1970

Auditory and audio-visual reception of words in noise by observers with normal and impaired hearing

Norman Phillip Erber

WASHINGTON UNIVERSITY

Department of Speech and Hearing

Dissertation Committee:

James D. Miller, Chairman
Ira J. Hirsh
Arthur F. Niemoeller
S. Richard Silverman

AUDITORY AND AUDIO-VISUAL RECEPTION OF WORDS IN NOISE BY OBSERVERS WITH NORMAL AND IMPAIRED HEARING

by

Norman Phillip Erber

A dissertation
presented to the Graduate School
of Arts and Sciences of Washington University
in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

January, 1970

Saint Louis, Missouri

Dept. of Speech & Hearing

764-13

ACKNOWLEDGMENTS

This research was supported in part by Program Project Grant NBO 3856 from the National Institute of Neurological Diseases and Stroke and in part by PHS pre-doctoral medical audiology training grant NBO 5344. Appreciation is expressed to the research staff of Central Institute for the Deaf, who provided helpful suggestions throughout the investigation.

TABLE OF CONTENTS

Page

Appendix

LIST OF FIGURES

LIST OF TABLES

"There may be more to lipreading...than meets the eye."

H.A. Leeper (1969),
after A.S. House (ca 1967)

GENERAL INTRODUCTION

STATEMENT OF PROBLEM

Many investigations have shown that provision of amplified acoustic cues to hearing-impaired children who watch the face of a talker results in higher word-recognition scores than does lipreading alone. These data were obtained mainly in quiet laboratories, although typical learning environments are noisy.

Adults with normal hearing are limited in their abilities to extract acoustic speech information from background noise. Children with impaired auditory mechanisms may be even more restricted in this capacity. The goal of this research was to determine some of the noise conditions under which amplified acoustic cues for speech can be beneficial to hearing-impaired children.

The findings of three studies of speech reception in noise are reported. The first experiment compares listening alone with listening and lipreading combined for recognition of words in noise by normal-hearing adults. The second study compares detection of words in noise by normal-hearing adults and profoundly deaf children. The third investigation compares normal-hearing, severely hearing-impaired, and profoundly deaf children on tasks requiring auditory detection, auditory recognition, and combined audio-visual recognition of words in noise.

REVIEW OF PRIOR RESEARCH

AUDITORY RECEPTION OF SPEECH IN NOISE

Students of human communication have recognized for a long time that acoustic noise can seriously degrade the intelligibility of speech. Many investigators have attempted to quantify the masking effects of noise on speech signals. In the basic experiment, speech material is presented acoustically in noise to listeners who are required to respond to some aspect of the speech stimulus. The decibel ratio between speech power and overall noise power is specified (dB S/N), and performance usually is described with respect to the S/N conditions under which the signals were presented. Most published data indicate rapid decrements in the intelligibility of speech as S/N ratio is diminished.

A paper by Hawkins and Stevens (1950) was one of the first to specify masked thresholds for speech. They presented wide-band noise and continuous speech over earphones to listeners, who adjusted an attenuator in the speech circuit. At noise sensation levels between 40-90 dB, the threshold of detection for speech (50%) was noted at -17 dB S/N, and the threshold of intelligibility (50%) was reached at -8 dB S/N.

Miller, Heise, and Lichten (1951) investigated the influence of context on the intelligibility of simple speech stimuli in noise. Speech level was held constant at 90 dB SPL, while wide-band noise was varied in intensity. For monosyllabic words, the threshold of intelligibility was systematically lowered about 18 dB S/N by decreasing the size of the message set from 1000 words (+4 dB S/N) through 256 words (-4 dB S/N) ....to 2 words (-14 dB S/N). The intelligibility of words spoken alone was compared with that of the same words presented in the context of a sentence. Placing a word in a sentence reduced its masked threshold for intelligibility about 6 dB S/N.

Hirsh and Bowman (1953) measured the masking effect of 12 different noise bands: wide-band and 11 others with center frequencies separated by equal pitch intervals. White noise was a more effective masker for spondaic words than was any narrower band tested, giving a threshold S/N ratio of about -15 dB. The most effective masking band was 670-1000 Hz (-25 dB S/N), with less masking resulting at similar levels for higher or lower frequency bands.

Hirsh, Reynolds, and Joseph (1954) obtained intelligibility scores in wide-band noise for nonsense syllables and for multisyllabic words. The authors manipulated not only the S/N ratio but also the cutoff frequency of low-pass and high-pass filters in the speech circuit. Eliminating the frequencies in speech either above or below 1600 Hz did not impair greatly the intelligibility of speech materials. Word intelligibility was related to the number of syllables in the word, and monosyllabic words were more intelligible than were nonsense syllables under similar masking conditions.

Miller and Nicely (1955) analyzed the consonant substitutions made by observers listening in a severe acoustic environment. Sixteen English consonants (in combination with /a/) were spoken over communication systems with frequency distortion and with wide-band noise. Listeners were forced to guess at every speech sound, and a count was made of all the errors that resulted when the spoken sound was confused with others. Five articulatory dimensions were used to describe the 16 consonants: voicing, nasality, affrication, duration, and place of articulation. Place cues were affected most by low-pass and noisy systems, while voicing and nasality cues were quite resistant to these unfavorable conditions.

These initial experiments served to define the major variables that limit transmission of speech information over a noisy communication channel. Shortly thereafter, a series of papers by J.M. Pickett and I. Pollack appeared, in which the authors described their attempts to extend these limits.

Pickett (1956) studied the influence of vocal effort on the intelligibility of monosyllabic words presented in wide-band noise. The range from 36-90 dB SPL at 1 m was examined. For constant S/N conditions, the range from 55-78 dB resulted in less than 5% change in intelligibility, while extremes of effort severely reduced communication effectiveness. The gains in S/N ratio achieved with voice levels above 80 dB were cancelled by the deterioration in intelligibility associated with the vocal distortions of shouting.

Pollack and Pickett (1958) investigated the reception of speech throughout a range of intense noise conditions. For constant S/N ratio, deterioration of word intelligibility occurred at noise levels above about 100 dB SPL. Word intelligibility in the quiet was found to be independent of speech level (99% correct from 80-125 dB SPL). Control tests demonstrated that the source of distortion was not the equipment and suggested that overloading of the auditory system by high noise levels was the determining factor.

Pollack (1958) speculated that voice communication in noise might be improved by proper design of communication procedures. Several techniques were compared under extremely unfavorable noise conditions (-12 to -22 dB S/N), and they were evaluated in terms of the channel time used in communicating at a given level of proficiency. A network procedure (in which successive binary selections were made) was compared with a repetition procedure (in which each word was repeated until received correctly). In terms of the joint time-accuracy criterion, the results clearly favored the network procedure, whose relative superiority increased as S/N deteriorated.

Pickett and Pollack (1958) even tested a small megaphone for use in direct voice communication against backgrounds of noise. Word intelligibility was determined for talkers speaking with and without the megaphone over a range of field voice levels (70-100 dB at 1 m). Nearly the entire overall acoustic gain produced by the megaphone (9.5 dB) was realized in improved intelligibility in noise. The acoustical gain of the megaphone permitted a lower voice level and delayed the deterioration in intelligibility associated with distortion due to shouting.

AUDIO-VISUAL RECEPTION OF SPEECH IN NOISE

In general, the preceding studies demonstrated that the human auditory system has certain limitations for the reception of acoustic cues for speech in a noise background. Further, the deficiencies of the ear can be overcome only partly by communication procedures that depend upon manipulation of stimulus parameters in the auditory mode. Recognizing these limitations, several investigators have examined the potential of visual cues (i.e., lipreading) for enhancement of oral speech communication in noise.

O'Neill (1954) determined the relative intelligibility of several units of speech under 5 conditions of S/N ratio and through 2 modes of reception: auditory only and combined audio-visual. For normal-hearing adults, the combined method was superior to auditory reception alone under all S/N conditions tested. Vision contributed most to the identification of consonants, less to the recognition of vowels and single words, and least to the comprehension of phrases.

Sumby and Pollack (1954) presented spondaic words to normal-hearing adults under several S/N conditions and via both auditory and audio-visual channels. Their data indicated that visual cues are relied upon to an increasing extent as the S/N ratio becomes poorer. Subjects' performance improved greatly when they were allowed to supplement their auditory input with visual cues.

Neely (1956) also investigated the contribution of lipreading to face-to-face communication in noise, examining the influence of angle and distance between observer and talker. The addition of visual cues improved speech intelligibility in noise. A 90° observation angle resulted in lower scores than did angles of 45° or 0°, but the viewing distance did not significantly influence audio-visual performance within the 3-9 ft limits tested.

Each of these studies clearly illustrated the advantages of providing visual cues to observers in environments that are unfavorable for auditory reception of speech. The superiority of a combined audio-visual method for speech communication in noise is most apparent under severe S/N ratio conditions in which acoustic cues for speech are effectively masked, and the ear is rendered useless as a channel for processing information.

STUDY I: AUDITORY AND AUDIO-VISUAL RECOGNITION OF SPONDAIC WORDS IN WIDE-BAND NOISE BY ADULTS WITH NORMAL HEARING

INTRODUCTION

When one observes a talker in quiet surroundings, he receives auditory and visual cues for speech that are closely related (i.e., particular oral configurations emit expected speech sounds). Under these conditions, audition and vision provide redundant information, and visual cues are superfluous for most individuals with normal hearing. When the acoustic S/N ratio is less than ideal, weaker phonemes are masked and they become inaudible. In this situation, the observer must rely relatively more on visual cues for satisfactory message reception. When one must decode speech under extremely poor S/N conditions where no meaningful auditory cues are available, lipreading can be the only source of information.

Papers by O'Neill (1954) and Sumby and Pollack (1954) have provided data to support these assumptions. Their findings indicate that audio-visual (A-V) recognition of single-word speech stimuli is more resistant to noise than is recognition under audition-only (A-only) conditions. That is, the observer always receives more speech information in noise when he watches the talker than when he does not. Each author also suggested that decrements in S/N ratio increase the visual contribution to A-V speech reception in noise. In each case, this trend was determined by comparing

A-V scores with A-only performance at successively poorer S/N ratios.

However, the results of these investigations did not clearly indicate minimum S/N ratios at which auditory cues might benefit an observer in a noisy environment. Also, neither study reported estimates of the variability in response between subjects. The present experiment repeated a portion of the Sumby and Pollack (1954) study to clarify these points and to gather additional data on modes of speech reception in acoustic noise.

METHOD

A test vocabulary of 250 words with spondaic stress pattern was selected for its familiarity to subjects and its lack of syllable repetition (i.e., 500 different syllables appeared in the sample). The vocabulary was divided arbitrarily into ten 25-word lists. At each of 10 testing sessions, each list was presented by the talker under a different condition. These conditions were: A-V at -36, -30, -24, -18, -12, and -6 dB S/N; A-only at -18, -12, -6, and 0 dB S/N. Words within lists were scrambled between sessions, and the 10 conditions of observation were counterbalanced in order to distribute possible learning effects.

A trained female talker monitored a VU meter as she spoke the words into a condenser microphone. Her speech was amplified, mixed electrically with white noise (low-pass filtered at 6000 Hz), amplified again, and then presented binaurally to observers via TDH-49 earphones in MX-41/AR cushions. An attenuator in the speech channel controlled the S/N ratio presented to the subjects. Speech level was defined as the average of the peaks indicated on a VU meter by the stimulus words. Overall noise level was maintained at 90 dB SPL in the ear during all conditions. An opaque screen was placed between the talker and the subjects during A-only conditions.

No carrier phrase was used by the talker. Instead, a warning light flashed 1 sec before presentation of each word. Subjects were allowed 10 sec in which to respond.

Five young adults (3 female, 2 male) of age range 20-23 years served as subjects. Each had normal hearing and corrected visual acuity. None was an experienced lipreader. They sat in 2 rows, facing the talker from a distance of 6-8 ft. The test room was well-lighted, quiet, and non-reverberant.

Each subject was supplied with an alphabetical list of the test vocabulary (Appendix 1), to which he was permitted to refer at any time. Subjects wrote their responses on answer sheets, and they were required to write only words contained in the 250-word list. They did not score their own papers and received no knowledge of results until conclusion of the experiment.

A response was scored "correct" if it matched phoneme-for-phoneme with the stimulus word that was presented. Misspelled words were not considered errors if they were phonemically correct. The raw data were converted into per cent words correctly recognized under each condition of observation. Group means and standard deviations are displayed graphically in Figure 1 and numerically in Appendix 2.

RESULTS AND DISCUSSION

Several comparisons can be made between A-V and A-only recognition of spondaic words in acoustic noise. Whereas A-only yields a sharply falling curve that descends from 80% to 20% within a range of 11 dB S/N, combined cues result in a function that also is high for favorable S/N ratios but does not drop abruptly under poorer acoustic conditions. Instead, A-V performance becomes asymptotic at about 50% for low signal levels, reflecting the ease with which this particular talker and vocabulary can be lipread in noise. These 2 curves approximate those obtained by Sumby and Pollack (1954) under similar conditions, if allowances are made for differences in talker and in noise bandwidth.

As S/N ratio is made less favorable, the variability between A-only scores increases slightly for moderate speech levels but diminishes again under more severe S/N conditions. However, the variability between observers' A-V responses steadily increases to asymptote (s.d. = 15%) from an initial magnitude that is comparable to the variability in A-only scores (s.d. = 6%). It is suggested that the greater variability between observers' A-V scores at poorer S/N ratios is

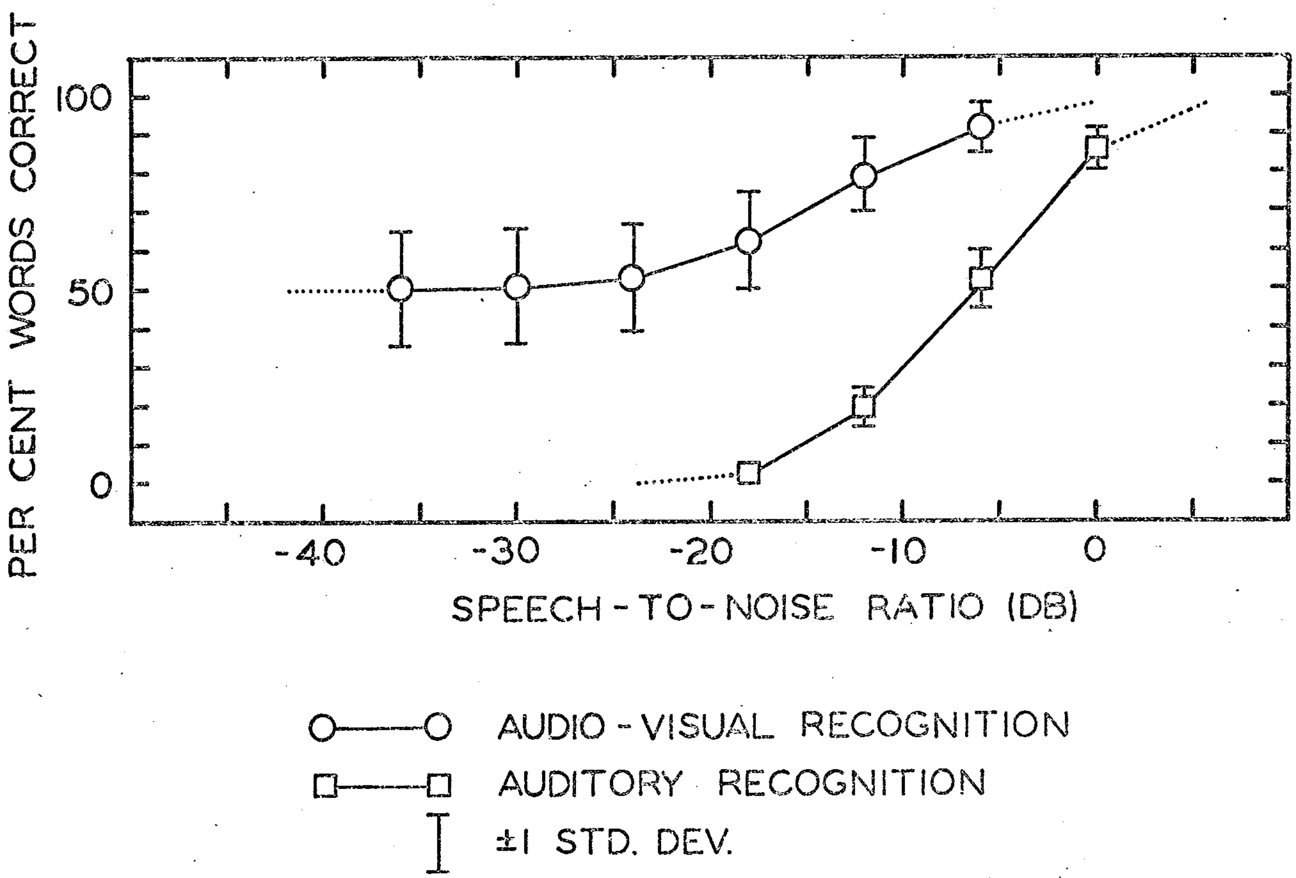


Figure 1. Recognition of 250 spondaic words in wide-band noise by 5 normal-hearing adults.

a function of differences in lipreading skill among untrained subjects. If that is true, then this result would support the notion that observers rely more on visual information as the acoustic speech signal is degraded.

The data indicate that acoustic S/N ratios below about -18 dB S/N are insufficient for A-only recognition above the chance level. Yet A-V recognition of the same speech material begins to rise above its asymptotic level (i.e., lipreading in noise) at about -24 dB S/N, and it continues to improve systematically under more favorable S/N conditions. At -18 dB S/N, for example, listening alone yields near-chance performance (2% correct), while the same S/N ratio results in a 12% improvement above the A-V asymptote when audition supplements vision.

It is suggested that at very low levels of the acoustic speech signal, the observer exploits his auditory channel for gross temporal information as well as for minimal discriminative cues. That is, the just-detectable vowel pulses may indicate to the observer the syllabic pattern of the speech stimulus and cue him to search the visual field (or short-term memory) for consonant information immediately following (or preceding) those bursts.

STUDY II: AUDITORY DETECTION OF SPONDAIC WORDS IN WIDE-BAND NOISE BY ADULTS WITH NORMAL HEARING AND BY CHILDREN WITH PROFOUND HEARING LOSSES

INTRODUCTION

The data presented by several authors (Miller, Heise, and Lichten, 1951; O'Neill, 1954; Sumby and Pollack, 1954; Hirsh, Reynolds, and Joseph, 1954; Miller and Nicely, 1955) indicate that for a wide-band random-noise masker, acoustic S/N ratios below about -15 to -20 dB can provide very little discriminative speech information to listeners with normal hearing. That is, poorer S/N ratios result in recognition performance near the chance level.

Study I also showed that acoustic S/N ratios below about -18 dB were insufficient for auditory recognition of spondaic words. Yet S/N ratios this low systematically improved observers' combined audio-visual recognition of the same words (Figure 1). It was concluded that observers with normal hearing can supplement lipreading with acoustic cues for speech that contain by themselves very little discriminative information. Similarly, Johnson (1963) demonstrated improvements in lipreading by normal-hearing adults when acoustic cues were amplified and delivered to their forearms by a tactile vibrator.

Gault (1928), Numbers and Hudgins (1948), Van Uden (1960), and Pickett (1963) all showed how the lipreading performance of deaf children can be improved also if amplified speech is made available to them via a hearing aid or a tactile vibrator. They demonstrated these improvements even for children whose recognition of simple words through their ears or skin alone was very poor.

The exact nature of the speech information available at low acoustic levels is unknown. Study I suggested that detection of vowel-peak energy in the stimulus words might be the minimal acoustic cue that improves lipreading under white-noise masking conditions. Perception of vowel pulses might assist the lipreader by indicating to him the syllabic pattern of the speech stimulus, thereby cuing him to search the talker's face (or his own short-term memory) for consonant information following (or preceding) those bursts.

This study determined whether improvement in audio-visual performance at low S/N ratios can be related to the detectability of similar speech stimuli presented under similar conditions of noise. If detection of burst patterns is indeed the source of lipreading improvement at low S/N ratios, then normal listeners' auditory detection of words in noise should begin to improve above chance at a similar point on the S/N dimension (i.e., about -22 dB S/N).

In general, profoundly deaf children are unable to recognize simple words by ear alone. The temporal and amplitude features of speech may be the only discriminative acoustic cues available to them, and these patterns must be clearly detectable to be used effectively (i.e., in combination with lipreading). The masking effects of noise on speech have never been determined quantitatively for profoundly deaf subjects. Because of deficiencies in their sensory mechanisms, deaf observers may require greater S/N ratios than do normals for comparable detection performance. For these reasons, this study also obtained preliminary data on S/N ratios required by deaf children for the detection of words presented in noise.

SUBJECTS

Ten subjects participated in the study. Half of these were young adults (1 female, 4 male) of age range 18-23 years. All demonstrated normal hearing as indicated by less than 20 dB HL at each of the audiometric octave test frequencies.

The other 5 subjects were deaf children (1 female, 4 male) enrolled in classes at Central Institute for the Deaf. They ranged in age from 9-12 years. Each had profound sensorineural hearing losses bilaterally (Table 1), and each had experienced amplified sound via group and wearable hearing aids for at least 4 years.

MATERIALS AND EQUIPMENT

Sixteen common words with spondaic stress pattern were selected that were familiar to all subjects (Appendix 3). Each word was recorded onto a separate Bell and Howell Language Master stimulus card (#111005) by the same female talker of Study I. Her speech was delivered to the recording

	subject		125	250	500	1000	2000	4000	8000 Hz	detection words	detection noise	discomfort words	discomfort noise
			monaural thresholds (frequency)							**binaural thresholds**			
◆	MH (9-f)	R	85	100	115	110	115	115↓	100↓	110	115	125	130↑
		L	80	110	120↓	120↓	125↓	115↓	100↓				
■	CN (11-m)	R	70	75	100	95	100	115	100↓	105	110	125	130
		L	55	90	105	105	110	105	100↓				
●	TS (10-m)	R	55	85	95	100	120	110	100↓	105	105	130↑	130↑
		L	75	90	95	105	120	110	100↓				
▽	MS (13-m)	R	55	85	100	100	110	110	100↓	100	100	125	125
		L	60	90	95	100	115	110	100↓				
△	JS (13-m)	R	85↓	80	95	100	115	115↓	100↓	100	105	130↑	130↑
		L	85↓	100	105	100	105	90	100↓				
			dB HL ISO							dB SPL		dB SPL	

Table 1. Audiometric thresholds of the 5 profoundly deaf subjects in Study II.

device through a condenser microphone and amplifier.

The Language Master (#711B) that was used in this study had been modified by the attachment of 2 microswitches. Insertion of a stimulus card caused the first of these to trigger a signal light that alerted the subjects about 1 sec before their observation interval. Deflection of the second switch gated on a burst of noise for the duration of the card's movement past the magnetic tape head (about 4 sec). This noise burst constituted the observation interval for the subjects.

The amplified output from each stimulus card was mixed electrically with the noise (low-pass filtered at 6000 Hz), amplified again, and then presented binaurally to the observers via TDH-49 earphones in MX-41/AR cushions. Overall noise level was held constant at 70 dB SPL in the ear for the normal-hearing subjects and at 120 dB SPL for those with profound hearing impairments. Both speech and noise levels were monitored on a VU meter (Daven #910-A). An attenuator in the speech channel controlled the S/N ratio presented to the subjects. Speech level was defined as the average of the peaks indicated on the VU meter by the 32 syllables in the stimulus vocabulary.

A block diagram of experimental equipment is shown in Figure 2.

PROCEDURE

At each test session, the 16 stimulus cards were

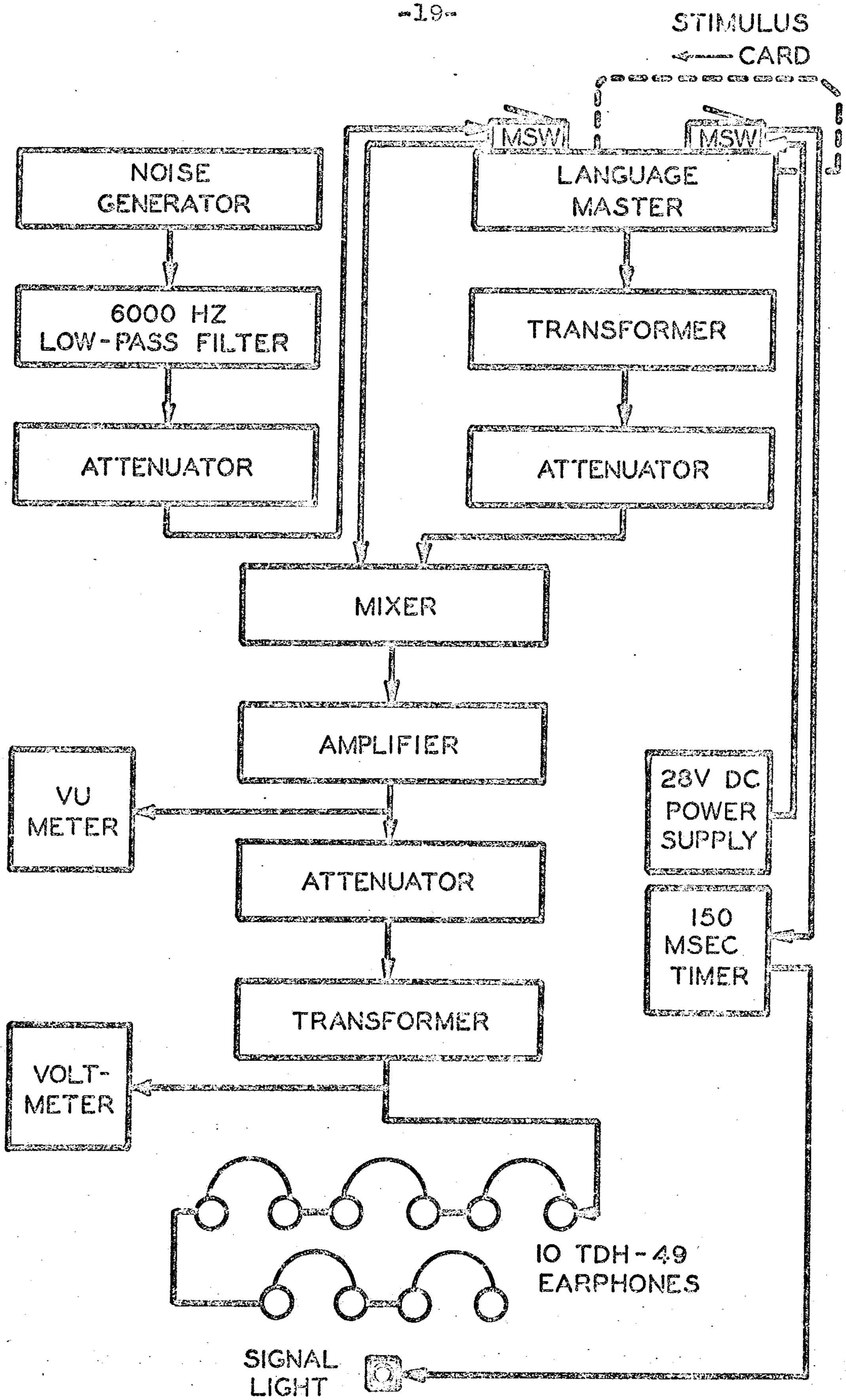


Figure 2. Block diagram of apparatus used to present recorded auditory speech stimuli in Study II.

shuffled with 16 blank cards and the deck of 32 was presented one at a time to observers under the following conditions: -28, -25, -22, -19, -16, -13, and -10 dB S/N for normal-hearing subjects; -13, -10, -7, -4, -1, and +2 dB S/N for deaf subjects. The 2 groups were tested separately.

The S/N ratio was held constant within each block of 32 trials. Cards were shuffled between blocks of trials. S/N conditions were presented to subjects in quasi-random order.

During each test session, subjects were required to judge each 4-sec noise burst for presence or absence of a speech signal (spondaic word) and to check either yes or no accordingly on their answer sheets. Subjects later scored their own responses.

All subjects were given practice for 1 week before formal testing began. Each S/N condition was presented 4 times during the experiment, and the 4 scores were averaged. The raw data were converted into per cent items identified correctly as signal or noise alone under each S/N condition.

RESULTS AND DISCUSSION

Detection of words in noise by normal-hearing adults is described in Figure 3, along with the curve from Study I that represents A-V recognition in noise for similar stimulus material. A comparison between the 2 curves indicates that auditory detection and A-V recognition of words in noise result in very similar functions at low S/N ratios. Each curve

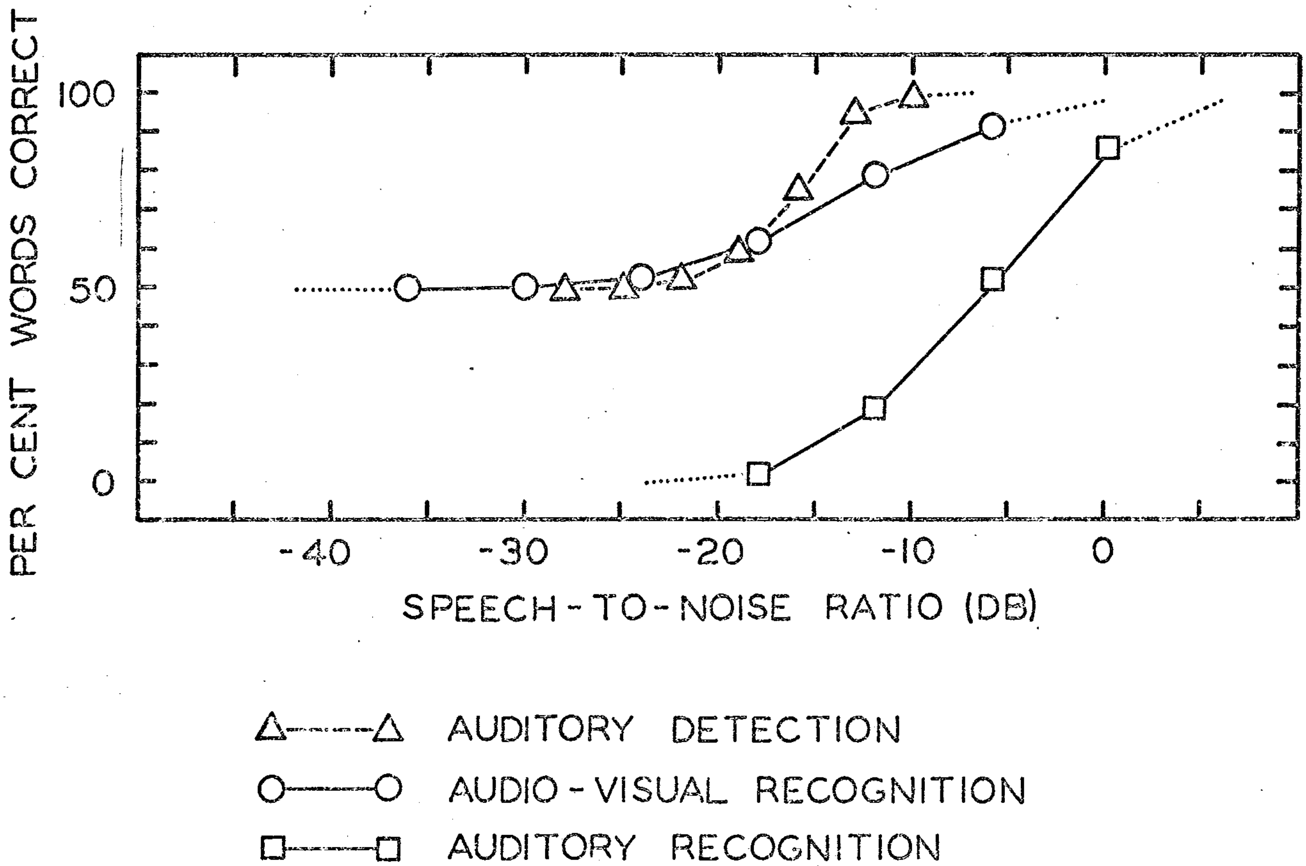


Figure 3. Detection and recognition of spondaic words in wide-band noise by normal-hearing adults.

begins to rise above its asymptotic level at about the same point on the S/N dimension (-22 dB S/N). This result suggests that A-V improvement at low S/N ratios is closely related to the detection of acoustic speech signals, and it implies that mere detection of speech can provide sufficient supplementary information to improve lipreading performance.

However, comparisons between these 2 curves must be viewed with caution because they describe the results of 2 separate experiments. Although the talker was the same in both cases, different word lists and different listeners were employed. Also, in one experiment the speech source was a live talker, while in the other study a Language Master reproduced the stimuli.

Detection of words in noise by the 5 deaf children is compared with that of the 5 normal-hearing adults in Figure 4 (see Appendix 4 for numerical data). Several relations between the 2 sets of curves may be noted: (1) the deaf subjects required about 8-9 dB greater S/N ratio than did the normals for 75%-correct detection performance; (2) there is no overlap between the 2 groups of data points; (3) the slopes of the 2 sets of functions are very similar. Another result, less formally obtained, was that none of the deaf subjects was able to recognize through ear alone (in quiet, at a comfortable listening level of 120 dB SPL) greater than a chance number of the 16 stimulus words. For this task, each child was supplied with a list of the test vocabulary for reference.

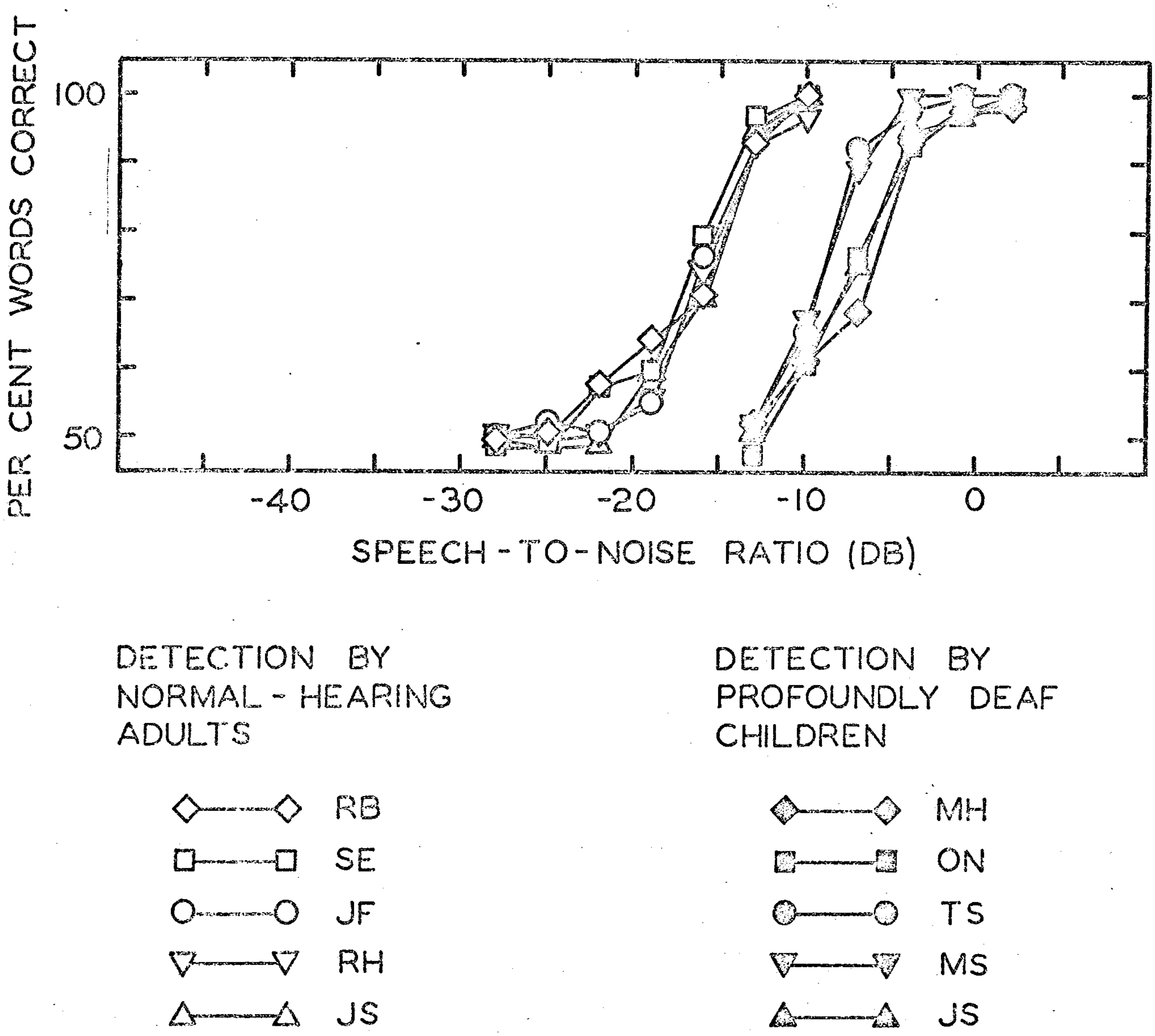


Figure 4. Detection of 16 spondaic words in wide-band noise by 5 normal-hearing adults and by 5 profoundly deaf children.

A direct comparison between the 2 sets of detection data may be questioned, as the profoundly deaf subjects received a 50 dB greater level of masking noise than the normals did. This difference in noise level was an unavoidable by-product of amplifying the stimuli to acoustic levels appropriate for the deaf subjects. The normal-hearing subjects probably would have detected words less well if they too had been presented noise at 120 dB SPL, due to overloading of their auditory systems (Pollack and Pickett, 1958).

SUPPLEMENTARY DATA

Shortly after these data were collected, the CID hearing clinic referred a subject (N.T., female, age 26) with a normal ear and a profoundly deaf ear (etiology unknown). Her audiometric data are shown in Table 2. During pure-tone testing of her impaired ear (with adequate contralateral masking), N.T. reported no sensation of hearing at any level of stimulation at any audiometric frequency. She did report feeling a vibration deep within her ear canal during testing, however, and her detection of these vibratory stimuli are indicated by the audiogram for that ear. Later, each ear was evaluated separately for its ability to detect words in noise. The procedure that was used has been described previously. Noise was presented to the normal ear at 70 dB SPL and to the deaf ear at 120 dB SPL. While the deaf ear was tested, the normal ear was masked by speech-shaped random noise (110 dB SPL) from a separate source.

	subject		monaural thresholds: frequency 125	250	500	1000	2000	4000	8000 Hz	binaural thresholds*: detection words	detection noise	discomfort words	discomfort noise
○	NT	R	10	-5	-5	-5	-5	0	15	DNT#	DNT	DNT	DNT
◎	(26-f)	L	85↓	90	105	115	125	115↓	100↓	105	105	125	130
△	DA	R	15	25	45	60	70	75	75	55	60	105	100
	(12-m)	L	15	30	50	55	70	80	65				
◇	WJ	R	70	80	95	100	110	110↓	90↓	100	100	120	120
	(12-m)	L	80	80	85	95	105	110↓	90↓				
⊡	JW	R	35	60	75	90	95	105	90↓	90	95	120	120
	(11-m)	L	40	70	80	95	100	110	90↓				
			dB HL ISO							dB SPL		dB SPL	

Table 2. Audiometric thresholds of the 4 supplementary subjects in Study II.

* NT, monaural

\# did not test

The results of these tests are shown in Figure 5, along with those obtained earlier from the 5 normal and the 5 deaf subjects. The performance of N.T.'s normal ear parallels closely that of other normal ears, while her deaf ear detects words in noise as poorly as the ears of the 5 profoundly deaf children. During the evaluation of her deaf ear, N.T. reported that she heard nothing but instead felt the words as increments of background vibration.

This comment suggested that the deaf children also might have responded to tactile stimulation. To test this possibility (indirectly), the words and noise (120 dB SPL) were presented over earphones to the hands of 5 different adults with normal hearing (see Nober, 1968 for a direct test of tactile response). Stimuli were not presented to the hands of the deaf children themselves, because their palms were too small to seal the earphone cushion properly. To prevent these normal-hearing subjects from receiving signal cues through their ears also, speech-shaped random masking noise (90 dB SPL) was delivered binaurally via earphones from a separate source.

The results of this tactile experiment are shown also in Figure 5. Detection of signals by hands and by profoundly deaf ears both begin to rise above the chance level (50%) at about the same point on the S/N dimension, but the data from hands follow a shallower slope. The shallower slope of detection by hands suggests greater variability in response by those subjects. They did report that occasionally they were

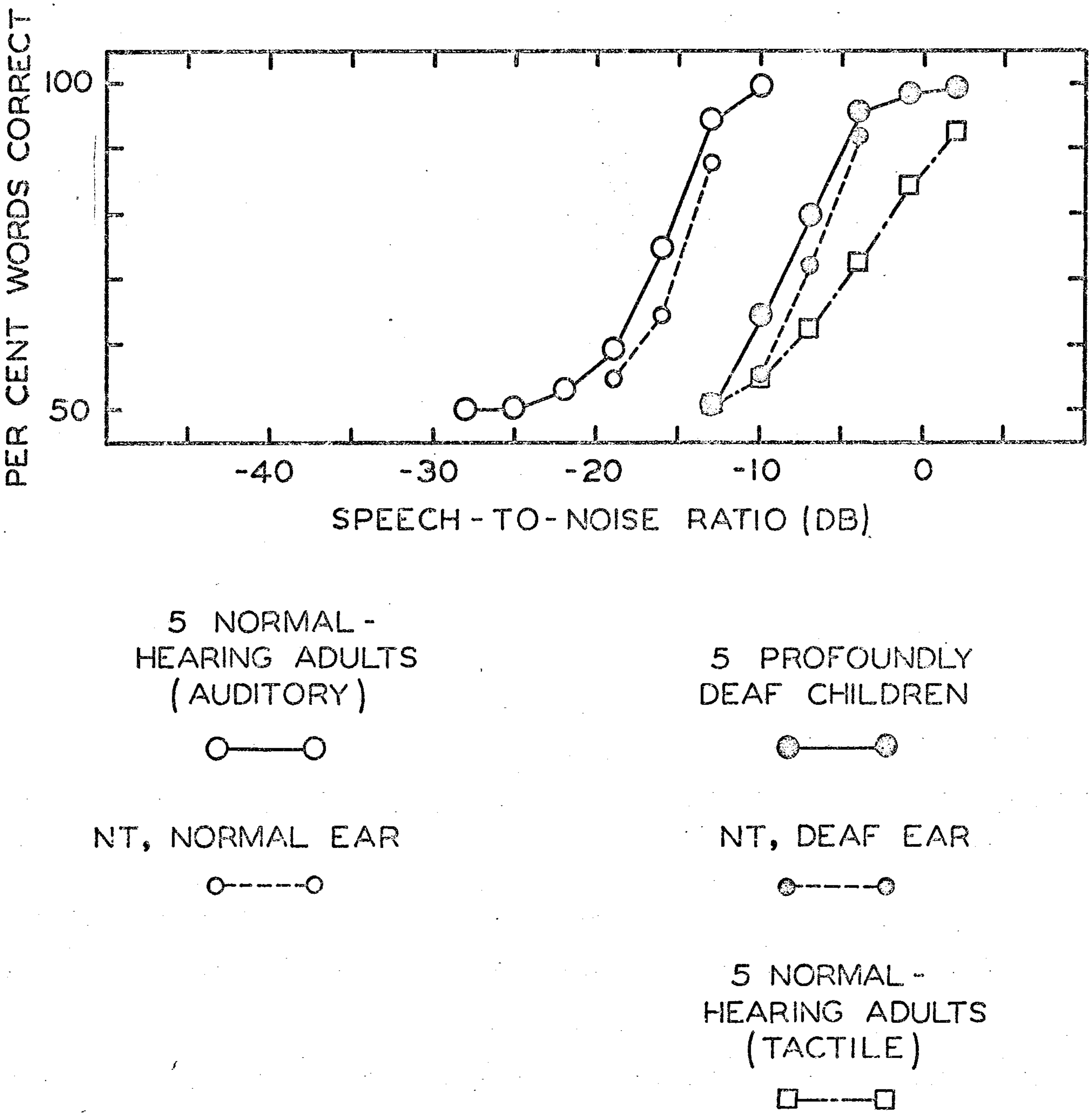


Figure 5. Detection of 16 spondaic words in wide-band noise by several groups of subjects.

unable to perceive even the noise burst because of hand movement or improper earphone position. Acoustic leaks may have occurred between the earphone cushion and the skin of the subjects' palms also. After a more suitable coupling device is developed for presenting acoustic signals both to an ear and to a skin surface, this experiment should be repeated.

Three other CID pupils with less severe sensorineural hearing losses also served as subjects. Their audiograms are given in Table 2. Words were presented in noise to them also as described before. The results are shown in Figure 6, along with detection data from the original 10 subjects.

D.A., whose pure-tone sensitivity indicates only a moderate hearing loss, is able to detect words in noise (90 dB SPL) nearly as well as subjects with normal hearing do. He recognized all 16 test words when they were presented in quiet at 90 dB SPL.

W.J. and J.W., whose hearing losses are severe, detect words in noise (110 dB SPL) a little better than did the 5 CID children with profound hearing losses. These 2 subjects also could identify 9 and 7 words, respectively, when the vocabulary was presented in quiet at 110 dB SPL.

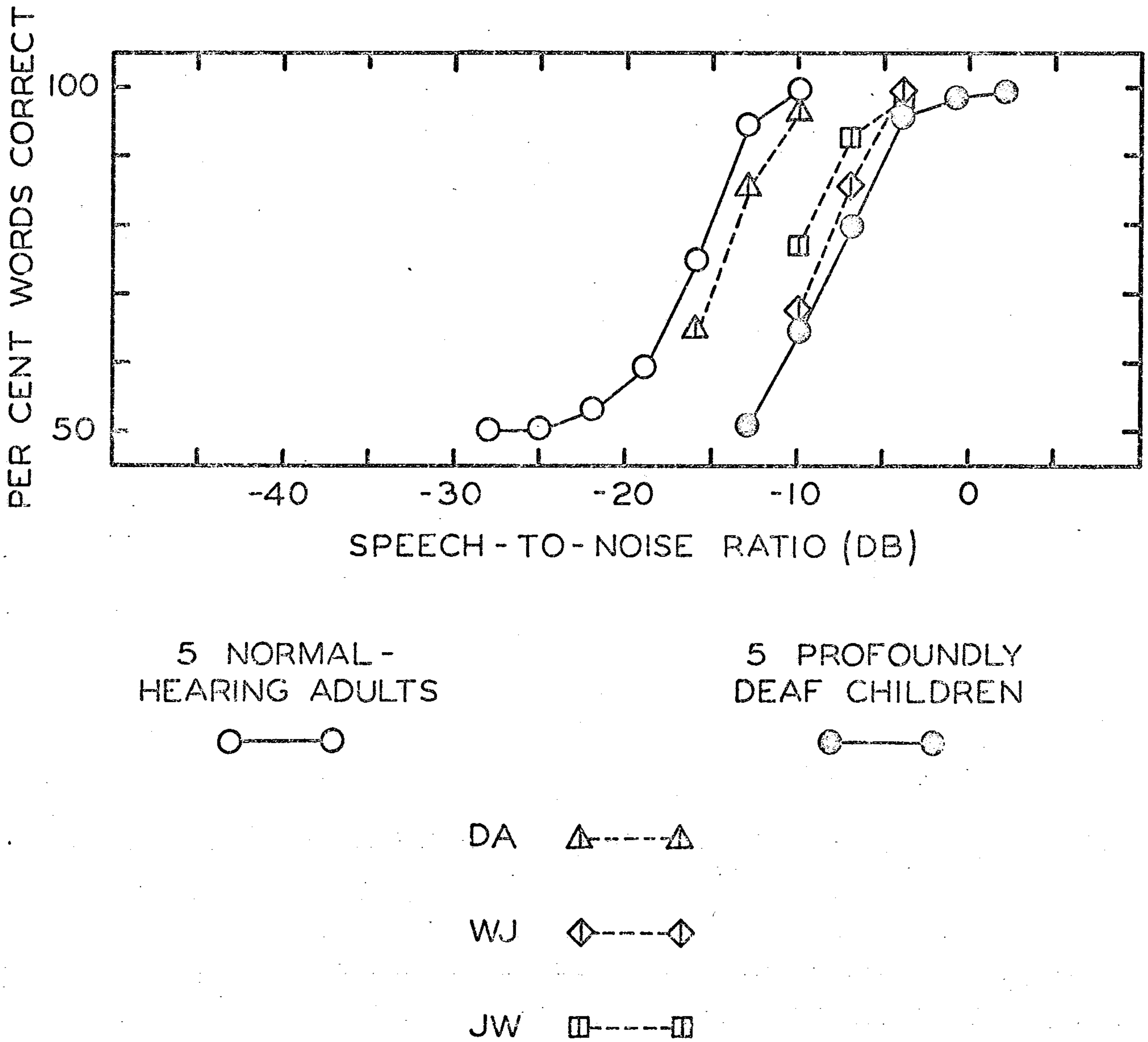


Figure 6. Detection of 16 spondaic words in wide-band noise by 5 normal-hearing adults, 5 profoundly deaf children, and 3 children with less severe hearing losses.

STUDY III: AUDITORY AND AUDIO-VISUAL RECEPTION OF WORDS IN LOW-FREQUENCY NOISE BY CHILDREN WITH NORMAL HEARING AND BY CHILDREN WITH IMPAIRED HEARING

INTRODUCTION

In general, educators of the oral deaf favor the provision of amplified acoustic cues for speech to all hearing-impaired children, even to those whose hearing is so poor that they must rely primarily upon lipreading for comprehension. Heider (1943), Numbers and Hudgins (1948), Hudgins (1953a,b), Hopkins and Hudgins (1953), Clarke (1957), Prall (1957), Hutton (1959), and Van Uden (1960) all have investigated the effect of combining multiple cues, and in general their results demonstrate the superiority of audio-visual (A-V) reception over lipreading alone. In each of these studies, those subjects who possessed better sensitivity for pure tones and greater ability to recognize words by ear also achieved more substantial gains by combining amplified acoustic information with visual cues. The increments were generally smaller for subjects whose profound hearing losses prevented them from recognizing amplified words by ear alone.

Basing their judgement on the above results, many authors (Hudgins, 1953b, 1954; Prall, 1957; O'Neill and Oyer, 1961; Watson, 1961; Whitehurst, 1964) have recommended a combined audio-visual approach for oral instruction to hearing-impaired children. They have stressed the employment of group amplification in the classroom and the use of wearable hearing aids at all other times.

Past research in quiet laboratories has demonstrated clearly that provision of amplified acoustic cues can improve the lipreading comprehension of speech material by children with severe hearing disorders. However, it is more common for those who depend upon hearing aids to use them under less ideal acoustic conditions (e.g., in classrooms, homes, or automobiles). The ambient noise in these everyday situations may have a seriously disrupting effect on the auditory reception of amplified speech by hearing-impaired children.

O'Neill (1954) and Sumby and Pollack (1954) have indicated some of the auditory and visual limitations of normal-hearing adults under unfavorable levels of background noise. Although Dale (1967), Watson (1964), and Ling (1964, 1966) have discussed the disruptive effects of noise on speech reception by hearing-impaired children, little quantitative research with this group has been reported. For this reason, very little is known about the influence of ambient noise on the auditory reception of amplified speech by children with severe hearing losses or about their ability to integrate visual information with acoustic cues for speech that are received under poor environmental conditions.

Severe hearing loss in children frequently involves a diminished range of sensitivity (Watson, 1961), a limited frequency bandwidth (Ling, 1964), underdeveloped listening skills (Whetnall, 1964), and incomplete knowledge of linguistic features (Hart and Rosenstein, 1964). Because of these limitations, there may be little relation between the abilities of normal-hearing and hearing-impaired children to extract acoustic speech information from a noise background. Definition of minimum S/N requirements for hearing-impaired children is essential. These data are necessary for acousticians to provide proper acoustic control of their listening and learning environments and for teachers to exploit maximally each child's residual sensitivity.

Several studies (Lightfoot, Carhart, and Gaeth, 1956; Jerger, Tillman, and Peterson, 1960) have indicated small differences in detection of noise-masked tones between adults with normal hearing and adults with moderate hearing impairments. Study II described large group differences in masked threshold for spondaic words (75% correct detection) between adults with normal hearing and children with profound hearing losses.

The present study compared the abilities of children with normal and impaired hearing to perceive orally presented speech stimuli under a range of S/N conditions, both through the ear alone and through the ear and eye simultaneously. The following variables were investigated: (1) how auditory detection and recognition of speech stimuli vary for each group as a function of S/N ratio; (2) with visual observation of the talker permitted, how A-V recognition of speech stimuli varies for each group as a function of S/N ratio; (3) how each group's lipreading performance (no acoustic input) compares with its A-V performance in noise; (4) within each group, how these several measures of speech reception relate to one another.

To summarize, the goal of this study was to define some of the S/N conditions under which provision of amplified sound can improve the reception of speech by hearing-impaired children who rely primarily upon lipreading. The investigation examined the interaction of several variables (degree of hearing loss, mode of observation, acoustic S/N ratio) and their resulting effects upon the reception of orally presented speech stimuli.

DESCRIPTION OF PERSONNEL

TALKER

A single female talker (age 22) presented the speech stimuli throughout the study. She spoke general American English with no gross abnormality in voice (f_o = 166 Hz) or articulation, as judged by a research audiologist and a clinical speech pathologist. Her speech also was considered to be very intelligible and easy to lipread. She was pre-trained to time her utterances with signals from an indicator lamp and to peak a VU meter within a maximum range of 6 dB.

OBSERVERS

Three children from each of the following audiometric groups participated as observers: normal hearing (10 dB HL ISO or better at each of the audiometric octave frequencies); severely hearing impaired (average HL at 500, 1000, and 2000 Hz between 75 and 85 dB ISO); profoundly deaf (average HL at 500, 1000, and 2000 Hz greater than 95 dB ISO). Only each subject's better ear was evaluated during the study. Their pure-tone audiograms and speech-discrimination scores are shown in Table 3. For the 6 hearing-impaired children, the pure-tone data represent an average of their 3 most recent audiograms. Subjects were chosen to achieve high audiometric homogeneity within groups.

The children with normal hearing were elementary school students of age range 9 yr - 3 mo to 10 yr - 9 mo. The severely hearing-impaired subjects ranged in age from 9 yr - 0 mo to 11 yr - 2 mo, while the profoundly deaf group ranged in age from 10 yr - 11 mo to 12 yr - 6 mo. The children with abnormal hearing had sustained their hearing losses before speech development and had been taught orally at Central Institute for the Deaf for at least the past 4 years. All of these CID students had been presented amplified speech sounds through group and wearable hearing aids for at least 5 years prior to their participation in this study. Before this investigation, none of the subjects was experienced at auditory or visual laboratory tasks.

With appropriate tests, all of the children were screened for normal visual acuity and for normal binocular fusion. All subjects possessed normal or above-average in-

* (Watson, 1957 - Test B, MJ/2)

\# did not test

monaural data

	subject	ear	125	250	500	1000	2000	4000	8000 Hz	detection words	detection noise	discomfort words	discomfort noise	speech discrim.*
□	SB (f) (9y,3m)	R	5	0	0	0	0	0	5	DNT#	DNT	DNT	DNT	96%
○	JK (f) (10y,0m)	R	10	5	0	0	0	0	5	DNT	DNT	DNT	DNT	100%
△	LM (f) (10y,9m)	R	5	0	0	0	0	0	0	DNT	DNT	DNT	DNT	100%
⊟	KB (f) (11y,2m)	L	45	65	80	85	85	75	70	85	85	120	120	92%
⊖	SH (f) (9y,1m)	L	60	65	80	80	70	75	70	90	90	125	120	92%
△	PT (m) (9y,0m)	L	65	70	80	80	90	85	90↓	90	90	120	120	64%
■	MK (m) (11y,9m)	L	65	75	90	100	105	100	100↓	100	100	130↑	130	20%
●	LS (f) (12y,6m)	L	70	85	85	95	110	115↓	100↓	100	100	130	130	4%
▲	JT (m) (10y,11m)	L	60	80	100	90	105	110	100↓	100	100	130	130	16%
					dB HL ISO					dB SPL		dB SPL		%

Table 3. Audiometric descriptions of the 9 subjects in Study III.

telligence, as indicated by their respective school records.

SPEECH STIMULI

The stimulus vocabulary consisted of 240 common nouns that had been selected from a preliminary list of about 1000 simple words. Of the final 240 words, 80 were monosyllabic, 80 were trochaic (i.e., with stress only on the first of 2 syllables), and 80 were spondaic in stress pattern (Appendix 5). Each stressed syllable appeared only once within the vocabulary, in order to minimize possible auditory and visual ambiguity. The chosen stimulus words could be defined by all subjects.

No overt attempt was made to phonemically balance the list with respect to available phoneme-count data obtained from samples of written material (Dewey, 1923). Nevertheless, the phonemic content of this large test vocabulary approximated closely that of the Dewey data (r_{rank} = 0.778).

Ease of lipreadability was not considered in the final selection of stimulus words. Rather, a wide range of difficulty among individual words was assumed to exist.

PROCEDURE

PREPARATION OF STIMULUS MATERIALS

A videotape recorder (Ampex VR-7000) and camera (Ampex CC-326), a microphone (Shure 545), and a 21" TV receiver (Ampex TR 821) were assembled for simultaneous audio and video recording. The camera's zoom lens was adjusted so that the image of the talker's face appeared life-size on the 21" TV monitor screen during playback. The recording studio was brightly lighted, quiet, and non-reverberant for satisfactory audio and video recording.

The stimulus vocabulary was divided into eight 30-word sublists. An oscillator, electronic timer, transformer, and PDR-10 transducer (mounted on the microphone case) were used to present an acoustic tone pulse (7000 Hz, 250-msec duration) every 13 sec. A light flash (1-sec duration) followed each tone pulse by 1.0 sec. While facing the camera (0^{o} angle, 3-ft distance), the talker spoke one word from the list during each signal from the light. She spoke with normal vocal effort and rhythm, as she monitored vowel peak levels on a VU meter (Daven #910-A). No carrier phrase was employed. The talker spoke all 240 words (8 sublists) 8 separate times, scrambling the order of words within sublists each time. These 8 sessions were recorded on video tape.

Table 4 describes the combinations of observation mode and S/N ratio under which each group of subjects was tested later. These ranges of S/N ratio were determined from the results of Studies I and II and other less formal observations. The test procedure required each group's A-V performance to be sampled at 8 different points along the S/N dimension. The 8 S/N ratios specified for each group were assigned in counterbalanced order to the 64 pre-recorded sublists of stimulus words.

dB S/N in audio channel	normal-hearing children: detection A	normal-hearing children: recognition A	normal-hearing children: recognition AV	normal-hearing children: recognition V	severely hearing-impaired children: detection A	severely hearing-impaired children: recognition A	severely hearing-impaired children: recognition AV	severely hearing-impaired children: recognition V	profoundly deaf children: detection A	profoundly deaf children: recognition A	profoundly deaf children: recognition AV	profoundly deaf children: recognition V
+ 5							X			X	X	
0						X	X			X	X	
- 5					X	X	X		X	X	X	
-10		X	X		X	X	X		X	X	X	
-15	X	X	X		X	X	X		X		X	
-20	X	X	X		X		X		X		X	
-25	X	X	X				X				X	
-30	X		X									
-35			X									
-40			X									
audio off				X				X				X

Table 4. Conditions under which each group of observers received speech stimuli in Study III.

In order to provide identical <u>acoustic</u> speech signals to subjects under both A-detection and A-V recognition conditions, secondary A-detection tapes were prepared from the original video tapes. This provision insured that the results describing subjects' A-detection and A-V recognition performance could be compared directly.

The 32 pre-recorded A-V blocks were identified that had been assigned S/N conditions scheduled also for A-detection testing. A-detection stimuli were prepared by recording the audio channels of these 32 blocks of video tape onto ¼" magnetic tape with a full-track audio tape recorder (Ampex Model 300). An additional 960 tone pulses were recorded also. The 32 recorded blocks of stimulus words and the tape containing only recorded tone pulses were cut-and-spliced to create 4 audio tapes, each containing 8 blocks of thirty 2-alternative forced-choice items. Half of the signals (i.e., words) appeared in the first position and half occurred as the second alternative (in random sequence). These audio tape recordings also were used during presentation of stimulus words for A-recognition testing.

<u>PRESENTATION OF STIMULUS MATERIALS</u>

A videotape recorder (Ampex VR-7000) was used to play back recorded stimuli during A-V recognition testing, while an audiotape recorder (Ampex Model 300) was used to play back tapes during A-detection and A-recognition sessions. Visual stimuli were presented over a 21" TV monitor (Ampex TR 821) placed 10 ft before the observers and 1 ft above their eye level. A small indicator lamp was placed atop the TV monitor. The observation area was dimly lighted for comfortable visual reception.

The recorded 7000-Hz tone pulse that preceded the onset of each stimulus word by about 1 sec triggered the gating of a 3-sec light flash (on the indicator lamp) to define for observers each observation interval (Figure 7). Except during visual testing without acoustic input (i.e., V-only), each tone pulse also triggered the gating of a 3-sec burst of random noise that was low-pass filtered according to the spectrum indicated in Figure 8. These octave-band levels approximate closely the "typical" outside noise spectrum described by Niemoeller (1968), and they are very similar also to the levels measured in CID classrooms and play areas.

Monaural presentation of the masking bursts and the stimulus words was through TDH-49 earphones mounted in MX-41/AR cushions. The noise was not on continuously during a block of trials to minimize auditory fatigue to subjects. Speech level was defined as the average of the peaks indicated on a VU meter by the recorded stimulus words. Overall noise level remained constant for each group throughout all blocks (except V-only): for normal-hearing, 70 dB SPL; for severely hearing-impaired, 110 dB SPL; for profoundly deaf, 120 dB SPL. An attenuator in the speech circuit controlled the S/N ratio presented to subjects during each block of trials. The

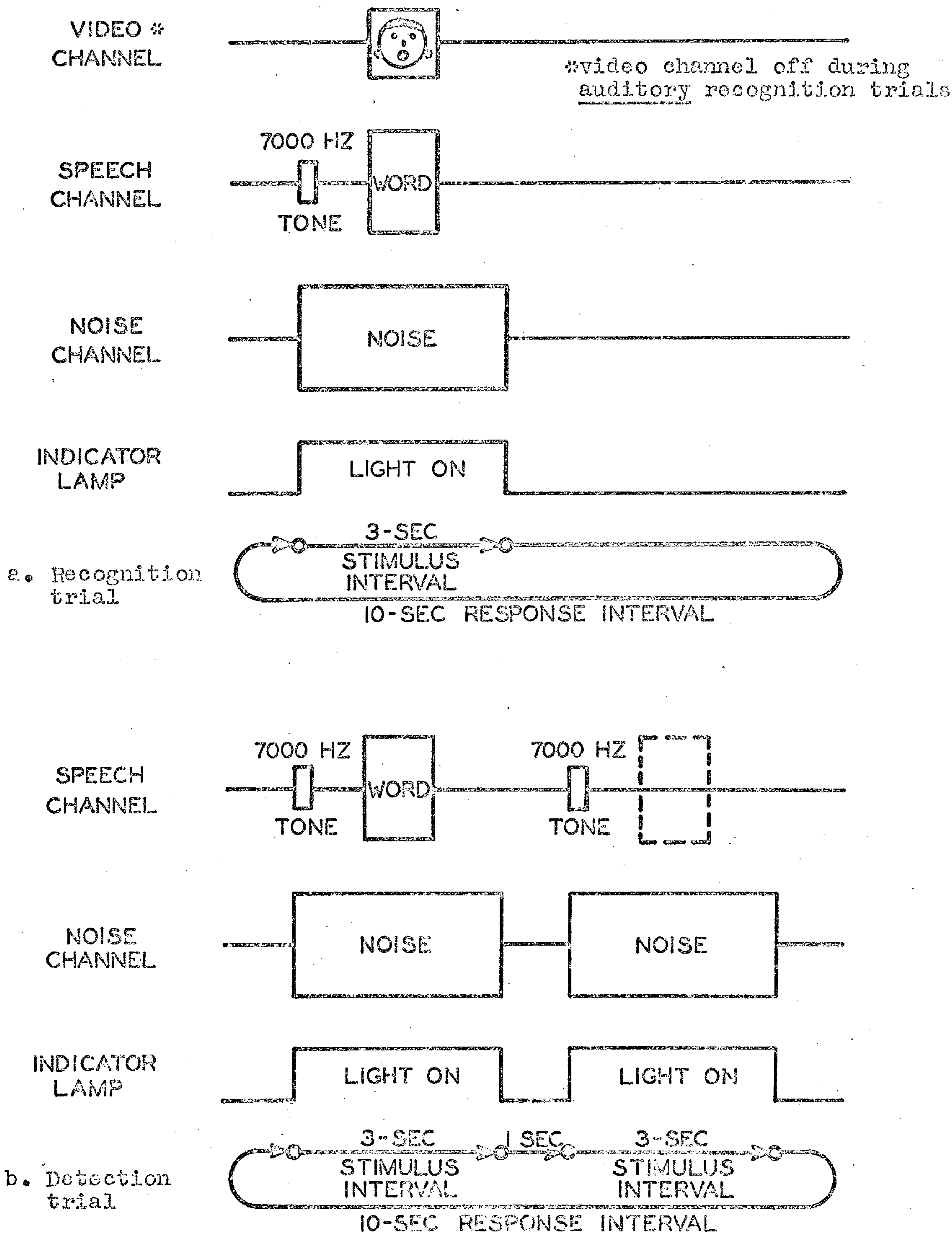
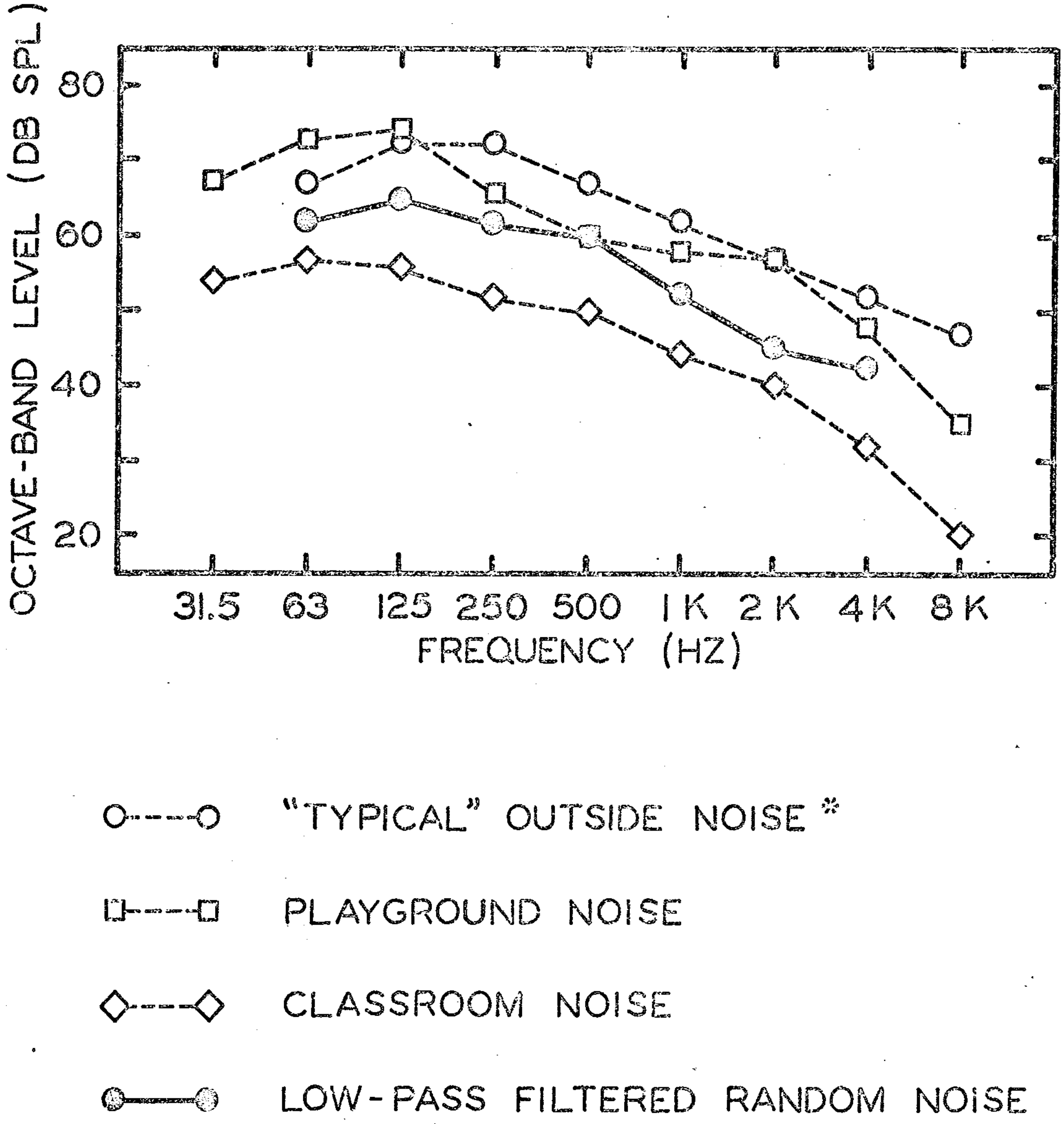


Figure 7. The temporal relation between stimulus and response events.



* Niemoeller (1968)

Figure 8. Octave-band spectra of some environmental noises measured at Central Institute for the Deaf and of the experimental masking noise.

effective range of the audio channel was 50-5000 Hz. A block diagram of the equipment is given in Figure 9.

At each of the 16 test sessions, 8 recorded sublists were presented to observers under S/N conditions that had been counterbalanced to distribute possible learning effects. Equipment availability dictated the following sequence of testing: (1) A-detection, (2) A-recognition, (3) A-V recognition. Each group of subjects was tested separately.

During each of 4 A-detection sessions, a different forced-choice stimulus tape was used to present by earphone 240 pairs of 3-sec noise bursts (each indicated by a 3-sec light flash). Subjects were required to select from each pair the noise burst that contained the recorded word and to check their responses 1st or 2nd, accordingly, on their answer sheets. They were given 10 sec in which to respond to each trial.

During each of 4 A-recognition and 8 A-V recognition sessions, observers received by earphone 240 3-sec noise bursts (each indicated by a 3-sec light flash). They were required to identify and write on their answer sheets the word that was presented during each masking burst (with or without the aid of visual cues on the TV monitor screen, depending upon mode of observation). For blocks during which subjects were required to lipread only (V-only), both audio channels (i.e., noise and speech) were switched off, and subjects were allowed only to view the indicator lamp

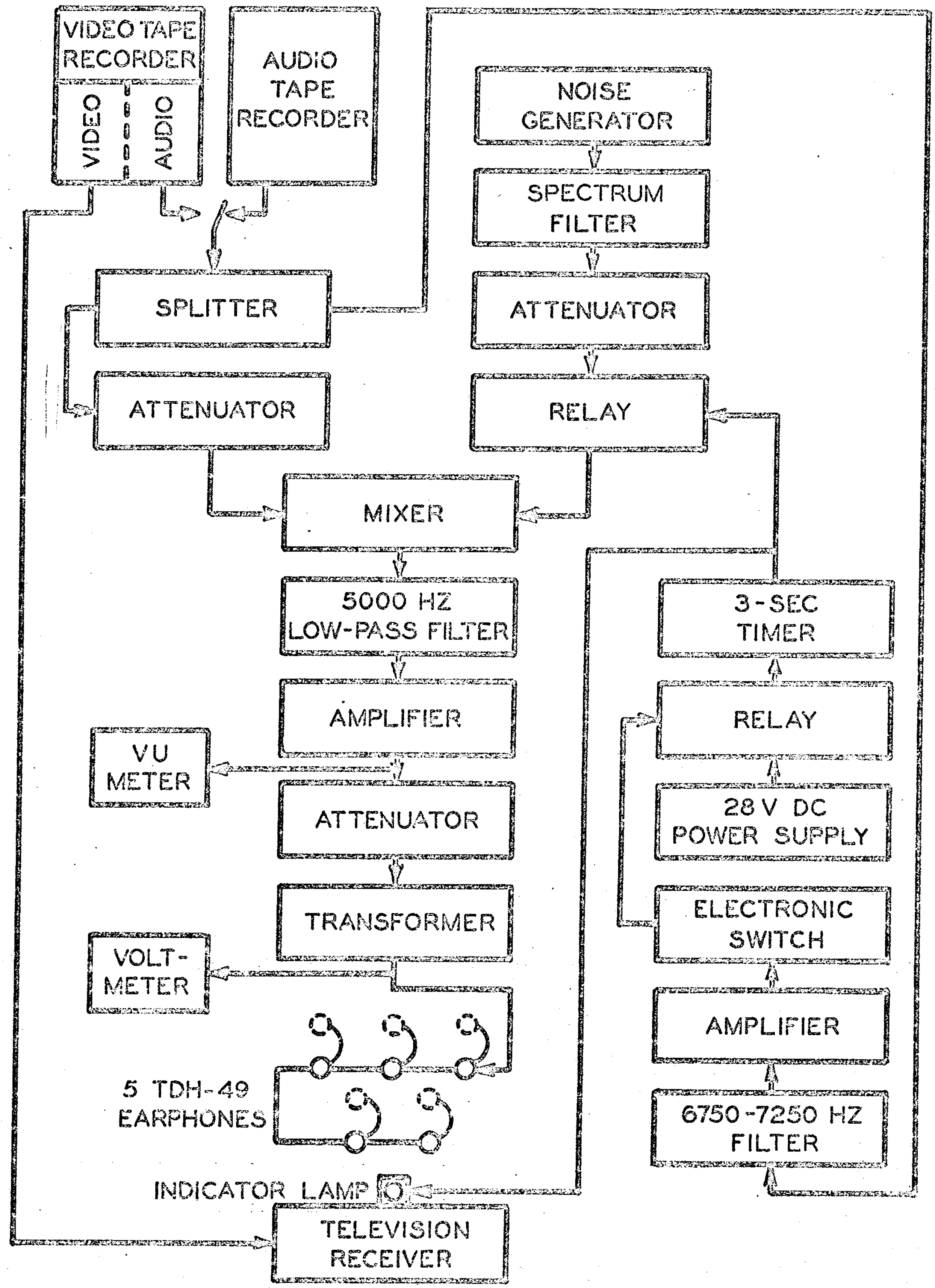


Figure 9. Block diagram of apparatus used to present recorded auditory and visual speech stimuli in Study III.

and the TV monitor screen. As before, they were given 10 sec in which to respond to each trial.

At all times during recognition sessions, subjects were allowed to refer to typed lists of the 240 stimulus words, which had been subdivided by stress pattern and alphabetized (Appendix 5). They were encouraged to respond to each presentation with a word from this vocabulary and to guess if necessary. Before each test session began, subjects were required to read the word list to re-acquaint themselves with the stimulus vocabulary.

A detection response was scored "correct" if it properly selected which of 2 paired noise bursts contained a stimulus word. A recognition response was scored "correct" if it matched phoneme-for-phoneme with the stimulus word that was presented. Misspelled words were not considered errors if they were phonemically correct. Subjects were allowed to examine their scored answer sheets following each test session. All data were converted into per cent items correctly detected or correctly recognized under each S/N condition.

RESULTS

The results of Study III are summarized in Figure 10. Figure 10a indicates that the 3 groups are dissimilar in their detection thresholds for acoustic speech signals (i.e., words) in low-frequency acoustic noise. Whereas children with normal hearing can detect words correctly 75% of the time at a S/N ratio of -22.5 to -25.5 dB, severely hearing-impaired children key to symbols

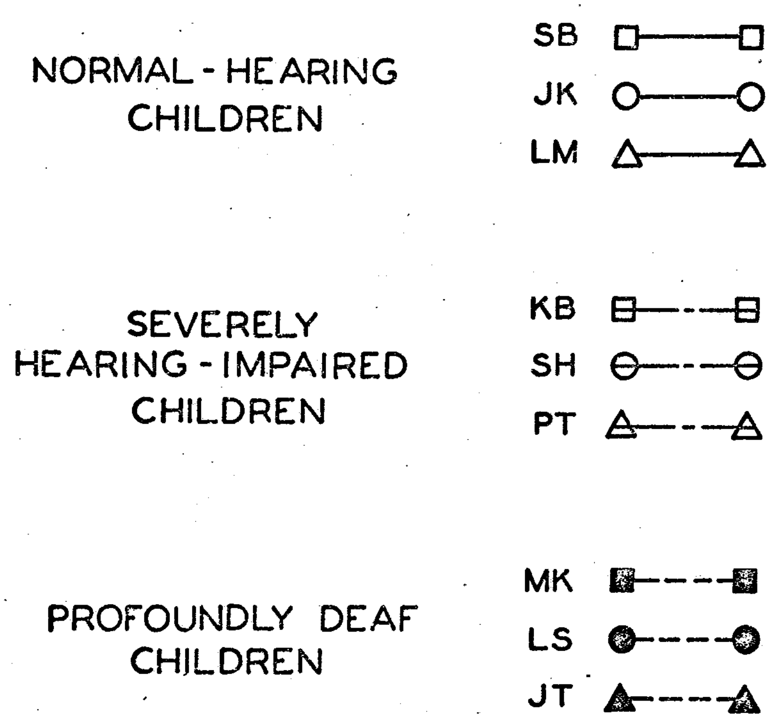
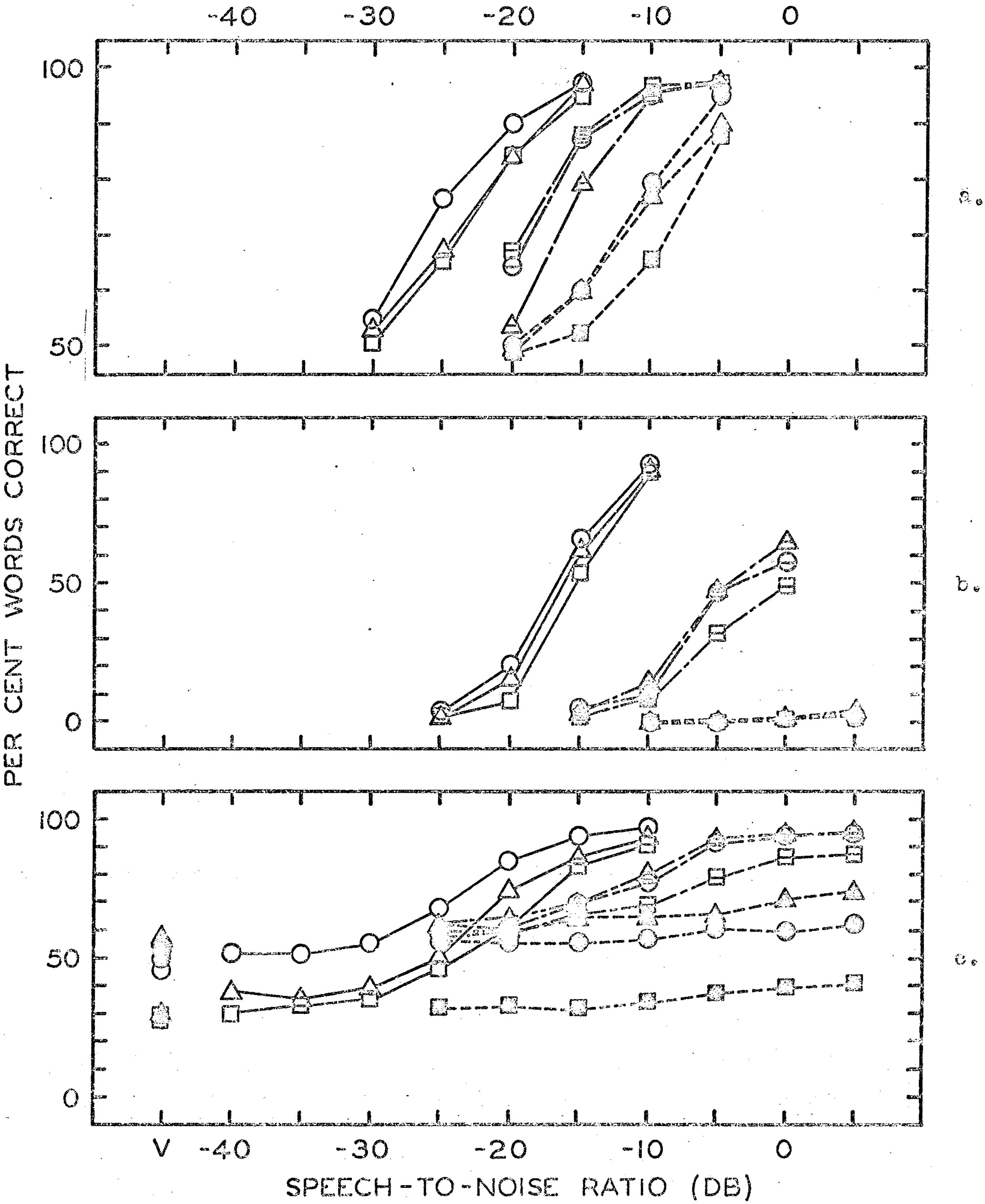


Figure 10. Auditory and audio-visual reception of 240 words in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children.

continued on next page

Figure 10
continued from
previous page



a. auditory detection
b. auditory recognition
c. audio-visual recognition

score 75% correct at -16 to -18 dB S/N, and profoundly deaf children require a S/N ratio of -8 to -11 dB for similar detection performance. The 3 sets of data points do not overlap, and the slopes of these curves are all very similar. One-way analysis of variance indicates significant group differences ($p < .001$) in S/N ratio required for 75% correct detection (i.e., threshold).

Figure 10b displays auditory word-recognition scores for the 3 groups as a function of S/N ratio. The group with normal hearing recognizes words near the chance level for S/N ratios below about -25 dB, but their A-recognition performance improves with increasing S/N ratio, and they recognize correctly 90-93% of the stimulus words at -10 dB S/N. These results are comparable to those reported by Hirsh and Bowman (1953) for speech discrimination in low-frequency noise by listeners with normal hearing. The severely hearing-impaired children cannot recognize words above the chance level at S/N ratios below about -15 dB, but their scores improve to 49-65% words correctly recognized at 0 dB S/N. Inspection of their errors at higher S/N ratios suggests difficulty in discriminating between phonemes with high-frequency components. They confuse mainly unvoiced consonants; vowel substitutions are less common. In contrast, the group of profoundly deaf children are unable to identify words correctly by ear alone above the chance level for any S/N ratio tested, although a slight uptrend in scores is apparent for higher S/N ratios. At higher S/N ratios, this group's errors are unpredictable, and they seem unable to distinguish even vowel sounds reliably.

A-V word-recognition scores for the 3 groups of children are shown in Figure 10c as a function of S/N ratio. In this diagram, each curve's low asymptote describes that subject's lipreading performance in noise. Each curve's high asymptote reflects the limit on A-V recognition imposed by that subject's hearing impairment. The A-V scores of normal-hearing children rise from a low asymptote (30-52%) for S/N ratios below about -30 dB to nearly perfect recognition (91-97%) at -10 dB S/N. The severely hearing-impaired children improve from a low asymptote (52-57%) below about -15 dB S/N to a high asymptote (about 88-96%) above about 0 dB S/N. The profoundly deaf children, however, do not demonstrate such large gains through addition of acoustic cues to lipreading in noise. Their A-V scores rise gradually from a low asymptote (33-62%) below about -5 dB S/N to a high asymptote (about 42-74%) for S/N ratios above +5 dB. Similar A-V performance has been obtained from the 2 groups of hearing-impaired children by Numbers and Hudgins (1948), who varied the <u>sensation level</u> of stimulus words and presented them in the <u>quiet</u>.

During A-detection and A-recognition sessions, each group received speech stimuli only at levels that were adequate for 100%-correct detection or near-maximum recognition, respectively, for those words <u>in the quiet</u>. Therefore, one can be reasonably certain that the data describe masking functions.

Each group of children received acoustic stimuli at a different sound pressure level (i.e., normal-hearing: noise at 70 dB SPL; severely hearing-impaired: noise at 110 dB SPL; profoundly deaf: noise at 120 dB SPL). These differences in noise level were unavoidable by-products of amplifying the stimuli to acoustic levels that are appropriate for each group. The normal-hearing children might have performed less well if they also had received words in noise of 110 or 120 dB SPL, because of overloading of their auditory systems (Pollack and Pickett, 1958).

A 2-way analysis of variance compared the lipreading scores of all 3 groups under the 2 background conditions that were investigated (i.e., quiet and noise). Scores obtained under the V-only condition were taken as lipreading in quiet, while the mean scores for the 2 lowest S/N ratios examined for each group (i.e., low A-V asymptotes) were taken as lipreading in noise (Figure 10c). Group differences in lipreading performance were found significant ($p < .05$) for scores pooled across quiet and noise conditions. However, neither the effect of background nor the interaction between group and background was shown to be significant.

DISCUSSION

A close relation exists between the S/N ratio at which each group's A-V recognition performance begins to improve above low asymptote (i.e., lipreading in noise) and the S/N ratio at which its corresponding A-detection performance begins to rise above the chance level (50%). Acoustic cues for speech must be detectable to supplement lipreading, and the results do demonstrate for all groups that A-V improvement at low S/N ratios is dependent upon the detection of acoustic speech signals (Figure 10a,c). It appears also that the poorer the pure-tone sensitivity of the group, the more detectable speech must be in noise for those observers to combine acoustic cues with lipreading for A-V improvement.

A comparison between each group's A-V recognition scores and its A-recognition scores reveals that each group can improve its lipreading performance at low S/N ratios with acoustic information that by itself is insufficient for A-recognition above the chance level (0.4%). It was suggested in Studies I and II that an observer might use just-detectable acoustic cues to indicate the syllabic pattern of the speech stimulus and to direct his search of the talker's face (or his own short-term memory) for discriminative information immediately following (or preceding) those minimal cues.

The data shown in Figure 10b indicate that both normal-hearing and severely hearing-impaired children can receive much discriminative speech information through their ears alone, if only appropriate amplification, adequate S/N conditions, and many years of prior auditory experience are provided. In contrast, the profoundly deaf group's A-recognition scores remain near the chance level throughout the range of S/N ratios tested. Because deaf children recognize words so poorly by ear alone, Figure 10b does not demonstrate clearly the influence of S/N ratio on speech discrimination by the 3 groups. To provide another kind of data for this comparison, the stress pattern of each response word was determined and tallied according to the stress pattern of the stimulus word that had been presented: monosyllabic, trochaic, or spondaic. The resulting confusion matrices (Table 5) illustrate the errors in 3-way categorization made by each group at each of the S/N ratios tested.

Figure 11 depicts these data plotted as per cent words correctly categorized by each subject at several S/N ratios. For better-than-chance A-categorization of stimulus words by stress pattern, normal-hearing children require about -20 dB S/N, severely hearing-impaired children need about -10 dB S/N, and profoundly deaf children require about -5 dB S/N. Both the normal-hearing (97-99%) and the severely hearing-impaired (84-90%) children achieve satisfactory word-categorization at higher S/N ratios; and under these same conditions they also can recognize by ear alone a large per cent of the stimulus vocabulary (Figure 10b). The group of profoundly deaf children, who cannot recognize these words by ear alone, also are relatively poor at distinguishing them by stress pattern (59-67% at +5 dB S/N). However, they achieve nearly 86% correct in dividing these words only between monosyllabic and disyllabic (i.e., trochaic and spondaic combined) categories. At higher

Normal-hearing children

S/N	Response	Stimulus M	Stimulus T	Stimulus S
-25 dB S/N	M	145	145	147
	T	63	67	56
	S	32	28	37
-20 dB S/N	M	180	119	97
	T	46	100	61
	S	14	21	82
-15 dB S/N	M	222	31	23
	T	10	193	44
	S	8	16	173
-10 dB S/N	M	235	4	0
	T	5	233	4
	S	0	3	236

Severely hearing-impaired children

S/N	Response	Stimulus M	Stimulus T	Stimulus S
-15 dB S/N	M	91	94	82
	T	95	91	82
	S	54	55	76
-10 dB S/N	M	151	76	43
	T	63	126	81
	S	26	38	116
-5 dB S/N	M	215	24	22
	T	14	194	41
	S	11	22	177
0 dB S/N	M	223	19	18
	T	13	209	28
	S	4	12	194

Profoundly deaf children

S/N	Response	Stimulus M	Stimulus T	Stimulus S
-10 dB S/N	M	102	105	76
	T	97	101	111
	S	41	34	53
-5 dB S/N	M	155	98	34
	T	64	112	137
	S	21	30	69
0 dB S/N	M	183	74	14
	T	49	141	129
	S	8	25	97
+5 dB S/N	M	197	52	9
	T	39	143	119
	S	4	45	112

Table 5. Auditory categorization of 240 words by stress pattern in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children (confusion matrices).

M = monosyllabic
T = trochaic
S = spondaic

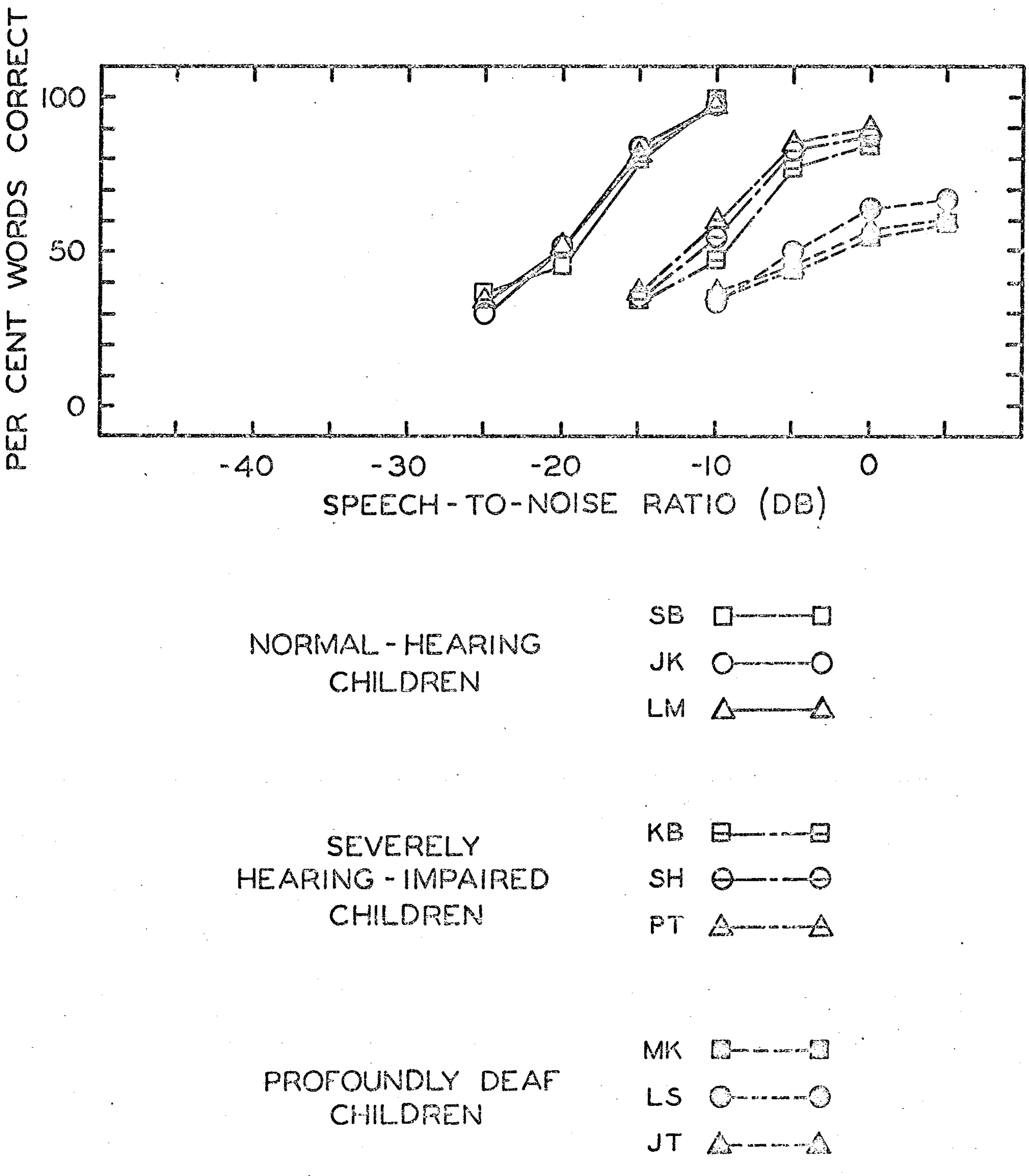


Figure 11. Auditory categorization of 240 words by stress pattern in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children.

S/N ratios, their errors consist mainly of monosyllabic responses to trochaic stimulus words and trochaic responses to spondaic stimuli (Table 5).

Ling (1966) also found that profoundly deaf children are poor at discerning by ear the stress pattern of multisyllabic words or phrases. He reported that syllable omissions occurred frequently for speech stimuli that contained an unstressed vowel (e.g., dʒækət) or contained voiced "boundary" consonants (e.g., rɛIlroUd). It is unknown why profoundly deaf children cannot perceive stress reliably by dictionary standards (i.e., as defined by persons with normal hearing), even in clearly detectable speech stimuli.

Extensive surveys are required to determine the discriminative limits of both the auditory and visual systems of profoundly deaf children. The results of those studies would indicate which phonemic (and linguistic) distinctions are possible for that group through either receptor alone, and they might also suggest how amplified acoustic cues can be delivered most effectively to deaf children for maximum contribution to lipreading.

The point on the S/N dimension at which each group's A-V recognition performance reaches high asymptote (Figure 10c) defines the lowest S/N ratio under which each group can make maximum use of acoustic cues in combination with lipreading. In this investigation, severely hearing-impaired and profoundly deaf children required about 0 and +5 dB S/N, respectively, for maximum A-V recognition of words in noise. However, because of the limited scope of this study, these S/N ratios cannot yet be considered adequate acoustic conditions for A-V reception of speech by hearing-impaired children.

The data reported here are based upon responses to single-word stimuli. Conversational speech consists of sentences and sentence fragments. Miller, Heise, and Lichten (1951) have indicated that listeners with normal hearing require higher S/N ratios to recognize words presented alone than they do to identify words given in a sentence context. It is questionable whether this relation would hold also for hearing-impaired children whose language competence is much less sophisticated. With regard to this problem, Morris (1944, cited by O'Neill and Oyer, 1961) reported that deaf subjects find words <u>harder to lipread</u> when the words are placed in long sentences than when they occur in shorter ones.

It is unknown how noises of other frequency spectra would influence A-V reception of speech by observers with normal or impaired hearing. For a given S/N ratio, word-intelligibility scores for normal-hearing adults vary considerably as a function of the masker spectrum (Miller, 1947; Hirsh and Bowman, 1953). However, fragmentary evidence (Studies I and II) suggests that for a given S/N ratio, word-<u>detection</u> by profoundly deaf children varies relatively little when the spectral characteristics of the masker are changed, provided that its bandwidth is not extended beyond the limits of the subjects' frequency range. Pickett and Martin (1968) also have reported that subjects with severe hearing impairments are poorer than normal at discriminating differences between noise spectra.

The low-pass filtered random noise used in this study was continuous during masking bursts. Miller (1947) has demonstrated for normal-hearing adults that a noise background of 1 or 2 rival voices (whose level and frequency content vary with time) is a less effective masker for desired speech than is the relatively continuous masking of 6 or 8 interfering talkers. Although they often must receive speech in competing-voice situations (e.g., at home or in the classroom), little data is available on the reception of speech by deaf children under background conditions that fluctuate in an unpredictable way (Watson, 1964).

In this study, each group of subjects listened to words in noise of only one level. Hawkins and Stevens (1950) have shown that the S/N ratio required by normal-hearing adults for detection of speech remains constant over a wide range of masking levels. It is unknown whether similar results for several noise levels can be obtained from hearing-impaired children, whose range of sensitivity is much narrower than normal.

Different speakers are heard with varying degrees of intelligibility both by adult listeners with normal hearing (Egan, 1944, cited by Dale, 1967) and by hearing-impaired children listening through hearing aids (Dale, 1967). O'Neill (1951, cited by O'Neill and Oyer, 1961) reported differences in both the auditory and visual (i.e., lipreading) intelligibility of a group of talkers for normal-hearing adults. The present data are based upon responses to the auditory and visual stimuli of a single talker. Whether talker differences would be reflected mainly in the asymptotes for A-V word recognition, or whether the relations involving S/N ratio would be affected also are questions for further research.

Leonard (cited by Oyer, 1964) reported that the lipreading performance of trained subjects is poorer in ambient noise, speech, or music than it is in the quiet. That similar results were not obtained in the present study may be attributed to the presentation of masking noise in 3-sec bursts. Several subjects stated that the onset of both noise burst and signal light (as in all masking trials) was superior to that of the light alone (as in all V-only trials) as an alerting stimulus. This effect may have depressed lipreading-in-quiet scores or elevated lipreading-in-noise scores to obscure any real differences between them (Figure 10c).

Murray (1951) and Lowell (1960) also compared the lipreading abilities of groups with unlike hearing levels, and both of their studies indicated similarly that observers with moderate-to-severe hearing losses are better lipreaders than are either normal-hearing or profoundly deaf subjects. Lowell attributed that group's superior performance to a combination of "learning opportunity" and "language facility". He described "learning opportunity" as a function of "motivation, practice, and knowledge of results". His data suggested also that group differences in lipreading skill are not fixed and may change with the age and experience of the subjects.

Hudgins (1954) described the gains achieved by a group of hearing-impaired children in auditory, visual, and audio-visual recognition of words throughout 6 years of oral instruction at Clarke School for the Deaf. He attributed these gains mainly to the increments in receptive vocabulary and in discriminative skill that result from intensive auditory training. The present study sampled performance during only a 1-month period. It is assumed that any learning to recognize words in noise that occurred during the test sessions was small relative to the long-term improvement that may be possible over a lifetime of practice. It is unlikely, however, that similar improvements in detection (e.g., pure-tone threshold) performance occur over a long period of time (Elliott, 1967).

CONCLUSION

The results of this investigation were obtained under the combined influence of many variables whose effects are only partly understood. Therefore, it would be unwise to propose at this time minimum S/N ratios (or maximum noise levels) for the learning environments of children with hearing losses. Until other corroborative studies are completed, one can only estimate that hearing-impaired children require <u>at least</u> about a 10-15 dB greater S/N ratio (i.e., about 0 to +5 dB S/N) than normal-hearing children need (i.e., about -10 dB S/N) for maximum intelligibility of speech through A-V reception in low-frequency noise.

In a low-frequency noise environment of 70 dB SPL overall, a classroom teacher who uses a group hearing aid should have little difficulty satisfying these S/N requirements if she speaks consistently within 6-9 inches of the microphone. In a similar noise background, a parent or teacher who speaks to a child wearing the microphone within his individual hearing aid can meet these criteria only marginally, even when he is as near as 3-4 feet from the child. Under these conditions, that child may receive acoustic cues for speech intermittently. For either greater noise levels or greater distances between the talker and the hearing-aid microphone, the child may be aware of amplified noise only. Favorable S/N ratios usually can be achieved by limiting the noise at its source or by decreasing the distance between the talker and the microphone (Niemoeller, 1968).

SUMMARY

This paper describes a study of auditory and visual reception of speech in the presence of background noise. The report is divided into four chapters.

Chapter I reviews, in chronological order, some prior studies that have examined the reception of speech in noisy environments. The initial investigations were mainly descriptive, while later ones attempted to overcome the limitations imposed by noise through manipulating acoustic variables. The most recent papers have investigated the potential value of visual cues (i.e., lipreading) for improving speech reception under unfavorable acoustic conditions.

Chapter II describes a preliminary investigation of reception of spondaic words in wide-band noise by adults with normal hearing. At different times, the subjects were required to listen to a talker or to listen to her and watch her face simultaneously. Combined audio-visual (A-V) recognition was found to be superior to auditory recognition alone; and a relation was suggested between improvement of A-V recognition of words at low S/N ratios and auditory detection of those words.

Chapter III presents a follow-up study in which auditory detection of spondaic words in wide-band noise was measured for both normal-hearing adults and for profoundly deaf children. For those with normal hearing, a close relation was found between the S/N ratio required for A-V recognition of

words (i.e., data from the previous experiment) and the S/N ratio required for auditory detection of similar materials. This result suggested that the mere detection of speech patterns may provide sufficient information to improve an observer's lipreading performance. The deaf subjects required a greater S/N ratio for 75% correct detection of words than did the normals. This finding implied that maximum communication effectiveness might occur for deaf children only at a higher S/N ratio than that required by normals.

Chapter IV describes a more inclusive study whose goal was to define some of the S/N conditions under which provision of amplified sound can improve speech reception by hearing-impaired children over lipreading alone. Common words (i.e., monosyllables, trochees, and spondees) were presented in low-frequency noise to children whose task was to detect or to recognize them under a range of S/N conditions and through several sensory channels (i.e., listening alone, lipreading alone, simultaneous listening and lipreading). Findings indicated that both profoundly deaf and severely hearing-impaired children require higher S/N ratios for 75% correct auditory detection of words than do children with normal hearing. The normal group was superior to the severely hearing-impaired group in auditory recognition of words in noise, while the deaf group was unable to recognize words by ear alone. A supplementary analysis of the data revealed that the deaf group was relatively poor even at categorizing the stimulus words by stress pattern. Provision of acoustic cues was found to improve greatly the lipreading performance of normal-hearing and severely hearing-impaired children; but profoundly deaf children did not demonstrate such large gains in A-V scores. Maximum A-V recognition resulted for the profoundly deaf group at a higher S/N ratio than that for the severely hearing-impaired group, who required a higher S/N ratio for maximum A-V performance than did the normal-hearing group. Improvement in A-V recognition at low S/N ratios was shown for all groups to depend upon the <u>detection</u> of acoustic cues for speech. Several topics for subsequent research were suggested, and tentative recommendations for minimum tolerable S/N ratios for A-V communication were proposed.

airplane
all right
anthill
armchair
ashtray
bagpipe
bank note
barbed wire
baseball
bathtub
bay leaf
bear hug
bedsheet
beehive
billfold
birthday
blacksmith
blind date
bloodstream
blowgun
blue jay
bobwhite
bonfire
bootstrap
boxcar
briefcase
broadcloth
buckwheat
bullfrog
cake mix
cardboard
catnip
caveman
cellblock
chalk dust
checkmate
childhood
chipmunk
clambake
clothesline
coal bin
cockpit
coin purse
corncob
cough drop
cowboy
crabgrass
creampuff
crew cut
cuff link
cub scout
dead end
dial tone
dime store
dish towel
doormat
doughnut
downstairs
drain plug
drawbridge
dry dock
duck pond
earmuff
earthquake
eggplant
eyebrow
fan belt
farmyard
fence post
filmstrip
flagship
footprint
forecast
Fort Knox
fox hunt
French fry
fullback
gangplank
gas tank
gearshift
girl friend
gold mine
golf club
grand slam
grapefruit
great dane
greenhorn
greyhound
grindstone
groundhog
hailstorm
hairnet
half mast
handshake
hangnail
hardware
hatband
haystack
head start
heartbeat
hindsight
home run
hoot owl
hopscotch
hot dog
hourglass
housefly
hubcap
ice cube
inchworm
infield
inkwell
jackknife
jailbreak
jawbone
jigsaw
joyride
June bug
keepsake
key punch
kinfolk
King Kong
knickknack
knockout
Kool Aid
lamb chop
lampshade
landlord
lawsuit
leap year
light bulb
lip-read
livestock
lukewarm
matchstick
meat hook
milkweed
moonshine
mousetrap
mouthwash
mushroom
muskrat
nearby
necktie
New York
nightmare
northwest
nursemaid
oatmeal
okay
old age
one way
oxbow
padlock
paintbrush
pantsleg
password
pawnshop
payroll
pea pod
phone booth
pickaxe
piecrust
pigpen
pinecone
Ping-Pong
pinwheel
pitchfork
pole vault
pup tent
racehorse
rag doll
railroad
ramrod
rawhide
Red Cross
reindeer
rib cage
roast beef
root beer
rosebud
round steak
rye bread
safeguard
St. Paul
sandbar
scapegoat
scarecrow
sea gull
shamrock
sheepskin
shirtsleeve
shoebench
shortstop
sidewalk
ski lift
slide rule
slingshot
slipknot
snare drum
snowplow
soapsuds
soup bowl
southeast
soybean
spare tire
spearmint
square dance
starfish
steamboat
sunburst
swan dive
Swiss cheese
switchblade
tailspin
takeoff
teacup
teamwork
test tube
textbook
throw rug
thumbtack
tightrope
time bomb
tin can
toadstool
topsoil
toothpaste
town hall
tow truck
trademark
train track
turnpike
upset
viewpoint
wardrobe
warpath
whiplash
whirlpool
whisk broom
wholesale
wigwam
wildlife
windmill
wolf pack
woodchuck
wristwatch
X ray
yuletide
zigzag

Appendix 1. The stimulus vocabulary of Study I.

		-36	-30	-24	-18	-12	-6	0	dB S/N
audio-visual recognition	BB	48.0	52.4	50.4	59.6	74.0	91.0		% correct
	JF	43.6	44.4	52.8	64.8	84.4	96.0		
	VM	77.6	75.6	76.0	81.6	92.8	99.2		
	JS	40.0	41.6	42.8	54.0	66.8	82.0		
	SS	42.0	40.0	42.0	51.2	75.2	87.6		
○	$\bar{x}$	50.2	50.8	52.8	62.2	78.6	91.2		
	sd	15.6	14.7	13.8	12.0	10.1	6.8		
auditory recognition	BB				3.6	22.0	51.2	86.8	
	JF				2.0	25.6	62.0	92.4	
	VM				1.6	20.0	59.2	92.8	
	JS				3.2	16.0	44.0	74.8	
	SS				1.2	13.6	45.6	83.6	
□	$\bar{x}$				2.3	19.4	52.4	86.1	
	sd				1.0	4.8	8.0	7.4	

Appendix 2. Audio-visual and auditory recognition of 250 spondaic words in wide-band noise by 5 normal-hearing adults.

Study I

airplane

baseball

bathtub

birthday

blackboard

boy scout

flagpole

girl friend

hot dog

ice cream

lunchroom

paintbrush

playground

school bus

sidewalk

toothpaste

Appendix 3. The stimulus vocabulary of Study II.

		-28	-25	-22	-19	-16	-13	-10	-7	-4	-1	+2	dB S/N
◇	RB	50.0	50.8	58.6	64.1	71.1	93.0	100.0					% correct
□	SE	48.4	48.4	57.8	59.4	78.9	96.9	100.0					
○	JF	50.8	52.3	50.8	55.5	76.6	93.7	99.2					
▽	RH	50.8	50.0	50.8	56.2	74.2	93.0	96.1					
△	JS	50.0	48.4	49.2	59.4	71.1	95.3	100.0					
	$\bar{x}$	50.0	50.0	53.4	58.9	74.4	94.4	99.1					
◇	MH						52.3	61.7	68.7	93.7	97.6	97.6	
□	ON						47.6	60.9	76.6	93.0	97.6	99.2	
○	TS						52.3	67.2	89.1	99.2	100.0	100.0	
▽	MS						52.3	66.4	91.4	97.6	100.0	100.0	
△	JS						51.6	62.5	75.8	95.3	96.9	99.2	
	$\bar{x}$						51.2	63.7	80.3	95.8	98.4	99.2	

Appendix 4a. Detection of 16 spondaic words in wide-band noise by 5 normal-hearing adults and by 5 profoundly deaf children.

Study II

		-28	-25	-22	-19	-16	-13	-10	-7	-4	-1	+2	dB S/N
○	NT-R							55.5	71.9	91.4			%
◉	NT-L				54.7	64.1	87.5						correct
	BD						50.0	57.0	69.5	81.2	87.5	93.0	
	RP						53.1	57.0	59.4	75.8	84.4	90.6	
	NW						48.4	55.5	58.6	75.0	85.1	96.1	
	AW						50.8	53.9	61.7	63.3	85.9	95.3	
	DZ						49.2	49.2	60.9	66.4	76.6	88.3	
□	$\bar{x}$						50.3	54.5	62.0	72.3	83.9	92.7	
△	DA					64.1	85.9	96.1					
◇	WJ							67.2	85.9	99.2			
⊞	JW							76.6	92.2	98.4			

Appendix 4b. Detection of 16 spondaic words in wide-band noise by 1 adult with a normal-hearing ear (L) and a deaf ear (R), by 5 normal-hearing adults (tactile response from palm of hand), and by 3 children with moderate-to-severe hearing losses.

Study II

MONOSYLLABIC WORDS		TROCHAIC WORDS		SPONDAIC WORDS	
arm	mouth	address	letter	airplane	mushroom
back	nail	answer	money	baseball	New York
bear	name	apple	morning	bathtub	northwest
bed	nest	April	mother	birthday	notebook
bell	night	arrow	mountain	blackboard	oatmeal
bird	nose	August	napkin	blue jay	outside
bug	page	baby	needle	bullfrog	paintbrush
cake	pan	barber	neighbor	caveman	Ping-Pong
car	pipe	Bible	nothing	chipmunk	playground
cat	plate	body	number	clothesline	popcorn
chair	queen	bottle	ocean	cough drop	pork chop
coat	rat	breakfast	office	cowboy	racehorse
desk	rock	brother	onion	cub scout	railroad
dish	roof	building	orange	doughnut	reindeer
door	rope	button	paper	downstairs	roast beef
duck	rug	camel	people	earphone	root beer
fire	salt	candy	picture	earthquake	runway
floor	sand	chicken	pillow	eggshell	rye bread
food	sheep	children	question	eyebrow	sailboat
fork	ship	chimney	rabbit	farmyard	school bus
game	shirt	Christmas	river	flagpole	shotgun
glass	shoe	circle	rubber	footprint	sidewalk
glove	smoke	city	sailor	French fry	snowflake
goat	snake	color	salad	gas tank	soup bowl
hand	speech	cotton	science	girl friend	southeast
hat	spoon	cousin	sister	gold mine	starfish
heart	stamp	diamond	soldier	golf club	stop sign
hill	state	doctor	spider	grapefruit	string bean
hole	stick	elbow	squirrel	greyhound	sunset
house	street	father	stomach	groundhog	Swiss cheese
king	tail	finger	sugar	haircut	takeoff
knife	tongue	flower	supper	headache	teacup
lake	tree	forest	table	home run	thumbtack
lamp	voice	garden	teacher	hot dog	toothpaste
leaf	week	husband	wagon	ice cube	tow truck
lunch	wheel	jacket	water	light bulb	train track
meat	wife	July	weather	lipread	whipped cream
month	wire	kitten	window	mailbox	windmill
moon	worm	language	winter	milkshake	wristwatch
mop	year	lemon	woman	mousetrap	X ray

Appendix 5. The stimulus vocabulary of Study III.

		-30	-25	-20	-15	-10	-5	dB S/N
□	SB	50.8	65.4	83.8	95.4			% correct
○	JK	55.4	76.7	90.8	97.5			
△	LM	53.3	67.5	83.8	97.5			
	$\bar{x}$	53.2	69.9	86.1	96.8			
◫	KB			67.5	87.9	97.0	97.5	
⦶	SH			64.5	87.5	95.4	97.5	
⚠	PT			54.1	79.1	95.4	97.9	
	$\bar{x}$			62.0	84.8	95.9	97.6	
■	MK			48.7	52.0	66.2	87.9	
●	LS			50.0	60.0	79.6	95.8	
▲	JT			49.1	60.0	77.0	91.2	
	$\bar{x}$			49.3	57.3	74.3	91.6	

Appendix 6a. Auditory detection (i.e., 2-alternative forced-choice) of 240 words in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children.

		-25	-20	-15	-10	-5	0	+5	dB S/N
□	SB	2.1	8.3	52.9	90.4				% correct
○	JK	3.7	21.3	65.8	93.3				
△	LM	2.5	15.8	63.3	92.1				
	$\bar{x}$	2.8	15.1	60.7	91.9				
◫	KB			2.1	8.7	32.1	49.2*		
⦶	SH			4.2	10.0	46.3	57.9		
◬	PT			2.9	15.4	47.9	64.6		
	$\bar{x}$			3.1	11.4	42.1	57.2		
■	MK				0.0	0.0	1.3	1.7	
●	IS				0.0	0.0	1.3	1.7	
▲	JT				0.4	0.8	2.5	4.6	
	$\bar{x}$				0.1	0.3	1.7	2.7	

Appendix 6b. Auditory recognition of 240 words in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children.

Study III

*

Even at the highest S/N ratios examined, the auditory-recognition scores of the 3 severely hearing-impaired subjects are lower than one might have predicted from their speech discrimination scores in the quiet (Table 3). Some possible causes for this discrepancy are: (1) unequal size of the response vocabulary (i.e., 240 vs 25 words, respectively) (Miller, Heise, and Lichten, 1951); (2) subjects' greater familiarity with the smaller list; (3) presentation by different talkers (Dale, 1967); (4) overload of the auditory system by high noise levels (Pollack and Pickett, 1958).

		V-only	-40	-35	-30	-25	-20	-15	-10	-5	0	+5	dB S/N
□	SB	27.9	30.4	33.7	35.8	47.1	62.1	82.5	90.8				% correct
○	JK	46.7	52.1	52.1	56.2	67.5	84.2	93.7	97.1				
△	LM	30.0	37.9	35.8	39.6	50.8	75.0	86.7	93.7				
	$\bar{x}$	34.9	40.1	40.5	43.9	55.1	73.8	87.6	93.9				
◫	KB	52.1				57.5	58.7	65.8	69.2	78.7	86.7	88.3	
⦶	SH	53.3				59.6	60.0	70.0	76.7	92.1	94.2	96.2	
⚠	PT	57.1				62.5	65.0	70.0	81.2	93.7	94.6	95.8	
	$\bar{x}$	54.2				59.9	61.2	68.6	75.7	88.2	91.8	93.4	
■	MK	28.7				32.9	34.2	33.7	35.0	38.3	40.0	42.1	
●	LS	51.2				56.2	55.8	55.8	57.5	61.2	60.0	62.9	
△	JT	57.1				62.5	61.2	64.2	64.6	66.2	70.8	73.7	
	$\bar{x}$	45.7				50.5	50.4	51.2	52.4	55.2	56.9	59.6	

Appendix 6c. Audio-visual recognition of 240 words in low-frequency noise by 3 normal-hearing children, 3 severely hearing-impaired children, and 3 profoundly deaf children.

Study III

BIBLIOGRAPHY

Clarke, B.R., Use of a group hearing aid by profoundly deaf children. In Ewing, A.W.G. (Ed.), Educational Guidance and the Deaf Child. Washington: Volta Bureau, 128-159 (1957).

Dale, D.M.C., Applied Audiology for Children. Springfield, Ill.: Thomas (1967).

Dewey, G., Relativ Frequency of English Speech Sounds. Cambridge: Harvard Univ. Press (1923).

Elliott, L.L., Descriptive analysis of audiometric and psychometric scores of students at a school for the deaf. J. Speech Hearing Dis., 10, 21-40 (1967).

Gault, R.H., Interpretation of spoken language when the feel of speech supplements vision of the speaking face. Volta Rev., 30, 379-386 (1928).

Hart, B.O. and Rosenstein, J., Examining the language behavior of deaf children. Volta Rev., 66, 679-682 (1964).

Hawkins, J.E. and Stevens, S.S., The masking of pure tones and of speech by white noise. J. acoust. Soc. Amer., 22, 6-13 (1950).

Heider, F., Acoustic training helps lipreading. Volta Rev., 45, 135, 180 (1943).

Hirsh, I.J. and Bowman, W.D., Masking of speech by bands of noise. J. acoust. Soc. Amer., 25, 1175-1180 (1953).

Hirsh, I.J., Reynolds, E.G., and Joseph, M., Intelligibility of different speech materials. J. acoust. Soc. Amer., 26, 530-538 (1954).

Hopkins, L.A. and Hudgins, C.V., The relationship between degree of deafness and response to acoustic training. Volta Rev., 55, 32-35 (1953).

Hudgins, C.V., Progress report on an acoustic training experiment. Volta Rev., 55, 35-38 (1953a).

Hudgins, C.V., The response of profoundly deaf children to auditory training. J. Speech Hearing Dis., 18, 273-288 (1953b).

Hudgins, C.V., Auditory training: Its possibilities and limitations. Volta Rev., 56, 339-349 (1954).

Hutton, C., Combining auditory and visual stimuli in aural rehabilitation. Volta Rev., 61, 316-319 (1959).

Jerger, J.F., Tillman, T.W., and Peterson, J.L., Masking by octave bands of noise in normal and impaired ears. J. acoust. Soc. Amer., 32, 385-390 (1960).

Johnson, G.F., The effects of cutaneous stimulation by speech on lipreading performance. Unpublished doctoral dissertation, Michigan State Univ. (1963).

Lightfoot, C., Carhart, R., and Gaeth, J.H., Masking of impaired ears by noise. J. Speech Hearing Dis., 21, 56-70 (1956).

Ling, D., An auditory approach to the education of deaf children. Audecibel, 13, 96-101 (1964).

Ling, D., The use of low frequency residual hearing in profoundly deaf children. Unpublished master's thesis, McGill Univ. (1966).

Lowell, E.L., Research in speech reading: Some relationships to language development and implications for the classroom teacher. Report of the Proceedings of the 39th Meeting of the Convention of American Instructors of the Deaf, 68-73 (1959).

Miller, G.A., The masking of speech. Psych. Bull., 44, 105-129 (1947).

Miller, G.A., Heise, G.A., and Lichten, W., The intelligibility of speech as a function of the context of the test materials. J. Exper. Psych., 41, 329-335 (1951).

Miller, G.A. and Nicely, P.E., An analysis of perceptual confusions among some English consonants. J. acoust. Soc. Amer., 27, 338-352 (1955).

Murray, N.E., Interim report on hearing aids and classification of deaf children. Report CAL-IR-2 (Australia) (1951).

Neely, K.K., Effect of visual factors on the intelligibility of speech. J. acoust. Soc. Amer., 28, 1275-1277 (1956).

Niemoeller, A.F., Acoustical design of classrooms for the deaf. Amer. Ann. Deaf, 113, 1040-1045 (1968).

Nober, E.H., Air and bone conduction thresholds of deaf and normal hearing subjects before and during the elimination of cutaneous-tactile interference with anesthesia. Final report: Project #6-3073, Syracuse Univ. (1968).

Numbers, M.E. and Hudgins, C.V., Speech perception in present-day education for deaf children. Volta Rev., 50, 449-456 (1948).

O'Neill, J.J., Contributions of the visual components of oral symbols to speech comprehension. J. Speech Hearing Dis., 19, 429-439 (1954).

O'Neill, J.J. and Oyer, H.J., Visual Communication for the Hard of Hearing. Englewood Cliffs, N.J.: Prentice-Hall (1961).

Oyer, H.J., An experimental approach to the study of lipreading. Report of the Proceedings of the 41st Meeting of the Convention of American Instructors of the Deaf, 322-326 (1963).

Pickett, J.M., Effects of vocal force on the intelligibility of speech sounds. J. acoust. Soc. Amer., 28, 902-905 (1956).

Pickett, J.M., Tactual communication of speech sounds to the deaf: Comparison with lipreading. J. Speech Hearing Dis., 28, 315-330 (1963).

Pickett, J.M. and Martin, E.S., Some comparative measurements of impaired discrimination for sound spectral differences. Amer. Ann. Deaf, 113, 259-267 (1968).

Pickett, J.M. and Pollack, I., Intelligibility at high voice levels and the use of a megaphone. J. acoust. Soc. Amer., 30, 1100-1104 (1958).

Pollack, I., Message procedures for unfavorable communication conditions. J. acoust. Soc. Amer., 30, 196-201 (1958).

Pollack, I. and Pickett, J.M., Masking of speech by noise at high sound levels. J. acoust. Soc. Amer., 30, 127-130 (1958).

Prall, J., Lipreading and hearing aids combine for better comprehension. Volta Rev., 59, 64-65 (1957).

Sumby, W.H. and Pollack, I., Visual contribution to speech intelligibility in noise. J. acoust. Soc. Amer., 26, 212-215 (1954).

Van Uden, A., A sound-perceptive method. In Ewing, A.W.G. (Ed.), The Modern Educational Treatment of Deafness. Washington: Volta Bureau, 19/3-19/12 (1960).

Watson, T.J., Speech audiometry for children. In Ewing, A.W.G. (Ed.), Educational Guidance and the Deaf Child. Washington: Volta Bureau, 278-296 (1957).

Watson, T.J., The use of residual hearing in the education of deaf children. Volta Rev., 63, 328-334, 385-392, 435-440, 487-492 (1961), 64, 31-38, 84-88 (1962).

Watson, T.J., The use of hearing aids by hearing-impaired pupils in ordinary schools. <u>Volta Rev.</u>, 66, 741-744, 787 (1964).

Whetnall, E., Use of hearing by deaf children. <u>Report of the Proceedings of the 41st Meeting of the Convention of American Instructors of the Deaf</u>, 271-279 (1963).

Whitehurst, M.W., Integration of lipreading, auditory training, and hearing aids. <u>Volta Rev.</u>, 66, 730-733 (1964).