An Analysis of Spondee Recognition Thresholds in Auditory-only and Audio-visual Conditions

Brenene Marie Brady-Herbst
Portland State University

Recommended Citation

Brady-Herbst, Brenene Marie, "An Analysis of Spondee Recognition Thresholds in Auditory-only and Audio-visual Conditions" (1996). *Dissertations and Theses.* Paper 5218.
https://doi.org/10.15760/etd.7094

THESIS APPROVAL

The abstract and thesis of Brenene Marie Brady-Herbst for the Master of

Science in Speech Communication: Speech and Hearing Science were

presented February 16, 1996, and accepted by the thesis committee and the

department.

COMMITTEE APPROVALS:

                    , Chair
Doug Martin

Mary Gordon-Brannon

William Fisher
Representative of the Office of GraduateSudies

DEPARTMENT APPROVAL:
Rhea Paul, Chair
Department of Speech Communication

****

ACCEPTED FOR PORTLAND STATE UNIVERSITY BY THE LIBRARY

by _____ on *7 March 1996*

An abstract of the thesis of Brenene Marie Brady-Herbst for the Master of

Science in Speech Communication: Speech and Hearing Science presented

February 16, 1996.

Title: An Analysis of Spondee Recognition Thresholds in Auditory-Only and

Auditory-Visual Conditions.

To date there are no acceptable speechreading tests with normative or

psychometric data indicating the test is a valid and reliable measure of

speechreading assessment. Middlewerd and Plomp (1987) completed a

study of speechreading assessment using sentences (auditory-only and

auditory-visual) in the presence of background noise. Results revealed

speech reception thresholds to be lower in the auditory-visual condition.

Montgomery and Demorest (1988) concurred that these results were

appealing, but unfortunately not efficient enough to be used clinically.

The purpose of this study was to develop a clinically valid and reliable

assessment of speechreading ability, following Middlewerd and Plomp's

(1987) framework to achieve this goal. The method of obtaining a valid

assessment tool was to define a group of stimuli that can be administered and

scored to produce reliable data efficiently. Because spondaic words are

accepted as a reliable method of clinically achieving speech reception

thresholds, they were chosen to be used as the stimuli in this study to develop an efficient clinical speechreading assessment tool.

Ten subjects were presented with spondaic words in each of two conditions, auditory-only and auditory-visual, in the presence of background noise. The spondee words were randomized for each presentation, to validate the data. A computerized presentation was used so that each subject received the identical input. The computer also produced a performance-intensity function for each spondaic word.

Results revealed an acceptable speech recognition threshold for 18 of the 36 spondee words in the auditory-only condition; 6 words were outside of one standard deviation; and the remaining 12 words did not produce obtainable thresholds. In the auditory-visual condition, all words except one had no obtainable threshold. Although these results invalidated the spondee words as an acceptable stimuli, the study does validate the foundation for further research to study different types of stimuli using this same framework.

AN ANALYSIS OF SPONDEE RECOGNITION THRESHOLDS IN

AUDITORY-ONLY AND AUDITORY-VISUAL CONDITIONS

by

BRENENE MARIE BRADY-HERBST

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
SPEECH COMMUNICATION:
SPEECH AND HEARING SCIENCE

Portland State University
1996

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

The use of modern hearing aids provides most listeners who are hearing-impaired with greatly improved speech detection and understanding. However, most listeners who are hearing impaired still find it necessary to use other skills, such as speechreading, to optimize their overall communication abilities (Montgomery & Demorest, 1988). Speechreading, as defined by Kaplan, Bally, and Garretson (1985), is "the ability to understand a speaker's thoughts by watching the movements of the face and body and by using information provided by the situation and the language" (p. 1). In other words, speechreading is the ability to recognize and understand speech based on both auditory and visual components using linguistic and contextual cues to assist in this understanding.

Test instruments and methods that are valid, reliable, and clinically efficient in assessing speechreading are vital for the clinical practice of aural rehabilitation (Montgomery & Demorest, 1988). There has been considerable research on the contribution to speech understanding of the auditory and visual components of speech (Summerfield, 1987). For instance, O'Neill, Sumby, and Pollack, (as cited in Silverman & Kricos 1990) reported on visual and auditory cues used in conjunction with speechreading during the 1950s. Since that time, many assessment procedures using the integration of auditory-visual cues have been attempted. Although numerous speechreading test paradigms have been proposed over the years, few

systematic investigations of their effectiveness have been undertaken (Montgomery & Demorest, 1988). As a result, there is currently a lack of reliable methods for accurately assessing speechreading abilities.

Middlewerd and Plomp (1987) have described a test that measures speech reception thresholds in the auditory and auditory-visual modalities. The speech reception thresholds were obtained for sentences in the presence of background noise. The auditory-visual modality revealed lower thresholds in both groups of subjects. Montgomery and Demorest (1988) described the problem with this testing procedure being the time involved to administer the test and difficulties in scoring; this calls into question the clinical effectiveness of this procedure. While limited additional work has been done with Middlewerd and Plomp's study , Montgomery and Demorest (1988), indicated that this measure may "yield an appealing and interesting measure of speechreading ability" (p.205). This suggests that the Middlewerd and Plomp study can be used as a model for the creation of a new speechreading test. There does, however, exist a need to define what would be an appropriate test stimuli for clinical effectiveness. This proposed study will look at one component of a new test using spondaic words rather then sentences as the test stimuli.

Statement of Purpose

This study is seen as the first in a series of research projects whose eventual goal is the development of a valid, reliable, and clinically efficient speechreading test. As a first step in development of this larger project, a need exists to define a group of words suitable for inclusion in the test. The

purpose of this study was to evaluate spondaic words presented in auditory-only and auditory-visual conditions to determine their suitability for use in this test. This was accomplished by determining groups of spondaic words that are homogenous for word recognition threshold in both auditory-only and auditory-visual conditions. To determine homogeneity of recognition threshold, complete performance intensity function across a range of signal-to-noise ratios were obtained for spondaic words presented in both auditory-only and auditory-visual conditions. Spondaic words that demonstrate recognition thresholds within one standard deviation of the group mean were considered to meet the criteria for homogeneity of recognition threshold. Therefore, the research question being investigated is: Can a group of Spondee words be defined which demonstrate homogeneity of recognition thresholds in both auditory-only and auditory-visual conditions when in the presence of background noise?

CHAPTER II

REVIEW OF THE LITERATURE

Why Speechreading Tests are Needed

The earliest speechreading tests grew out of the need to measure the results of speechreading training provided to children and adults enrolled in schools for the deaf and in early aural rehabilitation programs for the hearing-impaired (Silverman & Kricos, 1990). As the use of speechreading education grew, so did concern for assessing speechreading ability. As speechreading instruction has become an important component of overall aural rehabilitation programs (Montgomery & Demorest, 1988), the need for a valid and reliable test of speechreading has increased.

As Montgomery and Demorest (1988) asserted, such tests are necessary, (a) to determine speechreading performance prior to instruction, (b) to form an impression of potential for improvement as a result of instruction, and (c) to monitor progress as education is given. Additionally, because a cochlear implant often functions as an aid to speechreading, accurate speechreading tests are needed in the selection of candidates for cochlear implants (Pickett & McFarland, 1985) .

Current Speechreading Tests

Many speechreading tests are used today. There is, however, no single speechreading test which is used throughout the country, nor has any test been adopted as a norm to which other speechreading tests can be

compared. Speechreading tests vary widely in their content, method of presentation, and method of scoring. Some tests use words, while others use sentences or even stories. Some are on film or videotape, while others are "face-to-face" tests. Some are "silent" while others combine visual and auditory cues. Some were developed for children and some for adults.

The first actual filmed lipreading test was developed by Utley in 1946, (as cited in Hull, 1982). This test consisted of a silent film containing three segments: Segment one used words, segment two used sentences, and segment three consisted of a five-part story test. Scoring was based on the number of words correctly identified. Sentence and story tests were scored correct if their content was perceived with reasonable accuracy (Hull, 1982).

Another filmed speechreading test, the Keaster Film Test of Lipreading, was developed by Jacqueline Keaster of the Department of Otolaryngology, State University of Iowa. Her original test was "silent" and consisted of six forms of 10 sentences each, with each form presented by a different speaker (3 males and 3 females). These original films were later reassembled at the John Tracy Clinic into a two-form filmed test with each form consisting of 30 sentences given without voice. Scoring was based on the number of words correctly identified (Jeffers & Barley, 1971). Donnelly and Marshall later used material derived from the John Tracy "Film Test Of Lipreading" to develop a filmed test that was multiple-choice in format. It differed from previous filmed tests in that one form included sound (Hull, 1982).

Unlike the above described tests, the Barley Speechreading Test was not a filmed test, but rather was designed to be presented face-to-face but

without "voice." The Barley Test consists of sentences selected from the CID Everyday Speech Sentences lists, with scoring based on the number of words correctly identified (Jeffers & Barley, 1971).

A more recently developed test is the NAL/West Haven modification (Spitzer, Leder, Milner, Flevaris-Phillips, & Giolas 1987). This test is based on Australia's NAL speechreading stimuli but with American words substituted for certain Australian colloquialisms. In this test, the subject is given a topic with several related questions. Scoring is based on the number of correctly answered questions.

Problems in the Assessment of Speechreading Ability

Despite the many attempts to develop valid tests of speechreading, Montgomery and Demorest (1988) proposed six major problem areas with the current speechreading assessment methods. These problems interact to make accurate assessment of speechreading difficult, but not impossible.

Current speechreading tests suffer from the same difficulties associated with testing auditory speech reception. These difficulties, arising from the lack of an adequate theoretical model of speech perception, include (a) the difficulty of approximating real-life situations, (b) uncertainty over the most appropriate and sensitive material to use, and (c) controversy over whether it is best to test low-level perceptual skills using nonsense syllables or single words or to test high-level linguistic skills using sentences or connected speech (Montgomery & Demorest, 1988).

Current speechreading tests correlate very poorly with traits such as intelligence, verbal ability, visual memory, gender, education, and

socioeconomic status (Montgomery & Demorest, 1988). Montgomery and Demorest (1988) believe it is unclear whether these poor correlations are due to the poor quality of speech tests or the "the possibility that speechreading is a unique and independent skill unrelated to almost all other human activity"   (p. 196).

The measurement of change is another area in which current speechreading tests demonstrate difficulties. The nature and degree of the problem depends on the concepts, goals, assumptions, and statistics involved in the specific situation being examined. Montgomery and Demorest (1988) comment in detail on a number of typical clinical and research situations, pointing out the difficulties with each. Examples of these difficulties include (a) improvement after instruction, (b) relationship between pretest and improvement, (c) examination of extreme cases, (d) comparison of teaching methods, and (e) individual differences in improvement.

A fourth problem area relates to the lack of understanding of how visual and sensory information is integrated during the speechreading task. Many researchers, as cited by Rigo (1986), have studied the visual contribution to speech recognition, concluding that "visual cues improve speech comprehension when auditory processing is limited due to signal degradation or hearing loss" (p. 196). Erber (1975) demonstrated that the importance of the visual contribution to speech comprehension increased as the auditory channel is progressively impaired. In 1986, Massaro, Thompson, Barron, and Laren conducted a study demonstrating the bimodal use of auditory and visual information in speech perception among normal-hearing subjects. A positive correlation was found between speechreading

ability and the visual contribution magnitude (defined as the difference
between auditory-visual and auditory-only test scores). Despite these
studies, no valid tool has been described which adequately incorporates this
information into a clinical test procedure.

Citing numerous studies by researchers attempting to develop
auditory-visual tests, Montgomery and Demorest (1988) concluded that, until
the process of auditory and visual integration is better understood, good
tests of auditory-visual speech recognition will be difficult to devise. While
recognizing the difficulties associated with auditory-visual testing, they
concluded it is "obvious" that "realistic" speechreading tests be auditory-
visual in form. In real-life situations, both normal and hearing-impaired
individuals use a combination of visual and auditory cues to comprehend
speech.

The fifth problem area as defined by Montgomery and Demorest
(1988) is related to intertalker differences. The visual signal presented to a
speech reader is highly dependent on the specific talker employed.
Speakers vary in the amount of lip movements, amount of mouth opening,
use of nonstandard speech movements, and rate of speech. Differences
between talkers add to the difficulty of constructing a valid and reliable
speech test. Use of a single talker does not represent the real world, but the
use of more than one talker adds to the expense and effort involved in
developing and conducting speechreading tests.

A sixth problem area relates to scoring test results. Disagreements
exist as to the proper linguistic unit to use, e.g., phoneme, syllable, word, key
word in sentence, sentence meaning, and intelligibility of continuous

discourse. Added to this is the problem of visual similarity of various sounds, /p,b,m/ or /f,v/ for example, and the difficulties these present in scoring (Montgomery & Demorest, 1988). For example, in order for a subject to select the correct word from /bat/, /pat/, and /mat/ (homophones, words being visually identical, yet having distinctly different meanings), without the redundancy of a sentence, the subject would have to guess, hence the test would not be measuring speechreading ability.

Montgomery and Demorest (1988) concluded that, given the problems discussed above, it is not surprising that clinically usable tests of speechreading have not been developed. They stress, however, that the existence of these problem areas does not imply that current tests are inadequate, only that there is not sufficient psychometric data on individual tests to evaluate them. This situation, they contended, could be partially alleviated if speechreading tests were examined for internal-consistency reliability, retest reliability, interlist equivalence, and susceptibility to practice effects. They also emphasized that with the current technology available, both in computer video hardware and statistical software, there is every reason to believe that it is possible to develop improved speechreading assessment techniques that are reliable and valid.

One recently developed assessment method that Montgomery and Demorest (1988) feel holds much promise is a method described by Middlewerd and Plomp (1987) in which the speech reception threshold (SRT) for sentences in noise was determined in two modes: an auditory-visual mode (AV), and an auditory-only (A) mode. SRTs for the combined auditory-visual mode were 4 dB lower for young normal hearing adults and

4.6 dB lower for elderly presbycusic adults compared to auditory-only presentation. This difference in dB between the two modes (AV - A) represented the visual contribution to speech understanding.

The AV-A difference score "yields an appealing and meaningful measure of speechreading" (Montgomery & Demorest, 1988, p. 205) if this method can be shown to be reliable. This test can be considered a meaningful measure of speechreading because a positive correlation between speechreading ability and the AV-A difference was demonstrated by Massaro et al. (1986). To date, however, no study has been published which duplicates or further examines the reliability of the Middlewerd and Plomp (1987) method.

The use of sentences, the generation of a complete performance-intensity (P-I) function, and the use of open-set scoring make the Middlewerd and Plomp test very time consuming. For this reason, even if their method is found to be reliable, the time required to administer and score the test would make it impractical for clinical use.

The purpose of this study is to help lay the foundation for a valid, reliable and practical speechreading test. This test will be based on the project reported by Middlewerd and Plomp (1987) although it will differ in three significant ways. The first is in the method used for determining the SRT, the second is the linguistic unit used as stimuli, and the third is in the method of scoring. It is suggested that by employing these procedural changes this study will provide information about a method that will prove to be time-efficient as well as reliable.

Middlewerd and Plomp (1987) employed a complete performance-intensity function, with a total of 130 to 143 sentences presented to the subject in order to determine the SRTs for the two modes of presentation. This method is valid and reliable, but not feasible for clinical use.

To reduce the time needed to determine SRTs , the threshold search procedure described by Wilson, Morgan, and Dirks (1973) was used for this proposed test. This has been shown to be a valid and reliable method for achieving speech reception thresholds by Young and his associates , Young et al. (1982). This procedure is also easy to administer, with accurate threshold levels reached in 5 minutes or less.

Middlewerd and Plomp (1987) used sentences as their test stimuli. A total of 10 or 11 lists consisting of 13 sentences each were required for each subject. Four to 5 sentence lists were used in a preliminary test session to determine the proper signal-to-noise ratio to be used for the actual SRT measurement. Following the preliminary test, subjects were presented 3 sentence lists in auditory-only mode and 3 sentence lists in combined auditory-visual mode to determine the SRT for each mode. The use of sentences makes this method very time-consuming and difficult to score.

In this proposed test, spondaic words will be used as the test stimuli. Spondees are used in clinical audiology as a reliable method to determine speech reception thresholds. They are easily understandable, equally understandable, and quick and easy to use (Hodgson, 1980). In addition, Carhart, in 1946 (as cited in Hodgson, 1980)., found that SRTs obtained using spondees were equivalent to SRTs obtained using sentences.

In Middlewerd and Plomp's study (1987), SRT was defined as the signal-to-noise (S/N) ratio for which 50% of the syllables of a list of 13 sentences were correctly repeated. This open-set scoring method, with scores calculated on the number of correct syllables repeated by the subjects is very time-consuming and impractical for clinical use. A closed-set of responses has been chosen for this proposed test to reduce the time needed for scoring the test. A limited set of 36 spondaic words will be used, with all test stimuli to be drawn from this set. Subjects will choose from a closed-set of eight spondees for each stimuli presented.

To evaluate the feasibility of this approach, a brief pilot study was undertaken (Martin, personal communication, 1993). The threshold search procedure described by Wilson et al. (1973) was used to determine the average S/N ratio at which the subjects correctly identified 50% of the spondees presented. The carrier phrase "the word is _____" was used for presentation of each spondaic word. Test stimuli were presented live under two conditions: auditory-only and auditory-visual. Four subjects were tested.

The results of this pilot study revealed the average S/N ratio at which 50% correct responses were obtained were +8 dB for the auditory-only condition and -2 dB for the auditory-visual condition. An average AV-A difference of 6 to 8 dB was obtained. SRTs were always lower in the combined auditory-visual condition than in the auditory-only condition. This is consistent with the results reported by Middlewerd & Plomp (1987), suggesting that the proposed test may provide valuable information regarding the reliability of this method of speechreading assessment.

Summary

Speechreading, or the use of visual cues to assist in the recognition and understanding of speech, is a skill which is not entirely understood on a theoretical basis. It is, however, a skill that is used by normal hearing and hearing-impaired individuals alike to aid in speech recognition and understanding, especially in less than ideal listening environments. Valid and reliable tests for assessing speechreading ability are vital for the clinical practice of aural rehabilitation.

Over the years, many speechreading tests have been developed. Currently there is not a clinically efficient speechreading test that is considered valid and reliable. Current speechreading tests in clinical use suffer from a number of problems relating to their linguistic complexity, lack of correlation with psychometric traits, scoring methods, integration of auditory and visual sensory information, and mode of presentation.

An assessment method used by Middlewerd and Plomp (1987) appears to provide a meaningful measure of speechreading which resolves many of the problems cited above. Their procedure, however, is not clinically feasible and, because it has not been duplicated by other researchers, has not yet proven to be reliable.

The study presented here will address the issue of desirable stimuli for the proposed test. This proposed test is the first step in determining the feasibility of spondaic words as appropriate test stimuli. This was accomplished by obtaining information on the audibility of spondee words presented in auditory-only and auditory-visual modes. The information

derived was used to construct spondee word lists for the next phase of research. In effect, this study has been designed as the first in a series of studies which it is hoped will eventually result in a speechreading test which is valid, reliable, and clinically efficient.

CHAPTER III

METHODS

Subjects

Ten normal hearing subjects were tested in this study. All subjects demonstrated audiometric thresholds of <15 dB HL (re: ANSI 3.6-1969) at octave frequencies from 500 through 8000 Hz. Table 1 displays the subjects' audiometric thresholds. Subjects were between the ages of 19 and 33 years. Additionally, all subjects reported negative histories of otologic or neurologic dysfunction and reported vision that was normal or corrected to normal

Stimuli

Test stimuli consisted of the 36 spondaic words suggested for use in determining thresholds for speech by Hirsch, Davis, Reynolds, Eldert, and Benson (1952). These spondees are listed in Table 2. The spondees were presented with the carrier phrase "The word is _____." Because recorded tests offer controlled replication of stimuli in contrast to live presentation, the stimuli were videotaped, and subsequently pressed onto a videodisc.

The recording conditions used in the making of the videodisc were consistent with those described by Sims (1988). The speaker was a female judged to speak in general American dialect with normal articulation patterns. A lapel microphone was used by the speaker, and the recording was completed in a quiet environment. The speaker wore appropriate make-up to prevent the bright lighting from washing out her facial features.

Appropriate lighting was used to eliminate shadows on the speaker's face, and a small spotlight was directed into the speaker's mouth to illuminate the pharyngeal wall during articulation.

Table 1.

Mean pure tone thresholds (dB HL) for all subjects.

Frequency	Right ear	Left ear
250	6.5	5.0
500	3.5	2.0
1000	0.5	3.0
2000	1.0	1.5
4000	3.0	3.0
8000	4.5	6.0

Table 2.

The CID Spondees used in this study (Hirsch et al. 1952)

Airplane	Greyhound	Padlock
Armchair	Hardware	Pancake
Baseball	Headlight	Playground
Birthday	Horseshoe	Railroad
Cowboy	Hotdog	Schoolboy
Daybreak	Hothouse	Sidewalk
Doormat	Iceberg	Stairway
Drawbridge	Inkwell	Sunset
Duckpond	Mousetrap	Toothbrush
Eardrum	Mushroom	Whitewash
Farewell	Northwest	Woodwork
Grandson	Oatmeal	Workshop

The background behind the speaker, as well as the speaker's clothing, was neutral in color.

Upon completion of the final taping and editing, the taped stimuli was pressed into a master videodisc for use in actual administration of the test material. To create the necessary randomization for the repeated presentations of the test forms, the stimuli was administered via a computer-controlled videodisc system. This approach takes advantage of readily available technology to perform the necessary randomization of the test forms while avoiding the need to generate numerous test versions via cut-and-splice video tape techniques.

Procedures

Each subject was tested individually in a sound-treated room. The stimuli were presented to the subjects in auditory-only (videodisc monitor blacked out) and auditory-visual conditions. All auditory stimuli and noise were presented binaurally. In the auditory-only condition, the spondees were presented at a constant 65 dB HL and mixed with speech spectrum noise at levels from 50 to 64 dB HL, i.e., S/N ratios of +15 to +1 dB. In the auditory-visual condition, the spondees were presented at a constant 65 dB HL and mixed with speech spectrum noise at levels from 60 to 74 dB SPL, i.e., S/N ratios of +5 to -9 dB. The S/N ratio was adjusted in 2 dB steps from one condition to the next. The use of these signal-to noise ratios is based on information obtained in the pilot study. These S/N ratios represent a range of S/N levels from -7 dB to +7 dB from the mean S/N level at which the

subjects were able to identify correctly 50% of the spondee words presented to them.

Each subject's responses during their test sessions were tracked by the software program controlling the videodisc player. Subjects responded by touching their finger to the screen of the video monitor. Each subject was given a practice session to become familiar with this touch-to-select procedure.

The subjects were seated in an acoustically controlled test room in front of the video monitor. The TDH-49 earphones were placed in the appropriate position over the subject's ears. The 10 subjects were randomly divided into two groups. Group 1 was tested in the auditory-only condition first, followed by the auditory-visual condition, while Group 2 was tested first in the auditory-visual condition, followed by the auditory-only condition.

The subject was presented with each of the 36 spondees at each of the eight signal-to-noise ratios under each of the two test conditions. With the presentation of each test spondee, the subject was presented with a choice of the test spondee and seven response foils displayed on the monitor. The subject chose their response from this closed-set of 8 spondees. The 8 spondees were randomized by the computer for each presentation. The computer also scored each stimulus choice interval as either correct or incorrect.

Presentation of stimuli began at the most favorable S/N ratio for each condition and proceeded through each of the progressively more degraded S/N ratios. Each subject, therefore, was presented a total of 16 lists of 36 spondees each.

Instrumentation

This study was conducted in an acoustically treated test room (International Acoustics Corporation, Model 1403) located at Portland State University. A calibrated Grason Stadler audiometer (GSI 16) was used to gather audiometric information on each subject. This audiometer was also used to generate and deliver the speech spectrum noise and to route the audio test stimuli from the videodisc system to the subject via TDH-49 headphones.

A computer controlled videodisc system was used to administer the test stimuli. An Apple Centris 610 computer was used to run the software and to control a Panasonic V-2200 Laser Video Disc Player. The audio signal from the videodisc player was routed through the Grason Stadler audiometer (GSI 16) and to the subject via TDH-49 headphones. The video image was projected through a 25" Mitsubishi color monitor.

CHAPTER IV

RESULTS

Auditory Only Condition

Figures 1-36 show performance intensity functions for each of the words in the auditory only condition. The threshold is defined as the lowest level at which the word was identified as correct by 50% of the subjects. For the words Baseball, Birthday, Earthquake, Horseshoe, Ice-Cream, Mousetrap, Mushroom, Northwest, Outside, Stairway, Sunset, and Toothbrush, it was clearly not possible to determine a threshold as the worst recognition rate exceeded the 50% criteria. Of the words for which a threshold was identifiable, the mean signal-to-noise ratio for recognition threshold was 8.17 with a standard deviation of 1.2 dB, refer to table 3. Using the criteria of plus or minus one standard deviation from the mean (signal-to-noise ratio between 7-9 dB) to define suitable words for the speechreading assessment tool, the following words met the criteria: Armchair, Backbone, Blackboard, Cookbook, Cowboy, Doormat, Drawbridge, Duckpond, Eyebrow, Greyhound, Hardware, Hot-Dog, Inkwell, Oatmeal, Padlock, Pancake, Playground, and Railroad.

Auditory-Visual Condition

Figures 37-72 show the performance intensity function for each of the words in the auditory-visual condition. As can be noted, the results failed to consistently demonstrate any speech recognition threshold (rate never fell

below 50%) for all but one word, refer to table 4. The one word that did reveal a speech recognition threshold was the word headlight with a recognition threshold at a signal-to-noise ratio of +1 dB.

Table 3.Thresholds for Spondee words as a function of noise level (dB SPL). Auditory-Only condition

Words	Noise level	Words	Noise level
Airplane	+6	Hot Dog	+8
Armchair	+7	Ice Cream	N/A
Backbone	+8	Inkwell	+7
Baseball	N/A	Mousetrap	N/A
Birthday	N/A	Mushroom	N/A
Blackboard	+9	Northwest	N/A
Cookbook	+8	Nutmeg	+10
Cowboy	+7	Oatmeal	+8
Doormat	+8	Outside	N/A
Drawbridge	+9	Padlock	+7
Duckpond	+7	Pancake	+8
Eardrum	+6	Playground	+8
Earthquake	N/A	Railroad	+7
Eyebrow	+9	Stairway	N/A
Greyhound	+8	Sunset	N/A
Hardware	+7	Toothbrush	N/A
Headlight	+10	Whitewash	+10
Horseshoe	N/A	Woodwork	+6

Note. N/A-threshold could not be determined as recognition rate fell below 50%.
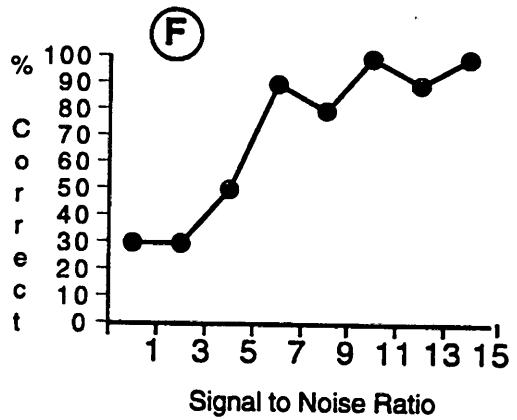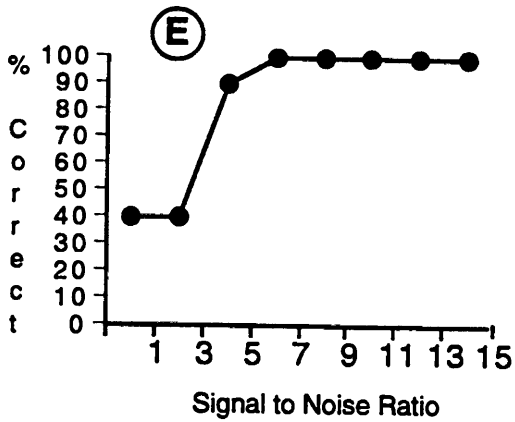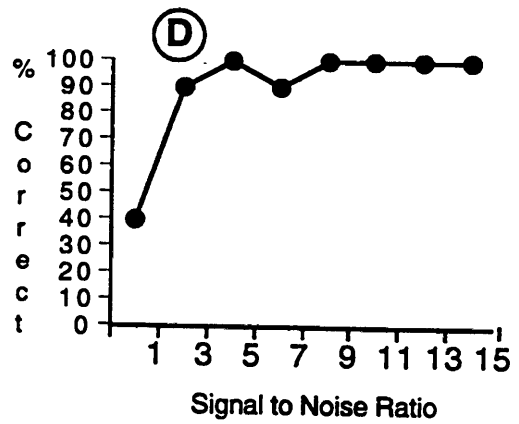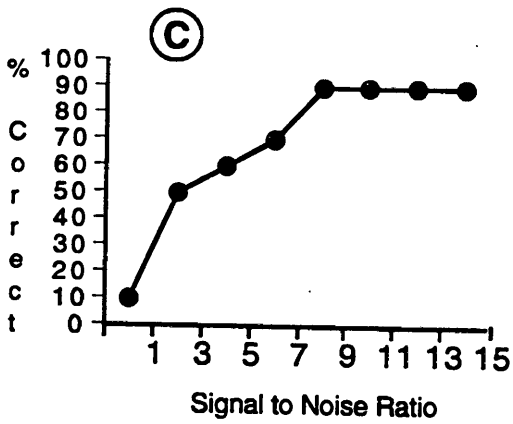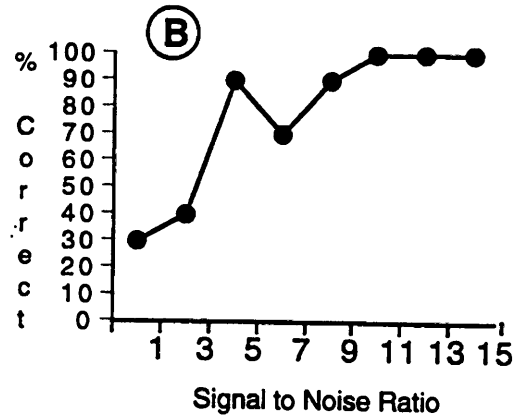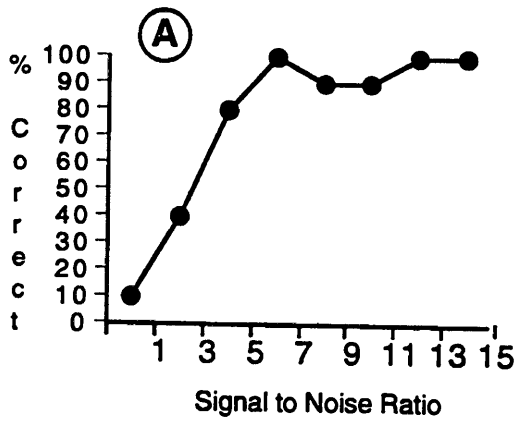
Figure 2. Performance-Intensity Functions for Doormat (A), Drawbridge (B), Duckpond (C), Eardrum (D), Farewell (E), and Grandson (F) in the Auditory-Only Mode.
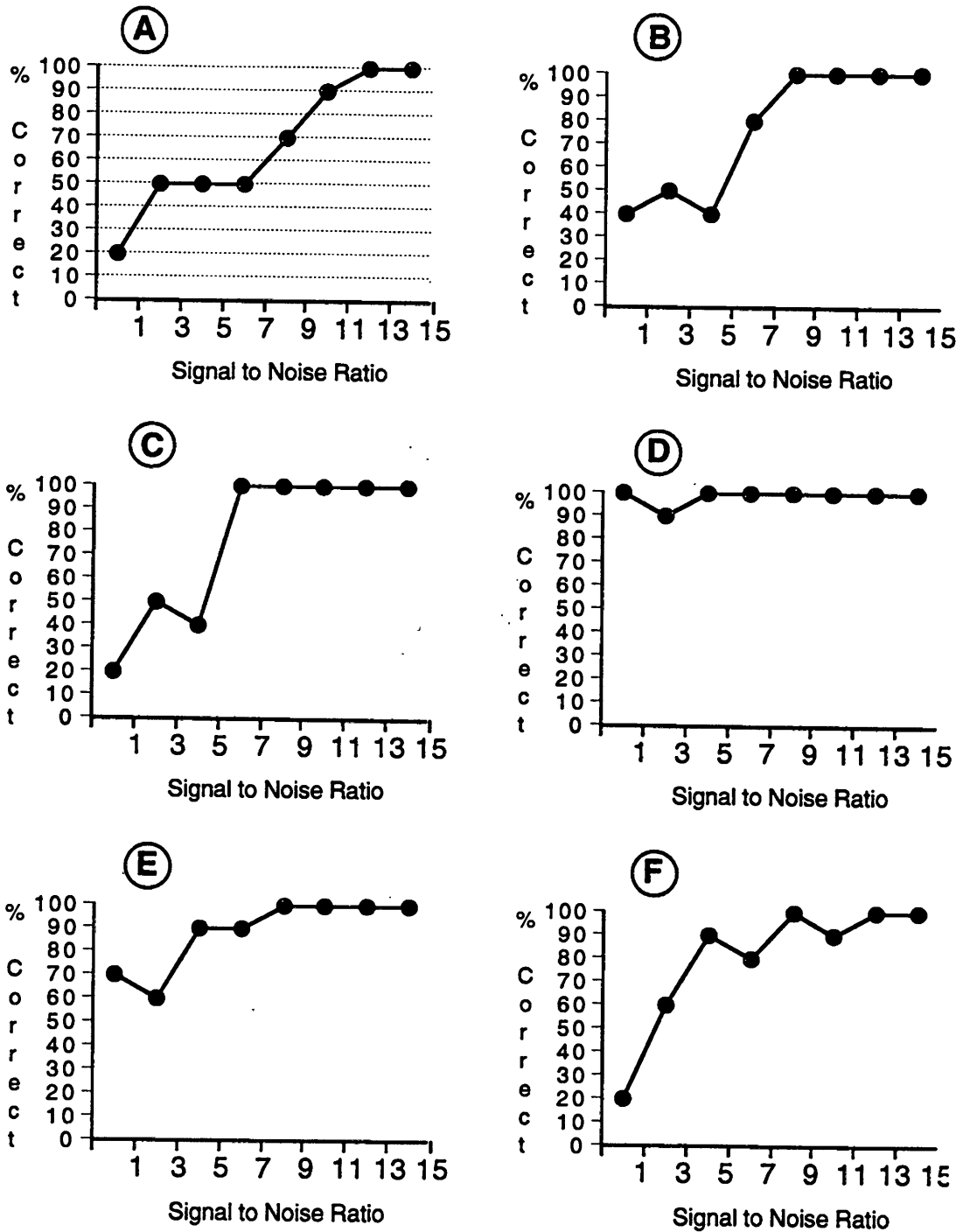
Figure 1. Performance-Intensity Functions for Airplane (A), Armchair (B), Baseball (C), Birthday (D), Cowboy (E), and Daybreak (F) in the Auditory-Only Mode.
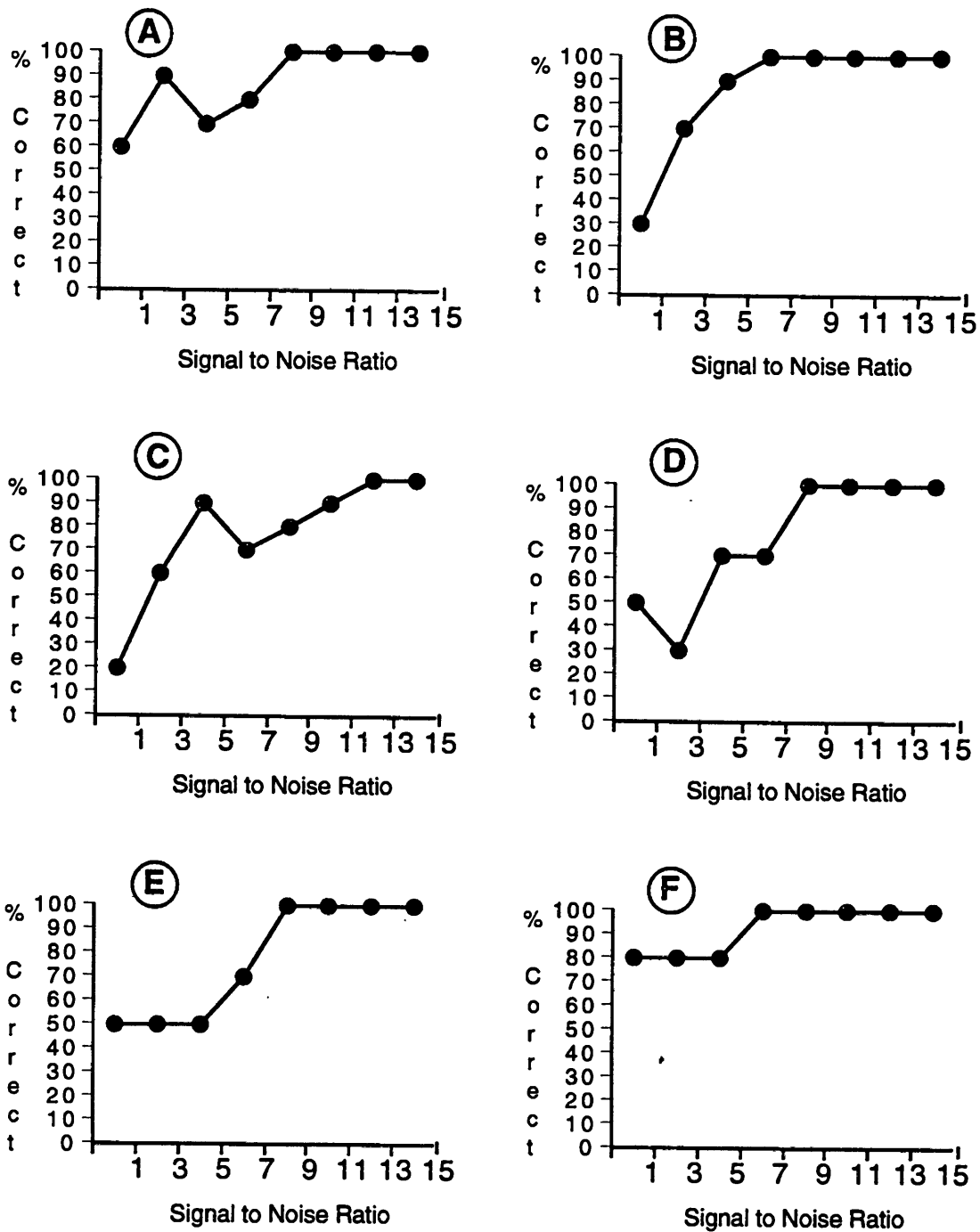
Figure 3. Performance-Intensity Functions for Greyhound (A), Hardware (B), Headlight (C), Horseshoe (D), Hot Dog (E), and Hothouse (F) in the Auditory-Only Mode.
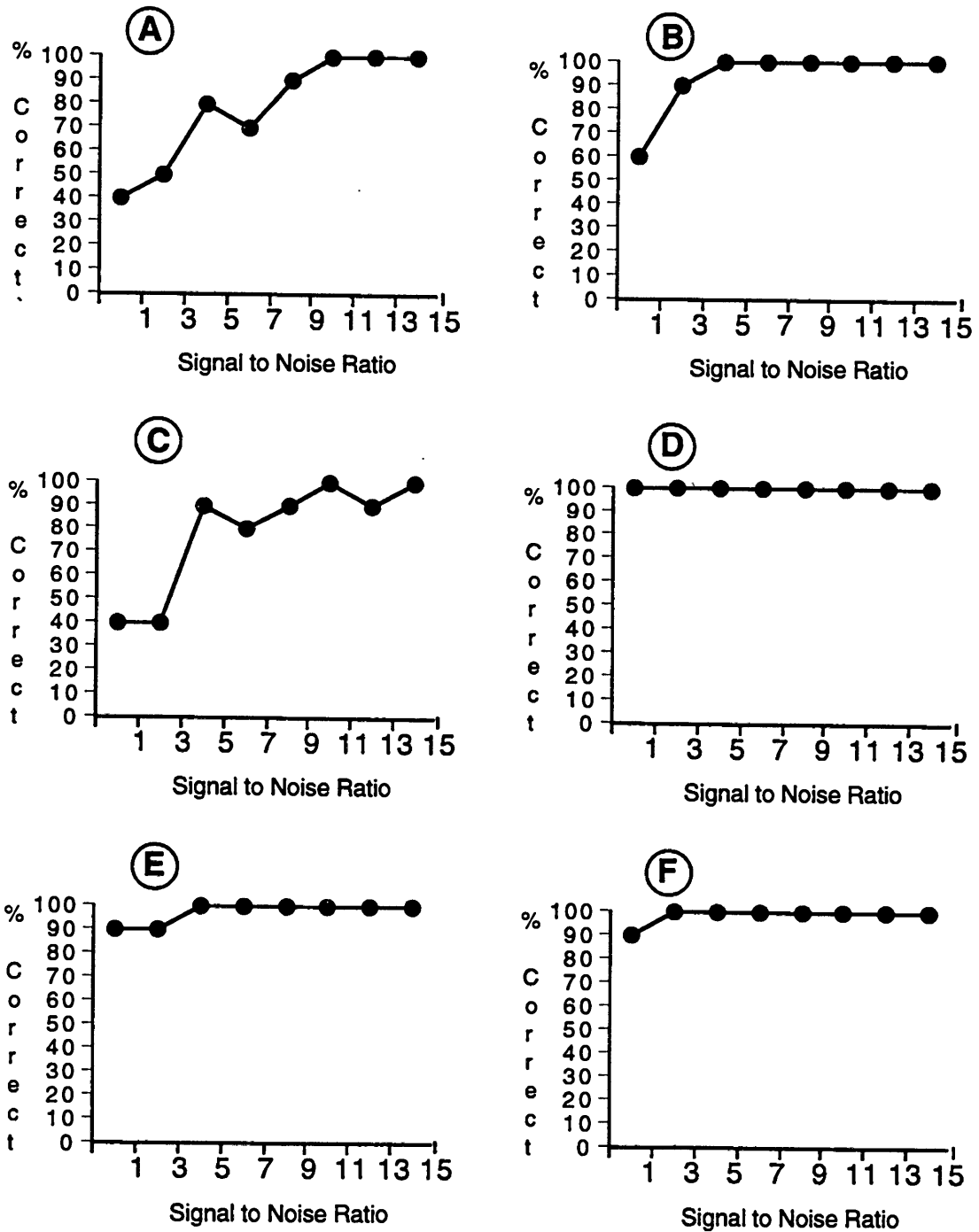
Figure 4. Performance-Intensity Functions for Iceberg (A), Inkwell (B), Mousetrap (C), Mushroom (D), Northwest (E), and Oatmeal (F) in the Auditory-Only Mode.
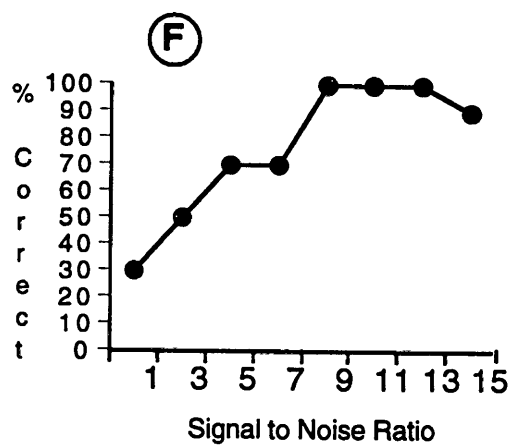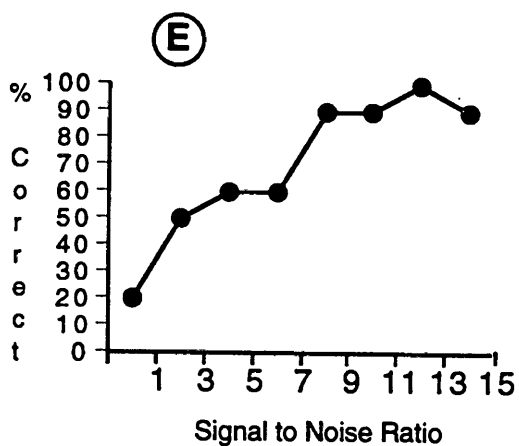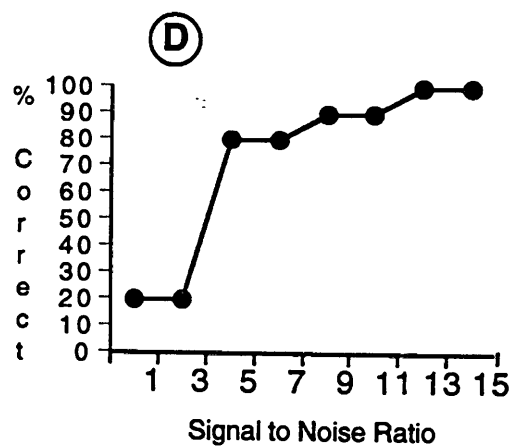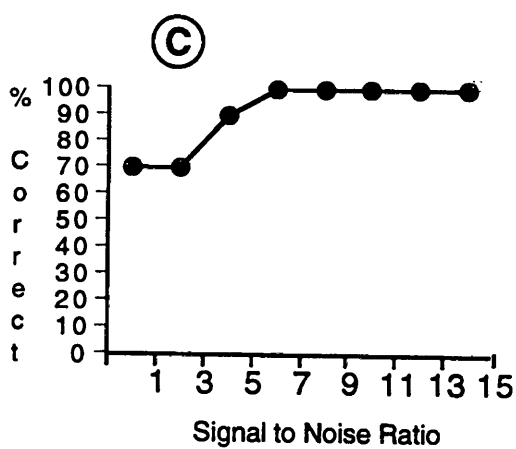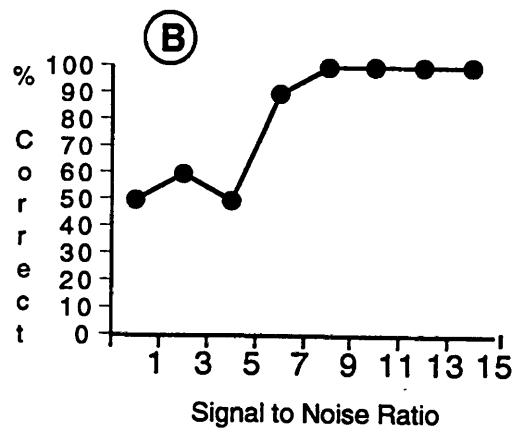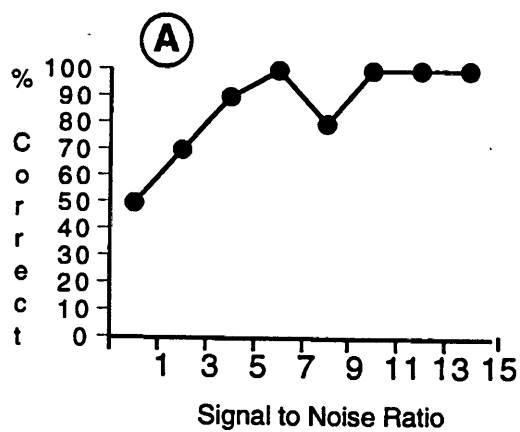
Figure 5. Performance-Intensity Functions for Padlock (A), Pancake (B), Playground (C), Railroad (D), Schoolboy (E), and Sidewalk (F) in the Auditory-Only Mode.
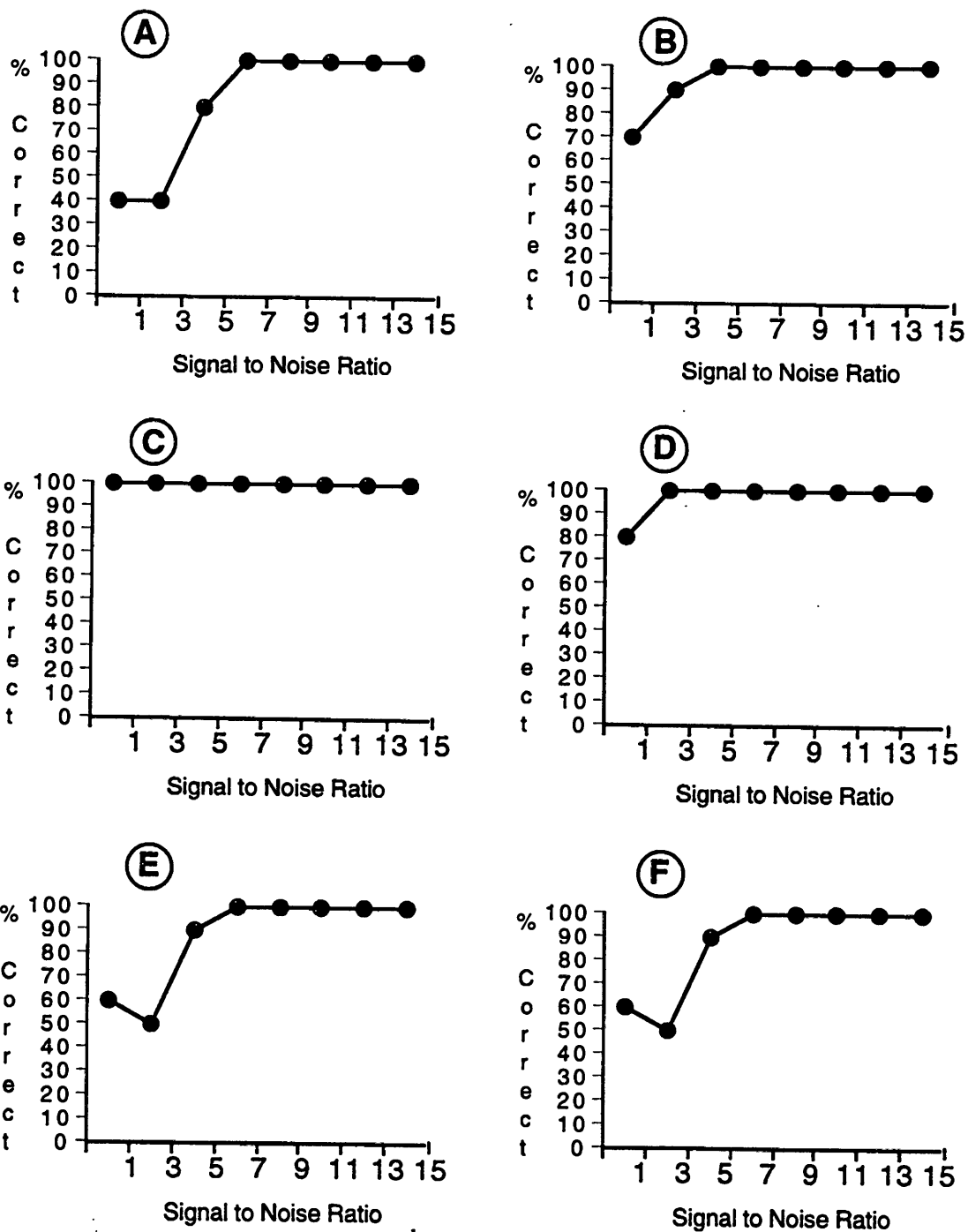
Figure 6. Performance-Intensity Functions for Stariway (A), Sunset (B), Toothbrush (C), Whitewash (D), Woodwork (E), and Workshop (F) in the Auditory-Only Mode.

Analysis of these results are somewhat interesting in that one might anticipate a correlation between the signal-to-noise ratio and the speech recognition threshold for a particular word. If the signal-to-noise ratio is the primary factor governing the intelligibility of the signal, then as the signal-to-noise ratio is degraded one might anticipate that the recognition rate of the spondee word will decrease. The inability to obtain consistently speech recognition thresholds and the observed lack of relationship between recognition rate and signal-to-noise ratio in many of the stimulus items might suggest that there is some other variable contributing to the intelligibility of these words.

Table 4.Thresholds for Spondee words as a function of noise level (dB SPL).
Auditory-Visual condition

Words	Noise level	Words	Noise level
Airplane	N/A	Hot Dog	N/A
Armchair	N/A	Ice Cream	N/A
Backbone	N/A	Inkwell	N/A
Baseball	N/A	Mousetrap	N/A
Birthday	N/A	Mushroom	N/A
Blackboard	N/A	Northwest	N/A
Cookbook	N/A	Nutmeg	N/A
Cowboy	N/A	Oatmeal	N/A
Doormat	N/A	Outside	N/A
Drawbridge	N/A	Padlock	N/A
Duckpond	N/A	Pancake	N/A
Eardrum	N/A	Playground	N/A
Earthquake	N/A	Railroad	N/A
Eyebrow	N/A	Stairway	N/A
Greyhound	N/A	Sunset	N/A
Hardware	N/A	Toothbrush	N/A
Headlight	-1	Whitewash	N/A
Horseshoe	N/A	Woodwork	N/A

Note. N/A-threshold could not be determined as recognition rate never fell
below 50%.
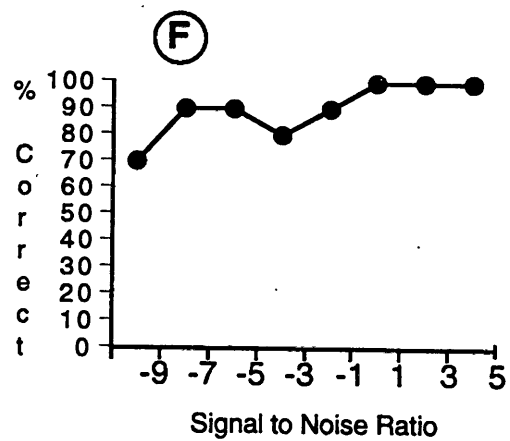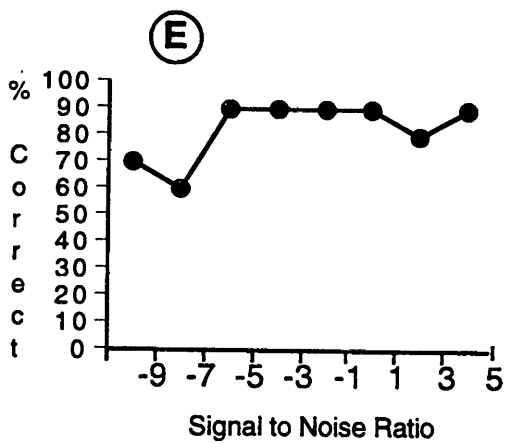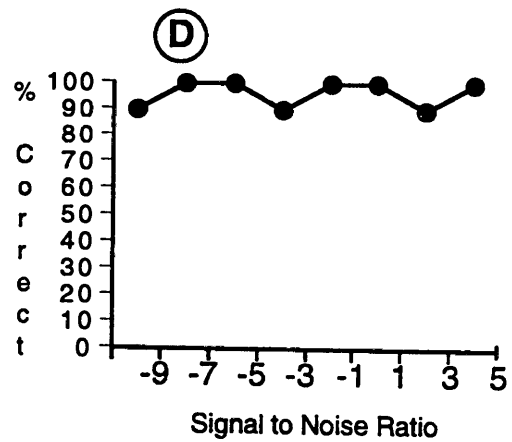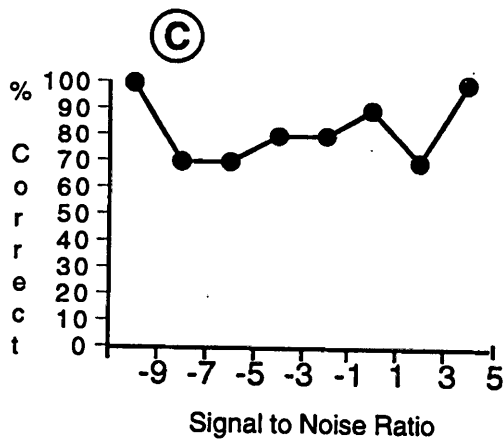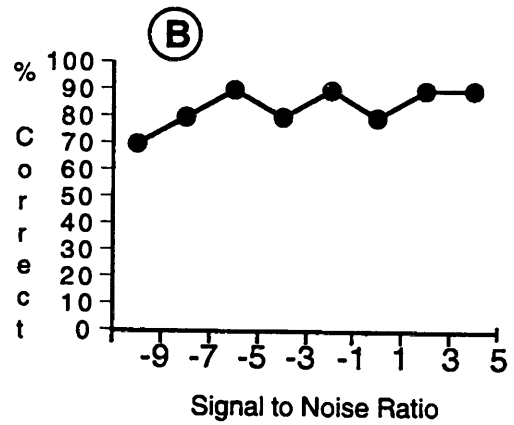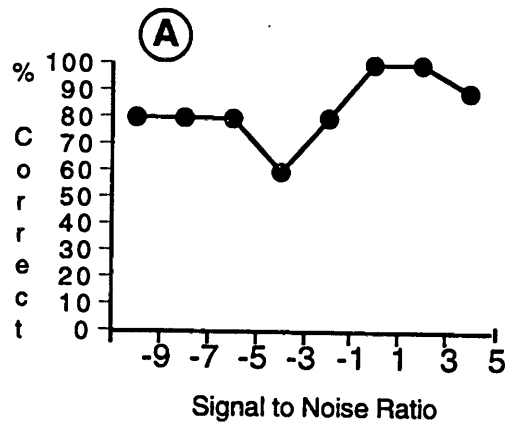
Figure 7. Performance-Intensity Functions for Airplane (A), Armchair (B), Baseball (C), Birthday (D), Cowboy (E), and Daybreak (F) in the Auditory-Visual Mode.
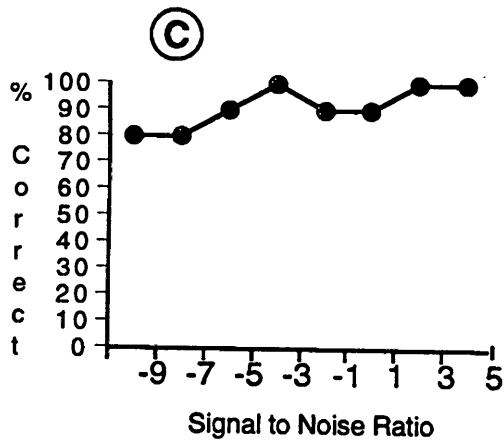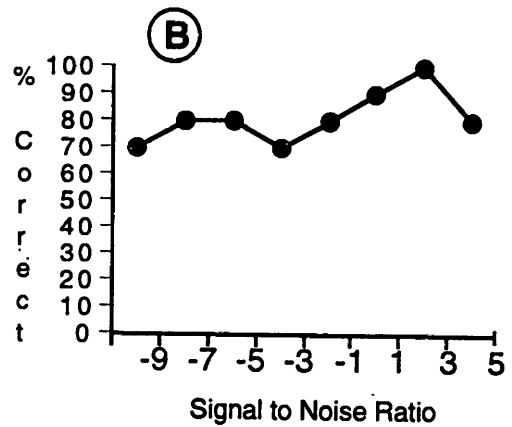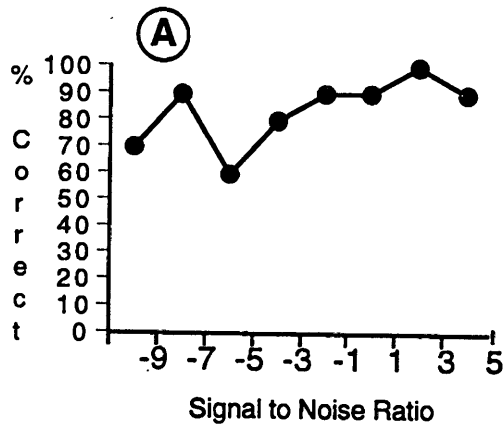
33
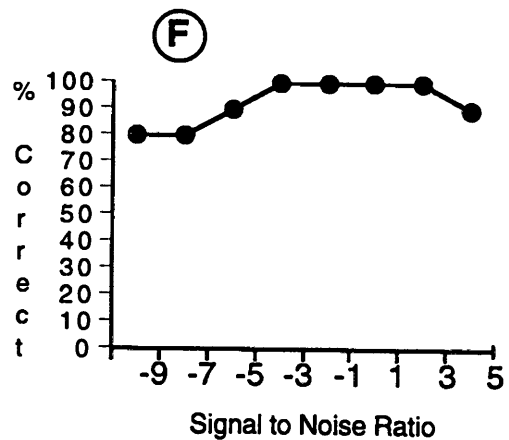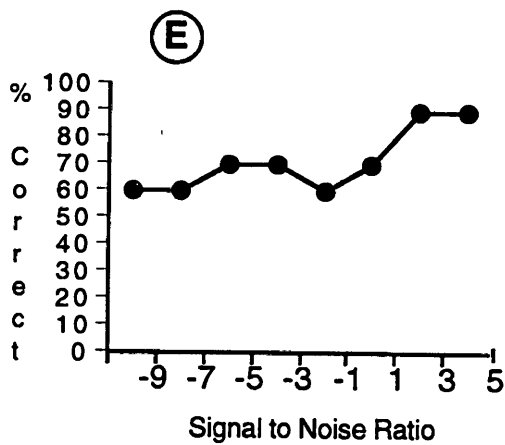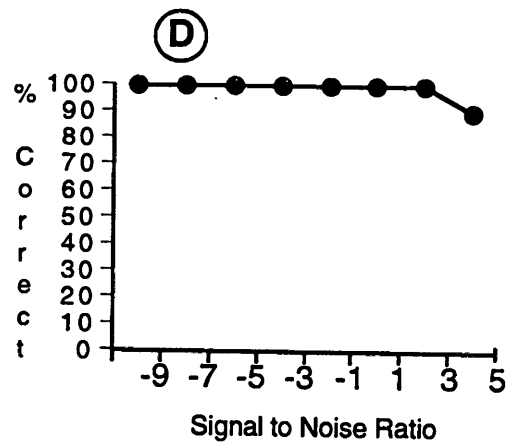
Figure 8. Performance-Intensity Functions for Doormat (A), Drawbridge (B), Duckpond (C), Eardrum (D), Farewell (E), and Grandson (F) in the Auditory-Visual Mode.
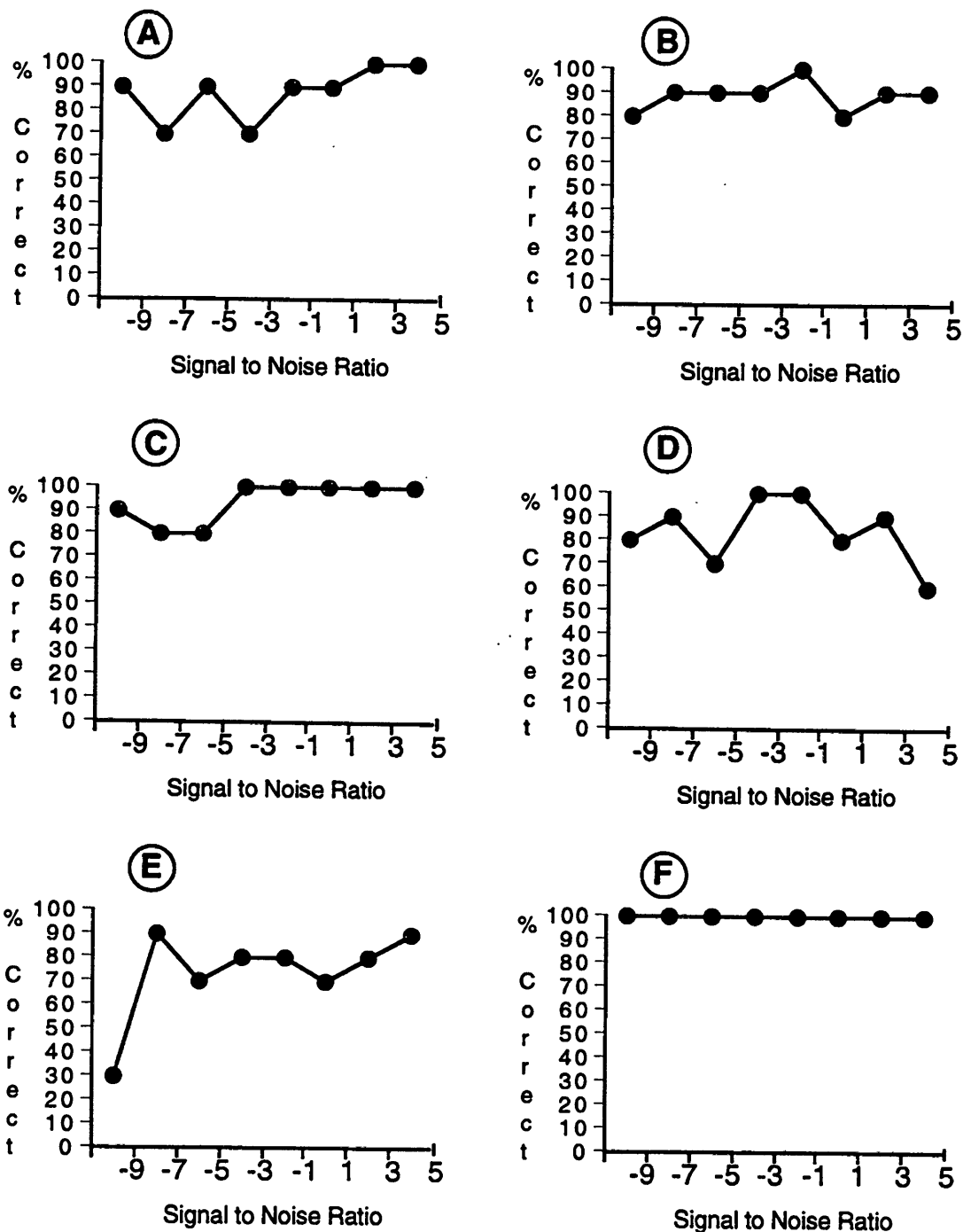
Figure 9. Performance-Intensity Functions for Greyhound (A), Hardware (B), Headlight (C), Horseshoe (D), Hot Dog (E), and Hothouse (F) in the Auditory-Visual Mode.
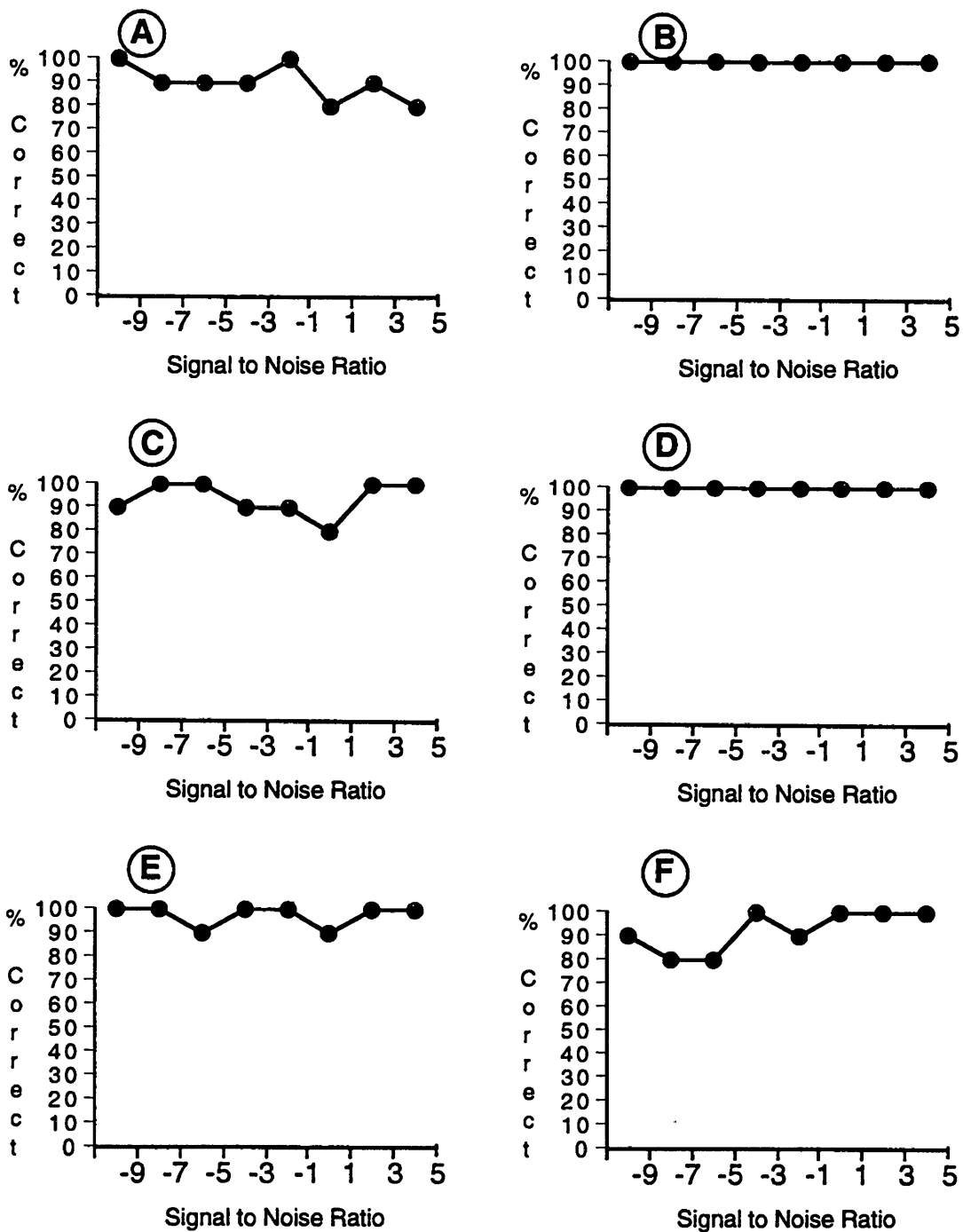
Figure 10. Performance-Intensity Functions for Iceberg (A), Inkwell (B), Mousetrap (C), Mushroom (D), Northwest (E), and Oatmeal (F) in the Auditory-Visual Mode.
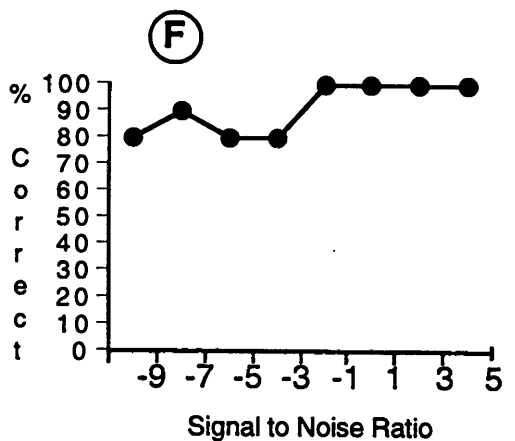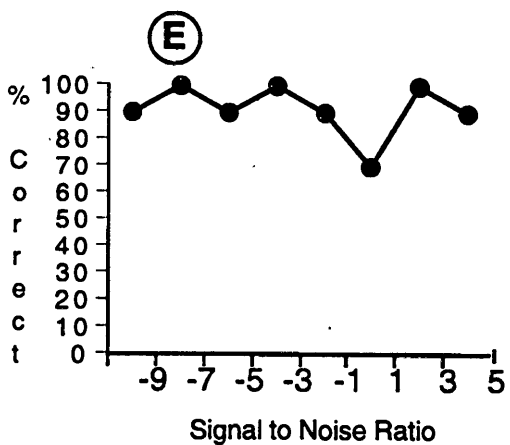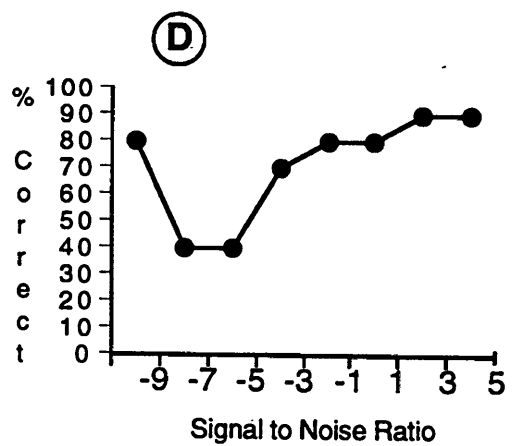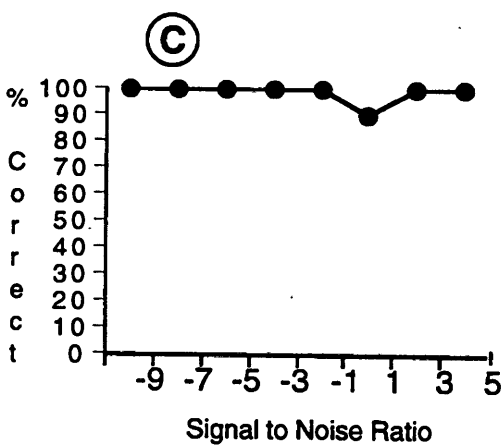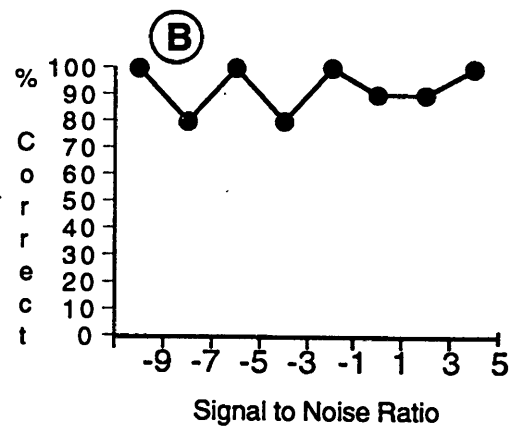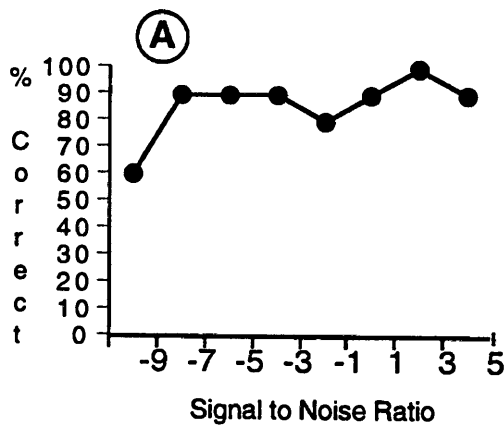
Figure 11. Performance-Intensity Functions for Padlock (A), Pancake (B), Playground (C), Railroad (D), Schoolboy (E), and Sidewalk (F) in the Auditory-Visual Mode.
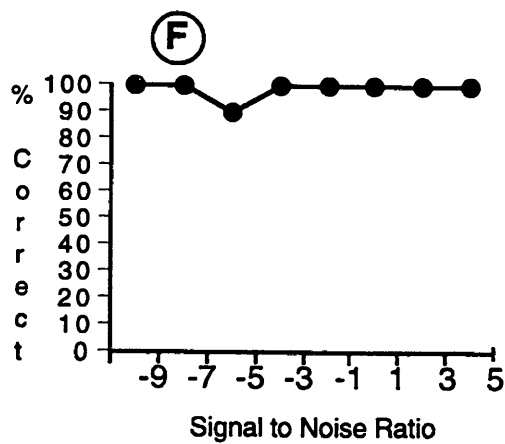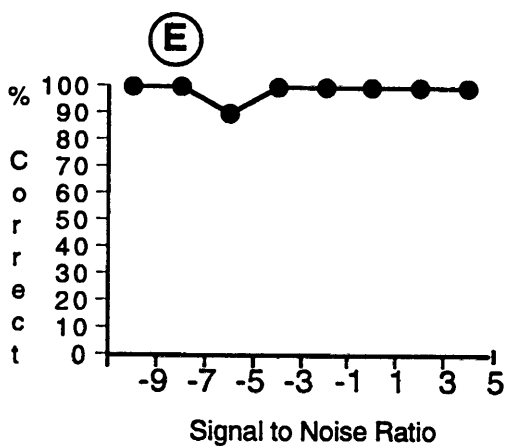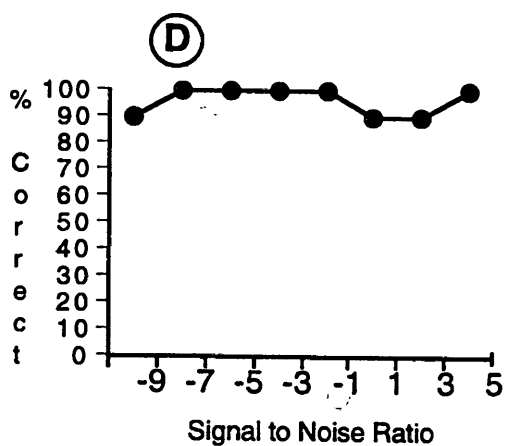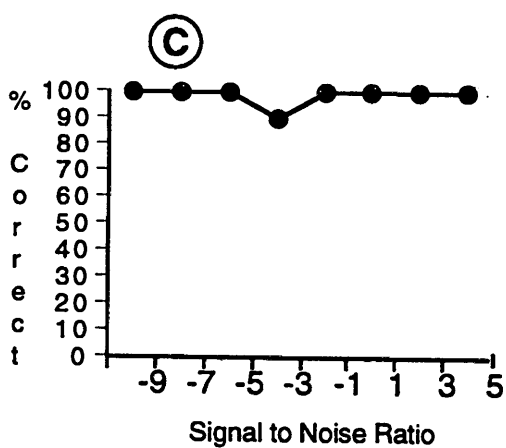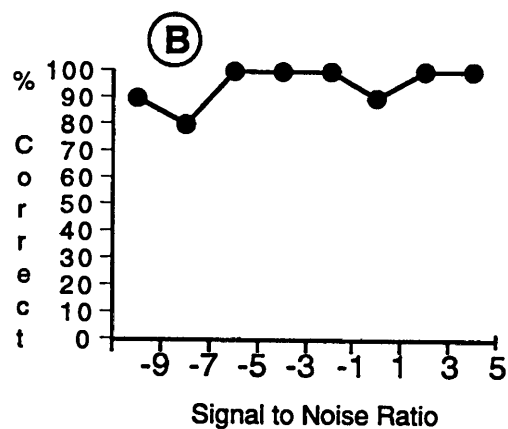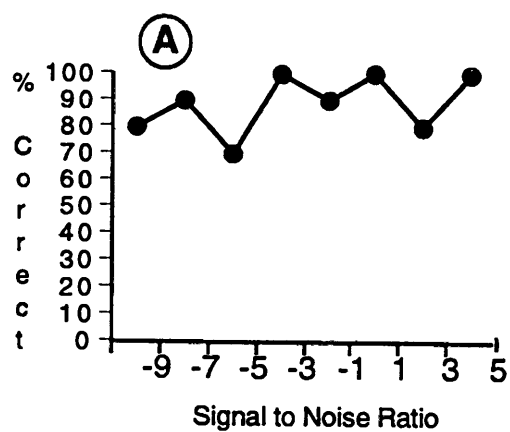
Figure 12. Performance-Intensity Functions for Stairway (A), Sunset (B), Toothbrush (C), Whitewash (D), Woodwork (E), and Workshop (F) in the Auditory-Visual Mode.

CHAPTER V

DISCUSSION

The purpose of this study was to determine if spondee words are an appropriate stimuli for a valid speechreading test. This was accomplished by developing performance intensity functions on spondaic words in auditory-only and auditory-visual conditions. In the auditory-only condition, 18 words satisfied the inclusion criteria (i.e., recognition threshold within one standard deviation of the group mean). However, in the auditory-visual modality, there was only one word that revealed a definable threshold. Therefore, spondaic words, as they were used in this study, do not appear to be a feasible stimuli for use in the proposed speechreading test.

The auditory-only condition yielded definable thresholds for half of the spondaic words. These results were to be expected based in part on the work of Wilson et al. (1973) and Young et al. (1982) who determined that the 36 spondaic words used in this study are auditorily homogenous. As this study presented these words in the presence of background noise, instead of in quiet as in previous studies, one could expect greater variance in the audibility of the word due to interaction with the noise signal. However, we were still able to document that at least a subset of the original 36 words retained homogeneity in recognition threshold when evaluated as a function of signal-to-noise ratio.

The auditory-visual portion of this study produced a ceiling effect on all but one word. In other words, the audibility of the words appeared to be

somewhat resistant to the effects of background noise suggesting that the contribution of visual cues may have overridden the contribution of audibility cues in terms of recognition threshold. While it is not surprising that the listeners would be more tolerant of background noise in the auditory-visual condition, this extreme effect on recognition threshold was not expected. There are several factors which may help explain why the audio-visual portion of this study did not produce the expected results. These factors may suggest options that can be explored with further research.

One possible factor may be the noise levels used in this study were not great enough to degrade the auditory signal adequately. However, it is unlikely this is the case as many of the subjects reported hearing only the noise, and not the words, at the poorest signal-to-noise ratios in the auditory-visual condition. Further research could investigate the use of less favorable signal-to-noise ratios.

A second factor that may have adversely affected results is that word lists were presented from a most-favorable to least-favorable signal-to-noise ratio. This may have resulted in a practice effect, such that by the time the subjects were presented the last lists of words, recognition of the words via their visual cues alone allowed near perfect intelligibility. Further research can easily investigate this potential by randomizing presentation order of the various lists.

A final factor in explaining the outcome of this study could be that the spondaic words as used in this study were simply too redundant in information content to present a challenging recognition task. While redundancy in information has been widely investigated from the auditory

perspective, considerably less is known about the redundancy inherent in integrated auditory and visual signal. If that were the case, an option for further research would be to consider different stimuli, such as nonsense syllables or monosyllabic words, which are considerably less redundant than the bisyllabic spondaic words used in this study. If a similar ceiling effect were to be found with these other stimuli, one might have to consider the likelihood that the use of simple stimuli within a speechreading assessment tool, as proposed in this study, might be difficult to realize. Our incomplete understanding of the relationship between auditory and visual cues may make it difficult to create a psychometrically reliable speechreading test, given the proposed test structure.

Summary, the purpose of this study was to determine the suitability of spondaic words for use in a test of speechreading assessment. The results of this study suggested that spondaic words, at least as presented in this study, do not represent feasible stimuli for the proposed speechreading test.

REFERENCES

Calhoun, A., Dancer, J., & Pittenger, J. (1988). A comparison of the Utley lipreading forms to three lists of Harris' revised CID everyday sentences as lipreading stimuli. The Volta Review, 90 149-154.

Erber, N. P. (1975). Auditory-visual perception of speech. Journal of Speech and Hearing Disorders, 40, 481-492.

Hirsch, I., Davis, H., Silverman, S., Reynolds, E., Eldert, E., & Benson, R. (1952). Development of the materials for speech audiometry. Journal of Speech and Hearing Disorders, 17, 321-337.

Hodgson, W. R. (1980). Basic audiologic evaluation. Baltimore: Williams & Wilkins.

Hull, R. H. (1982) Evaluation of success in aural rehabilitation. In R. Hull (Ed.), Rehabilitative audiology. New York: Grune & Stratton, Inc.

Jeffers, J., & Barley, M. (1971). Speechreading (lipreading). Springfield, Illinois: Charles C. Thomas.

Kaplan, H, Bally, S. J., & Garretson, C. (1985). Speechreading: A way to improve understanding. Washington, D. C.: Gallaudet College Press.

Massaro, D. W., Thompson, L. A. , Barron, B., & Laren, E. (1986). Developmental changes in visual and auditory contributions to speech perception. Journal of Experimental Psychology, 41, 93-113.

Middlewerd, M. J., & Plomp, R. (1987). The effects of speechreading on the speech-reception thresholds of sentences in noise. Journal of the Acoustical Society of America, 82(5), 2145-2147.

Montgomery, A. A., & Demorest, M. E. (1988). Issues and developments in the evaluation of speechreading. The Volta Review, 90(5), 193-214.

Pickett, J. M., & McFarland, W. (1985). Auditory implants and tactile aids for the profoundly deaf. Journal of Speech and Hearing Research, 28, 134-150.

Rigo, T. G. (1986). The relationship between the visual contribution to speech perception and lipreading ability during focused and divided attention. Ear and Hearing, 7(4), 266-271.

Silverman, S. R., & Kricos, P. B. (1990). Speechreading. The Volta Review, 92(4), P. 22-32.

Sims, D, G. (1982). Past present and future computer assisted c ommunication training at NTID. Academy of Rehabilitative Audiology, 15, 103-115.

Spitzer, J. B., Leder, S. B., Milner, P., Flevaris-Phillips, C., & Giolas, T. G. (1987). Standardization of four videotaped tests of speechreading ranging in task difficulty. Ear and Hearing, 8(4), 227-231.

Summerfield, Q. (1987). Some preliminaries to comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), Hearing by eye: The Psychology of lip-reading. London: Erlbaum.

Warren, Y., Dancer, J., Monfils,B., & Pitterger, J. (1989). The practice effect In speechreading distributed over five days: Same versus different CID Sentence lists. The Volts Review, 91(7) p. 321-325.

Wilson, R., Morgan, D., & Dirks, D. (1973). A proposed speech reception threshold procedure and its statistical precedent. Journal of Speech and Hearing Disorders, 38(2), 184-191.

Young, L., Dudley, B., & Gunter, M. (1982). Thresholds and psychometric function of the individual spondee words. Journal of Speech and Hearing Research, 25, 586-593.