded percentages, representing the 15,660 responses of the thirty-six subjects tested. Confusion matrix #1 is based on the 5,220 responses from the thirty-six subjects for the first condition, and confusion matrices #2-7 are each based on the 1,740 responses from the thirty-six subjects for each of the other conditions. The consonants at the top of each table represent the stimuli, and the same consonants listed on the left represent the responses.

The row at the bottom of each confusion matrix, labeled "BL." (blank), represents the frequency at which no responses was obtained for each consonant. It is interesting to note the sum of the percentages of no responses that occurred in each condition: beginning with 141 when the HB alone was presented, dropping to 71 and 59when

Table 1. Number of times each consonant[*] was presented in each condition.

CONSONANT	COND. #1 ONE EAR HB(TH)	CONDS. #2-6 SAME EAR OPP. EARS HB+LB
/p/	360	120
/t/	504	168
/k/	324	108
/f/	324	108
/s/	612	204
/b/	468	156
/d/	288	96
/g/	180	60
/m/	432	144
/n/	216	72
/w/	360	120
/r/	360	120
/l/	432	144
/h/	360	120
TOTAL	5,220	1,740

[*] /v/, /z/, /dʒ/, and /j/ were eliminated from analysis in this study since the writer felt that these four consonants did not appear often enough in the Fairbanks Rhyme test for their results to be meaningful.

Table 2. Confusion matrix, conditon #1: one ear, HB(TH).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	29%	7%	7%	3%	3%	2%	3%	4%	3%	2%			2%	4%
t	6	41	15	5	5	2	8	5	2	2	1	2		11
k	5	11	48	2	3	2	2	13	1		1	1	1	4
f	6	3	3	35	8	8			12	3		1	3	9
s	1	2	2	4	29	3	7	3	1	6		1	1	3
b	1	1		12	6	38	3		9	2	1	3	3	3
d	2	2		3	3	3	25	16		6		1	1	
g	1	2	2		2		3	38		2		1		
m	3	2	2	9	7	16	2	2	47	9		5	6	6
n	1	1	2	4	8	1	2	5	1	32		1		1
w	1	1	2	3	3	3	1		2	1	85	14	21	2
r	1	1	1	1	1	3	1	2	1	2	6	52	14	2
l	4	2	1	4	3	4	1		9		1	7	36	2
h	26	10	3	5	2	4	18	3	3	7	2	2	1	35
BL.	12	12	10	7	14	8	15	8	5	21	2	6	7	14
OTH.	1	2	1	2	3	2	7		3	3	1	3	3	3

Table 3. Confusion matrix, condition #2: opposite ears, HB(TH), LB(TH).

Stimuli

		p	t	k	f	s	b	d	g	m	n	w	r	l	h
	p	47%	8%	3%	1%	3%	3%	1%		6%	3%				4%
	t	8	51	18	6	4	1	8	3	1	1		2		8
	k	8	15	61	4	2	3	1	3	1			1		6
R	f	9	3	3	41	15	11	4	2	8		1		1	15
e	s	1	2	3	10	45	4	4		1	4		2		7
s	b				11	5	42	2		6		2	2	3	3
p	d		2		3	1	4	41	10	1	4		1		2
o	g		1	2	1	3		4	72		1		2		
n	m		1		7	2	14			57	14		1	7	2
s	n				1	3	2	2	5	1	47		1		
e	w				2	1	2			1	3	93	8	17	
s	r			1		1	1	1		2	1	3	65	11	1
	l	3		1	2	1	2	2		6		1	12	53	
	h	19	10	4	3	4	4	13	2	7	10	1	1	1	43
	BL.	5	7	4	6	6	3	11	3	1	10	1	3	3	8
	OTH.	1	1	2	2	2	3	5		1	1			4	2

Table 4. Confusion matrix, condition #3: same ear, HB(TH), LB(TH).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	48%	7%	2%	3%	3%	1%		2%	3%	6%		2%		3%
t	7	52	17	3	9	3	8		1	4				13
k	5	14	64	4	1			2				1		4
f	3	2	4	41	11	10	1		10					8
s	1	4	1	8	41	3	4		1	3		1	1	5
b	1	1		11	8	46	1	3	6			2	3	3
d		2			1	3	51	15	1	6				1
g	3	2	2	2	2		8	70		1		3		
m	2			6	2	18		2	58	21		2	7	3
n				4	1	1		2	2	46				
w				1	1				1		93	3	19	
r			2			1	3			1	5	74	12	
l	1			1	1	2	1		6			7	52	
h	23	9	4	10	8	3	8		8	6		1		51
BL.	5	5	4	4	6	4	8	3	3	6		1	2	8
OTH.	2	2		1	3	6	5	2		1	2	3	4	1

Table 5. Confusion matrix, condition #4: opposite ears, HB(TH), LB(20 dB SL).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	64%	11%	6%		3%	3%	1%	5%						6%
t	9	71	13	6	6	3	5	1	1			1		21
k	10	13	70	1	4	1						1		11
f	4		1	52	16	13	4							10
s	1			22	52	6	6	10				4		2
b				6	6	49	1	1	1		3	1	1	3
d	1			3		1	52	23	1					1
g			4		2	1	5	50				3		
m						3			85	22	1	4	3	1
n								1	6	69				
w						1	1				91	10	14	
r			1	1							1	53	3	
l									4		3	15	74	
h	9	2	4	2	2	7	7				1	1		38
BL.	1		1	4	3	7	6	5	1	6		2		5
OTH.	1			3	2	4	10	1	1	3	2	5	6	3

Table 6. Confusion matrix, condition #5: same ear, HB(TH), LB(20 dB SL).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	64%	12%	10%	6%	4%	4%						1%		10%
t	8	68	18	6	6	2	2							13
k	8	10	64	4	2		1					1		6
f	5	1	1	50	17	18	3		1	1	1			13
s	1	1		17	55	6	10	1				5		3
b	1	1		8	3	41		8			1	3	2	
d	1	1		3			49	22			1			1
g			1		1		2	63				3		
m	1			1		6			84	18	1	3	3	2
n		1			1			3	3	71		1		
w		1			1	1	1		1		83	8	9	
r				1		1	1			1	5	50	4	1
l	1				1	1			6		4	18	76	
h	4	4	4	3	3	10	7		1	1				47
BL.	3		2	2	3	8	10	1	3	4	3	8		3
OTH.	3		1	2	3	3	13		1	3	2	1	4	

Table 7. Confusion matrix, condition #6: opposite ears, HB(TH), LB(40 dB SL).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	63%	11%	4%		3%	3%	5%	5%	1%		1%	1%	1%	7%
t	11	65	15	6	6	3	3	1			3			15
k	8	14	65	2	1									7
f	3	1	3	56	18	21	4	5	1			1	1	12
s		1	2	14	48	10	5	8		1		2	1	7
b				4	4	41	8	7			3	1	2	3
d				1		1	28	17	1		1			
g	1			1	1	1	5	38				4		
m				1		3			77	17	2	3	2	
n								7	9	69		1	1	
w			1			1	1			3	79	32	21	
r			1	1	1	1	4	1			2	33	3	
l	1			1	1	1	2		10	3	4	15	60	
h	7	3	4	8	3	4	8	5						43
BL.	7	3	6	2	7	10	10	3		5	4	7	3	6
OTH.		1	1	2	3	2	15	1	1	1	2	2	3	2

Table 8. Confusion matrix, condition #7: same ear, HB(TH), LB(40 dB SL).

Stimuli

Responses

	p	t	k	f	s	b	d	g	m	n	w	r	l	h
p	49%	14%	12%	1%	6%	4%	2%	3%						9%
t	8	45	19	5	6	3	1	1			2			12
k	13	18	49	1	4	1	1							13
f	3	4	5	44	19	17	5	3			3	3	3	16
s	2		2	13	31	6	4	10		1	3	2	1	4
b	1	1	1	6	7	24	13	22	1	1	11	2	6	1
d	1			3	1	4	21	10	1		5		1	
g	1			2	1		3	23			3	4		
m				1		4		1	64	26	3	3	3	1
n	1			1		1	1	8	12	50	2	5	1	
w				2		1	3		2	6	26	16	17	
r				4	1	3	7	1		6	19	26	3	
l	2	1		3	1	3	3		15	3	12	29	47	
h	8	6	5	6	6	8	7	1						33
BL.	12	10	6	9	13	19	11	13	6	7	8	8	13	9
OTH.	2	1	3		2	2	17				4	2	3	2

the LB was added at threshold to the opposite and the same ear, respectively, dropping further to 41 and 50 when the LB was added at 20 dB SL to the opposite and the same ear, respectively, rising to 73 when the LB was added at 40 dB SL to the opposite ear, and rising sharply to 144 when the LB was added at 40 dB SL to the same ear. It can be seen that the effect on the percentages of no responses in each condition when the LB is added to the HB follows the same pattern as the effect of the LB on the articulation scores shown in Figure 1.

Miller and Nicely (1954) used the terms "audibility" and "confusibility" to explain the problems of high- and low-pass filtering, respectively. The decrease in the number of "no responses" when the LB was added to the HB would seem to indicate that audibility was the problem in the HB-alone condition. Only during the HB-alone presentations did any of the subjects motion to have the volume increased. However, when the LB was added at 40 dB SL to the same ear receiving the HB, the number of "no responses" returned to the original HB-alone level, and the writer suggests that confusibility was the problem in the final experimental condition. When the LB was added at 40 dB SL to the same ear receiving the HB, the subjects seemed to be particularly uncomfortable. When the testing procedure was completed, several subjects commented that certain tests were "hard to hear," and that one test in particular was "hard to understand."

The last row in the confusion matrix, labeled "OTH." (other), represents the responses that were different from the consonants used in this study, including such responses as /v/, /z/, /ð/, /dʒ/, /ʃ/, /tʃ/, /θ/, /j/, and assorted blends. Several subjects seemed to perceive two consonants when the LB was presented at a higher SL than the HB, particularly when the bands were split, one to each ear.

Rank Order of Consonants Correctly Identified

Table 9, which presents the relative standing of the consonants identified correctly in each condition as indicated by the confusion matrices, shows the following general trends:

1. The /k/, /m/, and /t/ appear in the top half of the rankings in all conditions.

2. The /w/ is #1 or #2 until the final condition where the LB is added at 40 dB SL to the same ear, dropping the /w/ to tenth place.

3. The /b/, which appears in the top half of the rankings only when the HB alone is presented, is replaced by the /l/ which remains in the top half for all six conditions where the LB is added.

Table 9. Relative standing of the consonants identified correctly in each condition as indicated by the confusion matrices.

STANDING	1 2 3 4 5 6 7	8 9 10 11 12-13 14
HB(TH)-ONE EAR	w r k m t g - b	l f - h n s - p d
HB(TH) LB(TH) - OPP. EARS	w g r k m l t	p - n s h b d - f
HB(TH) LB(TH) - SAME EAR	w r g k m l - t	h - d p b - n f - s
HB(TH)- LB(20 dB SL) - OPP. EARS	w m l t k n p	r f - s - d g b h
HB(TH) LB(20 dB SL) -SAME EAR	m w l n t p - k	g s f - r d h b
HB(TH) LB(40 dB SL) -OPP. EARS	w m n t - k p l	f s h b g r d
HB(TH) LB(40 dB SL) -SAME EAR	m n p - k l t f	h s w - r b g d

4. The /r/ and /g/ appear in the top half when the HB alone is presented and when the LB is added at threshold. These two phones then move to the lower half and are replaced by the /p/ and the /n/ which remain in the top half for the following four conditions, where the LB is added at both 20 dB and 40 dB SL.

The LB can be compared to noise when added at 20 dB and 40 dB SL, particularly when added to the same ear receiving the HB at threshold, and so the results of the present study may be compared to those reported by Fairbanks (1958) and House, et al. (1965) who measured articulation scores on the Rhyme Test in the presence of noise. In the present study, the standing of the voiceless plosives /p/, /t/, and /k/, the nasals /m/ and /n/, and the liquid /l/ is not affected when the LB is added at supra-threshold intensities. The voiced plosives /b/, /d/, and /g/ become the hardest to identify when the LB is added at 40 dB SL to the same ear.

It is interesting to note that Fairbanks (1958) found the highest scores for the nasals and higher scores for the voiced than the voiceless plosives, whereas House et al. (1965) found the lowest scores for the nasals and higher scores for the voiceless than the voiced plosives. House's suggestion that the shaped white noise used in his experiment would have less effect on voiceless plosives than the uniform white noise spectrum used by Fairbanks, could explain the similarity of the behavior of the plosives in this study and House's study, since the effect of the LB can be likened more to a shaped white noise than to a uniform white noise spectrum. In contrast, however, the behavior of the nasals in the present experiment is similar to their behavior in Fairbank's study. It is important to remember that Fairbanks' reported data were based on a -2 dB V/N ratio, but House's reported data were the average of six different S/N ratios: +4, 0, -4, -8, -12, and -16. House explained "(. . . but examination of the responses in detail shows that the general order of successful responses is maintained from level to level.)" (p. 163) This statement does not seem to be substantiated by the results of the present experiment which indicate that there is a change in the relative standing of "correct responses" as the LB is added at the different intensities.

Correct Responses

Figures 2-7 present the correct responses for each consonant for each condition expressed as rounded percentages. The graphs compare the effects on the HB-alone score when the LB is added to the HB at different sensation levels. The HB(TH) condition is always monaural. In Figures 2-7, the dashed line represents scores obtained when the LB is added to the same ear, and the solid

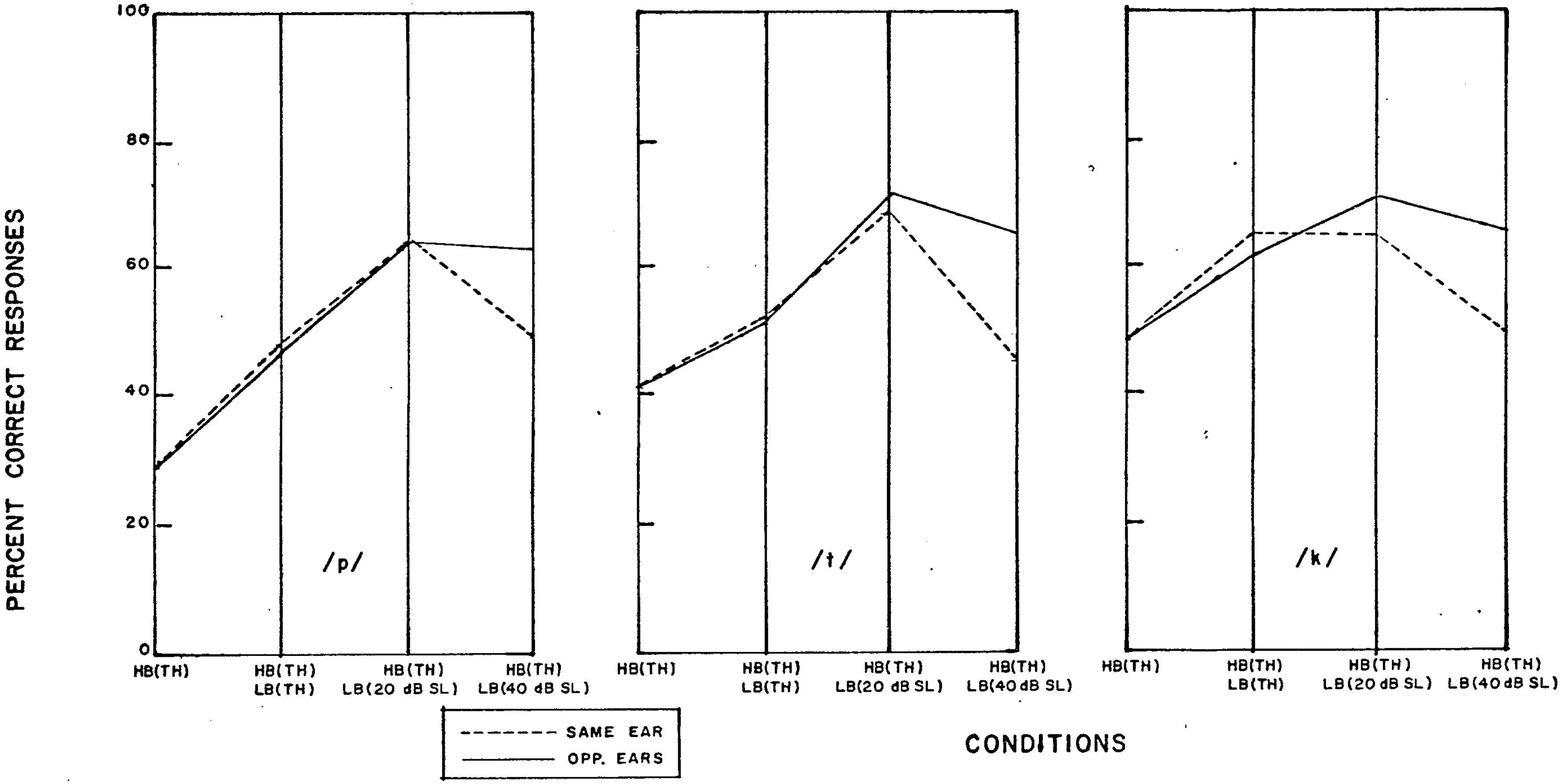

Figure 2. Percentage of correct responses for the stimuli /p/, /t/, and /k/.

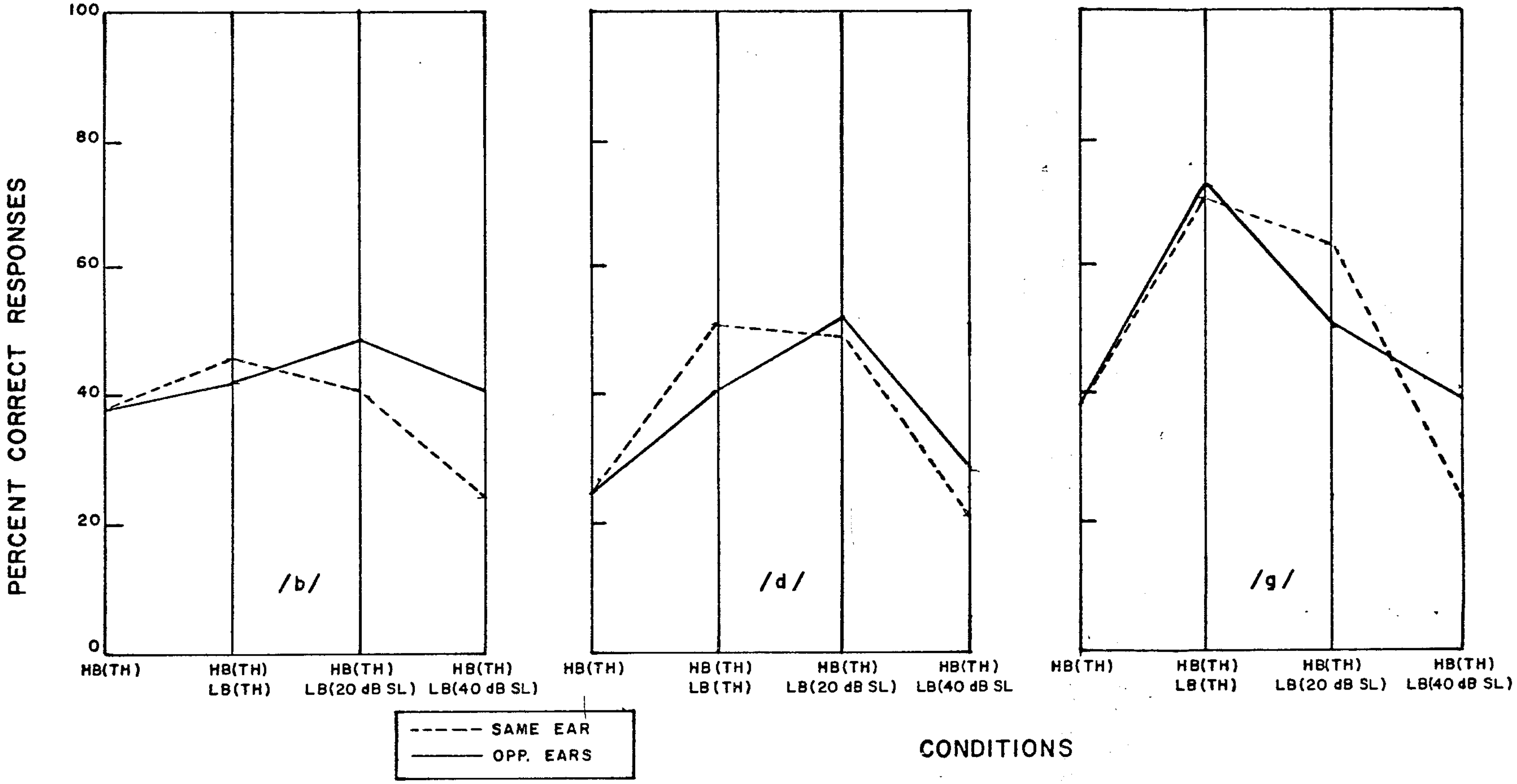

Figure 3. Percentage of correct responses for the stimuli /b/, /d/, and /g/.

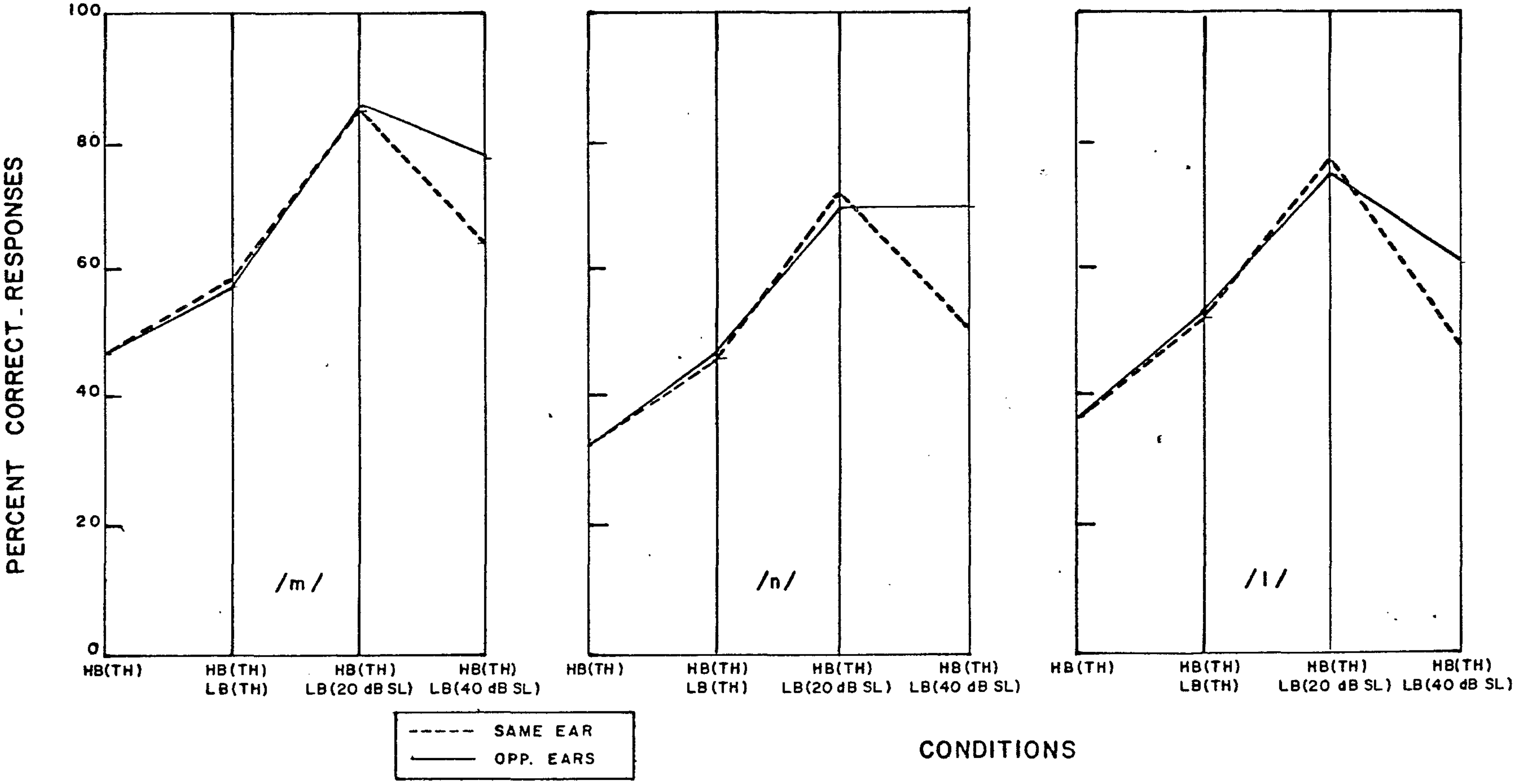

Figure 4. Percentage of correct responses for the stimuli /m/, /n/, and /l/.

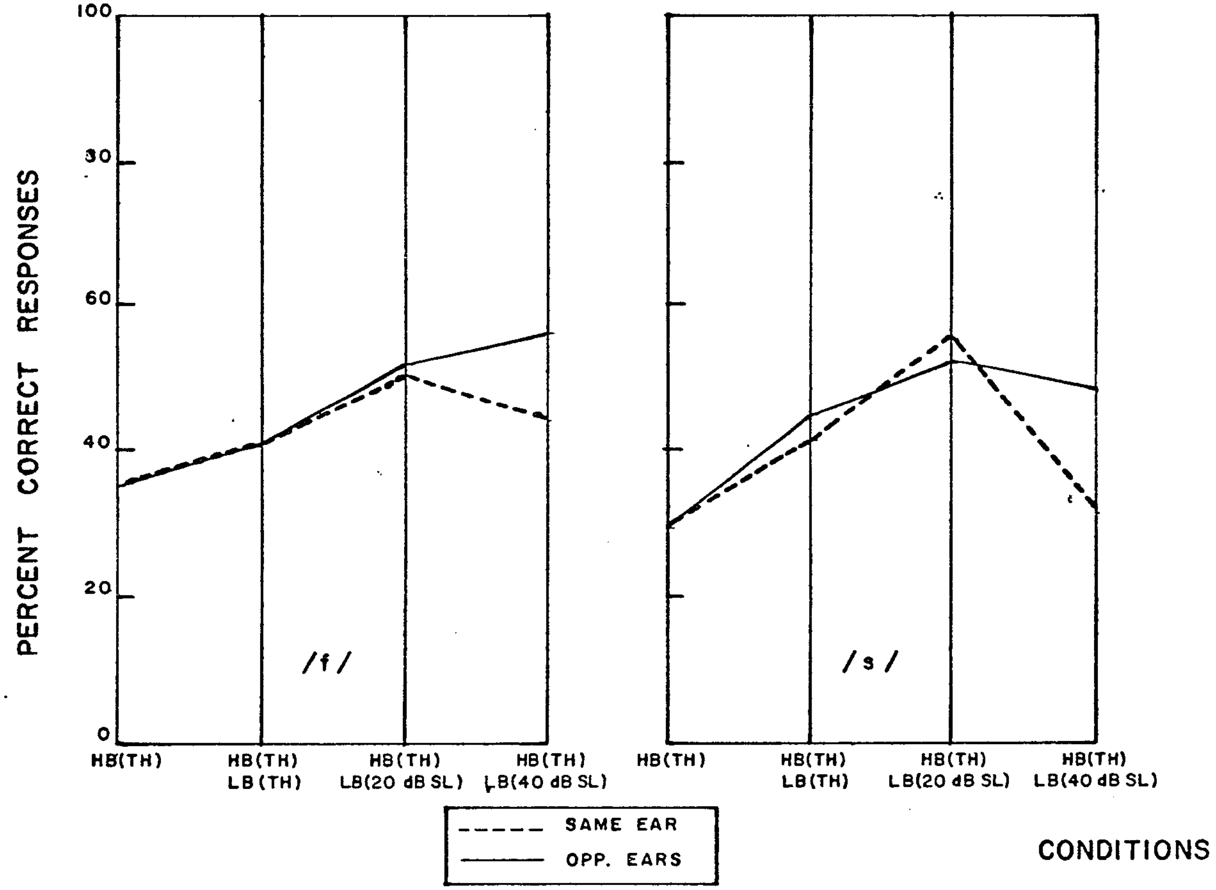

Figure 5. Percentage of correct responses for the stimuli /f/ and /s/.

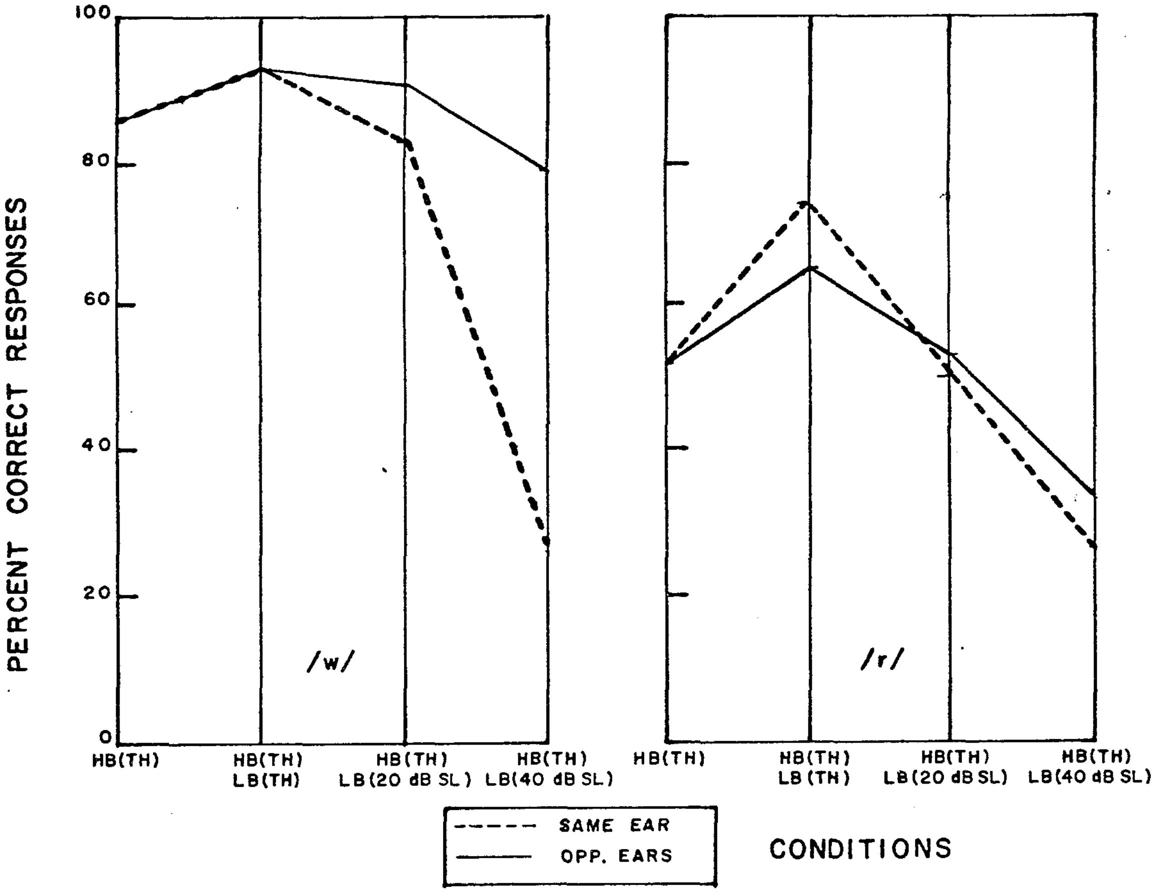

Figure 6. Percentage of correct responses for the stimuli /w/ and /r/.

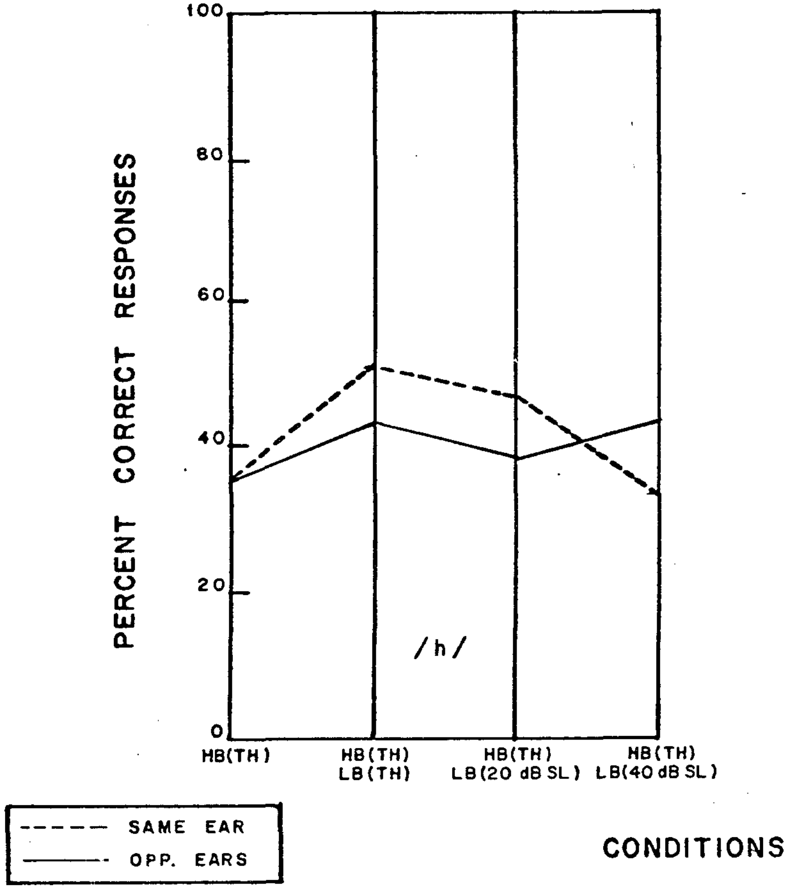

Figure 7. Percentage of correct responses for the stimulus /h/.

line represents scores obtained when the LB is added to the opposite ear. This information was obtained from the confusion matrices in Tables 2-8.

Plosives

Plosives (stops) have multiple cues, since they are identified both by their characteristic bursts and by their formant transitions. A comparison of the correct identification of the voiceless plosives in Figure 2 and the voiced plosives in Figure 3 shows that the LB, when added at 40 dB SL, particularly when added to the same ear, exerts a greater "masking" effect on the voiced plosives. The voiceless plosives have more energy than the voiced plosives in their bursts as well as in their transitions, which may explain why the voiceless plosives are more resistant to "masking" than their voiced counterparts.

Voiceless Plosives - /p/, /t/, and /k/

The voiceless stops shown in Figure 2 have a very similar pattern, showing an improvement in the HB-alone scores when the LB is added at threshold, with the maximum improvement occurring when the LB is added at 20 dB SL. The maximum increase over the score obtained with the HB alone is 35% points for the /p/, 30% points for the /t/, and 22% points for the /k/. There are negligible differences between the scores when the LB is added to the opposite ear and the same ear, both when the LB is added at threshold and at 20 dB SL. There is a difference, however, between the scores obtained when the LB is added at 40 dB SL to the opposite and the same ear. When the LB is added at 40 dB SL to the opposite ear there is only a slight decrease in the maximum scores which occurred at 20 dB SL. However, the scores show a substantial decrease when the LB is added at 40 dB SL to the same ear, with the scores for the /t/ and /k/ returning to their HB-alone level. The /p/ score also decreases but only to the same level where the LB was added at threshold.

Voiced Plosives - /b/, /d/, and /g/

The most noticeable feature of the voiced stops shown in Figure 3 is the great dissimilarity in their patterns, in sharp contrast to the consistency of the voiceless stop patterns. Of all the phones studied, the /b/ shows the least improvement in the score when the LB is added, with the maximum increase of 11% points occurring when the LB is added at 20 dB SL to the opposite ear. When the LB is added at threshold the score is slightly higher when added to the same ear, but when the LB is added at 20 dB SL the score is slightly higher when added to the opposite ear. When the

LB is added at 40 dB SL to the opposite ear the score decreases almost to the HB-alone level, but falls way below this level when added to the same ear. None of the phones previously discussed had a score in any of the six conditions where the LB was added that was lower than the score obtained with the HB alone.

Of all the phones in this study, the /d/ has the lowest score when the HB alone is presented. When the LB is added at threshold to the same ear the score is close to its maximum, remains at this level when the LB is increased to 20 dB SL, but falls slightly below the HB-alone level when the LB is added at 40 dB SL to the same ear. When the LB is added to the opposite ear at threshold the score is 10% points lower than when added to the same ear, representing the largest difference found in any phone in the present study between the scores when the LB is added at threshold to the same and the opposite ear. When the LB is added to the opposite ear at 20 dB SL, the maximum increase of 27% points over the HB-alone score is reached, and when added at 40 dB SL the score falls to its HB-alone level as did the /b/ score under the same conditions.

The /g/ differs from the phones already discussed in that the /g/ reaches its maximum increase of 34% points over the HB-alone score when the LB is presented at threshold. The scores begin to decrease when the LB is added at 20 dB SL, with a more severe drop when added to the opposite ear. When the LB is added at 40 dB SL to the opposite ear the score falls to the HB-alone level, but when added to the same ear the score falls way below this level, as did the /b/ score under the same conditions.

Nasals /m/ and /n/ and Liquid /l/

The nasals /m/ and /n/ shown in Figure 4 behave very much like the voiceless stops, which would seem to substantiate their classification by Malécot (1956) and Delattre et al. (1955) as "stops" rather than as "continuants." The nasals, like the plosives, have multiple cues, and are identified by their nasal resonances and by their formant transitions. One of the two major nasal resonances lies within the LB, which may explain why the /m/ and /n/ are resistant to "masking" when the LB is added at 40 dB SL. The /l/ pattern shown in Figure 4 in no way resembles the patterns of the /r/ and /w/, the phones with which the /l/ is usually classified, but instead presents a pattern that is similar to the "nasals" and the "stops," almost duplicating the pattern of the /m/.

The scores for the nasals improve when the LB is added at threshold with the maximum improvement occurring when the LB is added at 20 dB SL. The maximum increase in the score from that obtained with the HB alone is 38% points for the /m/ and 39%

points for the /n/. When the LB is added at 40 dB SL to the opposite ear, the score for the /n/ remains at its maximum point, but decreases slightly for the /m/. When the LB is added at 40 dB SL to the same ear, the scores decrease approximately 20% points from their maximum point, and are only slightly better than when the LB was added at threshold.

The /l/ also improves when the LB is added at threshold, with the maximum improvement of 40% points occurring when the LB is added at 20 dB SL. The score decreases when the LB is added at 40 dB SL, with the drop more pronounced when the LB is added to the same ear. The /l/ score drops slightly more than the /m/ score under the same conditions.

Voiceless Fricatives - /f/ and /s/

The /f/ and /s/ shown in Figure 5 behave somewhat the same, although the pattern for the /s/ more closely resembles the pattern for the /t/. The fricatives also have multiple cues and are identified both by a characteristic noise referred to as friction and by their formant transitions. If the friction cue is really the dominant cue for the identification of the /s/, the writer would have expected less of a "masking" effect when the LB was added at 40 dB SL to the same ear.

The /f/ score shows a steady improvement as the intensity of the LB is raised, and is the only phone in the study which reaches its maximum increase of 17% points when the LB is added at 40 dB SL to the opposite ear. However, the score drops, almost to the level where the LB was added at threshold, when the LB is added at 40 dB SL to the same ear. The /s/ score improves when the LB is added at threshold, and reaches its maximum increase of 26% points over the score obtained with the HB alone when the LB is added at 20 dB SL. The /s/ score, like the /t/ score, decreases only slightly when the LB is added at 40 dB SL to the opposite ear, but returns to the HB-alone level when added to the same ear.

Semi-Vowel /w/ and Liquid /r/

The /w/ and /r/ shown in Figure 6 do not have any additional cues such as bursts, friction, or nasal resonances to aid in their identification, which may explain the sharp drop in the scores when the LB is added at 40 dB SL to the same ear. The writer does not have a satisfactory explanation for the severe "masking" effect noted for the /r/ when the LB is added at 40 dB SL to the opposite ear, since the /r/ is the only consonant in the study which is affected in this manner. The transition for the /w/ must be essential to its identification since the second formant which lies at 600 Hz has been filtered out of the tape and the third formant lies well above the 2040 Hz high-frequency cut-off.

The /w/ and /r/ reach a maximum score when the LB is added at threshold. The maximum increase in the /w/ score of 8% points over the HB-alone score is slight, as would be expected since the HB-alone score is so very high for this phone. The /w/ score falls slightly below its HB-alone level when the LB is added at 20 dB SL to the same ear, but when the LB is added at 40 dB SL to the same ear the score now drops a dramatic 59% points from the HB-alone score. There is only a slight drop from the original score when the LB is added at 40 dB SL to the opposite ear, resulting in a difference of 53% points between the scores when the LB is added at 40 dB SL to the opposite and the same ear.

The HB-alone score for the /r/ is relatively high, second only to the score for the /w/. The maximum increase of 22% points is reached when the LB is added at threshold to the same ear, and the slightly higher score for the same-ear condition resembles the behavior of the /b/ and the /d/ under the same conditions. When the LB is added at 20 dB SL to the same ear, the /r/ score, as was the /w/ score under the same conditions, is already slightly below the HB-alone level. However, unlike the /w/ score which is only slightly affected when the LB is added at 40 dB SL to the opposite ear, the /r/ scores fall way below the HB-alone level when the LB is added at 40 dB SL to the opposite as well as the same ear.

Glottal - /h/

The changes in the /h/ score as shown in Figure 7 are slight when the LB is added at the various intensity levels to the opposite ear. There does seem to be some noticeable improvement in the score when the LB is added at threshold and at 20 dB SL to the same ear, with the maximum increase of 6% points over the HB-alone level occurring when the LB is added at threshold. When the LB is added at 40 dB SL to the same ear, the score falls slightly below the HB-alone level.

Perceptual Confusions

Consonants are usually described in terms of three distinctive features: (1) place-of-articulation -- where the major constriction in the vocal tract occurs, such as labial, alveolar, velar, etc., (2) manner-of-articulation -- what type of articulation occurs, such as plosive, fricative, nasal, etc., and (3) voiced-voiceless opposition -- whether the vocal cords vibrate or not. When two consonants are perceptually confused, three types of errors can result: (1) if they share a common place-of-articulation and vocal cord activity a "manner" error has been made; (2) if they share a common manner-of-articulation and vocal cord activity a "place" error has been made; and (3) if they share both a common manner- and place-of-articulation a "voicing" error has been made. A response that does not share a distinctive feature with the stimulus could be considered a "random" error.

The articulation curves for each sound indicate that when the LB is added at favorable HB/LB ratios, the number of correct responses is essentially the same, regardless of whether the LB is added to the same or opposite ear. It can be assumed that the same kind of information is contributed at favorable HB/LB ratios, since, as will be seen, the perceptual errors that occur are similar for the same and opposite ear conditions. However, when the LB is added at unfavorable HB/LB ratios, the articulation curves for each sound show that there is a greater "masking" effect when the LB is added to the same ear, and in addition to reducing the number of correct responses, it will be seen that there is a higher incidence of random perceptual errors than when the LB is added to the opposite ear.

The articulation curve is intercepted at the same point when the HB-alone is presented and when the LB is added at 40 dB SL to the same ear for the /t/, /k/, and /s/, and to the opposite ear for the /b/, /d/, and /g/. These points on the articulation curve which are equal quantitatively, are not always the same qualitatively, since there are generally more voicing errors in addition to a greater number of random perceptual errors when the HB-alone is presented, as compared to a higher percentage of place and manner errors at unfavorable HB/LB ratios. Miller and Nicely (1954) also noted that high-pass filtering produced random confusions whereas low-pass filtering left the phonemes similar in predictable ways.

Figures 8-21 present all perceptual confusions which occurred with an incidence of 5% or more[4], and an analysis was made of the specific information -- place, manner, and/or voicing -- that is supplied when the LB is added.

Voiceless Plosives - /p/, /t/, and /k/

As can be seen in Figures 8-10 almost all the errors are place errors: t/k, p/k, t/p, k/p, k/t, and p/t. The /h/ is substituted for both the /p/ and /t/, and it will be seen later that this is a reciprocal relationship, since the /p/ and /t/ as well as the

---

[4] The writer treated confusions which occurred less than 5% of the time as not significant. The perceptual confusions with less than a 5% incidence, in addition to all confusions presented in Figures 8-21, appear in the Confusion Matrices, Tables 2-8.

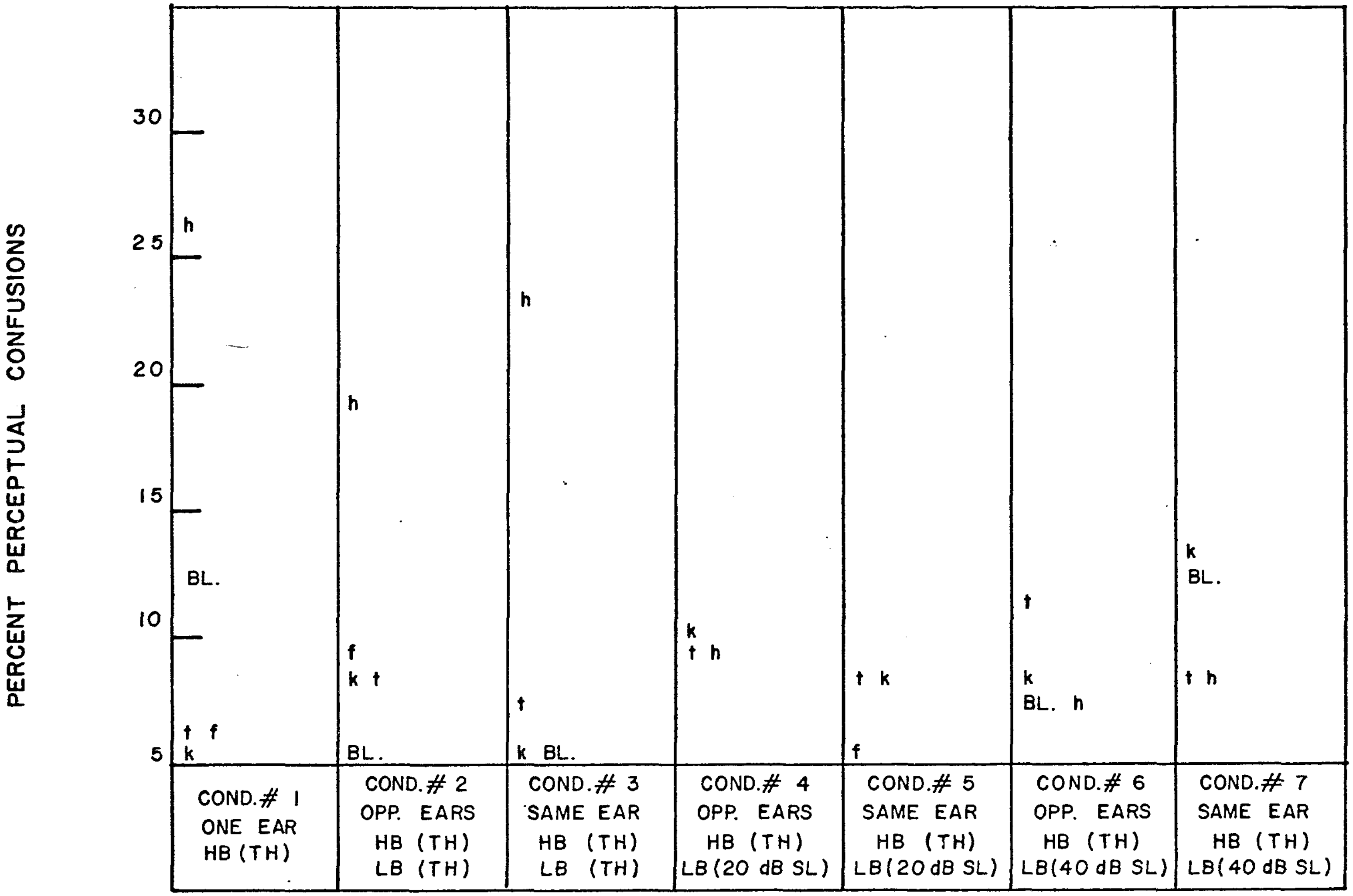

Figure 8. Perceptual confusions for the stimulus /p/.

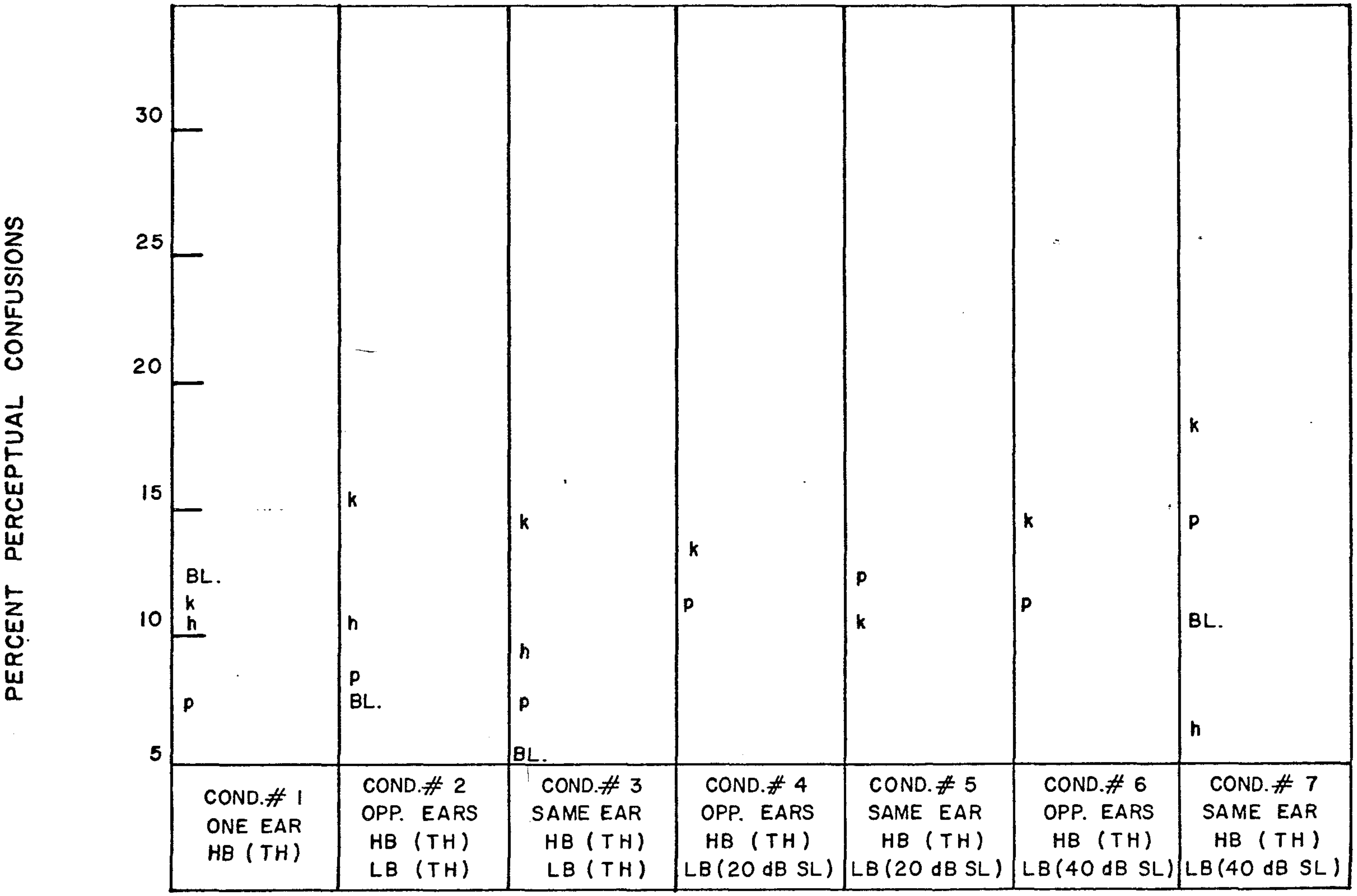

Figure 9. Perceptual confusions for the stimulus /t/.

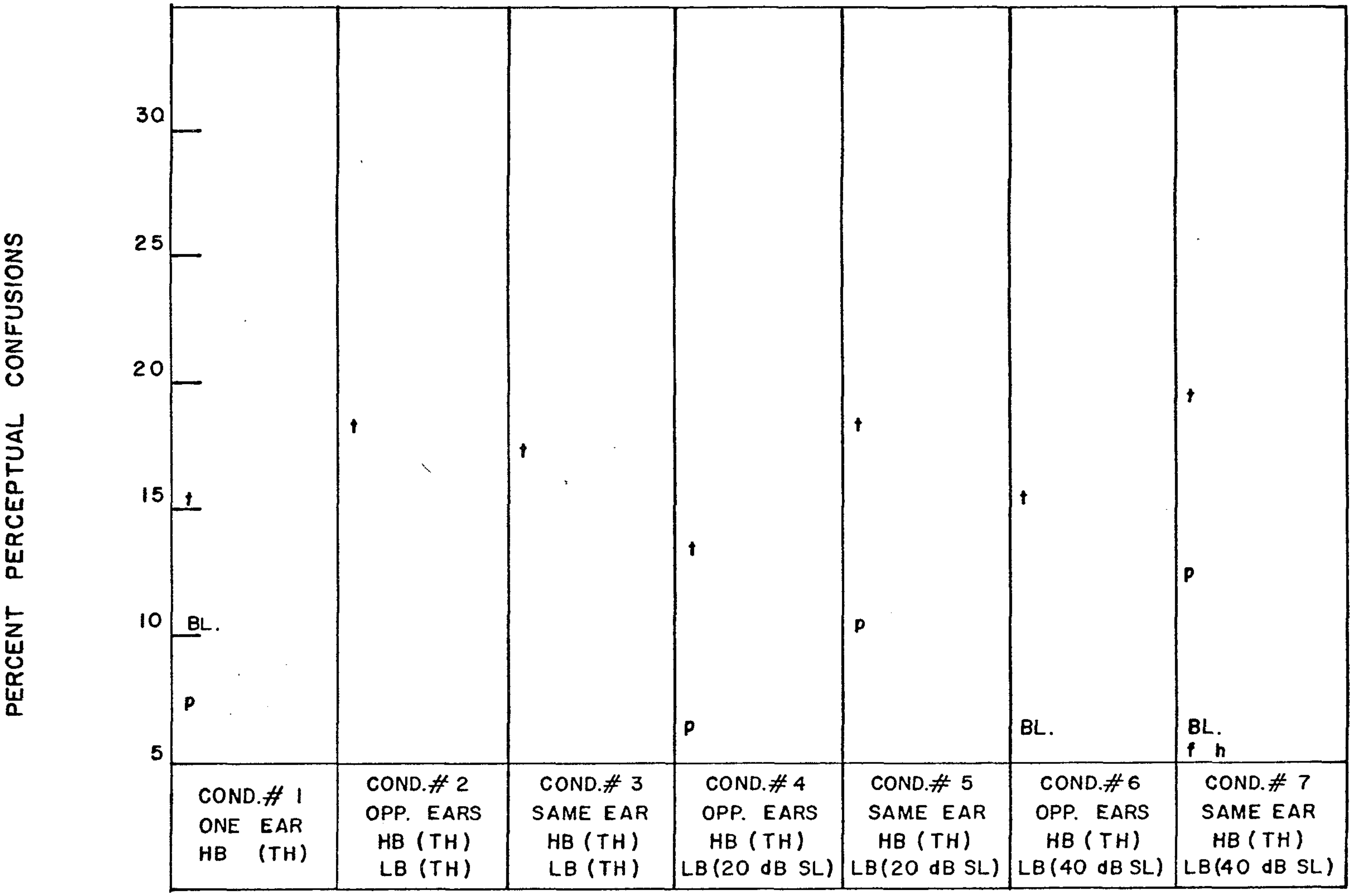

Figure 10. Perceptual confusions for the stimulus /k/.

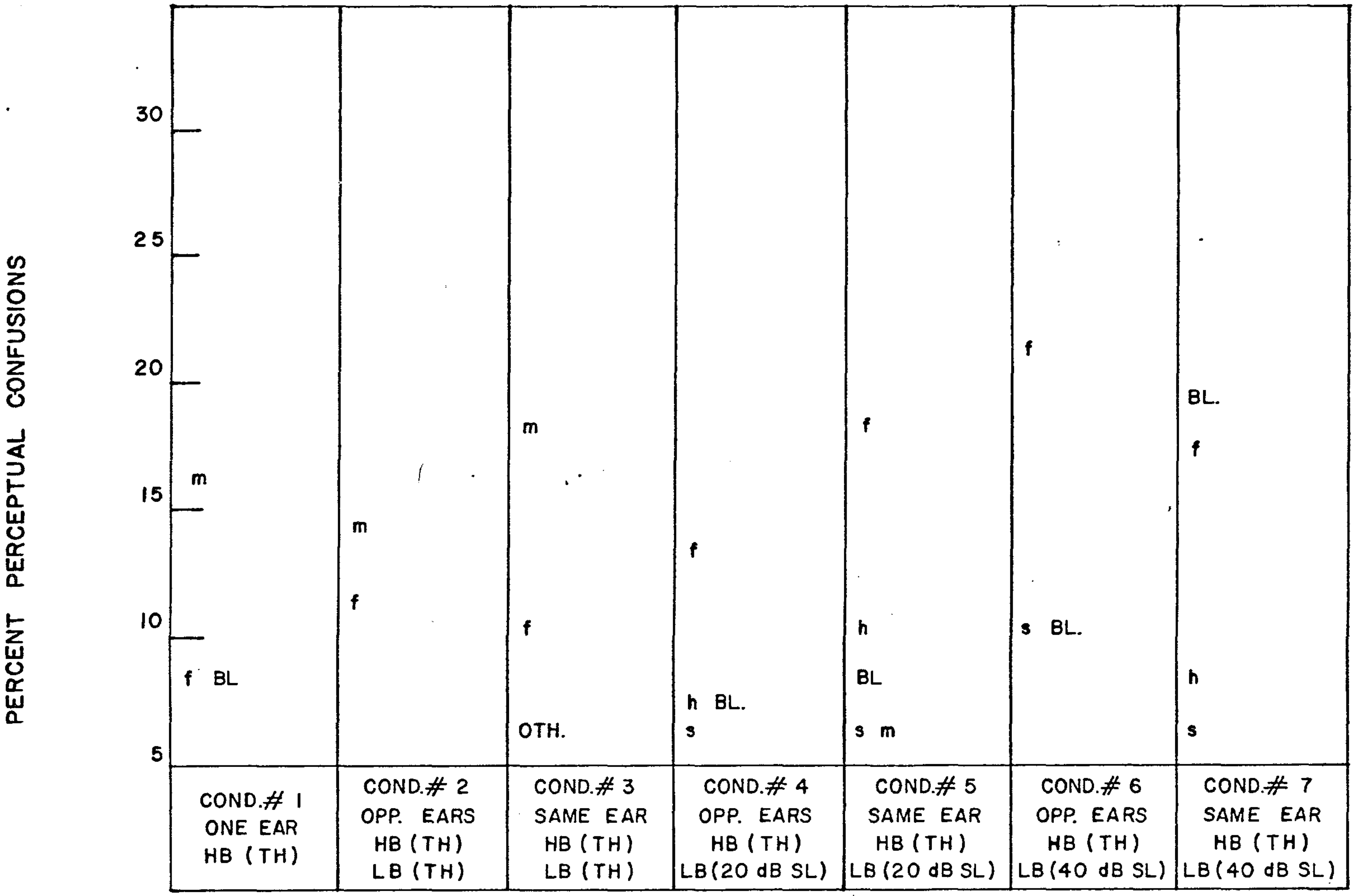

Figure 11. Perceptual confusions for the stimulus /b/.

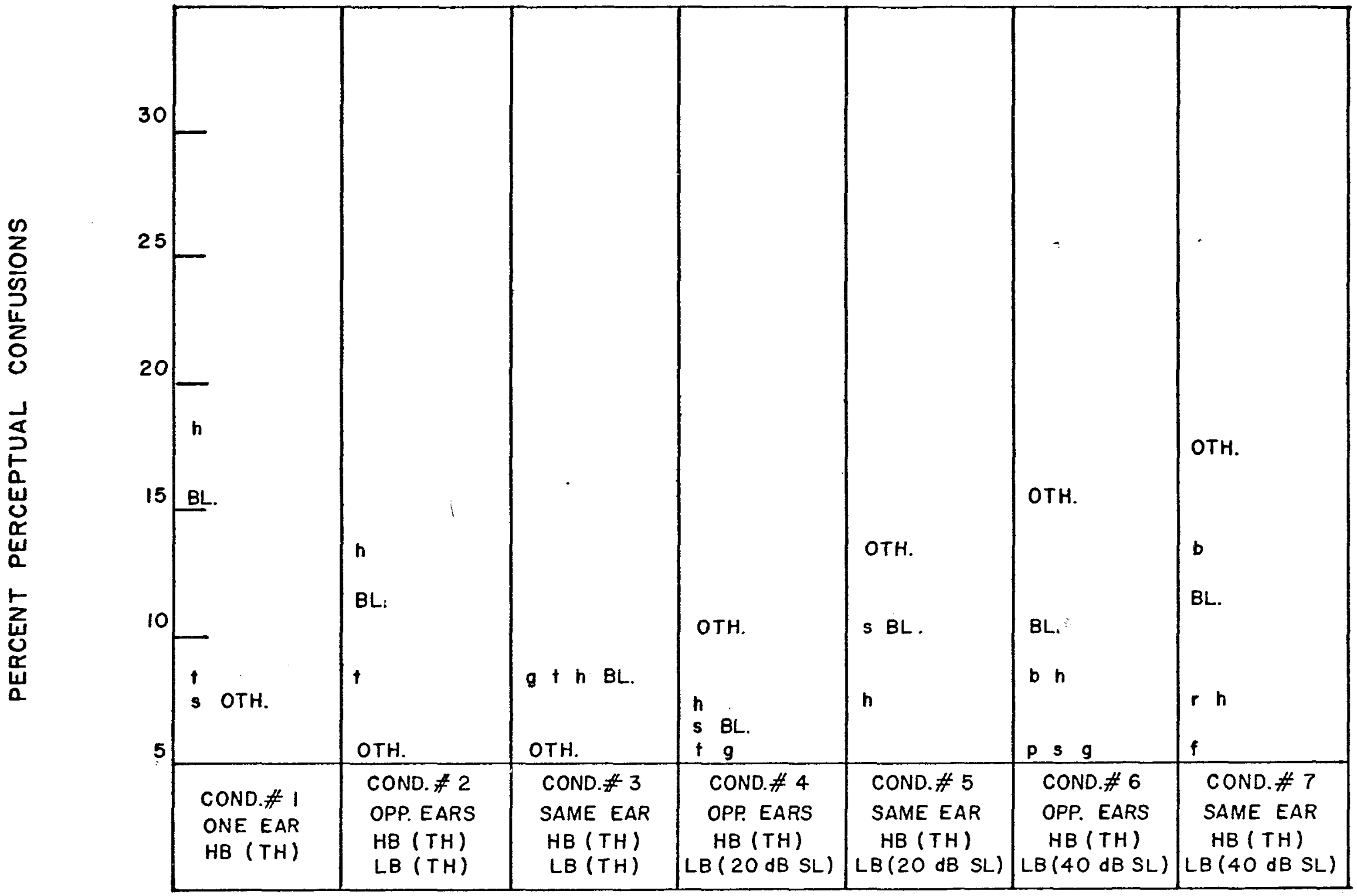

Figure 12. Perceptual confusions for the stimulus /d/.

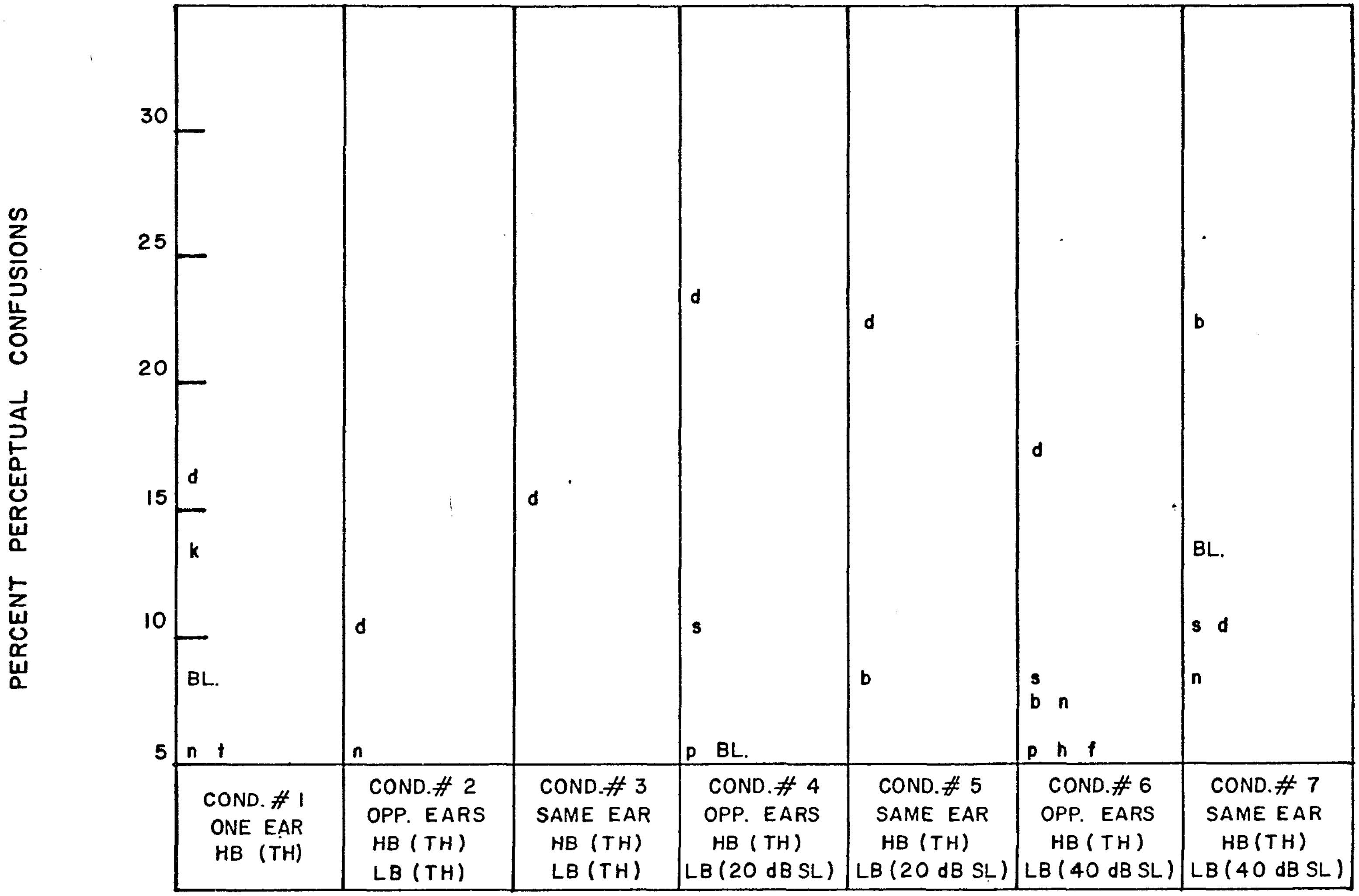

Figure 13. Perceptual confusions for the stimulus /g/.

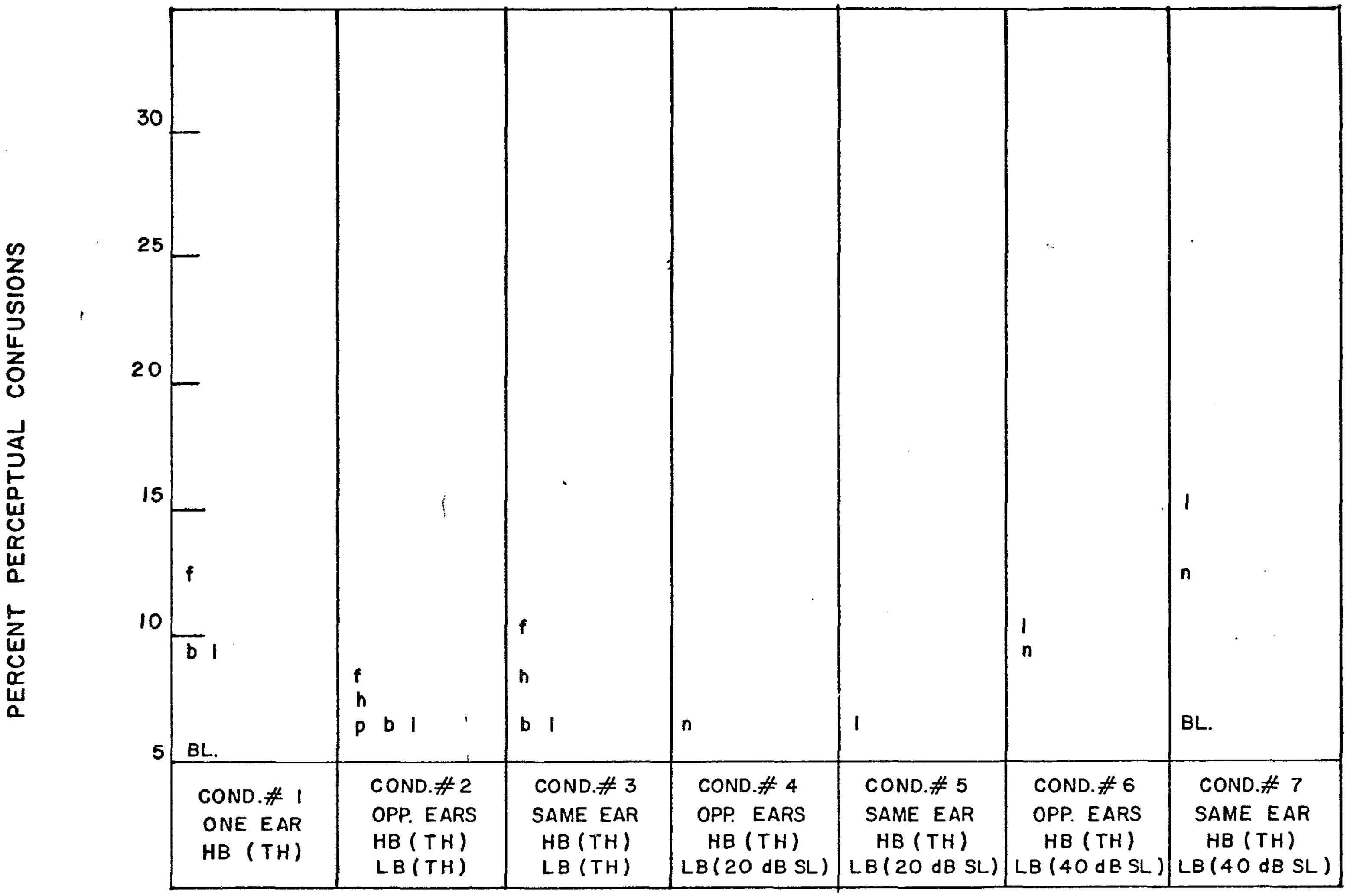

Figure 14. Perceptual confusions for the stimulus /m/.

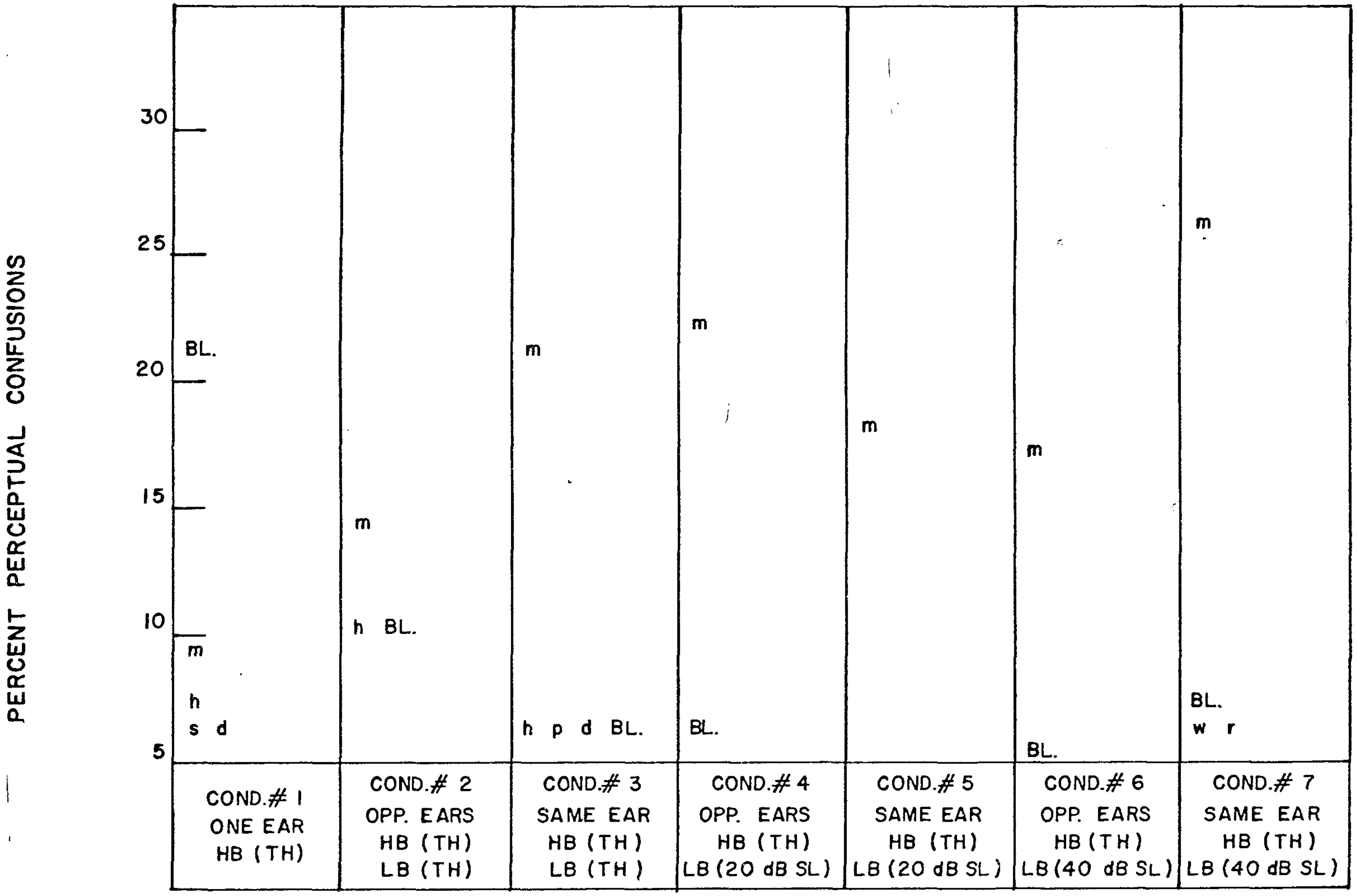

Figure 15. Perceptual confusions for the stimulus /n/.

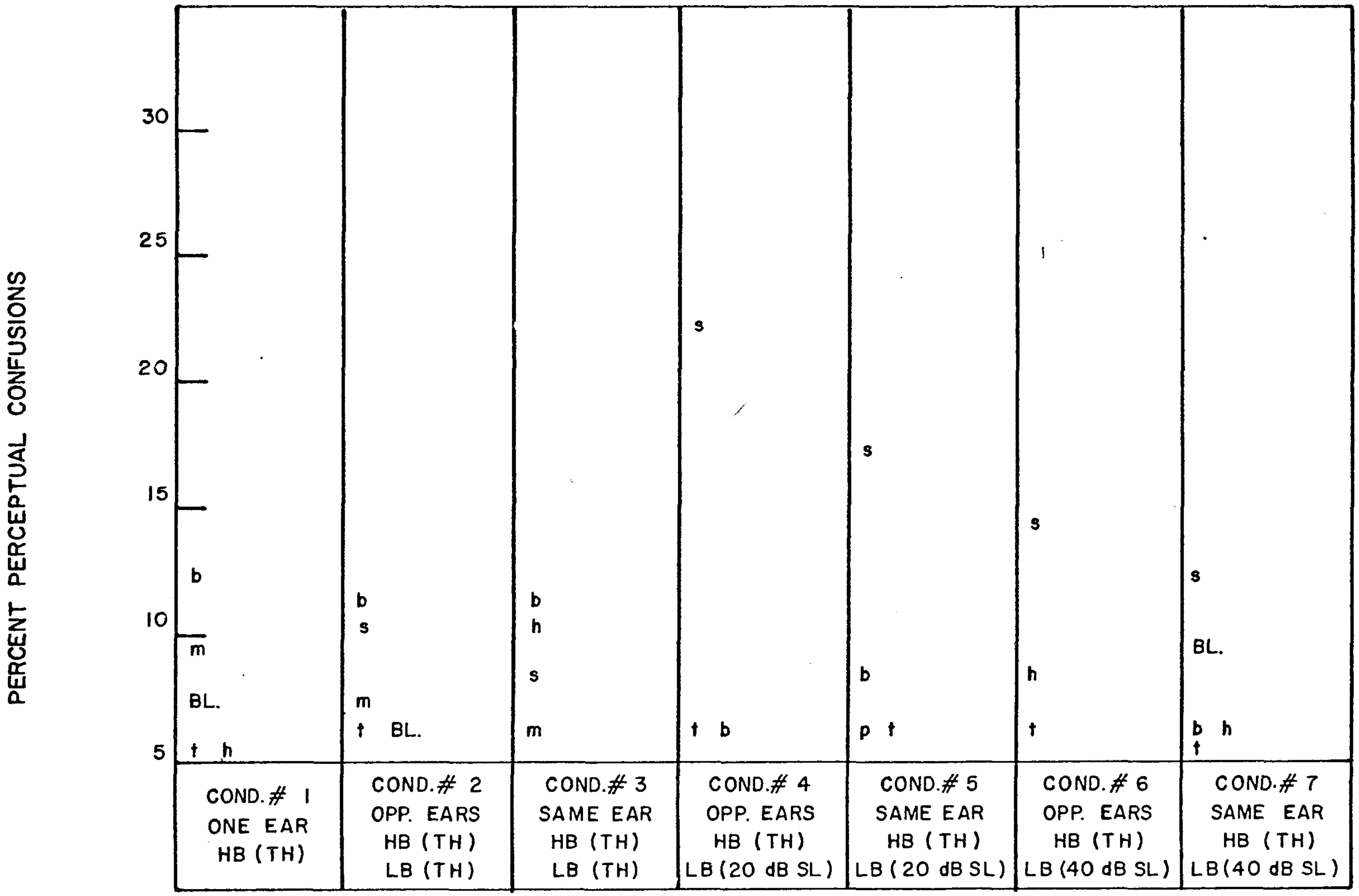

Figure 16. Perceptual confusions for the stimulus /f/.

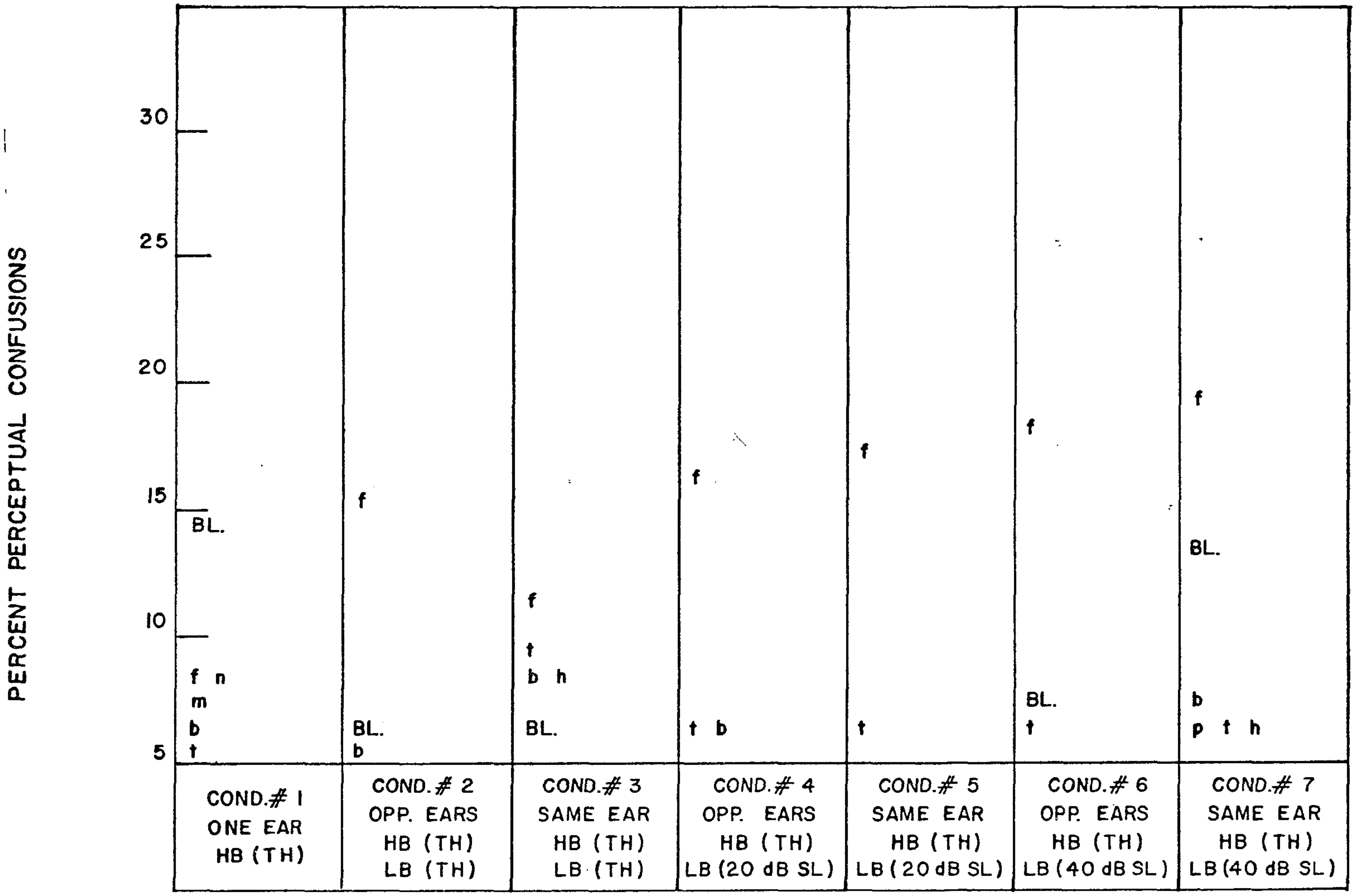

Figure 17. Perceptual confusions for the stimulus /s/.

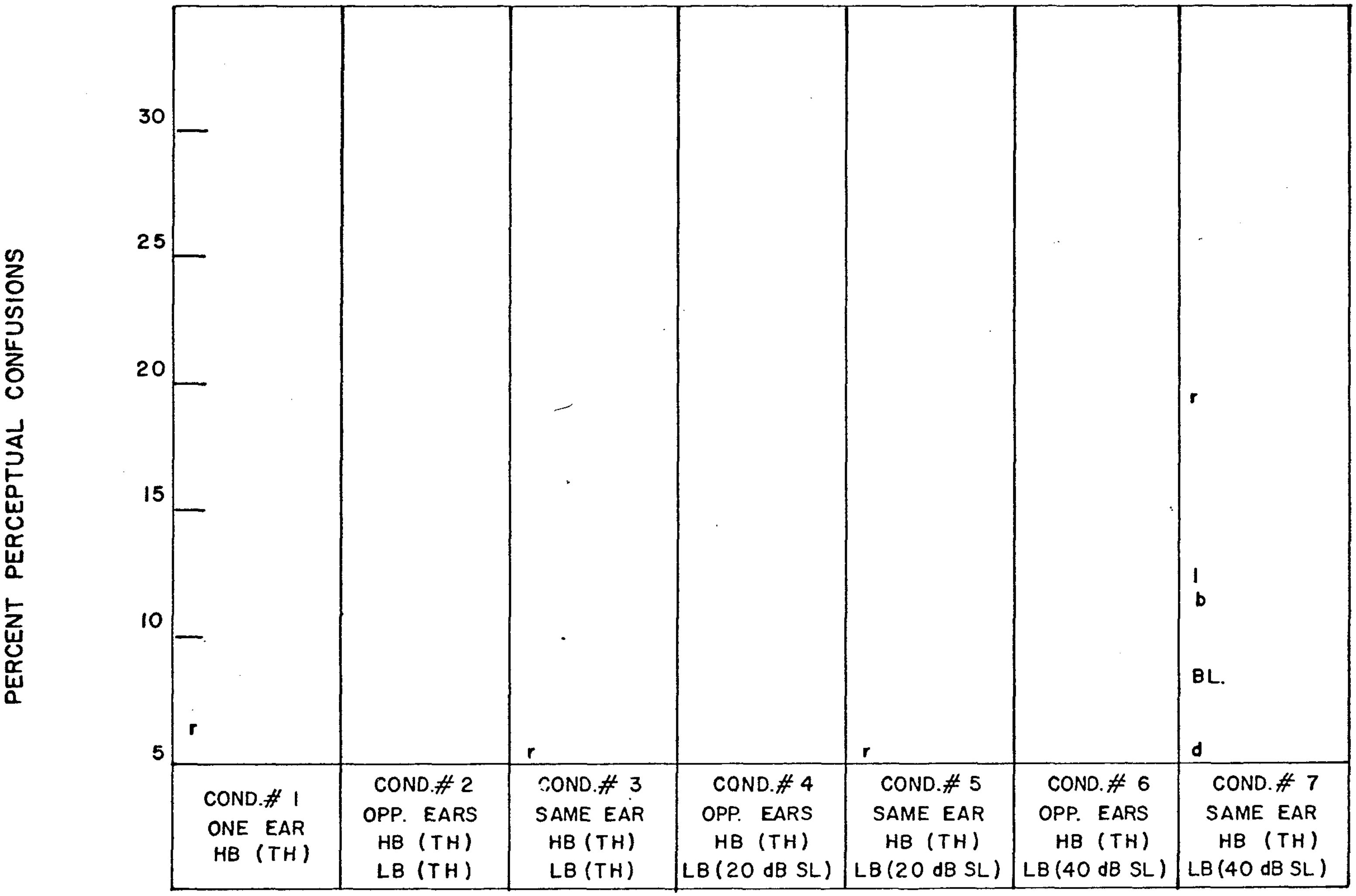

Figure 18. Perceptual confusions for the stimulus /w/.

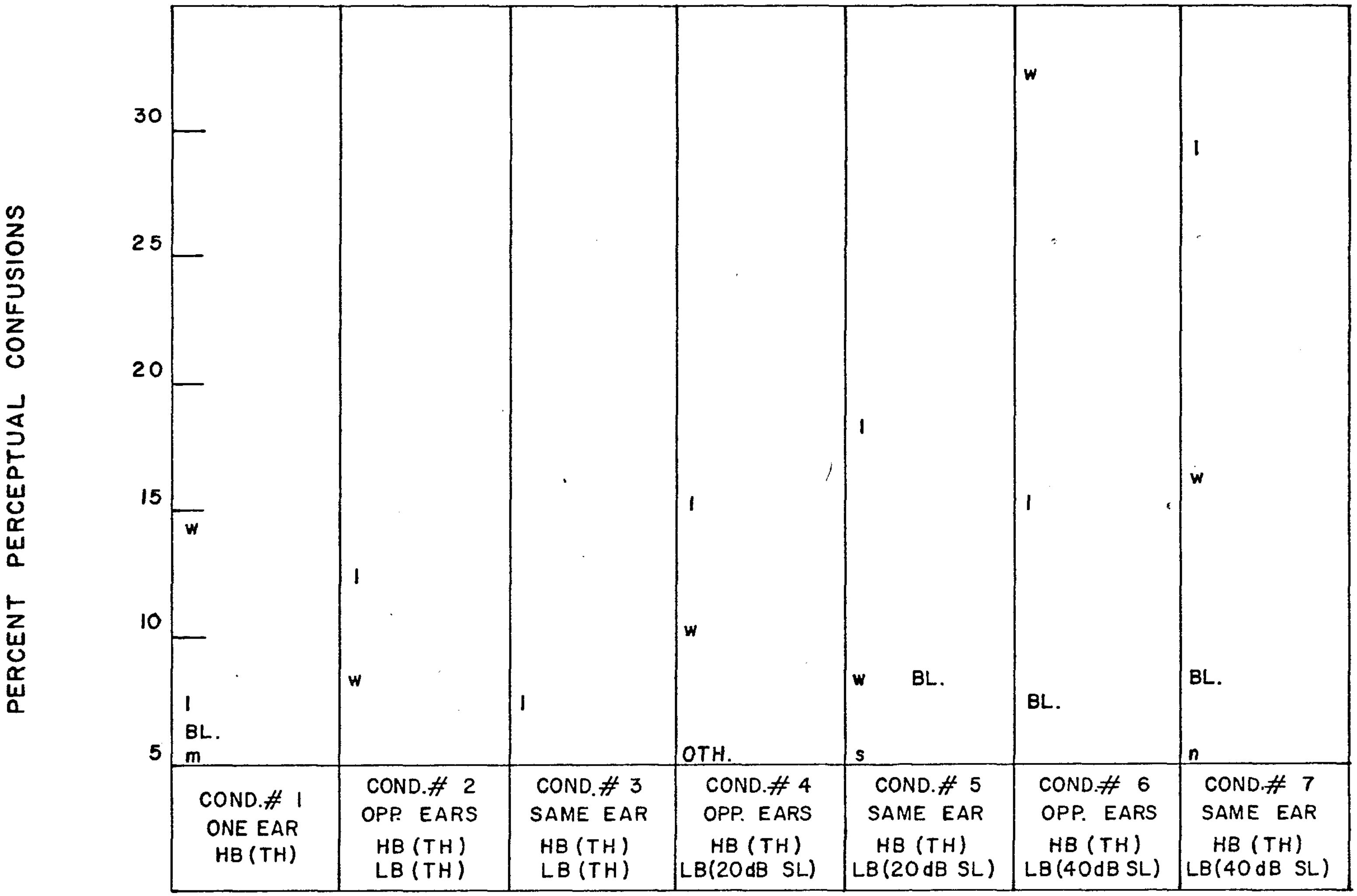

Figure 19. Perceptual confusions for the stimulus /r/.

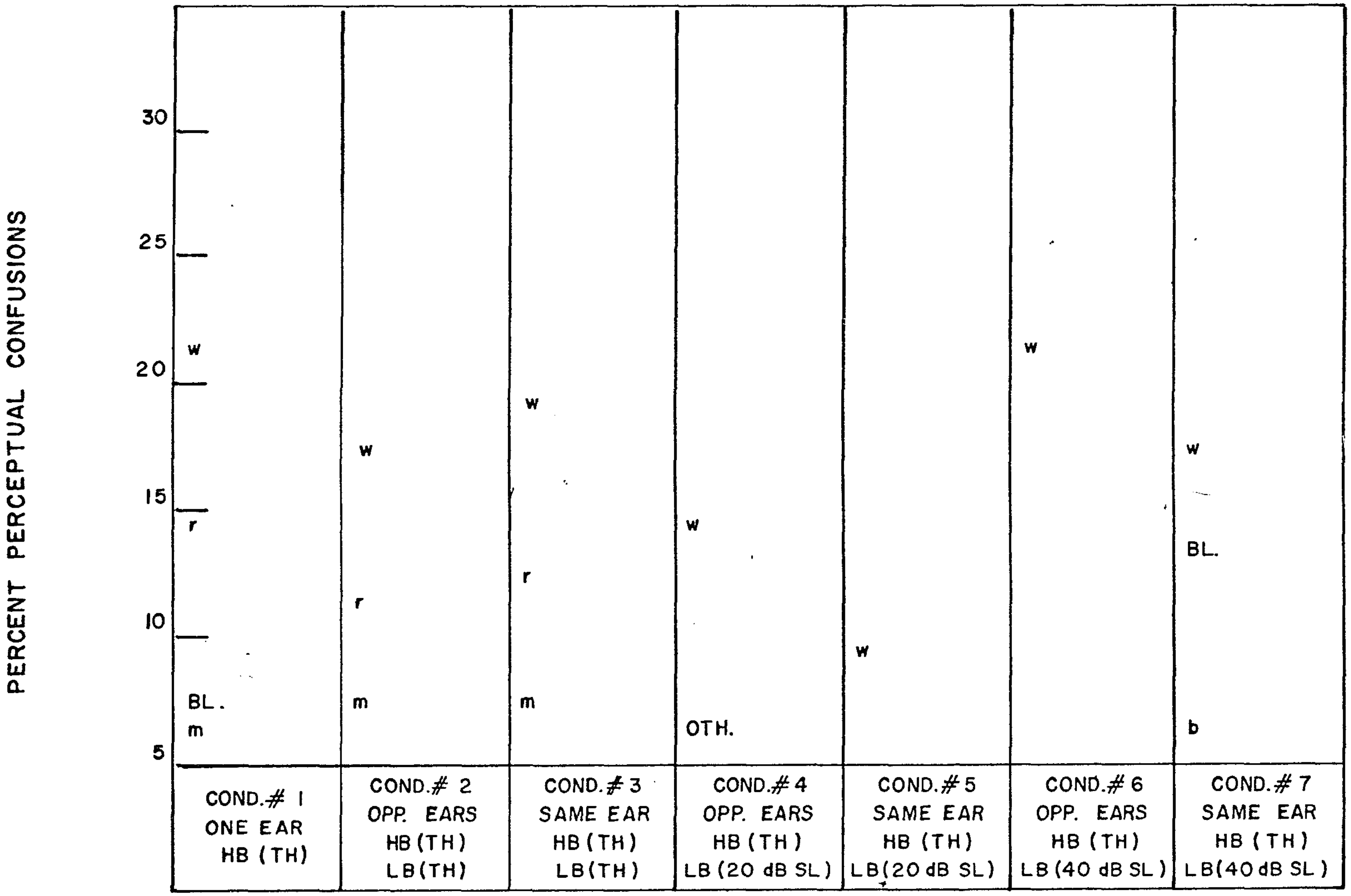

Figure 20. Perceptual confusions for the stimulus /l/.

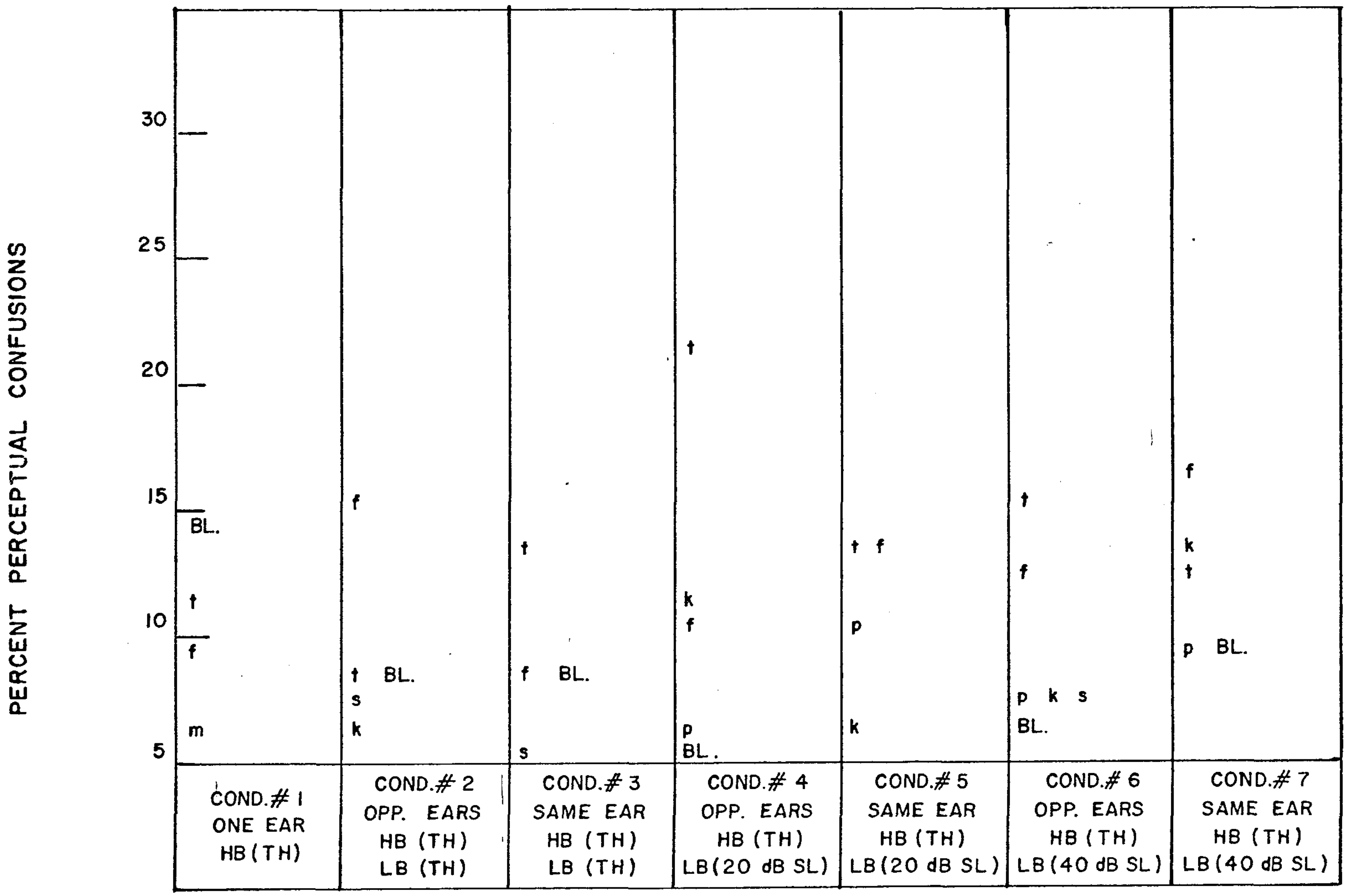

Figure 21. Perceptual confusions for the stimulus /h/.

/k/ are often substituted for the /h/.

/p/

1. There is an incidence of 26% of h/p substitutions when the HB alone is presented. When the LB is added at threshold to the opposite ear and then to the same ear, the substitutions occur 19% and 23% of the time, respectively, but in conditions #4, 6, and 7 these substitutions appear with a frequency of only 5-10%.

2. Both the t/p and k/p substitutions appear between 5-10% of the time in all conditions, with a slightly higher incidence of k/p substitutions when the LB is added at 40 dB SL to the same ear.

3. There is a 12% incidence of No Responses when the HB alone is presented, lowering to 5% when the LB is added at threshold. The incidence of No Responses falls below the 5% level when the LB is added at 20 dB SL and increases again to 7% and 12% when the LB is added at 40 dB SL to the opposite and the same ear, respectively.

4. There does not seem to be a pattern to the f/p substitutions which occur with a 6% incidence when the HB alone is presented and appear again with a 5-10% incidence only in conditions #2 and #5.

/t/

1. The k/t substitutions appear with an incidence between 10-15% in the first six conditions and rise to 18% when the LB is added at 40 dB SL to the same ear.

2. The p/t substitutions occur between 7-8% of the time when the HB is presented alone and when the LB is presented at threshold, rise to 11-12% in conditions #4, 5, and 6, and increase again to 14% when the LB is added at 40 dB SL to the same ear.

3. The h/t substitutions occur between 9-10% of the time when the HB is presented alone and when the LB is added at threshold but then do not appear again until the final condition, where there is an incidence of 6% when the LB is added at 40 dB SL to the same ear.

4. The number of No Responses is 12% when the HB alone is presented, decreases to 7% and 5% when the LB is added at threshold to the opposite and the same ear, respectively, then falls below the 5% level until the final condition where there is a 10% incidence of No Responses when the LB is added at 40 dB SL to the same ear.

/k/

1. The t/k substitutions occur with an incidence between

13-19% in all conditions.

2. The p/k substitutions appear with a 7% incidence when the HB alone is presented but fall below the 5% level when the LB is added at threshold. The substitutions reappear with a 6% and 10% incidence when the LB is presented at 20 dB SL to the opposite and the same ear, respectively, and with a 12% incidence when the LB is added at 40 dB SL to the same ear.

3. There is a 10% incidence of No Responses when the HB alone is presented. This category then falls below the 5% level for conditions #2 through 5 and reappears with a 6% incidence when the LB is added at 40 dB SL in conditions #6 and 7.

4. There does not seem to be a pattern to the 5% incidence of f/k and h/k substitutions noted in the final condition when the LB is added at 40 dB SL to the same ear.

Voiced Plosives - /b/, /d/, and /g/

As can be seen in Figures 11-13 the substitutions for the voiced plosives are more varied than for the voiceless plosives. In addition to the place errors of: d/g, b/g, g/d, and b/d, there is the manner error of m/b which will be seen later to be reciprocal. The incidence of b/m and m/b substitutions is sharply reduced when the LB is added at supra-threshold intensities. The substitution of /f/ for /b/, which is also reciprocal, can be considered a manner error (not an exact example of a manner error since the two phones do not share a common vocal cord activity), since both phones are articulated in the front of the mouth and share a common locus, 720-800 Hz. Another manner error is the /dʒ/ for /d/ substitution. The substitution of /t/ for /d/ is the only voicing error noted in the present study. Steinberg (1929) commented that under various high- and low-pass filter conditions, "...voiced consonants are rarely mistaken for unvoiced sounds." (p. 129) An interesting substitution is that of the /s/ for all three voiced plosives, and it will be seen later that the /s/ for /b/ substitution is reciprocal. The /s/ substitution, in addition to the substitutions of n/g and h/d, cannot be explained in terms of any distinctive features.

/b/

1. The f/b substitutions occur in all conditions, beginning with an incidence of 8% when the HB alone is presented, rising slightly to between 10-11% when the LB is presented at threshold, increasing again to 13% and 18% when the LB is added at 20 dB SL to the opposite and the same ear, respectively, reaching a peak of 21% when the LB is presented at 40 dB SL to the opposite ear, and falling again to 17% when the LB is presented to the same ear.

2. The m/b substitutions occur between 14-18% of the

time when the HB alone is presented and when the LB is added at threshold. These substitutions rise above the 5% level only once more -- when the LB is added at 20 dB SL to the same ear.

3. The s/b substitutions appear with an incidence between 6-10% only when the LB is added at both 20 dB and 40 dB SL.

4. The incidence of No Responses is 8% when the HB alone is presented, falls below the 5% level when the LB is added at threshold, reappears between 7-8% when the LB is added at 20 dB SL, rises slightly to 10% when the LB is added at 40 dB SL to the opposite ear, and reaches a peak of 19% when added to the same ear.

5. There does not seem to be a pattern to the incidence between 7-10% of h/b substitutions which appear when the LB is added at 20 dB, and also at 40 dB SL to the same ear.

/d/

1. The responses labeled "other" include an overall incidence of 50% for the /dʒ/ and 17% for the /z/. The incidence of "other" responses is between 5-7% when the HB alone is presented and when the LB is added at threshold. There is then a gradual increase from 10% to 17% as the LB is added at 20 dB and 40 dB SL, with the peak occurring in the final condition.

2. The h/d substitutions also occur in all conditions, with a maximum incidence of 18% when the HB alone is presented, falling to 13% when the LB is presented at threshold to the opposite ear, and occurring between 7-8% of the time in all other five conditions.

3. The t/d substitutions occur with an incidence between 5-8% in the first four conditions.

4. The s/d substitutions appear with a 7% incidence when the HB alone is presented, fall below the 5% level when the LB is added at threshold, appear again with a 6% and 10% incidence when the LB is added at 20 dB SL to the opposite and same ear, respectively, and occur with a 5% incidence when the LB is added at 40 dB SL to the opposite ear.

5. The b/d substitutions occur only when the LB is presented at 40 dB SL, with an 8% and 13% incidence for the opposite and the same ear, respectively.

6. The number of No Responses is at a maximum of 15% in the first condition when the HB alone is presented, and falls to 11% and 8% when the LB is added at threshold to the opposite and the same ear, respectively, reaches a minimum of 6% when the LB is added at 20 dB SL to the opposite ear, and rises again to an incidence between 10-11% for the last three conditions.

7. There are several substitutions which occur with a 5-8% incidence which do not seem to show a pattern: the g/d substitutions in conditions #3, 4, and 6, the p/d substitutions in condition #6, and the substitutions of /f/ and /r/ for the /d/ in the last condition.

/g/

1. The d/g substitutions are the only ones which occur in all conditions, with a 16% incidence when the HB alone is presented, falling to 10% when the LB is added at threshold to the opposite ear but rising to 15% when added to the same ear, rising sharply again to between 22-23% when the LB is added at 20 dB SL, falling again to 17% and then to 10% when the LB is added at 40 dB SL to the opposite and the same ear, respectively.

2. The b/g substitutions occur 7-8% of the time when the LB is added at 20 dB SL to the same ear and 40 dB SL to the opposite ear, reaching a peak incidence of 22% in the final condition. The /b/ and /d/ substitutions for the /g/ crisscross sharply in the last two conditions showing an inverse relationship.

3. The incidence of n/g substitutions is 5% for the first two conditions and rises above the 5% level again only when the LB is added at 40 dB SL.

4. There is an incidence between 8-10% of s/g in conditions #4, 6, and 7.

5. There is an incidence of No Responses of 8% in the first condition, 5% in condition #4, and 13% in the final condition.

6. There are several substitutions which do not seem to show a pattern: the 5% and 13% incidence of t/g and k/g, respectively, when the HB alone is presented; and the 5% incidence of p/g in conditions #4 and #6 and of h/g and f/g in condition #6.

Nasals - /m/ and /n/

As can be seen in Figures 14 and 15 the expected place errors of m/n and n/m occur. The substitutions of /l/ for /m/ will be seen later to be reciprocal. It should be noted that Cooper <u>et al.</u> (1952) also found l/m substitutions. The manner errors of b/m and f/m are both reciprocal (the f/m is not an exact example of a manner error since the two phones do not share a common vocal cord activity). As was noted earlier in the explanation of the f/b substitutions, the /f/, /b/, and /m/ can be considered to share a common place-of-articulation, and they have a common locus of 720-800 Hz. An interesting substitution which cannot be explained on the basis of any distinctive features is that of the /h/ for both the /m/ and the /n/.

/m/

1. The only consistent substitutions are l/m which occur in six out of the seven conditions, with an incidence between 6-9% when the HB alone is presented and when the LB is added at threshold, 6% when the LB is added at 20 dB SL to the same ear, rising to 10% and 15% when the LB is added at 40 dB SL to the opposite and the same ear, respectively.

2. The b/m substitutions occur 6-9% of the time only in the first three conditions, that is, when the HB alone is presented and when the LB is added at threshold.

3. The n/m substitutions almost seem to replace the b/m substitutions, since the n/m substitutions appear in the last three out of four conditions, with a 6% incidence when the LB is added at 20 dB SL to the opposite ear, rising to 9% and 12% when the LB is presented at 40 dB SL to the opposite and the same ear, respectively.

4. The f/m substitutions occur with a 12% incidence in the HB-alone condition, and with an 8% and 10% incidence when the LB is added at threshold to the opposite and same ear, respectively.

5. Several substitutions, with a 6-8% incidence, occur only when the LB is added at threshold. When the LB is added to the opposite ear the additional substitutions which occur are p/m and h/m, and when the LB is added to the same ear, only the h/m

substitutions occur.

6. The incidence of No Responses occurs between 5-6% of the time only when the HB alone is presented and when the LB is added at 40 dB SL to the same ear.

/n/

1. The predominant substitutions are m/n substitutions which correspond to Malécot's (1956) findings that there is a tendency for all nasals to be perceived as /m/. The incidence is 9% when the HB alone is presented and increases when the LB is added, occurring with a 14-22% incidence until the peak of 26% is reached when the LB is added at 40 dB SL to the same ear.

2. The number of No Responses is at a maximum of 21% when the HB alone is presented and decreases sharply when the LB is added, showing an inverse relationship to the m/n substitutions.

3. There is an incidence between 6-10% of h/n substitutions only when the HB alone is presented and when the LB is added at threshold.

4. There are several substitutions which occur with a 6% incidence which do not seem to show a pattern: d/n in conditions #1 and #3; s/n in the first condition; p/n in the third condition; and both the /w/ and /r/ for the /n/ in the final condition.

Voiceless Fricatives - /f/ and /s/

As seen in Figures 16 and 17 the expected place errors of f/s and s/f occur. The t/s substitution is a manner error and may be related to the t/f substitution. The reciprocal relationship of the b/f and m/f manner errors has already been discussed. The substitution of the nasals, particularly the /m/, for both the /f/ and /s/ almost disappears when the LB is added at suprathreshold levels. The /h/ is substituted for the /f/ and it will be seen later that both the /s/ and /f/ are substituted for the /h/. The reciprocal relationship of the b/s substitution also has been previously described.

/f/

1. There are no s/f substitutions when the HB alone is presented, but the substitutions appear in all conditions where the LB is added. There is an incidence of substitutions between 8-10% when the LB is added at threshold, rising sharply to 22% when the LB is added at 20 dB SL to the opposite ear and lowering slightly to 17% when the LB is added to the same ear, decreasing again to between 13-14% when the LB is added at 40 dB SL.

2. There is an incidence between 11-12% of b/f substitutions when the HB alone is presented and when the LB is presented at threshold. The incidence of substitutions decreases to 6% and

8% when the LB is added at 20 dB SL to the opposite and the same ear, respectively. The b/f substitutions also occur with a 6% incidence when the LB is added at 40 dB SL to the same ear.

3. The t/f substitutions occur between 5-6% of the time in all conditions except the third.

4. The m/f substitutions occur with an incidence between 6-9% only when the HB is presented alone and when the LB is added at threshold.

5. There is an incidence between 6-7% of No Responses when the HB alone is presented and when the LB is added at threshold to the opposite ear. The incidence of No Responses falls below the 5% level until the final condition where there is a 9% incidence when the LB is added at 40 dB SL to the same ear.

6. There is a 5% incidence of h/f substitutions when the HB alone is presented. These substitutions increase to 10% when the LB is added at threshold to the same ear and reappear with an incidence between 6-8% when the LB is presented at 40 dB SL.

7. The p/f substitutions do not show a pattern, since only when the LB is added at 20 dB SL to the same ear do these substitutions rise above the 5% level.

/s/

1. There is an 8% incidence of f/s substitutions when the HB alone is presented, and an incidence between 11-19% in all conditions where the LB is added.

2. The b/s substitutions appear with an incidence between 5-8% in five of the conditions: HB alone, LB added at threshold, LB added at 20 dB SL to the opposite ear, and LB added at 40 dB SL to the same ear.

3. The t/s substitutions occur between 5-9% of the time in all conditions except when the LB is added at threshold to the opposite ear.

4. The nasals /m/ and /n/ are substituted for the /s/ with a 7% and 8% incidence, respectively, only when the HB alone is presented.

5. There is a 14% incidence of No Responses when the HB alone is presented. This category falls to 6% when the LB is added at threshold, occurs below the 5% level when the LB is added at 20 dB SL, and reappears with a 7% and 13% incidence when the LB is added at 40 dB SL to the opposite and the same ear, respectively.

6. The following substitutions do not seem to show a pattern: the h/s substitutions rise above the 5% level only in conditions #3 and #7, and the p/s substitutions occur only in the final

condition.

Semi-Vowel - /w/ and Liquids - /r/ and /l/

Although the behavior of the liquid /l/ when the LB is added resembles the nasals /m/ and /n/, the perceptual confusions of the /l/ belong more to the semi-vowel and liquid categories. Figures 18-20 show the expected place errors: w/l, r/l, r/w, l/r, and w/r. The only non-place error is the m/l substitution, and the l/m substitution has already been described. It is interesting to note that the /l/ has characteristics of both the nasal and the semi-vowel and liquid categories.

<u>/w/</u>

1. The only substitutions which occur with any regularity are the r/w substitutions, with an incidence between 5-6% in conditions #1, 3, and 5, rising sharply to 19% in the final condition when the LB is added at 40 dB SL to the same ear.

2. Several substitutions which do not seem to show any pattern appear only in the final condition when the LB is added at 40 dB SL to the same ear: an incidence of 5% for d/w, 11% for b/w, 12% for l/w, and an 8% incidence of No Responses.

<u>/r/</u>

1. The l/r substitutions appear in all conditions, with a

7% incidence when the HB alone is presented. The incidence rises to 12% when the LB is presented at threshold to the opposite ear and falls slightly to 7% when presented to the same ear, rises to 15% and 18% when the LB is presented at 20 dB SL to the opposite and the same ear, respectively, falls slightly again to 15% when the LB is presented at 40 dB SL to the opposite ear, but reaches a peak of 29% when presented to the same ear.

2. The w/r substitutions occur in all conditions except the third, with a 14% incidence when the HB alone is presented, and an incidence between 8-10% when the LB is presented at threshold to the opposite ear and at 20 dB SL to both the opposite and the same ear. The incidence rises sharply to 32% when the LB is presented at 40 dB SL to the opposite ear and falls sharply to 16% in the final condition when the LB is presented to the same ear. The w/r and l/r substitutions exhibit an interesting pattern of crisscrossing, with the inverse relationship most pronounced in the last two conditions.

3. Several substitutions which appear with a 5% incidence do not seem to exhibit any pattern: m/r in the first condition, s/r in condition #5, and n/r in the final condition. There is also an incidence between 6-8% of No Responses in the first condition and again in the last three conditions.

/l/

1. The only substitutions which occur in all conditions are the w/l substitutions, with a 21% incidence when the HB alone is presented, the incidence falling slightly to 17% and 19% when the LB is added at threshold to the opposite and the same ear, respectively, falling sharply to 14% and 9% when the LB is added at 20 dB SL to the opposite and the same ear, respectively, rising sharply again to 21% when the LB is presented at 40 dB SL to the opposite ear, and falling slightly to 17% when presented to the same ear.

2. The r/l substitutions occur between 11-14% of the time only when the HB alone is presented and when the LB is added at threshold.

3. The m/l substitutions occur with an incidence between 6-7% in the same three conditions in which the r/l substitutions are found.

4. There does not seem to be a pattern for the 7% and 13% incidence of No Responses found in the first and the last condition, or for the b/l substitutions found only in the final condition.

Glottal - /h/

The /h/ has no locus of its own but takes on the formant structure of the adjoining vowels. It is understandable, therefore, why the substitutions for the /h/, as shown in Figure 21, include the voiceless plosives /p/, /t/, and /k/, and the voiceless fricatives /f/ and /s/.

1. The t/h substitutions occur in all conditions, with an 11% incidence when the HB alone is presented. The incidence falls slightly to 8% when the LB is added at threshold to the opposite ear and rises slightly to 13% when the LB is added to the same ear, peaks at 21% when the LB is added at 20 dB SL to the opposite ear, and falls again to between 12-15% for the last three conditions.

2. The f/h substitutions also occur in all conditions, with a 9% incidence when the HB alone is presented. The incidence rises to 15% when the LB is added at threshold to the opposite ear but falls to 8% when added to the same ear, then rises slowly from 10% to 16% in the last four conditions. There is an interesting inverse relationship occurring in all conditions between the /f/ and the /t/ substitutions for the /h/.

3. The k/h substitutions do not appear above the 5% level until the LB is added, with a 6% incidence when the LB is added at threshold to the opposite ear, an 11% and 6% incidence when the LB is added at 20 dB SL to the opposite and the same ear, respectively, and with a 7% and 13% incidence when the LB is added at 40 dB SL to the opposite and the same ear, respectively.

4. The p/h substitutions occur between 6-10% of the time only when the LB is added at 20 dB and 40 dB SL.

5. There is a 14% incidence of No Responses when the HB alone is presented, and a 5-9% incidence in all other conditions except #5.

6. The substitutions which do not seem to form a pattern are the 5-7% incidence of s/h in conditions #2, 3, and 6, and m/h in the first condition.

General Discussion

The average articulation score for initial consonants was 40% when the high-pass band of 1020-2040 Hz was presented at threshold to one ear. Palva (1965) reported only a 1% discrimination score for the 240-480 Hz band, and pre-testing in the present study also indicated that this low-frequency band contains negligible consonant information when presented alone. However, the observed improvement in the HB-alone articulation score, as much as 21-22% points when the low-frequency band of 240-480 Hz was added at 20 dB SL, would seem to indicate that this low-frequency band adds to intelligibility. At favorable HB/LB ratios, the number of correct responses is essentially the same, regardless of whether the LB is added to the same or the opposite ear.

The LB does not have the same effect on each of the fourteen consonants in this study, since the maximum articulation score occurred when the LB was added to the HB at:

1. threshold, either to the same or opposite ear, for the /g/, /w/, and /r/

2. 20 dB SL, either to the same or opposite ear, for the /p/, /t/, /k/, /b/, /d/, /g/, /m/, /n/, /l/, /s/, and /h/, and

3. 40 dB SL to the opposite ear for the /f/

At unfavorable HB/LB ratios the LB seems to act as a "masker." At favorable HB/LB ratios, the perceptual errors that occur are similar for the same and opposite ear conditions, but at unfavorable HB/LB ratios, there is a higher incidence of random perceptual errors when the LB is added to the same ear. When the LB is added at 40 dB SL to the same ear the articulation score drops from 61%, (LB added at 20 dB SL to the same ear) to 38%, 2% points below the HB-alone level. The "masking" effect was not the same for each consonant, and the sounds seem to fall into two distinct categories which the writer has designated Group I and Group II. The scores for the sounds in Group I -- the voiceless plosives /p/, /t/, and /k/, the voiceless fricatives /f/ and /s/, the nasals /m/ and /n/ and liquid /l/ -- do not drop as much as the scores for the sounds in Group II -- the voiced plosives /b/, /d/, and /g/, the semi-vowel /w/, and the liquid /r/.

The difference between the two groups could be explained on the basis of frequency, intensity, and/or multiple cues. If frequency was the predominant factor, each set of stops, /b/ and /p/, /d/ and /t/, and /g/ and /k/, should be in the same group since they have the same formant structure, *i. e.*, the frequency areas in which the acoustic energy is concentrated are the same. If intensity was the predominant factor, /p/ and /f/ should be in Group II, and /r/ in Group I, since the /p/ and /f/ appear at the end of Fletcher's (1929) list rating the consonants in terms of their relative phonetic power, whereas the /r/ is at the top of the list. The writer suggests that since the consonants in Group I are characterized by multiple cues, such as bursts, friction, or nasal resonances, it is this redundancy in the speech signal which is the predominant factor enabling these sounds to maintain some identity despite the "masking" effect of the LB at high intensity levels relative to the HB.

Implications of the Study

Palva (1965) experimented with various low- and high-pass band combinations selected from those frequency areas that contain the essential formants, 400-1000 Hz and 1200-3000 Hz, and obtained an articulation score for each combination. It remains for Palva's study to be repeated in terms of a quantitative and qualitative analysis of individual sounds for each set of low- and high-frequency band combinations.

Based on this type of research, the writer suggests that, in the future, hearing-aid evaluations could consist of selecting those two frequency bands that provide maximum discrimination for each hearing-impaired individual. Specific frequency bands could then be combined either monaurally, or each ear could be fitted with a hearing aid which will amplify one of the two selected bands. For example, when a hearing aid is worn by those individuals who have considerably more hearing at 250 and 500 Hz than in the rest of the speech range, there is the possibility of the low frequencies masking the rest of the speech signal. If the frequency spectrum were to be split so that low-frequency amplification could be supplied to one ear, with amplification of all or part of the remaining speech range supplied to the other ear, this would allow for independent adjustment of the gain on each -- thereby providing amplification over a wider range of an individual's residual hearing. This type of hearing-aid fitting, in addition to possibly reducing the gain required, might also eliminate or reduce the effect of recruitment which often occurs at specific frequency areas, assuming that recruitment was specific to a frequency or frequencies not essential to discrimination.

CHAPTER IV

SUMMARY AND CONCLUSIONS

The words in the Fairbanks Rhyme test, spoken by a trained male, were filtered into two bands, each with approximately 60-70 dB/octave attenuation: (1) a low-pass band (LB) of 240-480 Hz, and (2) a high-pass band (HB) of 1020-2040 Hz. The HB, chosen because of its high information content, was presented at threshold (0 dB SL) to one ear, and the LB, chosen because of its high energy and low information content when used in isolation, was added at threshold, 20 dB SL, and 40 dB SL, either to the ear receiving the HB (same ear) or to the opposite ear. Each subject's thresholds for a 1500 and a 360-Hz pulsed tone were used to establish 0 dB SL for the high- and low-pass bands, respectively. Lists I, II, and IV, each consisting of fifty different consonant-vowel-consonant (CVC) words, were presented in randomized sequence to each of thirty-six normal-hearing college students, ages seventeen to twenty-four.

The average articulation score (the Fairbanks test is concerned only with consonants in the initial position) was 40% when the high-pass band of 1020-2040 Hz was presented alone to one ear. The consistent improvement in the articulation score which occurred when the low-pass band of 240-480 Hz was added (except at 40 dB SL to the same ear) would seem to indicate that this low-frequency band contains consonant information. When the LB is added to the HB at threshold and at 20 dB SL, there is a negligible difference between the average articulation scores when the LB is added either to the same ear or to the opposite ear. When added at threshold there is an increase of 15 and 13% points, respectively, over the HB-alone score, and when added at 20 dB SL there is an increase of 21 and 22% points, respectively, over the HB-alone score. However, when the LB is added at 40 dB SL to the same ear the articulation score drops from 61% (maximum score obtained) occurring when the LB is added at 20 dB SL to the same ear, to 38%; but when the LB is added at 40 dB SL to the opposite ear, the score drops only to 54%. The LB thus seems to act as a "masker" at unfavorable HB/LB ratios when supplied to the same ear. The 7% decrease from the maximum score obtained that occurs when the LB is added at 40 dB SL to the opposite ear could be the result of "central masking."

The "masking" effect, however, was not the same for each consonant. When the LB is added at 40 dB SL to the same ear, the sounds seem to fall into two distinct categories which the writer has designated Group I and Group II. The scores for the sounds in

Group I -- the voiceless plosives /p/, /t/, and /k/, the voiceless fricatives /f/ and /s/, the nasals /m/ and /n/ and liquid /l/ -- do not drop as much as the scores for the sounds in Group II -- the voiced plosives /b/, /d/, and /g/, the semi-vowel /w/, and the liquid /r/. It was suggested by the writer that the sounds in Group I are characterized by multiple cues, such as bursts, friction, or nasal resonances, and this redundancy may enable these sounds to maintain some identity despite the "masking" effect of the LB at high intensity levels relative to the HB.

The LB does not have the same effect on each of the fourteen consonants studied, and therefore a quantitative analysis was made of changes from the HB-alone score when the LB is added at the three sensation levels:

1. Voiceless plosives - /p/, /t/, and /k/

The patterns of the three are similar, with maximum improvement over the HB-alone score occurring when the LB is added at 20 dB SL, from 29% to 64% and 64% for the /p/, from 41% to 71% and 68% for the /t/, and from 48% to 70% and 64% for the /k/, for the opposite and same ear, respectively. When the LB is added at 40 dB SL to the opposite ear, the scores remain at or near their maximum levels, but when the LB is added to the same ear, the scores drop considerably, with the /t/ and /k/ scores returning to their HB-alone level.

2. Voiced plosives - /b/, /d/, and /g/

The patterns of these sounds are not similar. The maximum improvement over the HB-alone score occurs when the LB is added at 20 dB SL, from 38% to 49% and 41% for the /b/, and from 25% to 52% and 49% for the /d/, for the opposite and same ear, respectively. However, the maximum improvement over the HB-alone score for the /g/ occurs when the LB is added at threshold, from 38% to 72% and 70%, for the opposite and same ear, respectively. When the LB is added at 40 dB SL to the opposite ear, the scores for all three sounds return to the HB-alone level, but when the LB is added to the same ear, the scores fall substantially below their HB-alone level.

3. Nasals - /m/ and /n/ and Liquid /l/

The /l/ scores are similar to the nasal scores when the LB is added, almost duplicating the /m/ pattern. The maximum improvement over the HB-alone score for all three sounds occurs when the LB is added at 20 dB SL, from 47% to 85% and 84% for the /m/, from 32% to 69% and 71% for the /n/, and from 36% to 74% and 76% for the /l/, for the opposite and same ear, respectively. When the LB is added at 40 dB SL to the opposite ear, the /n/ score is not affected, whereas the /m/ and /l/ scores drop somewhat, but when the LB is added to the same ear, the scores for all three drop considerably, although never as low as their HB-alone level.

4. Voiceless fricatives - /f/ and /s/

The maximum improvement over the HB-alone score for the /s/ occurs when the LB is added at 20 dB SL, from 29% to 52% and 55%, for the opposite and same ear, respectively. When the LB is added at 40 dB SL to the opposite ear, the score remains at its maximum, but when the LB is added to the same ear, the score returns to its HB-alone level. The /f/ is the only sound in the study for which a maximum score is obtained when the LB is added at 40 dB SL to the opposite ear, from the HB-alone score of 35% to 56%, but the score falls to 44% when the LB is added to the same ear.

5. Semi-vowel - /w/ and Liquid - /r/

The HB-alone score for both sounds is high, particularly the /w/, as compared to the other sounds. Both sounds showed maximum increase over their HB-alone level when the LB is added at threshold, from 52% to 65% and 74% for the /r/, and from 85% to 93% and 93% for the /w/, for the opposite and same ear, respectively. The /w/ score drops slightly when the LB is added at 20 dB and 40 dB SL to the opposite ear, but when the LB is added at 40 dB SL to the same ear, there is a dramatic drop of 59% points from its HB-alone level. The /r/ score returns to its HB-alone level when the LB is added at 20 dB SL, and falls considerably below the HB-alone level when the LB is added at 40 dB SL. The /r/ is the only sound in the study for which the scores drop considerably when the LB is added at 20 dB and 40 dB SL, regardless of whether the LB is added to the same or opposite ear.

6. Glottal - /h/

The changes in the /h/ score are slight, with maximum improvement over the HB-alone score occurring when the LB is added to the same ear at threshold, from 35% to 51%, returning to its HB-alone level when added at 40 dB SL to the same ear. The scores when the LB is added to the opposite ear do not change enough from the HB-alone score to form a meaningful pattern.

A qualitative analysis of the perceptual confusions indicated that most of the substitutions which occurred with an incidence of 5% or more can be explained on the basis of the distinctive features of voicing, manner, and place-of-articulation. At favorable HB/LB ratios, the perceptual errors that occur are similar for the same and opposite ear conditions, but at unfavorable HB/LB ratios, there is a higher incidence of random perceptual errors when the LB is added to the same ear. The following will be limited to place and manner errors, since the t/d substitution was the only voicing error noted in the study.

1. Voiceless plosives - /p/, /t/, and /k/

Almost all the errors are place errors: t/k, p/k, t/p, k/p, k/t, and p/t. The substitution of the /h/ for the /p/ and /t/ cannot be explained on the basis of distinctive features.

2. Voiced plosives - /b/, /d/, and /g/

There are more varied substitutions for these sounds than for the voiceless plosives, with both place errors: d/g, b/g, g/d, and b/d, and manner errors: m/b, f/b, and dʒ/d, occurring. The incidence of b/m and m/b substitutions is sharply reduced when the LB is added at supra-threshold intensities. The substitution of /s/ for all three voiced plosives cannot be explained on the basis of distinctive features.

3. Nasals - /m/ and /n/

Almost all the errors are either place errors: m/n and n/m, or manner errors: b/m and f/m. The l/m substitutions are not explainable on the basis of distinctive features, but the /l/ does behave like a nasal when the LB is added. Another substitution which cannot be explained in terms of distinctive features is /h/ for /m/ and /n/.

4. Voiceless fricatives - /f/ and /s/

The expected place errors of f/s and s/f occur. The b/f and m/f manner errors also occur, and the t/f substitutions may be related to the t/s manner error. The substitution of the nasals, particularly the /m/, for both the /f/ and /s/ almost disappears when the LB is added at supra-threshold intensities. Again, the /h/ is substituted for the /f/ and /s/, a perceptual error which cannot be explained on the basis of distinctive features.

5. Semi-Vowel - /w/ and Liquids - /r/ and /l/

Although the behavior of the /l/ when the LB is added resembles the nasals, the perceptual confusions belong more to the semi-vowel - liquid category. Except for the m/l substitutions, all the errors are place errors: w/l, r/l, r/w, l/r, and w/r.

6. Glottal - /h/

The substitutions include the voiceless plosives /p/, /t/, and /k/, and the voiceless fricatives /f/ and /s/.

Two further interesting phenomena were noted in the present study: (1) the /l/ seems to have properties of both the nasal and the semi-vowel - liquid categories, and (2) the articulation scores were higher for the right then the left ear, a finding in accord with the work of Milner, Kimura, and Shankweiler.

A possible practical application of this study concerns those hearing-impaired individuals who have considerably more hearing in the low frequencies than in the high frequencies. In a traditional hearing aid, the gain is not substantially different at each frequency, and unless the low frequencies are eliminated, there is the possibility of the low frequencies "masking" the high frequencies. However, if low-frequency amplification were to be supplied to one ear, with high-frequency amplification to the other ear, the intensity level of the two could be adjusted independently. The results of this study indicate that the low frequencies do contain information that contributes to intelligibility, and thus this type of binaural hearing aid fitting might improve an individual's discrimination by permitting useful amplification over a wider frequency range.

APPENDIX

Table 10. Fairbanks Rhyme test - lists I, II, and IV.

LIST I	LIST II	LIST IV
	Column A	
hot	got	pot
pay	may	way
top	hop	mop
peel	reel	heel
wake	take	cake
law	saw	paw
vile	mile	tile
neat	seat	heat
look	cook	took
fill	kill	till
	Column B	
tire	hire	fire
male	tale	pale
sent	rent	bent
moon	noon	boon
kick	sick	tick
same	fame	came
wide	tide	ride
rip	dip	hip
sore	bore	more
bang	hang	gang
	Column C	
men	den	pen
park	bark	mark
coil	foil	soil
big	wig	fig
rage	cage	sage
cast	past	last
gain	pain	rain
nest	west	best
gun	nun	sun
heel	deal	zeal

Table 10. Continued

LIST I	LIST II	LIST IV
	Column D	
sin	win	din
bust	just	rust
fine	mine	nine
mink	link	wink
sold	told	cold
hit	sit	fit
led	bed	wed
tend	send	lend
rid	bid	did
back	lack	jack
	Column E	
tail	sail	nail
fight	light	might
torn	worn	horn
rod	God	sod
dock	mock	lock
bump	pump	dump
date	rate	late
well	fell	bell
set	let	yet
luck	tuck	suck

Table 11. Age, sex and hearing levels* for each subject.

SEX	AGE	LB(TH)	HB(TH)
Subjects Who Received HB in Right Ear			
F	17	15 dB	5 dB
M	21	15	5
F	21	15	5
M	23	10	5
F	24	15	5
F	18	15	5
F	19	10	5
F	20	10	5
F	21	5	0
F	18	5	0
M	20	15	10
M	17	10	5
M	18	15	0
F	20	15	0
F	20	15	5
F	19	10	0
F	22	10	0
F	19	15	0
Subjects Who Received HB in Left Ear			
F	19	15	5
F	17	5	5
F	18	10	5
F	20	15	5
F	21	15	5
F	18	20	10
F	20	10	5
F	21	10	5
F	17	5	0
F	20	5	5
M	19	15	10
M	23	10	5
M	17	10	0
M	21	10	0
M	20	10	5
F	18	5	0
F	19	10	0
F	19	10	0

*The thresholds presented are audiometer dial readings.

Table 12. Summary of test order presentations for each subject.

KEY

High-frequency band	HB
Low-frequency band	LB
Threshold	(TH)
Right ear	R
Left ear	L

I. RIGHT EAR

A. 3 lists, HB-R(TH)
3 lists, HB-R(TH); LB-L: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)
3 lists, HB-R(TH); LB-R: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)

LISTS and INTENSITIES for the LB

SUBJECT			
	I	II	IV
#1	(TH)	20 dB SL	40 dB SL
#2	20 dB SL	40 dB SL	(TH)
#3	40 dB SL	(TH)	20 dB SL
	II	I	IV
#4	(TH)	20 dB SL	40 dB SL
#5	20 dB SL	40 dB SL	(TH)
#6	40 dB SL	(TH)	20 dB SL
	IV	I	II
#7	(TH)	20 dB SL	40 dB SL
#8	20 dB SL	40 dB SL	(TH)
#9	40 dB SL	(TH)	20 dB SL

Table 12. Continued

B.
3 lists, HB-R(TH)
3 lists, HB-R(TH); LB-R: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)
3 lists, HB-R(TH); LB-L: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)

LISTS and INTENSITIES for the LB

SUBJECT			
	I	IV	II
#10	(TH)	20 dB SL	40 dB SL
#11	20 dB SL	40 dB SL	(TH)
#12	40 dB SL	(TH)	20 dB SL
	II	IV	I
#13	(TH)	20 dB SL	40 dB SL
#14	20 dB SL	40 dB SL	(TH)
#15	40 dB SL	(TH)	20 dB SL
	IV	II	I
#16	(TH)	20 dB SL	40 dB SL
#17	20 dB SL	40 dB SL	(TH)
#18	40 dB SL	(TH)	20 dB SL

Table 12. Continued

II. LEFT EAR

A. 3 lists, HB-L(TH)
3 lists, HB-L(TH); LB-R: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)
3 lists, HB-L(TH); LB-L: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)

LISTS and INTENSITIES for the LB

SUBJECT			
	I	II	IV
#19	(TH)	20 dB SL	40 dB SL
#20	20 dB SL	40 dB SL	(TH)
#21	40 dB SL	(TH)	20 dB SL
	II	I	IV
#22	(TH)	20 dB SL	40 dB SL
#23	20 dB SL	40 dB SL	(TH)
#24	40 dB SL	(TH)	20 dB SL
	IV	I	II
#25	(TH)	20 dB SL	40 dB SL
#26	20 dB SL	40 dB SL	(TH)
#27	40 dB SL	(TH)	20 dB SL

Table 12. Continued

B. 3 lists, HB-L(TH)
3 lists, HB-L(TH);LB-L: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)
3 lists, HB-L(TH);LB-R: one list-(TH), one list-(20 dB SL), one list-(40 dB SL)

LIST and INTENSITIES for the LB

SUBJECT			
	I	IV	II
#28	(TH)	20 dB SL	40 dB SL
#29	20 dB SL	40 dB SL	(TH)
#30	40 dB SL	(TH)	20 dB SL
	II	IV	I
#31	(TH)	20 dB SL	40 dB SL
#32	20 dB SL	40 dB SL	(TH)
#33	40 dB SL	(TH)	20 dB SL
	IV	II	I
#34	(TH)	20 dB SL	40 dB SL
#35	20 dB SL	40 dB SL	(TH)
#36	40 dB SL	(TH)	20 dB SL

Table 13. Scores for subjects receiving HB in right ear.

	ONE EAR	OPP. EARS	SAME EAR	OPP. EARS	SAME EAR	OPP. EARS	SAME EAR
SUBJECT	HB(TH)	HB(TH) LB(TH)	HB(TH) LB(TH)	HB(TH) LB(20 dB SL)	HB(TH) LB(20 dB SL)	HB(TH) LB(40 dB SL)	HB(TH) LB(40 dB SL)
# 1	60%	68%	72%	72%	72%	58%	30%
# 2	41	36	56	66	72	60	42
# 3	56	66	72	74	74	74	44
# 4	34	48	40	62	52	44	28
# 5	45	60	64	68	54	68	46
# 6	55	80	70	72	68	74	60
# 7	41	34	46	64	56	58	36
# 8	45	62	68	58	58	74	38
# 9	39	60	54	56	54	62	36
#10	47	68	60	80	68	54	36
#11	61	56	58	76	74	72	50
#12	38	30	34	46	46	50	36
#13	43	46	62	60	54	70	42
#14	34	56	68	56	50	54	34
#15	26	46	44	58	54	50	42
#16	40	72	58	68	64	68	54
#17	47	64	66	48	54	46	28
#18	48	56	70	54	70	54	38
MEAN ARTIC. SCORES	44.4%	56.0%	59.0%	63.2%	60.8%	60.6%	40.0%

Table 14. Scores for subjects receiving HB in left ear.

SUBJECT	ONE EAR HB(TH)	OPP.EARS HB(TH) LB(TH)	SAME EAR HB(TH) LB(TH)	OPP.EARS HB(TH) LB(20 dB SL)	SAME EAR HB(TH) LB(20 dB SL)	OPP.EARS HB(TH) LB(40 dB SL)	SAME EAR HB(TH) LB(40 dB SL)
#19	36%	38%	60%	62%	64%	50%	24%
#20	41	36	40	68	68	44	34
#21	41	58	56	58	62	60	46
#22	35	38	50	62	56	44	36
#23	27	42	34	56	62	48	42
#24	24	30	40	36	36	18	20
#25	54	64	72	72	80	66	44
#26	25	50	50	58	60	40	52
#27	31	42	48	58	58	40	24
#28	34	54	50	46	54	38	20
#29	33	54	60	70	64	48	32
#30	37	48	40	64	66	62	44
#31	37	50	48	52	60	60	40
#32	36	70	54	62	76	48	34
#33	28	52	62	68	60	28	24
#34	45	46	52	70	56	62	52
#35	48	70	54	68	66	66	40
#36	34	42	36	60	54	36	34
MEAN ARTIC. SCORES	35.9%	49.1%	50.3%	60.6%	61.2%	47.7%	35.7%

LIST OF REFERENCES

Black and Hast, "Speech Reception with Altering Signal," J. Speech Hear. Res., Vol. 5, No. 21, Mar. 1962, pp. 70-75.

Bocca, "Binaural Hearing: Another Approach," Laryngoscope, Vol. 65, 1955, pp. 1164-1171.

Bocca, Calearo, and Cassinari, "A New Method for Testing Hearing in Temporal Lobe Tumours," Acta Oto-Laryngol., Vol. 44, 1954, pp. 219-221.

Bocca, Calearo, Cassinari, and Migliavacca, "Testing 'Cortical' Hearing in Temporal Lobe Tumours," Acta Oto-Laryngol., Vol. 45, 1955, pp. 289-303.

Broadbent and Ladefoged, "On the Fusion of Sounds Reaching Different Sense Organs," J. Acoust. Soc. Amer., Vol. 29, June 1957, pp. 708-710.

Calearo, "Binaural Summation in Lesions of the Temporal Lobe," Acta Oto-Laryngol., Vol. 47, 1957, pp. 392-395.

Cooper, Delattre, Liberman, Borst, and Gerstman, "Some Experiments on the Perception of Synthetic Speech Sounds," J. Acoust. Soc. Amer., Vol. 24, No. 6, Nov. 1952, pp. 597-606.

David, Guttman, and van Bergeijk, "On the Mechanism of Binaural Fusion," (Letters to the Editor), J. Acoust. Soc. Amer., Vol. 30, No. 8, Aug. 1958, pp. 801-802.

Deatherage, "Examination of Binaural Interaction," J. Acoust. Soc. Amer., Vol. 39, 1966, pp. 232-249.

Delattre, Liberman, and Cooper, "Acoustic Loci and Transitional Cues for Consonants," J. Acoust. Soc. Amer., Vol. 27, No. 4, July, 1955, pp. 769-773.

Egan, "Articulation Testing Methods," Laryngoscope, Vol. 58, 1948, pp. 955-991.

Egan, Carterette and Thwing, "Some Factors Affecting Multi-Channel Listening," J. Acoust. Soc. Amer., Vol. 26, Sept. 1954, pp. 774-782.

Fairbanks, "Test of Phonemic Differentiation: The Rhyme Test," J. Acoust. Soc. Amer., Vol. 30, No. 7, July, 1958, pp. 596-599.

Fletcher, Speech and Hearing in Communication, O. Van Nostrand Co., Inc., 1929 and 1953 editions.

French and Steinberg, "Factors Governing the Intelligibility of Speech Sounds," J. Acoust. Soc. Amer., Vol. 19, No. 1, Jan., 1947, pp. 90-119

Harris, K., "Cues for the Discrimination of American English Fricatives in Spoken Syllables," Lang. Speech, Vol. 1, Part 1, Jan.-Mar. 1958, pp. 1-7.

Hayashi, Ohta, and Morimoto, "Binaural Fusion Test: A Diagnostic Approach to the Central Auditory Disorders," Int. Audiol., Vol V, No. 2, June, 1966, pp. 133-135.

Hirsh and Bowman, "Masking of Speech by Bands of Noise," J. Acoust. Soc. Amer., Vol. 25, No. 6, Nov. 1953, pp. 1175-1180.

House, Williams, Hecker, and Kryter, "Articulation-Testing Methods: Consonantal Differentiation with a Closed-Response Set," J. Acoust. Soc. Amer., Vol. 37, No. 1, Jan., 1965, pp. 158-166.

Huizing, "Triplet Speech Audiometry," Int. Audiol., Vol. 1, No. 2, Sept., 1962, pp. 189-190.

Huizing, Kruisinga, and Taselaar, "Triplet Audiometry: An Analysis of Band Discrimination in Speech Reception," Acta Oto-Laryngol., Vol. 51, 1960, pp. 256-259.

Huizing and Taselaar, "Experiments on Binaural Hearing, *Acta Oto-Laryngol.*, Vol. 53, Mar.-Apr., 1961, pp. 151-153.

Jerger, "Observations on Auditory Behavior in Lesions of the Central Auditory Pathways," *Arch. Otolaryngol.*, Vol. 71, May, 1960, pp. 65/797-74/806.

Kimura, "Some Effects of Temporal-Lobe Damage on Auditory Perception," *Canad. J. Psychol.*, Vol. 15, 1961^{a}, pp. 156-165.

Kimura, "Cerebral Dominance and the Perception of Verbal Stimuli," *Canad. J. Psychol.*, Vol. 15, 1961^{b}, pp. 166-171.

Kimura, "Left-Right Differences in the Perception of Melodies," *Quart. J. Exp. Psychol.*, Vol. 15-16, 1963-64, pp. 355-358.

Kryter, "Speech Bandwidth Compression through Spectrum Selection," *J. Acoust. Soc. Amer.*, Vol. 32, May, 1960, pp. 547-556.

Lehiste and Peterson, "Linguistic Considerations in the Study of Speech Intelligibility," *J. Acoust. Soc. Amer.*, Vol. 31, 1959, pp. 280-286.

Liberman, Delattre, and Cooper, "Some Cues for the Distinction Between Voiced and Voiceless Stops in Initial Position," *Lang. Speech*, Vol. 1, Part 3, July-Sept., 1958, pp. 153-167.

Licklider, "The Influence of Interaural Phase Relations Upon the Masking of Speech by White Noise," *J. Acoust. Soc. Amer.*, Vol. 20, Mar., 1948, pp. 150-159.

Linden, "Distorted Speech and Binaural Speech Resynthesis Tests," *Acta Oto-Laryngol.*, Vol. 58, 1964, pp. 32-47.

Lisker, "Minimal Cues for Separating /w, r,   ,   / in Intervocalic Position," *Word*, Vol. 13, No. 2, Aug., 1957, pp. 256-267.

Malécot, "Acoustic Cues for Nasal Consonants An Experimental Study Involving a Tape-Splicing Technique," *Language*, Vol. 32, No. 2, Apr., 1956, pp. 274-284.

Matzker, "The Binaural Test," Int. Audiol., Vol. 1, No. 2, Sept., 1962, pp. 209-211.

Miller, "The Masking of Speech," Psychol. Bull., Vol. 44, No. 2, Mar., 1947, pp. 105-129.

Miller and Nicely, "An Analysis of Perceptual Confusions Among Some English Consonants," J. Acoust. Soc. Amer., Vol. 27, No. 2, Mar., 1954, pp. 338-352.

Milner, "Laterality Effects in Audition," Interhemispheric Relations and Cerebral Dominance, Mountcastle, Baltimore, John Hopkins, 1962, pp. 177-195.

Ohta, Hayashi, and Morimoto, "Differential Diagnosis of Retrocochlear Deafness: Binaural Fusion and Binaural Separation Test," Int. Audiol., Vol. IV, No. 1, July, 1967, pp. 58-62.

Palva, "Filtered Speech Audiometry, I Basic Studies with Finnish Speech Towards the Creation of a Method for the Diagnosis of Central Hearing Disorders," Acta Oto-Laryngol. Supplementum 210, 1965, pp. 1-86.

Pollack, "Effects of High Pass and Low Pass Filtering on the Intelligibility of Speech in Noise," J. Acoust. Soc. Amer., Vol. 20, No. 3, 1948, pp. 259-266.

Shankweiler and Studdert-Kennedy, "Identification of Consonants and Vowels Presented to Left and Right Ears," Quart. J. Exp. Psychol., Vol. XIX, Part 1, Feb., 1967, pp. 59-62.

Speaks, "Intelligibility of Filtered Speech Synthetic Sentences," J. Speech Hear. Res., Vol. 10, No. 2, June, 1967, pp. 289-298.

Steinberg, "Effects of Distortion Upon the Recognition of Speech Sounds," J. Acoust. Soc. Amer., Oct., 1929, pp. 121-137.

The Young Deaf Child: Identification and Management, Acta Oto-Laryngol. Supplementum 206, 1965.

Tillman and Carhart, "Monaural Versus Binaural Discrimination for Filtered CNC Materials, The Normal Auditory Mechanism," USAF School of Aerospace Medicine, Aerospace Medical Division, Brooks Air Force Base, Texas, SAM-TR-65-79, Dec., 1965.

Webster and Klumpp, "Articulation Index and Average Curve-Fitting Methods of Predicting Speech Interference," J. Acoust. Soc. Amer., Vol. 35, No. 9, Sept. 1963, pp. 1339-1344.

Zwislocki, "Acoustic Attenuation between the Ears," J. Acoust. Soc. Amer., Vol. 25, No. 4, 1953, pp. 752-759.