

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

A

THE EFFECTS OF AGE, INPUT MODALITY, AND EVALUATION
METHOD ON THE MEASURED SPEECH IMITATION PERFORMANCE OF
PROFOUNDLY HEARING-IMPAIRED CHILDREN

by

ORNA ERAN

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

1996

UMI Number: 9707086

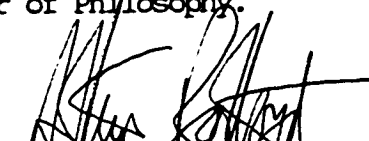
UMI Microform 9707086
Copyright 1996, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

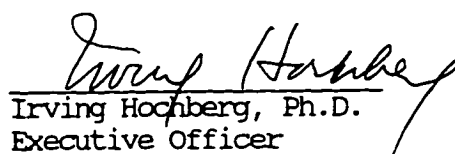
UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted by the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

August 6, 1996
Date


Arthur Boothroyd, Ph.D.
Chair of Examining Committee

August 6, 1996
Date


Irving Hochberg, Ph.D.
Executive Officer

Supervisory Committee:

Arthur Boothroyd, Ph.D.

Katherine Harris, Ph.D.

Irving Hochberg, Ph.D.

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECTS OF AGE, INPUT MODALITY, AND EVALUATION
METHOD ON THE MEASURED SPEECH IMITATION PERFORMANCE OF
PROFOUNDLY HEARING-IMPAIRED CHILDREN
by

Orna Eran

Advisor: Professor Arthur Boothroyd

The recorded imitations of syllables by 12 prelingually, profoundly deaf children were evaluated by three methods: a) behaviorally, using a closed-set identification task, b) behaviorally, using an open-set recognition task, and c) acoustically, in terms of vowel formants. Performance on the three measures was examined in relation to age (3-5 years versus 13-15 years) and input modality (auditory-visual versus auditory-only). Imitation of auditory-visual input was considerably better in older than younger subjects, indicating maturation-related and/or training-related improvements of phonetic-level speech production. In contrast, there was no evidence of age-related changes in imitation of auditory-only input. This last finding is consistent with the hypothesis that speech production measures based on imitation of a purely auditory input can provide a valid index of auditory speech perception

capacity in children as young as three years of age. It is also in keeping with the hypothesis that auditory capacity is a fixed entity, determined by cochlear pathology and unresponsive to the effects of age, listening experience, or training. A second finding was that younger children's imitations were not dramatically different for auditory-visual and auditory-only input. This was in contrast to the older children, who showed a large difference. This finding is in keeping with the hypothesis that the quality of speech production in these young hearing-impaired children is driven primarily by auditory capacity. These results were, essentially, independent of the assessment method. These findings, increase confidence in the inferences of fixed auditory capacity and auditory-driven production in younger children, and the validity of imitative production measures as indices of auditory speech perception capacity.

To Yagil, Yuval, and Yorai

Acknowledgments

I would like to thank my committee members, Irv Hochberg and Katherine Harris for their guidance and many helpful suggestions. My greatest debt goes to the chair of my committee, Arthur Boothroyd who provided crucial direction, encouragement, and constructive criticism throughout the doctoral process.

I am very thankful to Mark Ross for serving as my external examiner and for years of affection.

I am most grateful to Robin Waldstein, Tina Cartusciello, Anita Haravon, and Elizabeth Galletta who served as subjects for this research.

I would like to express my gratitude to many friends and colleagues for their support, collaboration, and encouragement over the years.

Finally, I wish to thank my family for their love, patience and constant support.

This work was partially supported by NIH grant #5P50DC00178.

Table of contents	Page
Chapter 1. Introduction	1
Chapter 2. Review of the literature	
The effects of profound sensorineural hearing loss on speech intelligibility of hearing- impaired children.....	11
Auditory speech perception by profoundly hearing-impaired children.....	19
Methods of measuring speech perception performance in young hearing-impaired children.....	29
Chapter 3. The IMitative test of Speech PAttern Contrast perception (IMSPAC)	60
Chapter 4. Methodology	85
Chapter 5. Results	95
Chapter 6. Acoustic analysis of vowels	151
Chapter 7. Discussion	193
Clinical implications and future research	209
Appendix A	215
Appendix B	216
Appendix C	217
Appendix D	218
Appendix E	219
Appendix F	220
Appendix G	228
Appendix H	230

Appendix I	232
Bibliography	233

List of tables

3.1	Five subtests of the phonologically significant contrasts represented in the IMSPAC test, together with examples for the contrasts and the response alternatives	65
5.1	Original and new composite IMSPAC scores for the auditory and the auditory-visual input modalities for the 12 hearing-impaired subject	97
5.2	Analysis of variance: effects of AGE (A), TEST (ORIGINAL/NEW IMSPAC) (T), and INPUT MODALITY (M) on the IMSPAC scores of the 12 hearing-impaired subjects.	100
5.3	Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the new IMSPAC data.	102
5.4	Results of a post-hoc analysis with planned comparison applied to the means of the arcsine transforms of the new IMSPAC data.	103
5.5	Mean scores on 10 speech-pattern contrasts of the IMSPAC test for the older and younger subjects for the auditory-only input modality.	106
5.6	Mean scores on 10 speech-pattern contrasts of the IMSPAC test for the older and younger subjects for the auditory-visual input modality.	107
5.7	Analysis of variance: effects of AGE (A), INPUT MODALITY (M), LISTENER (L), CONTRAST (C) on the IMSPAC scores of the 12 hearing-impaired subjects.	108
5.8	Percent recognition of phonemes in syllables produced by imitation, by 12 hearing-impaired subjects as a function of age group and input modality.	116

5.9	Analysis of variance: effects of AGE (A) and INPUT MODALITY (M) on the Open-set scores of the 12 hearing-impaired subjects.	118
5.10	Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data.	120
5.11	Open-set scores and new segmental closed-set scores for 12 hearing-impaired subjects as a function of age group and input modality.	122
5.12	Analysis of variance: effects of AGE (A), TEST (OPEN VS. NEW CLOSED-SET IMSPAC) (T), and INPUT MODALITY (M) on the composite scores of the 12 hearing-impaired subjects.	125
5.13	Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data and the new closed-set data (mean of eight segmental speech contrasts corrected for guessing).	128
5.14	Analysis of variance: effects of AGE (A), INPUT MODALITY (M), LISTENER (L) on the open-set scores of the 12 hearing-impaired subjects.	138
5.15	Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data.	141
5.16	Test and retest composite IMSPAC scores, in percent, for the auditory and the auditory-visual input modalities for four of the hearing-impaired subjects. Differences between test and retest scores are also shown.	144
5.17	Test and retest open-set scores, in percent, for the audition and the auditory-visual input modalities for four of the hearing-impaired subjects. Differences between test and retest scores are also shown.	145

5.18	Analysis of variance: effects of TEST REPLICATION (R), and INPUT MODALITY (M) on the closed-set scores of the four hearing-impaired subjects.	146
5.19	Analysis of variance: effects of TEST REPLICATION (R), and INPUT MODALITY (M) on the open-set scores of the four hearing-impaired subjects.	146
6.1	Mean values for the frequencies of the first (F1) and second (F2) formants and number of vowels that were correctly identified by all four listeners in the auditory-visual condition on the open-set listening task for each of the 12 hearing-impaired subjects.	159
6.2	Mean frequency values and standard deviations (SD) for the first (F1) and second (F2) formants for the auditory-visual condition for both younger and older subjects.	172
6.3	Mean frequency values and standard deviations (SD) for the first (F1) and second (F2) formants for the auditory-only condition for both younger and older subjects.	175
6.4	d' values for frequencies of the second (F2) and first (F1) formants for the auditory-visual condition for both younger and older subjects.	182
6.5	d' values for frequencies of the second (F2) and first (F1) formants for the auditory-only condition for both younger and older subjects	183

List of figures

3.1	The IMSPAC test procedure	68
3.2	Composite auditory-visual IMSPAC scores as a function of auditory-only scores for 22 implant users and 46 hearing aid users.	74
3.3	Test-retest data for composite auditory-visual and auditory-only IMSPAC scores of 18 hearing aid users.	76
3.4	Mean auditory-visual and auditory-only IMSPAC scores of 83 severely and profoundly hearing-impaired, aided children as a function of age	77
5.1	Group mean (+/- 1 std. error) original and new IMSPAC scores for 6 younger and 6 older children.	98
5.2	The auditory and auditory-plus-visual IMSPAC contrasts profile for the younger group.	109
5.3	The auditory and auditory-plus-visual IMSPAC contrasts profile for the older group.	110
5.4	Group mean IMSPAC auditory scores as a function of speech contrasts and age group.	113
5.5	Group mean IMSPAC auditory-plus-visual scores as a function of speech contrasts and age group.	114
5.6	Group mean (+/- 1 std. error) open-set scores for 6 younger and 6 older children.	117
5.7	Group mean (+/- 1 std. error) new IMSPAC and open-set scores for 6 younger and 6 older children.	124
5.8	Mean new closed-set IMSPAC and open-set scores for the younger group as a function of input modality.	126

5.9	Group mean new IMSPAC and open-set scores for the auditory-only condition.	130
5.10	Mean new closed-set IMSPAC and open-set scores for the older group as a function of input modality.	131
5.11	Open-set scores as a function of closed-set scores for the auditory-only condition.	133
5.12	Open-set scores as a function of closed-set scores for the auditory-plus-visual condition.	134
5.13	Auditory-plus-visual and auditory-only closed-set IMSPAC scores collapsed across the 12 hearing-impaired subjects as a function of listener.	136
5.14	Auditory-plus-visual and auditory-only IMSPAC scores, as a function of listener and speech contrast.	137
5.15	Auditory-plus-visual and auditory-only open-set scores collapsed across the 12 hearing-impaired subjects as a function of listener.	140
6.1	Waveform and wide-band spectrogram for the utterance "eed", produced by a hearing-impaired child in the auditory-visual condition.	155
6.2	The same spectrogram and waveform as in Figure 6.1, but with an LPC spectrum displayed in window "C"..	157
6.3	Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-visual condition.	161
6.4	Second versus first format frequency for the vowel templates produced by the 12 hearing-impaired subjects, and for the Peterson and Barney (1952) data for children, women, and men.	163

6.5	Second versus first formant frequency for four vowels produced by 30 hearing children. Data was taken from Eguchi and Hirsh (1969).....	166
6.6	Distributions of first and second formant frequencies of correctly identified vowels in the auditory-visual imitations of 6 younger hearing-impaired children.	168
6.7	Distributions of first and second formant frequencies of correctly identified vowels in the auditory-visual imitations of 6 older hearing-impaired children.	169
6.8	Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-visual condition.	173
6.9	Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-only condition.	176
6.10	The value of d' for a pair of vowels is determined by the ratio between the separations of their mean formant values and the standard deviation of the probability density distribution.	178
6.11	Assume that a listener identifies any formant whose frequency is above the decision criterion as belonging to vowel (y) and any formant whose frequency is below the decision criterion as belonging to vowel (x). Assume also, that the decision criterion is mid-way between the two means. Then the probability of a correct response is given by the shaded area to the left of the decision criterion. This value can be determined from z tables.	180
6.12	Obtained closed-set scores (% re chance) and predicted vowel scores (% re chance) for the younger and older subjects. The predicted scores were derived from d' analysis.	185

- 6.13 Obtained closed-set scores (% re chance)
and predicted vowel scores (% re chance)
as a function of age for auditory-only and
auditory-visual modalities. The predicted
scores were derived from d' analysis. 186
- 6.14 Combined predicted vowel scores for the
younger and older subjects. Scores were
derived from a d' analysis. 187

Chapter 1

Introduction

The research to be described deals with speech production and perception in children with prelingually acquired, profound hearing loss. Specifically, it deals with the role of peripheral auditory speech perception capacity in the acquisition of speech production skills and our ability to infer that capacity from measures of imitative production.

For the purpose of the present study, peripheral auditory speech perception capacity is defined as the ability of the peripheral auditory system (including the cochlea, auditory nerve and, possibly lower auditory centers in the brain stem) to transmit information about acoustic speech signals to higher centers in the brain. The term "transmit information" refers both to information about the presence of acoustic patterns that contribute to the recognition of phonemes (audibility) and information about the details of those acoustic patterns that are needed to perceive changes of meaning (discrimination).

Speech perception is a loosely defined term that includes both the transmission of information (referred

to by Boothroyd, 1993 as "sensory evidence") and the use of that information for such things as phoneme recognition, word recognition, sentence perception, and the comprehension of spoken messages. Speech perception as so defined requires not only sensory evidence but also linguistic knowledge, world knowledge, and perceptual skills (Boothroyd, 1993).

Much of the impact of a sensorineural prelingually acquired hearing loss on a developing child depends on the extent of the loss and its effect on perception of spoken language, both self-generated and generated by others. Knowledge of the hearing-impaired child's peripheral auditory speech perception capacity at an early age is crucial for planning and managing an appropriate habilitation program. One of the primary goals of habilitation is to enhance spoken language perception. To pursue this goal, a series of decisions needs to be made regarding the selection of sensory aids, educational placement, and intervention programs. The first avenue of attack is to increase the ability to hear spoken language by using hearing aids or cochlear-implants. The second is to ensure that the resulting peripheral auditory speech perception capacity is used

to maximum advantage in the development of receptive and expressive language skills.

Unfortunately, we have no direct measures of peripheral auditory speech perception capacity but must infer it from the results of speech perception tests. The problem is that the results are likely to reflect not only peripheral auditory speech perception capacity but also the child's knowledge and skill. The relative contributions are usually unknown and it is, therefore, difficult to make inferences about either of them. For many application, this is unimportant because the purpose of the test is to measure overall speech perception performance under conditions that reflect a typical communication situation. When studying sensory aids (be they hearing aids, cochlear implants, or tactile aids), however, the distinction becomes crucial. Sensory aids have only one purpose--to increase the amount of sensory evidence available from the peripheral system. If, therefore, we wish to measure their effects across subjects, it is important that the effects of knowledge and skill be minimized so that the results can reflect mainly the peripheral auditory speech perception capacity of the child, as determined by the combination of the physical characteristics of his peripheral auditory system and the properties of the sensory aid.

Therefore, there is a need for valid and reliable tests of peripheral auditory speech perception capacity whose results are not confounded by non-auditory factors such as reading ability, linguistic knowledge, cognitive function, or attention.

Between 1990 and 1995, research on peripheral auditory speech perception capacity was carried out at the City University of New York. One of the general goals of this research was to develop actuarial methods for assigning probabilities of the outcomes of amplification and implantation in young deaf children. More specific goals included the development of pediatric tests of sensory speech perception capacity, and the development of a "normative" data base of peripheral auditory speech perception capacity in hearing-impaired children who wear hearing aids. To address these goals, data on aided and implanted children were collected using two outcome tests. Both tests were designed to estimate the probability of perception of phonologically significant contrasts presented in a varying phonetic context (Boothroyd, 1991a).

The test of interest in the present study is the Imitative test of Speech Pattern Contrast perception

(IMSPAC). The child is asked to imitate a series of nonsense syllables, under two modality inputs--auditory-visual and auditory-only. The child's imitations are tape recorded and subsequently presented to a team of normal hearing listeners. For each utterance, the task of the listeners is to determine in a four-alternative, forced-choice context what he/she believes to have been the model the child was trying to imitate. The listeners' responses provide information about the child's ability to convey phonologically contrastive information in his speech output. The extent to which he can do so, given only an acoustic input, is taken as evidence of his peripheral auditory speech perception capacity.

Previous findings in aided hearing-impaired children were as follow:

1. The ability of hearing-impaired children to imitate syllables presented by hearing and vision improves considerably with age.
2. In contrast, their ability to imitate syllables presented by hearing alone does not improve with age.
3. The auditory-visual imitations of younger hearing-impaired children (3-5 years of age) are only a little better than their auditory-only imitations.

4. In contrast, the auditory-visual imitations of older hearing-impaired children (13-15 years of age) are considerably better than their auditory-only imitations.

The absence of an age effect for the auditory scores, together with the findings of an age effect for the auditory-visual scores and better auditory-visual than auditory performance for both age groups suggests two things. First, that information about peripheral auditory speech perception capacity can be inferred from measures of imitation of auditory-only input. Second, that although speech production improves over time, auditory capacities remain unchanged. These implications are based on two lines of reasoning. First, if the auditory-only IMSPAC score reflected the influence of cognitive and/or speech production factors, they would be expected to improve both with age and with the observed improved production skills. Second, if peripheral auditory speech perception capacity improved over time, this should also be reflected by an increase in the auditory-only score. Thus the absence of an age effect provides evidence in support of the conclusions that auditory imitations do indeed reflect auditory capacity and that this capacity remains unchanged at least from age four to age fourteen.

These two conclusions are quite striking and, potentially, controversial. Both hinge on the failure to demonstrate an age-related change in auditory imitation performance.

The conclusion that there is no age-related change in the auditory imitations should, however, be taken somewhat cautiously for several reasons:

1) The children included in the study represent a heterogeneous sample that varied greatly with respect to age at onset of deafness and degree of hearing loss. We have no guarantee, therefore that the older and younger children were identical in all respects except age.

2) The listener variable was not controlled. Different groups of listeners were used to assess the imitations of older and younger children (though not to compare input modality for each child). Thus, if the listeners who audited the younger children were "better" than those auditing the older children, an age effect could have been masked.

3) A closed-set technique, used to assess the children's imitations, may not be a sufficiently sensitive tool to demonstrate important changes that may exist in the auditory-only condition.

Therefore, the primary purpose of the present study was to investigate the effects of age and evaluation method on measured auditory speech perception performance in profoundly hearing-impaired children. More specifically, this study reevaluated existing recordings of a carefully selected sample of younger and older hearing-impaired children to determine whether the original findings remain intact, or whether they simply reflected the effects of low power of either the test instruments or the design. Specific experimental questions were as follows:

1. Are the original findings still supported when the same listeners are used across all children and conditions and when the older and younger children are matched in terms of pure tone average hearing loss, age at onset of deafness and, educational environment.
2. Is there an effect of age on the auditory-only score that is revealed when a more traditional evaluation technique, such as open-set recognition, is used?
3. Is the effect of input modality in younger children increased when an open-set evaluation technique is used?
4. What is the relationship between scores obtained with the closed-set and open-set evaluation techniques?

5. Which of the evaluated speech contrasts are most affected by age and/or input modality?

6. Are there age-related changes in the ability to imitate auditory input that are revealed when imitations are evaluated in terms of their acoustic properties?

7. Are the behavioral results predictable from acoustic measures of performance?

These questions were addressed by the following means:

1. A reexamination of the existing data using matched samples of younger and older children.

2. An evaluation of recorded utterances using the same group of listeners for all children and conditions.

3. An assessment of the children's utterances with both closed-set and open-set evaluation techniques.

4. An acoustic analysis of vowels--specifically, the frequencies of the first two formants.

Replication of the previous findings, that is, of no age effect for the auditory-only condition, but a significant age effect for the auditory-visual condition, would provide support for the following conclusions:

1. That the peripheral auditory speech perception capacity does not change for the better as a result of maturation, experience and training.

2. That information about peripheral auditory speech perception capacity can be reliably inferred from results of auditory imitation.

Chapter 2

Review of the literature

The following review consists of three parts. The first part deals with research on the speech production abilities and characteristics of children with prelingually acquired hearing loss. The second part deals with auditory speech perception in the same population. The third part is a review of the formal and informal methods that have been used to assess speech perception performance and/or capacity in hearing-impaired children.

The effects of profound sensorineural hearing loss on speech intelligibility of hearing- impaired children

The following review of the literature deals with the listener's perception of the speech of the deaf, or speech intelligibility. Speech intelligibility is typically defined as the degree to which units within the speech of a hearing-impaired person are recognized by a normal hearing listener. The units can be words, phonemes, or phonetic features, such as voicing or

place of articulation. Phonemes may be spoken in nonsense syllables or words. Words may be spoken in isolation or in phrases and sentences.

To a certain extent, the measure of speech intelligibility is a measure of the interactions among at least four major components: 1. the characteristics of the hearing-impaired individual whose speech is under study (e.g., age of onset of hearing loss, current age, degree of hearing loss, mode of communication); 2. the task that elicited the production (e.g., reading, spontaneous speech, imitation) and the speech materials to be spoken (e.g., phonemes, words, sentences); 3. the listener's previous experience with deaf speech; and 4. the listener's task (e.g., stimulus recognition in open-set format, stimulus recognition in closed-set format, interval scaling technique). All of these factors have been shown to influence intelligibility scores, and the degree to which they do so has been the subject of extensive research (Harris, 1990; Osberger, 1992).

The focus of the present review addresses speech intelligibility in relation to the characteristics of the hearing-impaired individual. Degree of hearing loss has been found to be the most significant factor affecting speech intelligibility (Boothroyd, 1969,

1984, 1985b; Boothroyd and Lambert, 1980; Markides, 1970; Smith, 1975; Jensema, Karchmer, and Trybus, 1978; Gold, 1978). A negative correlation exists between degree of hearing loss and intelligibility scores. Thus, as severity of hearing loss increases, speech intelligibility decreases. Smith (1975), for example, found a correlation of -0.65 between word intelligibility¹ scores and average hearing level at 125, 500, and 1000 Hz. She reported a somewhat lower correlation ($r = -0.59$) when the average of 500, 1000, and 2000 Hz was taken. Smith had argued that selecting the average of 125, 500, and 1000 was justified because 24 out of her 40 children did not have measurable hearing above 1000 Hz. She did find, however, that the most intelligible children in her study had measurable hearing up to 3000 Hz.

Monsen (1978) explored the relationship between word intelligibility scores and a variety of audiometric configurations for 67 hearing-impaired children. The highest correlation of $r = -0.62$ was obtained between intelligibility and the average of 1000, 2000, and 4000 Hz. This correlation was only

¹ Recognition of monosyllabic words spoken in isolation in response to written input.

slightly higher than the correlation obtained for the average of 500, 1000, and 2000 Hz ($r = -0.6$).

Similarly, Boothroyd (1984) reported a significant correlation ($r = -0.62$) between phoneme intelligibility² and pure tone average of 500, 1000, and 2000 Hz for 120 hearing-impaired students (ages 11-18) at an oral school for the deaf. In an earlier study, Boothroyd (1969) reported on the relationship between phoneme intelligibility scores and single frequencies. He found significant correlations with percent intelligibility at all frequencies but they were strongest at 1000 and 2000 Hz.

The above-cited studies agree that approximately 40% of the variance in intelligibility scores is attributable to hearing sensitivity. However, it is not clear which frequency or combination of frequencies (pure tone average) gives the most predictable relationship between speech intelligibility and hearing loss.

Despite the general acknowledgment that degree of hearing loss is an important variable affecting speech intelligibility, the correlation between these two

² Recognition of phonemes in CVC monosyllable spoken in response to a written input.

factors is not perfect. Monsen (1978) reported that the children in his study with hearing losses of 95 dB or less had intelligible speech, whereas the children with losses greater than 95 dB did not always have unintelligible speech. Monsen pointed out that although 65% of his profoundly hearing-impaired subjects (n = 67) demonstrated quite similar audiometric thresholds, their speech intelligibility scores varied over a wide range. He concluded that "a 'good' audiogram is a fairly reliable indicator of intelligible speech, but a 'bad' audiogram does not necessarily always predict poor or unintelligible speech" (Monsen, 1978 p.215).

Other investigators have made similar observations. Boothroyd (1969), for example, reported that although four students at the Clarke School for the Deaf had little or no measurable hearing (in excess of 110 dB), they obtained high scores on the Magner speech intelligibility test (words in sentences, written input). In contrast, two other students with thresholds of 90 dB or less at 1000 and 2000 Hz obtained poor intelligibility scores.

Geers and Moog (1992) examined the relationship between speech production skills and communication mode for matched groups of hearing-impaired students from

two different educational settings. The two groups, one from Oral schools for the deaf and the other from Total Communication programs, were compared on a variety of measures of speech perception and production skills. Word intelligibility for students with hearing losses greater than 90 dB HL ranged from 73% to 87% for the Oral group and below 60% (unintelligible range) for the Total Communication group. Geers and Moog concluded that even profoundly hearing-impaired children have the potential to develop intelligible speech given an appropriate habilitation program. They agreed, however, that speech intelligibility is, to a certain extent, related to degree of hearing loss.

Researchers have hypothesized that differences in speech intelligibility among profoundly hearing-impaired children may actually reflect differences in their ability to use effectively the acoustic cues available to them, rather than differences of pure tone threshold per se (Smith, 1975; Stark and Levitt, 1974; Monsen, 1978; Boothroyd, 1984). If this is true, a stronger relationship should be found between speech production and auditory speech perception than between speech production and pure tone threshold.

Support for the above hypothesis is provided by Smith (1975), who reported that for her subjects, auditory perception of speech features was more highly correlated with speech intelligibility ($r = 0.78$) than with hearing level. Similarly, Monsen (1978) reported a higher correlation between speech intelligibility and five of the nine acoustically measured speech variables he had used ($r = 0.67-0.82$) (measurements of consonant production, vowel production, and suprasegmentals) than between speech intelligibility and the hearing level reported above.

Boothroyd (1984), in the study mentioned earlier, also examined the speech intelligibility (recognition of words in sentences) of his hearing-impaired subjects as a function of their auditory speech contrast perception performance. The correlation between speech intelligibility and the receptive measure was -0.68 . This is somewhat higher than the correlation of -0.62 reported between intelligibility and hearing level.

Stark and Levitt (1974) made an important observation on the relationship between degree of hearing loss and production skills. In their study, profoundly hearing-impaired children who scored well on perception of suprasegmental speech features did not always score well on production of the same features.

Children who scored high on production of suprasegmental features, however, consistently scored average to above average on the perceptual tests. Thus, Stark and Levitt concluded that good perceivers are not always good producers, but good producers are generally good perceivers.

Osberger et al., (1993), also found that profoundly hearing-impaired children with the highest speech intelligibility scores demonstrated relatively good auditory speech perception skills. There were some subjects, however, that demonstrated good perception skills, but whose speech intelligibility was poor. Thus, it appears that profoundly hearing-impaired children who demonstrate good speech production skills will also demonstrate good auditory speech perception skills. However, the converse is not always true.

In summary, numerous studies have shown a substantial correlation between speech intelligibility and the degree of hearing loss in subjects with prelingually acquired severe and profound hearing losses. On average, speech intelligibility performance tends to deteriorate with increasing hearing loss. Moreover, the strength of the correlations has been consistent across studies (approximately 0.6). Stronger

correlations, however, exist between speech intelligibility and auditory speech perception (approximately 0.7); that is, a hearing-impaired child tends to produce best those aspects of speech that he/she can perceive best via the sense of hearing. This suggests that speech production skills provide a good guide to auditory speech perception skills.

Auditory speech perception by profoundly hearing-impaired children

A review of the literature on the auditory speech perception of severely and profoundly hearing-impaired children reveals two obvious trends. First, overall speech perception performance declines as a function of increasing hearing loss, and second, wide variations in performances are present, mainly within the profound hearing loss range, even with almost identical audiometric configurations.

Erber (1974) demonstrated, for example, that auditory spondee recognition scores for children with average hearing losses better than about 85 dB are high

(70-100% correct), while spondee recognition scores for children with hearing losses poorer than about 100 dB are low (0-30% correct). For hearing losses between 85 and 100 dB, however, scores are widely spread (0 to 96% correct).

In a later study, Erber and Alencewicz (1976) collected data from 126 ears using the Monosyllable-Trochee-Spondee test (MTS). The test assesses the child's ability to identify words and word stress patterns presented in a closed set of 12 words in three stress categories (i.e., monosyllable, trochee, or spondee). Again, the data roughly divided into three groups of hearing losses. Subjects with hearing losses better than 95 dB scored between 70-100% correct for word identification and could clearly access spectral components of speech. Subjects with hearing losses above 110 dB scored between 0-25% for word identification, which implies purely random guessing within stress categories. However, subjects with hearing losses between 95 and 110 dB scored from near zero to near perfect. The authors argued that because pure tone threshold is not a good predictor of speech perception ability for children with average hearing levels between

about 91 and 110 dB, direct measures of speech recognition are needed.

It should be noted, however, that to a certain extent, the findings of Erber and Alencewicz reflect the use of percent correct as a measure, with its attendant floor and ceiling effects. An arcsine transform or a Z transform would work against the tendency for the variance to be reduced near the floor and the ceiling. Therefore, increases the homogeneity of the variance.

Typically, however, the findings indicate that in a group of subjects covering a hearing loss range of roughly 70 to 110 dB, about 50% of the variance in speech perception performance is attributable to the degree of the hearing loss. Consequently, the speech perception of individuals cannot be precisely predicted from knowledge of their hearing losses.

Another important aspect of sensorineural hearing loss is that it affects the perception of some speech features more than others. A considerable amount of research has been devoted to investigating the effect of various degrees of hearing loss on perception of segmental speech features. Pickett et al. (1972), for example, analyzed data obtained with the Modified Rhyme

Test (MRT) for 99 severely and profoundly hearing-impaired students. The test assesses perception of vowels and consonants in initial and final positions. Subjects were divided into four quartiles based on their average performance on the test. Across quartiles, vowels were perceived better than initial consonants which, in turn, were perceived somewhat better than final consonants. Overall, a significant correlation ($r = -0.66$) was reported between MRT score and average hearing loss.

The perception of specific components of speech has also been studied by Erber (1982) and Caplan-Hack and Erber (1982). They reported that children with severe to profound hearing losses have a tendency to confuse front and back vowels having first formants in similar frequency regions. The authors suggested that subjects perceptually confuse front and back vowels because they have difficulty accessing the acoustic information in the second formant region, which normally allows discrimination between these vowels.

In an earlier study, Erber (1972) assessed auditory discrimination of manner, voicing and place distinctions within a set of stop and nasal consonants in /aCa/ context. He found that severely hearing-impaired

subjects were able to discriminate voiced from voiceless consonants and nasal from non-nasal consonants (voicing and manner distinction). These subjects, however, could not discriminate within each voicing and nasality category (place distinction). Erber suggested that his subjects, who had hearing in the lower frequency regions, could perceive voicing and manner contrasts because many of the acoustic cues for these speech features are conveyed by the lower frequencies.

Boothroyd (1984) used a four-alternative, forced-choice procedure to measure the perception of seven speech contrasts. His subjects were school-aged children with average hearing losses ranging from 55-124 dB HL. Boothroyd showed that, overall, auditory access to all speech contrasts decreased with increasing hearing loss, but that some of the contrasts were more susceptible than others. Using a 50% performance criterion re chance, Boothroyd demonstrated that specific speech contrasts are auditorily accessible up to the following average hearing losses: 75 dB HL for consonant place of articulation, 85 dB HL for initial consonant continuance, 90 dB HL for initial consonant voicing, 100 dB HL for vowel place (front/back dimension), 105 dB HL for talker's gender, 115 dB HL for syllabic pattern of a

phrase, and in excess of 115 dB HL for vowel height. Boothroyd suggested that contrasts of vowel height are accompanied by changes of intensity. Thus, the data on vowel height and syllabic pattern are consistent with the hypothesis that some subjects with hearing losses greater than 110 dB HL can still have access to time and intensity cues, even if they do not perceive spectral cues.

Plant (1984) developed a diagnostic speech perception test (PLOT) that consists of nine subtests aimed at assessing perception of both segmental and suprasegmental speech features of increasing difficulty. Plant administered the test to 73 hearing-impaired children aged 6-18 years, with hearing losses ranging from 70-118 dB HL. The scores obtained reflect the general trend in the literature: performance declines as hearing loss increases, with the highest level of difficulty being the initial consonant place of articulation subtest.

The results of the above noted experiments indicate that among the segments, vowels are perceived better than consonants. Within the vowels, vowel height contrasts are perceived better than vowel place

contrasts (front/back dimension). Among the consonants, initial consonants are perceived better than final consonants, and voicing and manner contrasts are perceived better than place of articulation contrasts. Consonant place of articulation (for example, bilabial, alveolar, velar) is reported to be the speech feature most affected by sensorineural hearing loss. This probably occurs because important acoustic cues for distinction of place contrasts, especially rapid second formant transition to the neighboring vowels, are carried in the higher frequency region (Godfrey & Millay, 1978). The higher frequency region is usually inaccessible to individuals with severe to profound hearing losses.

Compared to the data on perception of segmental features, only a relatively small amount of quantitative data is available on the perception of suprasegmental features by severely and profoundly hearing-impaired children. Stark and Levitt (1974) and Levitt and Resnick (1978) developed a prosodic features reception test that was used in perception and production studies. The test assesses recognition of prosodic contrasts conveyed by means of simple sentences. The prosodic contrasts include intonation (question vs. statement), stress

(early vs. late position in a sentence), and pause (early vs. late position in a sentence). The test was administered to 93 profoundly hearing-impaired children. Pause was the most easily identified feature. Early stress was more accurately identified than late stress. Identification of intonation pattern was the most difficult task. The authors, however, felt that "the problem with the test is that many deaf children are not familiar with the concept of suprasegmental features and that some training is required before stable test results are obtained" (p. 125).

Boothroyd (1984), in the second part of the study described earlier, also assessed perception of talker gender and intonation. The task for his subjects (with hearing losses of 75-114 dB HL) was to decide, in a forced-choice procedure, whether a short phrase was spoken with a rising or falling intonation contour by a male or a female talker. The response alternatives took the form of cartoon faces: a male and a female "happy" face to represent rising contour and a male and a female "sad" face to represent falling contour. Subjects were given examples of these symbolic representations before the test began. According to a 75% performance criterion re chance, results indicated that talker gender was

perceivable by subjects with hearing losses up to 105 dB HL. Most of the subjects, however, performed at chance level on the intonation contrast. Puzzled by this finding, Boothroyd suggested that poor performance was task-dependent rather than perception-dependent. Similarly to Levitt and Resnick (1978), he associated poor performance with lack of experience in making judgments about intonation contours. Boothroyd argued that many of these subjects have appropriate intonation contours in their spontaneous speech which provides evidence of auditory access to this acoustic feature.

Most (1985) further explored perception of intonation by subjects with severe to profound sensorineural hearing loss. Most's subjects (ages 3-10 years) were presented with synthetic speech stimuli consisting of flat, rising, and falling intonation contours. The same stimuli were used in two evaluation procedures: the first involved an imitation task, and the second a forced-choice task. Overall, the data showed that all of the subjects with severe hearing losses, and about 65% of the subjects with profound hearing losses up to 110 dB, were able to perceive intonation contours. The subjects with hearing losses in excess of 110 dB (n=4), however, could not perceive intonation patterns.

From the review of the literature on the auditory speech perception of hearing-impaired children, it appears that pure tone threshold, although an important determinant of performance, is an imperfect predictor of auditory speech perception capacity, particularly for individuals with profound hearing loss. Children with similar hearing sensitivity have been shown to demonstrate very different levels of speech perception performance. These observations highlight the importance of assessing speech perception directly by administering speech tests rather than by inferring capacities from measures of hearing sensitivity.

An important question remains, however: to what extent can it be assumed that auditory performance on speech perception tests reflects true auditory capacity? In some of the studies reported earlier, repeated testing yielded improvement in performance (see, for example, Cramer and Erber, 1974). If we are truly measuring auditory capacity, then stable and consistent performance would be expected. Furthermore, linguistic proficiency and cognitive level are some of the variables known to affect speech scores. To what extent, then, are data reported here influenced by these variables, and is it possible to separate acquired skills from capacities?

**Methods of measuring speech perception
performance in young hearing-impaired children.**

The purpose of this section is twofold: first, to raise some issues that need to be addressed when designing and evaluating speech perception tests; second, to review some of the tests most commonly used to assess the auditory speech perception performance of hearing-impaired children.

Much of the impact of a sensorineural hearing loss on a developing child depends on the extent of the loss and its effect on the perception of spoken language. Knowledge of the hearing-impaired child's auditory speech perception capacity at an early age is crucial for planning and managing an appropriate habilitation program. One of the primary goals of habilitation is to enhance spoken language perception. To pursue this goal a series of decisions needs to be made regarding the selection of sensory aids, educational placement, and intervention programs. These decisions are not easily addressed, and the difficulties become even more acute when the hearing loss is profound. As discussed earlier, the profoundly deaf (i.e., those with losses in excess of 90 dB) present a wide range of performances for

auditory speech perception that cannot be accurately predicted from their pure tone threshold values. Hence there is a need for valid, reliable, sensitive, and efficient speech perception tests.

Unfortunately, designing good speech perception tests for young, profoundly deaf children is not an easy matter. The child's linguistic knowledge, cognitive status, auditory experience, and attention span are only a few of the variables that make it a difficult task. Therefore, it is not always apparent whether reduced performance on a test reflects a child's speech perception capacities or the inability to measure those capacities with appropriate, valid and reliable measures (Boothroyd, 1991a). Nevertheless, many speech-perception tests have been developed.

Any test needs first to be clear in defining its purpose and the property it is intended to measure. For many applications the interest of speech perception tests is in assessing sensory capacity. Examples are when assessing the need for prosthetic listening devices (e.g., hearing aids, cochlear implants), evaluating the effectiveness of such devices, or making decisions about candidacy for implants. These applications require speech perception tests that are maximally dependent on

sensory evidence and minimally dependent on variables such as context, cognitive status, and linguistic knowledge and skills. At the same time, however, the test should measure sensory capacity in a manner that is predictive of potential for the development of higher-level speech perception skills (Boothroyd, 1991a, 1991b; Tyler, 1993).

Other applications may call for tests that are sensitive to all aspects of speech perception, including cognitive and linguistic factors. Examples include tests of everyday function and measures designed to address academic needs and placement.

For purposes of the present study the focus is on peripheral auditory speech perception capacity. It is important to distinguish between peripheral auditory speech perception capacity and overall performance. A sensory aid may be able to provide sufficient auditory evidence for speech to be recognized. However, speech may not be understood due to limited linguistic knowledge and cognitive functioning. Furthermore, capacity is a hypothetical construct which cannot be measured directly. However, performance can be measured. We expect performance to improve over time, approaching some asymptote that is determined by peripheral auditory capacity. If we can be sure that the asymptote has been

reached then we have an indirect measure of peripheral auditory speech perception capacity. Thus, for present purposes, a test of peripheral auditory speech perception capacity is one in which cognitive, linguistic, and other age-dependent factors play no role. It is also a test in which the child may be assumed to be performing as well as he ever will.

In some disciplines, such as psychology, education, and language science, the use of standardized tests is a rule rather than an exception. A standardized test is one for which comparative norms on a large and representative sample of the target population have been derived, validity and reliability have been established, and comprehensive instructions for administration and scoring are provided. In speech and hearing sciences, however, this goal is rarely accomplished. The length and complexity of the process of standardization, as well as its high cost, are probably the main reasons. To the knowledge of this writer, there is no standardized speech perception test specifically designed for profoundly hearing-impaired children. In fact, most of the existing speech perception tests have been developed

in specific centers to address some of their own urgent clinical or research needs.

Even without complete standardization, however, there are two essential requirements that every test should meet: validity and reliability. Validity refers to the extent to which the test measures what it is intended to measure. Reliability refers to the consistency of the test in measuring whatever it measures (Ary et al., 1985, Walden, 1984).

Several aspects of validity, covering the basic purposes for which tests are used, have been outlined. These aspects include face validity, content validity, construct validity, and criterion-related validity. Face, content and construct validities are rather weak. They refer, respectively, to the extent to which a test appears to measure what is intended to measure, the extent to which its content represents the entire domain it is intended to measure, and the extent to which its underlying theoretical constructs are expressed in its content. These types of validity are assessed through subjective judgments. The fourth type of validity, criterion-related validity, is the most quantitative measure. Criterion-related validity refers to the relationship between performance on the test and performance on some independent variable (criterion)

that is known to measure what the test under evaluation is intended to measure. It is assessed empirically by obtaining coefficients of correlation between the test measure and the independent criterion. The difficulty in speech perception testing, however, is in finding such an independent criterion. By default, test developers tend to validate new tests against other tests that are commonly used but have not necessarily been validated themselves. It is possible, however, to use a test with high face, content, and construct validities as the external criterion. For example, in speech perception assessment, a nonsense syllable recognition test may be validated against a test of word recognition in everyday sentences.

Reliability, as stated earlier, is a measure of the stability and precision of a given score. There are two ways to assess the reliability of a test: by reliability coefficient, or by standard error of measurement. The reliability coefficient is obtained by correlating performance on a test administered to the same group of subjects on two separate occasions. The resulting correlation coefficient expresses the extent to which each subject maintains the same relative position within the group. Standard error of measurement is the standard

deviation of repeated measurement for a single subject, from which confidence limits for a single score can be derived. The standard error of measurement gives an estimate of what would be the spread of scores for a single subject if he were to be tested repeatedly. A reliable test has a high reliability coefficient and a low standard error of measurement. The first measure is important for tests that distinguish among subjects. The second is important for tests designed to compare conditions or test occasions within a single subject.

The foregoing discussion has been about important issues that need to be considered when evaluating any new tool of assessment. There are several important issues, as well, that need to be considered specifically when designing speech perception tests.

When the purpose of a test has been clearly defined and the target population identified, it is time to select the stimuli and determine the task. In speech perception, the unit being assessed can be sentence, word, syllable, phoneme, or speech feature. The speech context can be syllable, word, or sentence. The test context can be verbal, pictorial, object or print. The perceptual task can be discrimination (closed set of two

options, or detection of a difference), identification (closed-set), recognition (open-set), or comprehension. The number of items or lists of items to be included in the test should be considered. A large number of items increases reliability but also increases testing time. Mode of presentation can be either live voice or recordings. The response task can be oral, written, carrying out a series of actions, pointing to (or selecting) a picture, an object, a printed item, or a position (in an oddity task). The resulting performance can be expressed in the form of a single numerical value, profile (several numerical values representing different stimuli, contexts, or tasks), or classification of subjects into categories (Erber, 1979; Boothroyd, 1991a; Tyler, 1993).

When designing a speech perception test for the sole purpose of assessing the adequacy of sensory evidence, it is necessary to minimize the confounding effects of other factors. To accomplish this purpose one may choose to use, for example, nonsense syllables as the stimuli to be perceived, thus eliminating the confounding effects of sentence-level comprehension and lexical knowledge. However, to a certain extent, the effect of phonological knowledge remains (Boothroyd, 1995). The nonsense-syllable test can be presented as an

oddity task, thus eliminating the need for reading skills and the ambiguity that can result from presentation of pictures.

The following is a review of some of the tests most commonly used to assess various aspects of the auditory speech perception performance of hearing-impaired children. The purpose of the review is to examine the ways in which test developers chose to address and overcome the problems of designing and evaluating speech perception tests. The tests are grouped according to their primary purpose, although some of them can be used to assess several different aspects of speech perception.

Envelope And Stress Perception Tests

Various measures have been developed to distinguish children who can perceive words on the basis of fine temporal and spectral information from those who can only perceive gross syllabic patterns of speech.

Auditory Number Test (ANT) (Erber, 1980)

The ANT's stimulus materials consist of five cards each having a number from one to five and a picture of

the corresponding number of ants. The child is trained to point to the correct card based on detection of syllable number. For example, "One-two" is presented to elicit a response to "Two". During the test, either a sequence of numbers (such as "One-two-three"), or a single number (such as "Three") is presented. The score is the number of correct identifications only of the single-number presentations. Chance score is 20%. Selecting the correct single number suggests gross spectral perception. In contrast, selecting always the number "One" card when a single-number is presented suggests pattern perception. The test requires basic number concepts. No evaluative data are available on the test.

Monosyllable-Trochee-Spondee (MTS) Test (Erber and Alencewicz, 1976)

The MTS test assesses stress-pattern categorization and word identification. It consists of 12 pictured words with three different stress patterns, four monosyllables, four trochees, and four spondees. Each word is presented twice, and the child's pointing responses are scored in term of percentage correct stress-pattern categorization and word identification. The score for each subject is averaged across the two

trials. The stress-pattern score depends on whether the word selected comes from the correct syllable-stress category, with a chance score of 33%. A word score is determined by the number of words identified correctly, with a chance score of 8.3% (or 25% if the child perceives differences in stress pattern correctly).

The requirement of this test is that the words be within the child's vocabulary. Erber and Alencewicz stated that the test is appropriate for most hearing-impaired children aged five years and older. In their data, however, they included children aged eight to 16 years.

Erber and Alencewicz administered the test to 67 prelingually hearing-impaired children with hearing levels between 55-125 dB. A test-retest reliability assessment obtained for 85 ears yielded a high reliability coefficient ($r = 0.87$).

The ANT and the MTS test served as a basis for the Early Speech Perception Test battery described next.

Early Speech Perception (ESP) Test (Moog & Geers, 1990)

The ESP test was designed to measure the ability of profoundly hearing-impaired children to choose a pictured word based on selected auditory distinctions.

The test is used to place children in four speech perception categories:

1. No pattern perception.
2. Pattern perception.
3. Some word identification.
4. Consistent word identification.

The ESP battery consists of a pattern perception subtest and two word identification subtests. The criteria for passing score and speech category assignment vary across subtests. The Pattern Perception subtest includes 12 words that differ in number of syllables or stress pattern (three monosyllables, three trochees, three spondees, and three three-syllable words). A response to any word with the same stress pattern as the test word counts as correct. Chance performance is 25%. The Spondee Identification subtest uses 12 spondees, and good performance requires access to spectral information. Chance performance is 8%. The Monosyllable Identification subtest includes 12 monosyllabic words, all starting with the letter /b/. Words differ from one another primarily in the vowel sound. Chance performance is 8%.

A Low Verbal version of the ESP test has been developed for younger children with limited vocabulary and attention span.

The test's manual indicates that the predictive validity of the ESP was established by correlating performance on the test with performance on the WIPI test (Ross & Lerman, 1970). The WIPI, however, was never validated. In addition, while the WIPI was designed for children with moderate to severe losses, Moog and Geers collected data from a group of profoundly hearing-impaired children.

Twenty-three profoundly hearing-impaired children, aged seven to 15, were tested twice with the ESP test over a 30-day period. Test-retest correlation coefficients were 0.78 for pattern perception, and 0.89 for the spondee and monosyllable combined score. Moog and Geers also reported a high reliability coefficient for category placement ($r = 0.94$).

Closed-Set and Open-Set Word Identification Tests

Both closed and open-set word identification tests, have been used as clinical tools (Penrod, 1994). They represent a traditional way of measuring speech perception performance. The purpose of these tests was never clearly defined. One may assume, however, that the attempt is to measure auditory capacity rather than

communicative competence. Word identification tests use words in isolation, thus eliminating confounding effects of syntactic knowledge and contextual evidence. Other effects, however, such as those of lexical and phonological knowledge, are maintained. The use of age-appropriate word lists is preferable so as to minimize such confounding effects on test scores. Identification of words in a closed-set format is easier than in an open-set format. In a closed set, all alternatives are present and should be known to the child. In an open-set format, the possible response alternatives are the subject's entire vocabulary. Vocabulary, however, may be an additional confounding factor, and items of equivalent difficulty are difficult to develop.

Closed-Set Word Identification Tests

Word Intelligibility by Picture Identification

(WIPI) (Ross & Lerman, 1970)

This closed-set identification test requires the child to identify a monosyllabic word from a six-choice picture set. The test items share phonetic similarities with the response alternatives, and primarily require consonant discrimination ability (e.g., hair, pear, stair, bear, and chair). Chance performance is 17%. WIPI

is recommended for children of age 5-6 with moderate hearing loss, and up to age 7-8 with severe hearing loss. WIPI is considered more difficult than the majority of other closed-set identification tests, first because the vocabulary level is higher (e.g., seal, screen), and second because it evaluates consonant discrimination, which is more difficult than vowel discrimination. In addition, clinicians often complain of either poor or old pictorial representation. Nevertheless it is a commonly used test.

Four test lists were administered twice to 61 children with moderate to severe hearing loss. Their ages ranged from 4 years, 7 months, to 13 years, 9 months, with a mean of 10 years, 2 months. The test-retest reliability coefficients range from 0.87 to 0.94 (Ross and Lerman, 1970).

Test of Auditory Comprehension (TAC) (Trammell & Owens, 1980)

The TAC evaluates a wide range of auditory abilities. It consists of ten subtests, ordered in a hierarchy of task complexity ranging from discrimination between environmental sounds and speech to comprehension of stories in a background of competing speech. The ten subtests are presented in sequence until two consecutive

subtests are failed. The number of items per subtest, the number of alternatives per item, and the percent correct passing criterion vary from one subtest to another. The child hears recorded stimuli and responds by pointing to a picture.

The TAC test was standardized on children between the ages of 4 and 17 years with varying degrees of hearing loss. However, it does not provide normative data for profoundly hearing-impaired children. Six of the subtests require high levels of language and cognitive skills, and thus seem inappropriate for young children.

Northwestern University Children's Speech
Perception Test (NU-CHIPS) (Elliott & Katz, 1980)

This test assesses identification of monosyllabic words in a four-choice picture-pointing format. The test requires both vowel and consonant discrimination ability in order to select the test word from its foils. A total of 50 items are presented on audiotape or via live voice. Chance performance is 25%. Elliott and Katz indicated that the test uses words that are usually within the vocabulary of seven-year-old hearing-impaired children. Most of the data they have presented, however, were obtained on normal hearing children.

Performance on the test was correlated with performance on the WIPI and with pure tone thresholds. As discussed earlier, however, the WIPI was never validated. Pure tone threshold values, again, are not a perfect predictor of speech perception performance. In fact Eliot and Katz's sample included only a small number of hearing-impaired children in each of their four hearing status categories. However, in the absence of a direct criterion to establish validation, additional empirical evidence of performance on other tests (e.g. a sentence-level test) would add support to their data.

An estimated reliability was obtained for the four test forms using the Spearman-Brown prophecy formula and the Kuder-Richardson formula 20 (Eliott and Katz, 1980). It is a less desirable form of reliability which assumes the uniformity of the test and tends to overestimate its real reliability (Ary et al., 1985). The estimated reliability for the hearing-impaired subjects ranges from 0.83 to 0.96.

Open-Set Word Identification

Phonetically Balanced Kindergarten Test (PBK)

(H.A., Haskins, 1949)

This open-set word identification test has four phonetically balanced lists of 50 monosyllabic words. The lists were geared to the vocabulary level of kindergarten-age children (six years and older). Performance is scored as the number of words and the number of phonemes repeated correctly. A child who can repeat two or more words from a set of 50 words is judged to have developed some open-set word identification. The disadvantage of this test is that word familiarity can affect performance on the test.

No published evaluative data are available on the PBK test. Nevertheless it is a commonly used test.

Sentence Tests

The purpose of sentence-level tests is to provide a global estimate of speech perception performance. Sentence tests should realistically represent the performance of a subject in natural conversation situations. Performance on a sentence test reflects the

influence of cognitive and linguistic variables such as memory, vocabulary, and syntax.

Glendonald Auditory Screening Procedure (GASP)
(Erber, 1982)

The sentence recognition subtest of the GASP consists of ten simple questions (e.g., "What is your name?", "What color is your shoe?") presented in an open-set format without visual cues, and requires the child to answer the question using his or her preferred communication mode. A sentence is scored correct only if the child responds to the question appropriately. This subjective scoring technique could affect validity, reliability or both.

No evaluative data are available on the test.

Common Phrases Test (Osberger et al. 1991b)

This test was developed to assess performance in a setting where the child has some contextual information. Before the test begins the child is shown a set of 10 pictured items (e.g., door, birthday cake) and is told that the following sentences will be about the pictures. The pictures are removed when testing begins. There are six forms with ten phrases per form. Each phrase is a simple command or question consisting of two to six

words. The child's task is to repeat (by speaking or signing) the entire phrase. The test is presented in an audition-plus-vision condition first. If the child repeats correctly at least two of the phrases, testing is performed in an audition-alone condition. A disadvantage is that the child is required to repeat short, highly predictable and familiar phrases. In addition, the ability to remember topics from the pretest pictures could differ across children and affect the results.

No evaluative data are available on the test.

Speech Contrast Perception

Several tests have been developed in an attempt to obtain detailed information about relative access to various phonetic contrasts. The purpose is to measure the amount of sensory evidence that is available to a given subject. To accomplish this purpose it is necessary to minimize the confounding effects of linguistic knowledge. Presumably these tests of speech contrast perception come closest to meeting the need for a test of "peripheral auditory capacity".

Change/No Change Test (Osberger et al. 1991a & b)

The Change/No Change test assesses the ability to perceive change in a sequence of nonsense syllables. There are seven subtests, each of which contrasts one particular feature of speech: syllable length, intonation, fundamental frequency, talker gender, vowel height, vowel place, and consonant manner. Each subtest has 15 trials, each consisting of a string of ten nonsense syllables. Five of the trials are "no change" trials in which the ten syllables are the same within a trial (e.g., dadadadadadadadadada), and ten are change trials in which the last five stimuli differ from the first five (e.g., dadadadadazazazazaza). The child is instructed to raise his hand when a change trial occurs, or to respond "same" or "different" after each trial.

The advantages of the test are the use of nonsense syllables and the simplicity of the task and the response (which can be either verbal or nonverbal). A disadvantage is the high chance performance level, which reduces the efficiency of the test.

No evaluative data are available on the test.

Speech PAttern Contrast Test (SPAC) (Boothroyd, 1984)

This test is designed to estimate the child's ability to perceive phonologically significant speech contrasts in a varying phonetic context. It involves the identification of words and phrases using a forced-choice procedure in a printed format. SPAC evaluates the perception of ten speech contrasts: vowel height, vowel place, initial and final consonant place, voicing, and continuance, word stress, and pitch rise or fall. There are 12 trials within each subtest. The test generates two measures: a profile of individual speech contrast perception, and a single composite score (based on the average of all ten contrasts).

SPAC has the advantage of efficiency because data on several speech contrasts can be obtained in a relatively short period of time. The main disadvantage of SPAC is that it requires both vocabulary knowledge and reading ability. Therefore, the test is most appropriate for postlingually hearing-impaired children ten years and older, who have good reading skills.

Criterion-related validity was established between SPAC scores and word recognition in sentences of known topic (Boothroyd, 1991a), but only for normally hearing adults listening to distorted speech. Performance on

SPAC was also correlated with the speech intelligibility of 120 hearing-impaired students and with their open-set phoneme recognition of CVC words. The coefficients of correlation between SPAC composite score and speech intelligibility and phoneme recognition were 0.68 and 0.81, respectively.

THree-Interval Forced-Choice Test of Speech Pattern Contrast Perception (THRIFT) (Boothroyd et al., 1988)

The effects of vocabulary knowledge and reading skills are eliminated in the THRIFT, a three-interval oddity test in which the subject must choose the odd-man-out from a sequence of three Consonant-Vowel (CV) or Vowel-Consonant (VC) nonsense syllables. In a given subtest, the odd-man-out differs from the other two along a single, phonologically significant dimension. The child hears, for example, "soo, zoo, soo", and must say whether the first, second, or third syllable differs from the other two. Within each subtest, the contrast remains constant, but the phonetic context changes from trial to trial. THRIFT evaluates the perception of nine contrasts: intonation, vowel height, vowel place, initial and final consonant place, voicing, and continuance. Similarly to SPAC, THRIFT provides a profile of individual speech contrast perception as well

as a single composite score based on the average of all nine subtests.

THRIFT's greater advantage is the use of nonsense syllables, which greatly reduces the effect of linguistic knowledge. The disadvantage is that it is cognitively and motivationally demanding, and thus not recommended for hearing-impaired children below the age of eight.

Criterion-related validity was established between THRIFT scores and phoneme recognition in CVC words and CVC words in a carrier phrase, and word recognition in sentences of known topic (Boothroyd, 1991b), but only for normally hearing adults with simulated hearing losses. THRIFT scores were also correlated with pure tone thresholds of 71 hearing-impaired children with hearing losses from 75 to 122 dB HL ($r = -0.65$). THRIFT was administered twice to 110 hearing-impaired children and yielded a high reliability coefficient ($r = 0.81$) (Boothroyd, 1991b, 1995).

Table 2.1 summarizes the evaluation procedures accomplished for each of the above listed tests. The second column indicates whether the test has been completely standardized; and if not, the third and

fourth columns indicate whether criterion-related validity and reliability have at least been established.

Table 2.1. Summary of evaluative data available for each of the reviewed speech perception tests.

Test	Standardized	AT LEAST Criterion-Related	
		Reliability	Validity
Auditory Number Test (ANT) (Erber, 1980).	NO	NO	NO
Monosyllable-Trochee-Spondee (MTS) Test (Erber and Alencewicz, 1976).	NO	NO	YES
Early Speech Perception Tests (ESP) (Moog & Geers, 1990).	NO	YES	YES*
Word Intelligibility by Picture Identification (WIPI) (Ross & Lerman, 1970).	NO	NO	YES
Test of Auditory Comprehension (TAC) (Trammell & Owens, 1980).	YES*		
Northwestern University Children's Speech Perception Test (NU-CHIPS) (Elliott & Katz, 1980).	NO	YES*	YES*
Phonetically Balanced Kindergarten Test (PBK) (H.A., Haskins, 1949).	UNKNOWN	UNKNOWN	UNKNOWN
Glendonald Auditory Screening Procedure (GASP), Sentence subtest (Erber, 1982).	NO	NO	NO
Common Phrases Test (Osberger et. al, 1991b).	NO	NO	NO
Change/No Change Test (Osberger et al, 1991a&b).	NO	NO	NO
Speech Pattern Contrast Test (SPAC) (Boothroyd, 1984).	NO	NO	YES*
THree-Interval Forced-Choice Test of Speech Pattern Contrast. (Boothroyd et. al., 1986).	NO	YES	YES*

* Refer to the text for further details

From the foregoing review it is clear that despite years of creative test development there is a shortage of evaluative data concerning these tests. Moreover, the purposes of many of these tests are not always clearly defined nor even stated. Therefore, a clinician or researcher who may need a tool of assessment will encounter great difficulty in determining how well a given test measures the properties it intended to measure, with what consistency, and for which population. Not surprisingly, many tests have been developed in attempts to overcome the problems outlined above, so as to satisfy specific needs. Somewhat ironic, however, is the fact that constructing new tests based on awareness of the limitations of other tests does not necessarily result in confronting the difficulties.

Despite the foregoing criticism it is acknowledged that auditory speech perception is a difficult entity to assess, particularly with young prelingually hearing-impaired children. It follows that neither designing nor evaluating speech perception tests is an easy task. Nevertheless, there are some guidelines that need to be closely followed when creating a new speech perception test. First, the purpose of the test ought to be clearly stated and translated into properties to be measured. Second, the target population needs to be identified.

And finally, empirical data in support of test validity and reliability must be produced. The issue of validity must be considered in relation to the purpose of the test. Each of the four types of validities outlined earlier is concerned with a different aspect of the general question 'Does the test measure what it is supposed to measure?' In the absence of data to establish validity directly, it is proposed that several indirect relations be demonstrated with other measures of properties assumed to be correlated with what the test is intended to measure. How consistently does a test measure whatever it does measure? This is the problem of reliability.

The last two tests described above, SPAC and THRIFT, were created as part of the effort to develop measures of peripheral auditory speech perception capacity in hearing-impaired children. Both tests are based on the concept of speech pattern contrast testing. The goal of speech pattern contrast testing is to estimate the probability of perception of phonologically significant speech contrasts in a varying phonetic context (Boothroyd, 1991b). SPAC was developed first and was found unsuitable for use with young children. Next,

THRIFT was designed to avoid the need for vocabulary knowledge and reading skills. Although meeting its goal, THRIFT was found to be cognitively demanding, and thus is recommended for children older than eight. In a continuing effort to meet the need for an appropriate test in younger children, the IMitative test of Speech PAttern Contrast perception (IMSPAC) (Boothroyd, 1991a) was developed. The IMSPAC test is described in greater detail in the next chapter.

Discussion of the Literature

Prelingual hearing loss affects both speech perception and speech production. It seems logical to assume that the cause-effect relationship is sequential. That is, hearing loss affects the ability to hear speech which in turn causes a loss of ability to produce speech. This assumption is supported by findings reported in the literature just reviewed. The correlations between measures of perception and production are reported to be higher than those between either of these and pure tone threshold. Speech features that are most difficult to perceive tend to be also the ones that are most difficult to produce. This implies a strong relationship between auditory perception and speech production, although not a perfect one.

Although it seems that speech production skills provide a good guide to auditory speech perception skills, the literature does not provide a more definitive account of this relationship. One of the goals of the present work is to explore this relationship.

Several attempts have been made to assess peripheral auditory speech perception capacity in young prelingually deaf children fitted with hearing aids or

cochlear-implants. In general, the exact purpose of these tests has not been well articulated in relation to the distinction between auditory capacity and auditory performance. In the absence of a clear definition of purpose, and of pre-existing methods of meeting that purpose, it is not surprising that evaluative data, particularly on the validity of these tests, are virtually non-existent. Judgment of the tests therefore rests on face validity, which is subject to the knowledge, beliefs, and biases of the judge.

Two main conclusions can be drawn from the reviewed literature. First, there is a need for valid and reliable tests of peripheral auditory speech perception capacity whose results are not confounded by linguistic and cognitive factors. Second, because of the strong link between auditory speech perception and speech production, there is a possibility that production can be used as an indirect means of assessing perception.

These conclusions led to the development of an imitative test of speech pattern contrast perception, the details of which are to be presented in the next chapter. The study which follows represents an attempt to address the validity of this test as a measure of peripheral auditory speech perception capacity.

Chapter 3
The Imitative Test of Speech Pattern
Contrast Perception (IMSPAC)

The Imitative Test of Speech Pattern Contrast Perception (IMSPAC) grows out of a continuing effort to meet the need for an appropriate test of sensory capacity for young pre-phonological hearing-impaired children. The following chapter consists of three parts. The first part focuses on describing the IMSPAC test, including a discussion of its rationale, purpose and validity. The second part includes a detailed description of the IMSPAC procedure, from the initial stage of recording to the final stage of auditing. The third part of the chapter reviews some of the relevant findings obtained so far with the IMSPAC test as part of the research project on the efficacy of cochlear implants in children at the Graduate Center of the City University of New York (Boothroyd et al., 1995; Boothroyd and Eran, 1994). The data reported here are not part of the data collected for this dissertation, but serve as the starting point. The data are reported at length because the findings are directly relevant to the interpretation of the data to be presented in Chapter 5. The chapter proceeds with a discussion of the

main findings and ends with a summary of the specific goals of the present study.

One of the general goals of the research project on the efficacy of cochlear implants in children was to develop actuarial methods for assigning probabilities of the outcomes of amplification and implantation in young deaf children. More specific goals included the development of pediatric tests for the assessment of peripheral auditory speech perception capacity and the establishment of a database of peripheral auditory speech perception capacity in hearing-impaired children who wear hearing aids.

To address these goals, the vocal imitations of 97 aided children were tape recorded using the IMSPAC test. The children's imitations were digitized, edited and subsequently evaluated for their ability to convey phonetic contrasts to teams of normally hearing listeners.

IMSPAC was designed to provide a measure of auditory speech perception that is maximally sensitive to inherent peripheral auditory speech perception capacity and minimally sensitive to the effects of

lexical and other higher linguistic factors (Boothroyd, 1991a, 1991b; Boothroyd and Eran, 1994).

Performance on the IMSPAC test is an estimate of the child's ability to produce, by imitation, phonologically contrastive utterances. Since the utterances are produced by imitation of an acoustic input only, the results may be taken as a lower estimate of the child's peripheral auditory perception of these phonetic contrasts. If further testing shows the child to be capable of producing the contrasts at a higher level of performance, then the audition score may be taken as a valid indicator of perceptual capacity.

Part I: The Imitative test of Speech Pattern Contrast perception (IMSPAC)

IMSPAC estimates the probability of the perception of phonologically significant contrasts presented in a varying phonetic context. A child is asked to imitate a series of consonant-vowel (CV) or vowel-consonant (VC) syllables, such as "boo" or "oob." Both the tester's models and the child's imitations are tape recorded. The child's utterances are then extracted in digital form and, subsequently, presented to a group of four normally hearing listeners. For each utterance, the task of the

listeners is to determine in a four-alternative, forced-choice context what he/she believes to have been the model the child was trying to imitate. For example, for imitation of the syllable "zaw," the listener must decide whether the model was "faw," "vaw," "saw," or "zaw." Responses averaged over several listeners reveal the child's ability to convey information about two phonetic contrasts—initial consonant place of articulation and initial consonant voicing. A response of "saw," for example, is scored correct for initial consonant place of articulation, but incorrect for initial consonant voicing. A response of "vaw" is scored correct for voicing and incorrect for place. A response of "faw" is scored incorrect for both contrasts. Only a response of "zaw" is scored correct for both initial consonant voicing and initial consonant place. To increase face validity of the test as a predictor of performance at the sentence level, the phonetic context within a subtest changes from trial to trial. Following the example just given, the next test stimulus might be "shee" with response alternatives of "zhee," "see," "shee," and "zee."

IMSPAC assesses performance on ten phonologically significant speech contrasts. Two are suprasegmental

contrasts: intonation and syllable number. Eight are segmental contrasts: vowel height (high vs. low), vowel place (front vs. back), initial consonant voicing, initial consonant continuance (stop vs. continuant), initial consonant place, final consonant voicing, final consonant continuance, and final consonant place. As indicated above, the ten contrasts are presented in pairs to give a total of five subtests. The allocation of the contrasts to subtests is outlined in Table 3.1 together with examples of the individual contrasts and response alternatives for each subtest. A single subtest consists of eight syllables, giving a total of 40 syllables for a complete test form. There are four equivalent forms of the test: A, B, C, and D. Each form involves a different selection of the test stimuli from the response alternatives presented to the listeners.

Table 3.1 Five subtests of the IMSPAC test, together with examples of the contrasts and the response alternatives.

Sub-test	Contrast	Example	Response Alternatives
1	Syllable Number (SYL)	"ah" versus "ah-ah"	ah!/ah?/ah-ah!/ah-ah?
2	Intonation (INT)	"ah?" versus "ah!"	
	Vowel Height (VH)	"bee" versus "beh"	mee/meh/bee/beh
3	Initial Consonant Continuance (ICC)	"beh" versus "meh"	
	Vowel Place (VP)	"oosh" versus "eesh"	oosh/ooch/eesh/eech
4	Final Consonant Continuance (FCC)	"eesh" versus "eech"	
	Initial Consonant Voicing (ICV)	"faw" versus "vaw"	faw/vaw/saw/zaw
5	Initial Consonant Place (ICP)	"faw" versus "saw"	
	Final Consonant Voicing (FCV)	"eed" versus "eet"	eed/eeek/eeg/eet
	Final Consonant Place (FCP)	"eed" versus "eeg"	

The test generates two types of score: individual contrast scores, each based on eight binary decisions, and a composite score (mean of the ten contrasts) based on 80 binary decisions.

Administration of the test is done under two input modalities: first by auditory-visual input, and then by auditory-only input. In the first presentation, the child is given every possible input. The child is encouraged to listen, speechread, and read the syllables (if he/she can read) either in print or finger-spelled (if he/she knows that code). The main goal of testing with every possible input is to demonstrate that production is not a limiting factor for the auditory imitation task. The assumption, here, is that if a child is provided with optimal conditions, the child will give his/her "best" production performance. There are two objectives in testing the hearing-impaired child **initially** in the auditory-visual condition. The first is to increase familiarity with the task and to reduce any anxiety that may be provoked by "putting him on the spot" when testing his/her weaker ability. The second is to optimize the learning of the task.

Following the first test presentation, all of the visual cues are removed, and the child is required to imitate the tester's models given only auditory input. If the auditory-visual score is significantly higher than the auditory-only score, the difference is attributable to a limitation on the part of the peripheral auditory mechanism. The auditory-only score may then be taken as a measure of the child's peripheral auditory speech perception capacity. If the auditory-visual score is no better than the auditory-only score, then the auditory-only score may be taken as a lower estimate of peripheral auditory speech perception capacity. If both scores are close to zero, then the information provided by the auditory-only score is of negligible value (Boothroyd and Eran, 1995).

Part II: The IMSPAC Procedure

The IMSPAC test procedure consists of three stages: recording, editing and auditing, as illustrated in Figure 3.1.

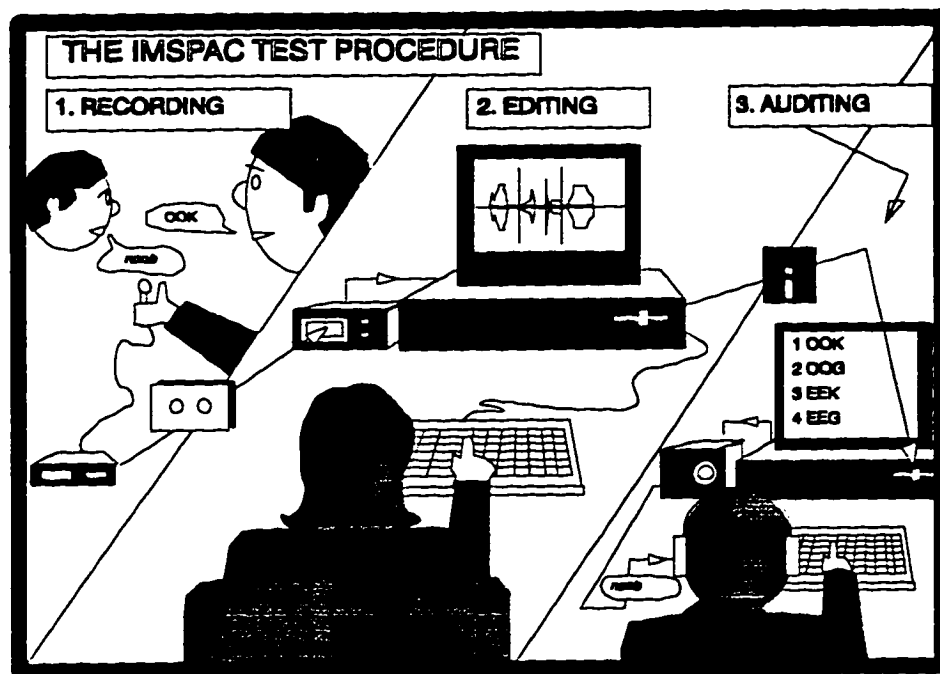


Figure 3.1. In the IMSPAC procedure:

1. The child is first required to imitate a series of syllables and his imitations are recorded.
2. The recordings are digitized, edited to remove all but the child's utterances, and saved as a series of files.
3. The edited utterances are then presented to normally hearing listeners who are required to respond in a four-alternative, forced-choice format.

The complete test involves two presentations. In the first, the child receives input by both audition and vision (speechreading, print, etc.). In the second, he/she receives input by audition only. The first score provides an estimate of the child's best production ability, given all possible inputs. The second provides an estimate of how well he/she can produce the syllables when only allowed to hear them. The present study is designed to test the hypothesis that this second, auditory-only, score provides an estimate of inherent peripheral auditory speech perception capacity, uncontaminated by limited production skills (Boothroyd et al., 1995).

The Recording Stage

Instrumentation

In collecting the data to be described, auditory stimuli were presented to the subjects monaurally under TDH50 headphones. Input was by live voice from a Realistic Microphone (Model no. 33-1052) and a two-channel Stimulus Presentation System (SPS) (Yeung et al., 1991). The SPS consists of input preamplifiers, a peak level indicator, an optional 1 KHz high pass filter, and power amplifiers. Passive output attenuators control presentation level in 2 dB steps up to a maximum undistorted level of 136 dB SPL (rms for pure tones). The bandwidth of the system is approximately 80-8000 Hz.

The subjects' productions were picked up by a Shure head-worn dynamic microphone (Model SM10A), placed approximately two inches from the mouth. The subjects were recorded with a Panasonic portable digital audio tape recorder (Model no. SV-255). The recordings were made on both channels of the tape recorder, but the level of the left channel was set to 10 dB below that of the right channel. In case overloading occurred on the primary channel, there was a good chance of an undistorted sample on the second channel. During testing, monitor headphones (Sennheiser HD414) were used by the tester to monitor the recorded signal.

Procedure

Speech awareness threshold (SAT) and highest acceptable listening levels for live voice stimuli were obtained using headphones via the SPS. A test ear was selected using the following criteria: a) if the child was monaurally aided, the aided ear was selected; b) if binaurally aided, the better ear (lower pure-tone-average) was selected; or c) if binaurally aided and the loss was symmetrical, the subject chose the preferred ear. Test stimuli were presented at the child's highest acceptable listening level, or at 30 dB SL, whichever was lower. If a sensation level greater than 10 dB could not be obtained, testing was done in a soundfield using the child's personal amplification.

One IMSPAC form per input condition (auditory-visual and auditory-only) was administered, for a total of 80 utterances. The form number (out of four) and the order of presentation of the subtests were randomly assigned for each child to reduce order effects. In the auditory-visual condition, the child faced the tester who produced a series of syllables the child was required to imitate. Both the tester's models and the child's imitations were tape recorded. If less than perfect imitation was produced, the tester would repeat

the model and ask for additional imitation. This procedure was repeated before advancing to the next item if the tester felt the child was capable of better performance. As discussed earlier, the tester employed various strategies in the auditory-visual condition to elicit the best production. Strategies included finger spelling the test item, pointing to the printed syllable, emphasizing the contrast ("that was a /b/, not a /p/"), and even providing brief training, especially with more difficult subtests such as intonation and syllable number.

In the auditory-only condition, the tester re-positioned the child to remove any visual cues and used a different form of the test. As in the auditory-visual input condition, every stimulus not imitated correctly was repeated. Here, however, no additional strategies were used beyond encouraging careful listening.

The Editing Stage

The analog test recordings were first converted to digital form using the Qdisplay digital wave-form editing software (Yeung, 1995). The recordings were digitized at 20,000 Hz sampling rate and 12-bit quantization. Next, the recordings were edited to remove

everything but the subject's best utterances. The edited utterances—a total of 40 syllables per condition—were then stored on floppy disks.

The Listening Stage

The edited utterances were presented to a group of four normally hearing, native English-speaking listeners. The task of the listeners was to determine, in a four-alternative, forced-choice context, what they believed to be the model the subject was trying to imitate.

Each listener was seated at a listening station that consists of a portable IBM-compatible computer, a custom-built box that contained antialiasing filters, a microphone amplifier and a power amplifier for play back, and Sennheiser HD414 headphones. Stimulus presentation and scoring were controlled by software. Each successive utterance was triggered by the listener's keying in the response. Performance on the listening task was averaged across the four listeners, reducing the contribution of listener differences to error variance.

Part III: Evaluative Data

Perception, Production, and Communication Mode

Figure 3.2 illustrates the auditory-visual and auditory-only relationship for 22 implanted and 46 aided children (Eran and Boothroyd, 1993). The horizontal axis shows the auditory-only performance, and the vertical axis shows the auditory-visual performance. Each data point is averaged across 10 speech contrasts and four listeners. The solid line shows the linear regression function. There is a significant correlation of $r=0.80$ between the two input modalities. The data show that for most of these children, performance with auditory-only input was not limited by production abilities, and thus their auditory scores can be taken as a measure of peripheral auditory speech perception capacity. In only a few cases were the two scores indistinguishable. In these cases, we may assume that peripheral auditory speech perception capacity is at least as good as that shown by auditory-only performance.

Although not reported here, Boothroyd et al., (1995), have shown that orally-trained children perform better on the auditory-plus-visual task than children in total communication (TC) programs. Performance of the two groups on the auditory-only task, however, was

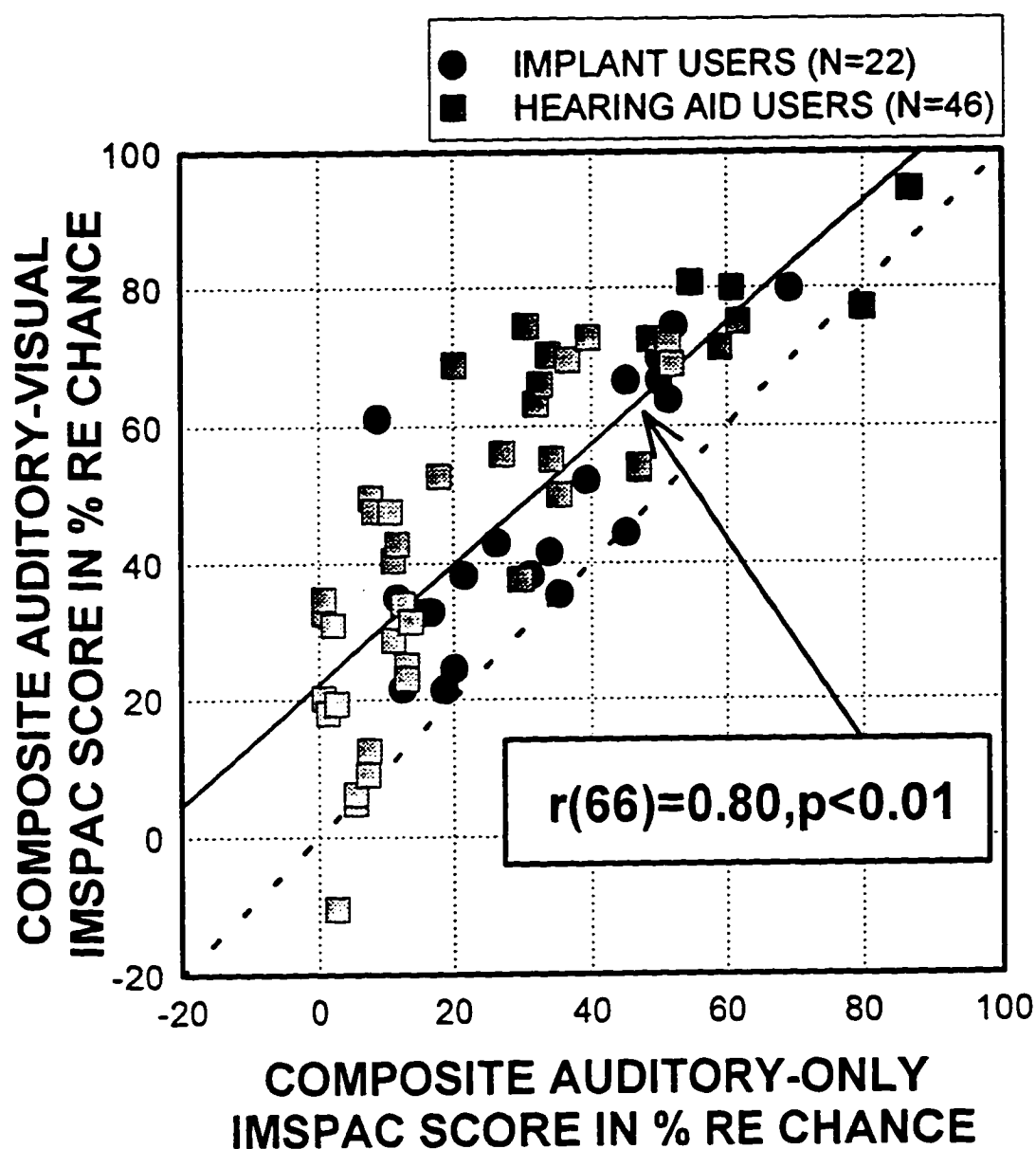


Figure 3.2. Composite auditory-visual IMSPAC scores as a function of auditory-only scores for 22 implant users and 46 hearing aid users. Each data point is averaged across 10 speech contrasts and four listeners. The composite scores are expressed in percent correct (re chance). The solid line shows linear regression function. The dotted line shows the expected relationship when the two scores are equal.

indistinguishable. This finding further supports the notion that the auditory-only score provides a valid measure not only of performance, but also of capacity.

Test-retest Reliability

The issue of test-retest reliability was addressed by administering the IMSPAC test twice to 18 aided children. Test-retest reliability for composite auditory-visual and auditory-only IMSPAC scores was 0.89. Figure 3.3 illustrates the relationship between the first and second administration of the test for both input modalities.

Effect of Age on Speech Perception and Production

The three panels in Figure 3.4. show the linear regression functions of IMSPAC scores on age for auditory-visual and auditory-only input modalities for 83 severely and profoundly hearing-impaired, aided children. The upper and middle panels were included to illustrate the scatter of the individual data for each input modality. The regressions were derived from the composite contrast scores (the mean of ten contrasts averaged across four listeners, in percent correct re guessing). Performance improved significantly ($p < 0.001$)

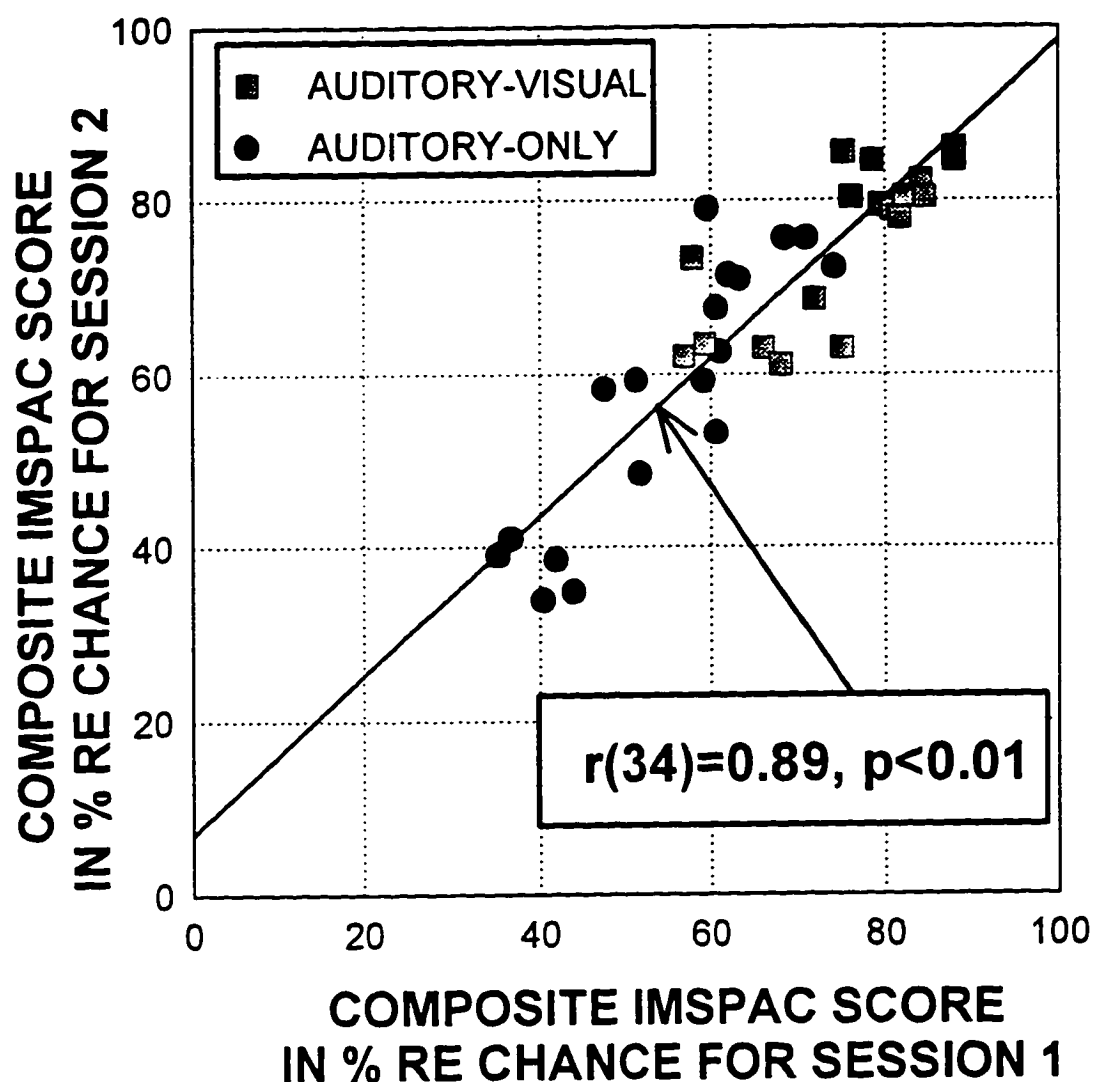


Figure 3.3. Test-retest data for composite auditory-visual and auditory-only IMSPAC scores of 18 hearing aid users. Each data point is averaged across 10 speech contrasts and four listeners. The solid line shows the linear regression function for session one on session two.

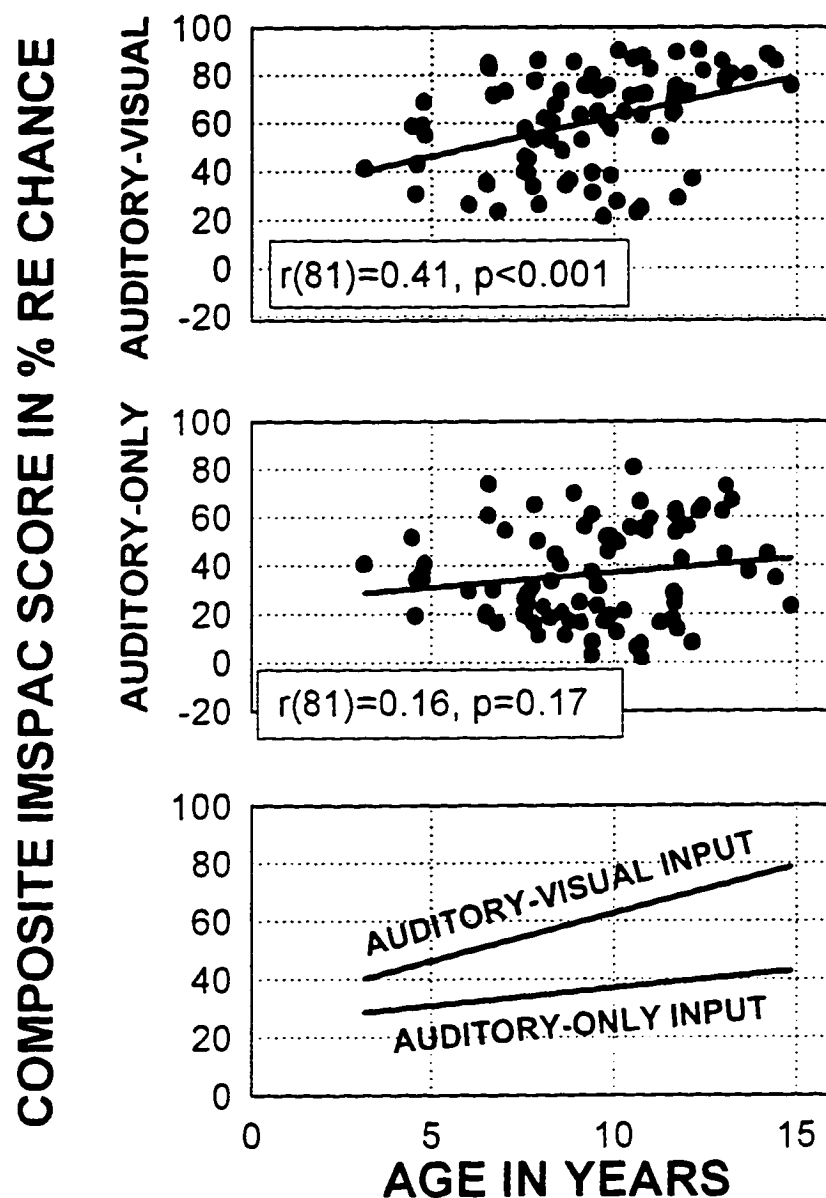


Figure 3.4. Mean auditory-visual and auditory-only IMSPAC scores of 83 severely and profoundly hearing-impaired, aided children as a function of age. The upper panel shows their mean auditory-visual scores, and the middle panel shows their mean auditory-only scores. The solid lines, in both panels, show the linear regression functions of IMSPAC scores on age. The same regression functions, for both input modalities, are also shown in the lower panel. The regressions were derived from the composite contrast scores (in percent correct re. guessing).

as a function of age, but only for the auditory-visual input modality. The coefficients of correlation between composite scores and age were $r(81)=0.41$; ($p<0.0001$) for the auditory-visual condition and $r(81)=0.16$; ($p=0.17$) for the auditory-only condition. The above findings revealed the following: (1) An absent or weak age effect for the auditory-only scores. (2) Presence of a strong age effect for the auditory-visual scores. (3) Only a small difference in performance between the auditory-only and auditory-visual modalities for the younger children. (4) A large difference in performance between the two input modalities for the older children.

Discussion

The absence of a strong age effect for the auditory-only scores suggests that peripheral auditory speech perception capacity does not change as a result of maturation, experience and training. Thus, the imitative auditory-only scores appear to be a valid measure of peripheral auditory speech perception capacity as defined in the introduction.

The strong age effect for the auditory-visual scores is not surprising. It is a common finding that

production performance continues to improve as the hearing-impaired child learns to integrate multiple sensory inputs, in response to experience and formal training (Boothroyd, 1985b; Geers and Tobey, 1992).

The small difference in performance between the auditory-only and auditory-visual modalities for the younger children is in keeping with the hypothesis that in young hearing-impaired children, speech production performance is driven primarily by auditory capacity (Montgomery, 1967; Erber, 1974; Smith, 1975). In other words, the young hearing-impaired child imitates what he/she hears without taking advantage of other available sensory inputs, such as visual cues. Since the hearing does not improve over time, the auditory imitations should remain unchanged. It is further expected, however, that performance on the auditory-visual task will improve progressively over time given appropriate training and experience.

A large difference in performance between the two input modalities for the older children is, again, in keeping with the notion of continued improvement in production. The auditory-visual scores of the older children could be assumed to reflect their total

perception capacity, i.e., the combination of auditory and visual capacity as well as the effects of experience and formal training on phonological and motor-speech knowledge and skills.

In summary, the data discussed in this section suggest that although speech production improves over time, peripheral auditory speech perception capacities remain unchanged. Thus, when the visual modality is taken away from the older children, they are left with the same audition they had as youngsters.

Rationale for the present study

The interpretation of these findings should be viewed cautiously for several reasons. It is possible, for example, that the data emerging from the study are not sensitive to small age effects because of inherent design issues. Recall, first, that the design of the original study called for random testing of a large sample of children. The selection criteria for the tested children, although defined, depended on availability. Thus, the sample varied greatly with respect to such independent variables as age at onset of deafness and degree of hearing loss. To the extent that

these variables affect performance on the test, they increase error variance and reduce the power of the analysis. Therefore, the present study reexamined the existing data with careful control over selected independent variables (degree of hearing loss, age of onset of deafness, and educational placement) for matched samples of younger and older children.

Second, note that the listener variable was not carefully controlled in the earlier study. Specifically, different groups of listeners were used to assess the speech recordings of older and younger children (though not for the input modality comparison for each child), thus increasing the error variance for inter-group comparison. Hence, the recorded utterances of the selected samples of younger and older children were evaluated here using the same group of listeners.

A third issue is the use of a forced-choice technique for evaluating the children's productions. This technique may not be sufficiently sensitive to changes in the auditory-only condition. More specifically, even in the auditory-only condition, older children may actually produce tokens that are closer to target. This would be expected as motor-speech skills

improve. Thus, the present study additionally assessed the children's utterances with an open-set evaluation technique using the same listeners.

As was discussed, performance on IMSPAC requires the child to have production skills. To confirm that this condition is met, the test is first administered using both auditory and visual inputs. The test is then administered a second time using only auditory input. If the auditory-visual score exceeds the auditory-only score, it is taken as evidence that the latter is being limited, not by production skills, but by peripheral auditory speech perception capacity. It could be argued, however, that the child is less comfortable with the auditory-only condition and more reluctant to respond. If this is the case, then the difference between the auditory-visual and the auditory-only conditions would be the same for all contrasts. If, however, the addition of vision provides the extra sensory input which permits improved production, then the difference should be most apparent for those contrasts that are more visually accessible, such as consonant and vowel place, and absent for those contrasts that are invisible, such as intonation. To address this issue, the present study examined which of the evaluated speech contrasts, if any, are most affected by age and/or input modality, and

whether these contrasts are predictable on the basis of the known relationships between speech articulation and speech acoustics.

In summary, the purpose of the present study was to investigate the effects of age and evaluation method on measured auditory speech perception performance in profoundly hearing-impaired children. Specifically, the study reevaluated existing IMSPAC recordings of a carefully selected sample of younger and older hearing-impaired children using three evaluation procedures, to determine whether the original findings remain intact, or whether they simply reflected the effects of low power of either the test instruments or the design. Specific experimental questions were as follows:

1. Are the original findings still supported when the same listeners are used across all children and conditions and when the older and younger children are matched in terms of pure tone average hearing loss, age at onset of deafness and, educational environment.

2. Is there an effect of age on the auditory-only score that is revealed when a more traditional evaluation technique, such as open-set recognition, is used?

3. Is the effect of input modality in younger children increased when an open-set evaluation technique is used?

4. What is the relationship between scores obtained with the closed-set and open-set evaluation techniques?

5. Which of the evaluated speech contrasts are most affected by age and/or input modality?

6. Are there age-related changes in the ability to imitate auditory input that are revealed when imitations are evaluated in terms of their acoustic properties?

7. Are the behavioral results predictable from acoustic measures of performance?

This last question was added to provide information about the acoustic correlates of any observed behavioral changes.

Chapter 4

Methodology

As previously stated, the primary purpose of this study was to investigate the effects of age and evaluation method on measured auditory speech perception performance in profoundly hearing-impaired children. More specifically, this study reexamined existing IMSPAC recordings of representative matched samples of younger and older children, using closed-set and open-set evaluation procedures. In addition, an acoustic analysis of vowels was carried out and will be reported in Chapter 6.

Subjects

Twelve children with profound sensorineural hearing loss served as subjects for the present study. These children were selected from the pool of 94 hearing-impaired children tested previously with IMSPAC as reported in chapter 3. The criteria for their selection were:

1. Age. Because the main purpose of the study was to examine speech perception performance as a function of age, two age groups were selected, representing both

ends of the age continuum taken from the hearing-aid user database.

2. Hearing status. All children had a profound sensorineural hearing loss, with no conductive components as reported by the school's audiologist, and three frequencies (500 Hz, 1 KHz, 2 KHz) pure tone averages of 90 dB HL or higher in the better ear. Within the hearing-impaired population, the profoundly deaf, specifically within the range of 91-110 dB HL, represent a unique group that demonstrates a wide range of speech perception performances. This makes it difficult to predict accurately their peripheral auditory speech perception capacities based solely upon their pure tone audiograms. Moreover, cochlear implants are presently being fit to young, profoundly deaf children. This is in contrast to the more strict "totally deaf" criterion previously used. Therefore, there is a pressing need for early and accurate quantification of auditory capacities in this population.

3. Age at onset. All subjects were prelingually deaf. Age at onset of deafness ranged from birth to two years. Prelingually deaf children are an especially

interesting group because they must acquire speech without the benefit of normal hearing and with little or no auditory memory. Prelingually deaf children also represent the majority in schools for the deaf.

4. Educational environment. The children came from either oral programs or total communication (TC) programs with a strong auditory component. This criterion helped ensure that all the children had received appropriate auditory and speech training.

5. Other impairments. Subjects had no reported handicapping conditions (other than hearing loss) which might have confounding effects upon performance on the imitative task.

The selected children were divided into two age groups. The "younger" group consisted of six children between 3.2 to 4.8 years old (Mean = 4.4). Their pure tone average (500 Hz, 1000 Hz, 2000 Hz) hearing loss ranged from 93 to 102 dB (Mean = 97.3 dB) in the better, tested ear. The "older" group consisted of six children between 12.7 and 14.8 years old (Mean = 13.8). Their pure tone average ranged from 95 to 110 dB (Mean = 100.3 dB) in the better, tested ear.

There were no statistically significant differences between the two groups with respect to their pure tone average ($t(10)=1.13$, $p>0.05$), age at onset of deafness ($t(10)=1.17$, $p>0.05$), and age when amplification began ($t(10)=0.49$, $p>0.05$). Ten of the subjects attended an oral school for the deaf. Two children in the younger group attended a total communication (TC) program with a strong emphasis on aural and oral training. The demographic and audiometric characteristics of the subjects appear in Tables 4.1 and 4.2, respectively. As can be seen from Table 4.2, there were no statistically significant differences between the two groups with respect to audiometric configuration except for 1000 Hz ($t(10)=2.36$, $p=0.04$) in favor of the younger group¹.

Listeners

Four normally hearing adults, native speakers of English, served as listeners in the study and audited all 12 subjects. All were recruited from the Ph.D. program in Speech and Hearing Sciences at the City

¹ Although the two groups were quite well matched for hearing loss, it will be seen that there was a tendency for the younger subjects to have somewhat better thresholds (3 dB difference in pure tone average and "flatter" audiograms). Reference will be made to this in the discussion chapter.

Table 4.1 Demographic characteristics of subjects.

Subject Number	Group	Age (Years)	Age at onset (Years)	Age aided (Years)	Gender	Etiology	Mode of communication
H21	Young	4.8	0.0 ^a	0.9	M	Unknown	Oral
H44	Young	4.6	0.0	0.4	M	Genetic	Oral
H46	Young	4.8	0.0	1.5	F	Genetic	Oral
H47	Young	4.6	1.5	1.8	M	Meningitis	Oral ^b
H66	Young	4.5	0.0	1.1	M	Genetic	TC
H70	Young	3.2	0.0	1.5	M	Unknown	TC
H23	Old	12.7	0.0	0.5	F	Unknown	Oral
H78	Old	14.8	2.0	2.0	F	Unknown	Oral
H80	Old	14.4	0.0	0.8	M	Trauma	Oral
H81	Old	13.6	0.0	2.0	M	Genetic	Oral
H84	Old	14.1	2.0	2.2	M	Unknown	Oral
H106	Old	13.1	0.8	0.8	M	Unknown	Oral
Mean	Young	4.40	0.25	1.20			
SD	Young	0.61	0.61	0.51			
Mean	Old	13.80	0.80	1.38			
SD	Old	0.79 ^c	0.98	0.76			
t-test		22.86 ^c	1.17	0.49			

^a 0.0 corresponds to congenital hearing loss.^b TC = Total Communication^c $t(10)=22.86$, $p<0.00005$

Table 4.2. Audiometric characteristics of subjects.

Subject Number	Group	PTA ^a (dB HL)	F R E Q U E N C Y (Hz)								Test Ear	SAT (dB SPL)	Sensation Level (dB)	Other
			250	500	1000	2000	4000	8000						
H21	Young	97	80	95	100	95	105	115+ ^b	L	114	14			
H44	Young	102	85	90	105	110	100	100	FF ^c	-99	-99		better ear-R	
H46	Young	93	105	100	95	85	75	55	FF	-99	-99		better ear-R	
H47	Young	97	95	95	95	100	100	115+	FF	-99	-99		better ear-L	
H66	Young	98	75	90	100	105	90	80	R	102	30			
H70	Young	97	-99 ^d	90	100	100	115	-99	R	114	22			
H23	Old	110	95	100	115	115	120	-99	R	118	12			
H78	Old	98	90	90	100	105	115+	115+	L	106	26			
H80	Old	97	70	90	100	100	110	115+	L	110	24			
H81	Old	95	80	85	105	95	90	95	R	112	16			
H84	Old	105	95	100	115	100	100	105	L	122	14			
H106	Old	97	75	85	105	100	100	100	R	108	24			
Mean	Young	97.3	88.0	93.3	99.2	99.2	97.5	93.0						
SD	Young	2.87	12.0	4.1	3.8	8.6	13.7	25.6						
Mean	Old	100.3	84.2	91.7	106.7	102.5	105.8	106.0						
SD	Old	5.85	10.7	6.8	6.8	6.9	11.1	8.9						
t-test		1.13	0.56	0.51	2.36 ^e	0.74	1.15	1.1						

^a PTA - Pure Tone Average (500, 1000, 2000 Hz)^b 115+ - 5 dB above measured no audiometric response^c FF - tested in Free Field^d -99 - missing data^e t(10)=2.36, p=0.04

University of New York. They had at least one year of experience in broad transcription. None of the listeners had previously participated in listening to IMSPAC stimuli, nor were they familiar with the children's speech, their background, or their audiological status.

Before data collection began, the listeners were presented with practice recordings for both testing formats in order to familiarize them with the procedure and the listening task and to establish inter-listener reliability. Three practice recordings were used. Two were of hearing-impaired children taken from the hearing-aid-users database (refer to the data presented in Chapter 3). One of these two recordings was used for practice purposes throughout our database collection. The third recording was of a normal hearing speaker and was used only with the closed-set listening task. This last recording was also part of the protocol for the data presented in Chapter 3. The criterion for inter-listener reliability for each of the three recordings was a composite IMSPAC score or open-set score within $\pm 10\%$ points of each other. All four listeners scored within the target criterion, and, as a result, additional training was not provided.

Instrumentation

The instrumentation used for listening in the present study is described in detail in Chapter 3.

Evaluation procedure

The listeners were asked to make two types of decision regarding the children's utterances. The first used an open-set evaluation procedure. The second used the original forced-choice evaluation procedure. In each of the evaluation procedures, listeners were presented with 24 sets of recordings (12 sets of auditory-only recordings and 12 sets of auditory-visual recordings).

Listeners were presented first with the open-set evaluation procedure and then with the closed-set evaluation procedure. This order was selected so that the listeners would avoid learning from exposure to the written response alternatives in the closed-set task. Randomization of the recordings for both evaluation procedures and for the four listeners was controlled by the software. However, the order in which the speech contrasts were represented for all recordings remained fixed.

Each listener participated in a total of 14 45-minute listening sessions. Stimuli were presented binaurally at a comfortable output level through headphones in a quiet room.

Each listener was shown a sheet with instructions (Appendices A and C) that described the nature of the listening task. On the open-set task, the listener was instructed to broadly transcribe each utterance on a form provided (see Appendix B). On the closed-set task, the listener was asked to determine—in a four-alternative, forced-choice context—which syllable corresponded to the child's utterance. The listener triggered the next utterance by keying in the response. Unlimited listening repetitions of the utterances were permitted on both listening tasks.

Test-retest reliability

Within-listener test-retest reliability was established for both closed-set and open-set evaluation procedures. Eight selected recordings (four recordings from each age group) were re-administered to all four listeners approximately six months after completing the first administration. The results, described in details

in the next chapter, revealed little or no learning effect on behalf of the listeners.

Chapter 5

Results

Analysis of the closed-set listening task

It will be recalled that the first question of the study was whether the original findings, that is, no age effect for the auditory-only condition but a significant age effect for the auditory-visual condition, are still supported when the same listeners are used across all subjects and conditions. To answer this question, the previous and new listeners' responses to the IMSPAC recordings were analyzed for each subject in terms of the proportion of correct responses for each contrast for the auditory-only and the auditory-visual input modalities. For each input modality, a composite score, the average of ten contrasts, was computed for each subject and each listener. Composite scores were then averaged across the four listeners and represented one data point for each subject. A scoring form that represents the process of obtaining one data point for one subject is included in Appendix E. Individual contrast scores obtained by each listener for each of the 12 subjects,

as well as composite contrast scores, can be found in Appendix F.

Table 5.1 shows the original and the new composite IMSPAC scores. It will be seen that the original and the new IMSPAC scores provided four data points for each of the 12 hearing-impaired subjects. The mean scores and standard errors for each condition and age group are also presented in this table. Group means (± 1 standard error) are illustrated in Figure 5.1. The left panel shows performance for the original IMSPAC data, and the right panel for the new IMSPAC data. Measured scores were corrected for the effect of guessing, using the following equation (Boothroyd, 1988):

$$C = (M - E) / (100 - E) * 100 \dots \dots \dots (1)$$

where: C = corrected score in %

M = measured raw score in %

E = expected chance score (50% in this case)

The expected chance score for IMSPAC is 50%. Using this equation the range from 50% to 100% is transformed to become 0% to 100%.

Table 5.1. Original and new composite IMSPAC scores, in percent, for the auditory-only and the auditory-visual input modalities for the 12 hearing-impaired subjects. The mean scores and standard errors for each condition and age group are also shown.

AGE GROUP	SUBJECT NUMBER	LISTENING TASK AND CONDITION			
		Original Closed-Set Audition (%)	Original Closed-Set Audition+ Vision (%)	New Closed-Set Audition (%)	New Closed-Set Audition+ Vision (%)
YOUNGER	1	67.21	84.38	76.91	87.16
YOUNGER	2	59.71	65.32	64.55	68.35
YOUNGER	3	70.33	77.52	69.56	80.60
YOUNGER	4	67.20	71.58	71.39	73.64
YOUNGER	5	74.07	83.44	74.59	84.98
YOUNGER	6	70.98	75.34	73.34	73.96
MEAN		68.25	76.26	71.72	78.11
SE		2.0	2.95	1.76	2.98
OLDER	7	64.08	85.64	63.13	86.89
OLDER	8	63.45	87.21	65.01	90.33
OLDER	9	68.18	92.21	69.07	94.38
OLDER	10	71.58	87.81	67.20	89.07
OLDER	11	75.33	95.32	70.32	94.07
OLDER	12	70.64	88.76	70.33	88.76
MEAN		68.87	89.49	67.51	90.58
SE		1.87	1.47	1.20	1.23

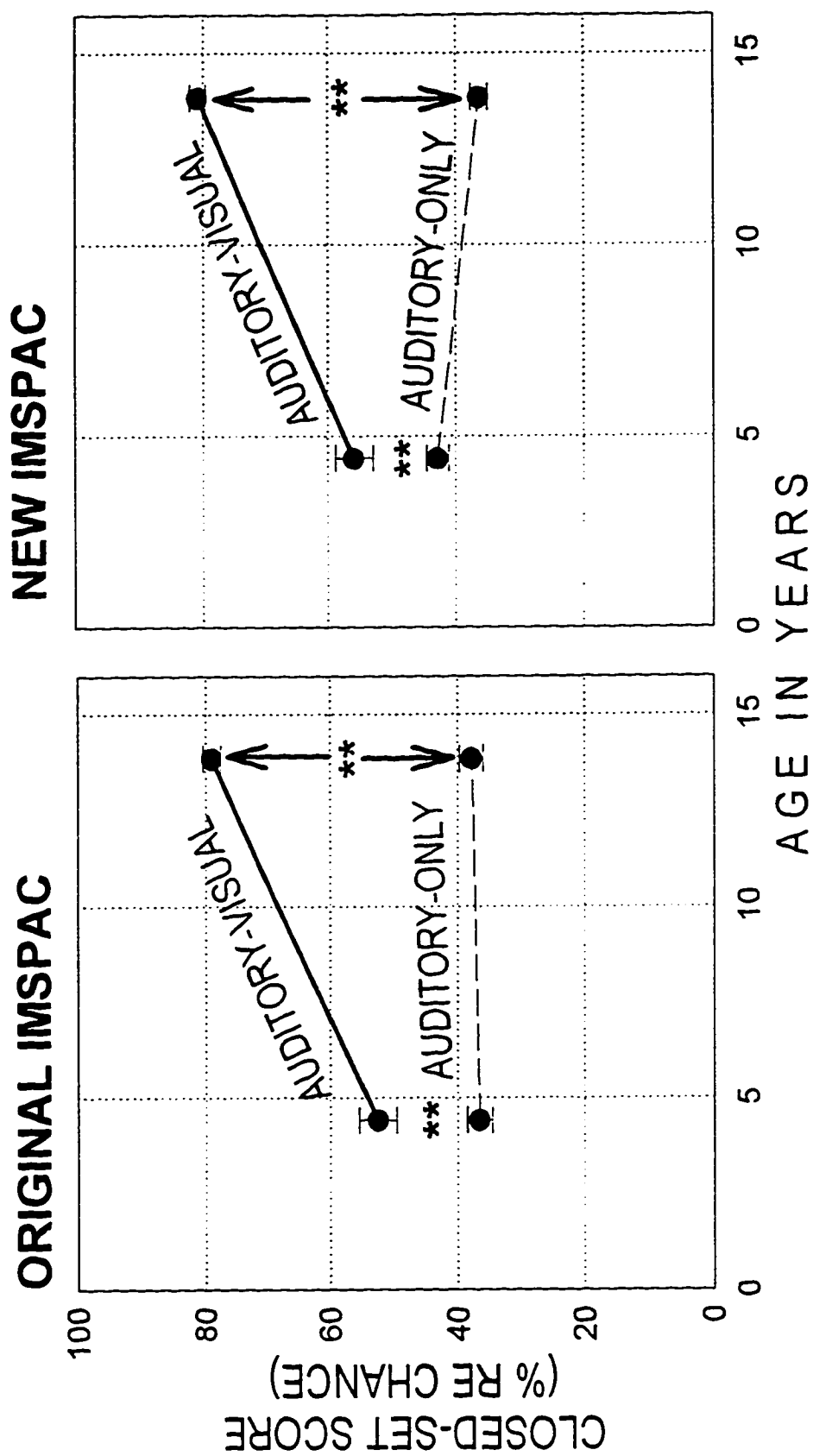


Figure 5.1. Group mean (± 1 std. error) original and new IMSPAC scores for 6 younger and 6 older children. Heavy lines and double asterisks show significance at the 1% level.

The arcsine transform (Brownlee, 1965) was applied to the raw data to increase homogeneity of variance. A two-factor repeated-measures analysis of variance was performed separately on the arcsine-transformed original and new IMSPAC data (see Appendix G). The results showed that the error variance was not significantly different in the two analyses suggesting that the use of the same group of listeners for all subjects did not improve the sensitivity of the design. In view of this finding the data were treated as a whole and a three-factor repeated-measures analysis of variance was performed on the arcsine-transformed original and new IMSPAC data. The results are shown in Table 5.2.

Table 5.2. Analysis of variance: effects of AGE (A), TEST (ORIGINAL/NEW IMSPAC) (T), and INPUT MODALITY (M) on the IMSPAC scores of the 12 hearing-impaired subjects

Effect	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
Age	1	318.54	10	47.45401	6.712606	0.02691*
Test	1	11.90956	10	2.724551	4.3712	0.06305
Modality	1	1592.41	10	10.39256	153.226	0.000000 ¹
AxT	1	7.248102	10	2.724551	2.660292	0.133934
AxM	1	446.7291	10	10.39256	42.98547	0.000064
TxM	1	0.193682	10	1.01183	0.191418	0.671037
AxTxM	1	4.344666	10	1.01183	4.293871	0.065047

*Significant effects are highlighted

The important finding of this analysis is that there is no significant test effect either as a main effect or in interaction with other effects. Although there is a tendency for new IMSPAC scores to be higher than original IMSPAC scores, differences are small. This finding does not support the conclusion that there

¹ Because of rounding issues probability values were calculated down to six decimal places and were 0.000000 appears it should be interpreted as being less than 0.0000005

is a difference between the original IMSPAC and the new IMSPAC testing. Therefore, the data further support the conclusion that the original findings are not a function of the use of different listeners for different subjects.

It will be recalled that the findings reported for the original IMSPAC testing were the absence of an age effect for the auditory-only scores but the presence of a strong age effect for the auditory-visual scores. There was a small significant difference in performance between the auditory-only and auditory-visual modalities for the younger subjects but a large significant difference in performance between the two input modalities for the older subjects.

These findings are illustrated in the left panel of Figure 5.1. The heavy lines and double asterisks show significance at the 1% level. To investigate further whether these original findings are replicated with the new IMSPAC scores, a post-hoc analysis using the Scheffe test was applied to the means of the arcsine transforms of the raw data. The results are shown in Table 5.3.

Table 5.3 Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the new IMSPAC data.

Age Group	Input Modality	Means (%)	(1)	(2)	(3)	(4)
			64.19524	69.13873	61.86710	80.21687
YOUNGER	Audition	(1)		.037779*	.466540	.000005
YOUNGER	Audition +Vision	(2)	.037779		.003486	.000128
OLDER	Audition	(3)	.466540	.003486		.000001
OLDER	Audition +Vision	(4)	.000005	.000128	.000001	

*Significant effects are highlighted

The same pattern of findings was found with the new IMSPAC testing as with the original data. A significant age effect was found for the auditory-visual scores ($p=0.0001$), but not for the auditory-only scores ($p=0.47$). A significant difference in performance was also found between the auditory-only and auditory-visual modalities for both younger ($p=0.04$) and older subjects ($p=0.000001$). These findings are illustrated in the right panel of Figure 5.1. The use of a post-hoc analysis with the Scheffe

test is a somewhat more conservative procedure that may not actually show small existing differences in the auditory-only condition. Thus, a more lax, planned comparison procedure was applied to the data. The results are shown in table 5.4.

Table 5.4 Results of a post-hoc analysis with planned comparison applied to the means of the arcsine transforms of the new IMSPAC data.

Age Group	Input Modality	(1)	(2)	(3)	(4)
		64.19524	69.13873	61.86710	80.21687
YOUNGER	Audition (1)		.005491	.127863	.000000
YOUNGER	Audition (2)	.005491		.000410	.000013
	+Vision				
OLDER	Audition (3)	.127863	.000410		.000000
OLDER	Audition (4)	.000000	.000013	.000000	
	+Vision				
*Significant effects are highlighted					

Here again no significant age effect was found for the auditory-only scores ($p=0.13$).

The results of both the Scheffe and the planned comparison post-hoc analyses support the conclusion that the original and the new closed-set findings are

not significantly different and thus, theoretically, could be collapsed for further analysis of the data. However, because the purpose of the study was to assess performance using different evaluation procedures with the same listeners, the new IMSPAC data will be treated separately from the original IMSPAC data. As a matter of fact, from now on only the new IMSPAC data will be analyzed.

Contrast profiles

To address the issue of perception of individual speech contrasts, performance for each input modality was averaged across the six subjects within each group for each contrast assessed by the closed-set evaluation procedure. Results are shown for the auditory-only and auditory-visual input modalities in Tables 5.5 and 5.6, respectively.

A four-factor analysis of variance was performed on the arcsine transformed contrast scores for the two age groups. The results are shown in Table 5.7. The important findings of this analysis are that there is a significant main effect of contrast ($F(9,90)=27.07$, $p < 0.01$), and significant interactions between age and contrast ($F(9,90)=2.64$, $p < 0.01$), and input modality and contrast ($F(9,90)=4.105$, $p < 0.01$).

Figures 5.2 and 5.3 show the means for the ten speech contrasts under the two input modalities for the younger group and the older group, respectively. The contrasts have been arranged in descending order based on the mean scores for the auditory-only condition. It will be seen that for the younger group (Figure 5.2) the benefit of visual cues, although limited, is evident mainly for the place contrasts.

Table 5.5. Mean scores on 10 speech-pattern contrasts of the IMSPAC test for the older and younger subjects for the auditory-only input modality.

GROUP	AGE Syllable Number	Intona -tion	Vowel Height	Vowel Place	Initial Voicing		Final Voicing		Initial Contin -uance		Final Contin -uance		Initial Place	Final Place
					Initial	Final	Initial	Final	Initial	Final	Initial	Final		
YOUNGER	100.0	90.6	96.9	81.1	71.8	62.3	62.3	74.9	55.9	71.8	59.1	59.1	71.8	59.1
YOUNGER	96.9	50.0	71.6	56.0	49.9	49.8	49.8	75.0	71.8	68.8	55.9	55.9	68.8	55.9
YOUNGER	100.0	53.0	78.0	81.3	59.3	59.3	59.3	68.5	78.1	52.9	65.4	65.4	52.9	65.4
YOUNGER	100.0	62.0	90.6	62.5	87.5	74.8	74.8	52.9	65.5	65.3	52.9	52.9	65.3	52.9
YOUNGER	100.0	68.5	96.9	78.1	78.1	78.1	78.1	49.8	78.0	68.5	49.9	49.9	68.5	49.9
YOUNGER	100.0	96.9	100.0	74.9	59.0	62.0	62.0	46.9	62.5	84.4	46.9	46.9	84.4	46.9
MEAN	99.48	70.12	89.0	72.32	67.6	64.38	64.38	61.33	68.63	68.62	55.02	55.02	68.62	55.02
S.D.	1.27	19.51	11.59	10.59	14.0	10.44	10.44	12.92	8.91	10.17	6.66	6.66	10.17	6.66
OLDER	100	56.3	65.6	50.0	59.4	56.3	56.3	62.5	78.1	53.1	50.0	50.0	53.1	50.0
OLDER	96.9	50.0	96.9	46.9	46.9	81.3	81.3	87.5	65.6	40.6	37.5	37.5	40.6	37.5
OLDER	100	78.1	93.8	62.5	65.6	87.5	87.5	50.0	43.8	43.8	65.6	65.6	43.8	65.6
OLDER	96.9	75.0	81.3	75.0	53.1	50.0	50.0	78.1	75.0	43.8	43.8	43.8	43.8	43.8
OLDER	96.9	62.5	84.4	75.0	75.0	59.4	59.4	65.6	46.9	78.1	59.4	59.4	78.1	59.4
OLDER	100	68.8	100	62.5	53.1	81.3	81.3	68.8	65.6	56.3	46.9	46.9	56.3	46.9
MEAN	98.45	65.12	87.0	61.98	58.85	69.3	69.3	68.75	62.5	52.62	50.53	50.53	52.62	50.53
S.D.	1.69	10.89	12.73	11.92	10.15	15.87	15.87	12.96	14.23	13.88	10.34	10.34	13.88	10.34

Table 5.6. Mean scores on 10 speech-pattern contrasts of the IMSPAC test for the older and younger subjects for the auditory-visual input modality.

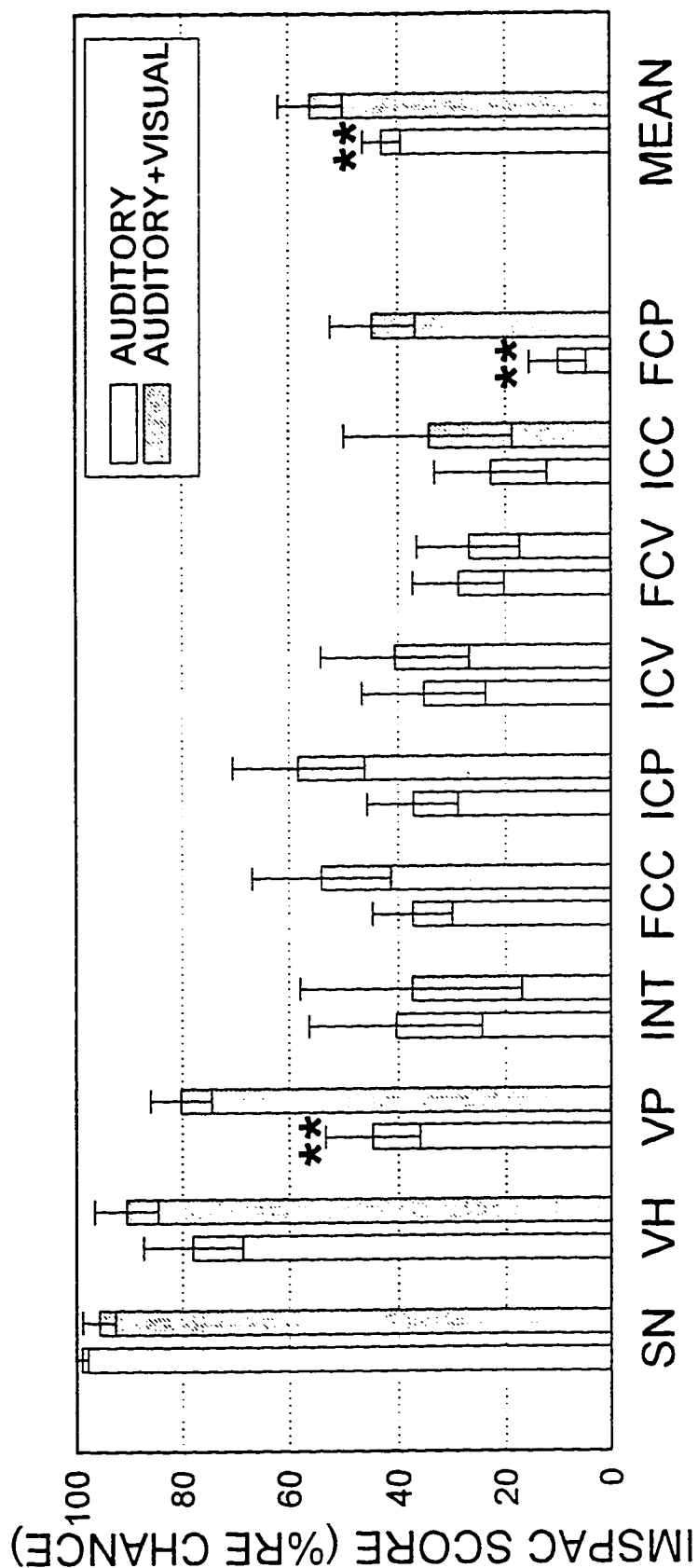
AGE GROUP	Syllable Number	Intona -tion	Vowel Height	Vowel Place	Initial Voicing	Final Voicing	Initial Continu -ance	Final Continu -ance	Initial Place	Final Place
YOUNGER	100.0	81.3	96.9	90.6	93.8	62.3	93.8	75.0	90.6	87.5
YOUNGER	90.6	31.3	93.8	87.5	75.0	71.9	49.9	68.5	56.0	59.1
YOUNGER	100.0	53.0	81.3	84.4	81.3	65.5	87.5	93.8	84.4	75.0
YOUNGER	100.0	59.1	100.0	81.3	59.3	59.0	53.0	71.8	84.4	68.6
YOUNGER	100.0	96.9	100.0	96.9	46.9	78.1	65.5	96.9	93.8	74.9
YOUNGER	96.9	90.6	100.0	100.0	65.3	43.8	53.0	56.0	65.6	68.5
MEAN	97.92	68.7	95.33	90.12	70.27	63.43	67.12	77.0	79.13	72.27
S.D.	3.79	25.18	7.31	7.23	16.67	11.81	19.10	15.64	14.97	9.46
OLDER	96.9	81.3	100.0	90.6	87.5	90.6	93.8	87.5	71.9	68.8
OLDER	96.9	81.3	100.0	96.9	90.6	87.5	96.9	96.9	68.8	87.5
OLDER	100.0	90.6	100.0	90.6	84.4	100.0	100.0	100.0	81.3	96.9
OLDER	100.0	71.9	100.0	75.0	93.8	84.4	100.0	87.5	87.5	90.6
OLDER	100.0	78.1	100.0	96.9	100.0	81.3	100.0	100.0	96.9	87.5
OLDER	96.9	81.3	100.0	87.5	96.9	87.5	87.5	87.5	90.6	71.9
MEAN	98.45	80.75	100.0	89.58	92.2	88.55	96.37	93.23	82.83	83.87
S.D.	1.70	6.05	0.0	8.08	5.85	6.44	5.00	6.38	10.94	11.06

Table 5.7. Analysis of variance: effects of AGE (A), INPUT MODALITY (M), LISTENER (L), CONTRAST (C) on the IMSPAC scores of the 12 hearing-impaired subjects.

	df Effect	MS Effect	df Error	MS Error	F	p-level
Age	1	5712.664	9	1346.284	4.243283	0.069489
Modality	1	47343.31	9	378.9632	124.9285	0.000001*
Listener	3	302.0768	30	67.57927	4.469963	0.010394
Contrast	9	12737.98	90	470.5587	27.0699	0.000000
A x M	1	14423.05	9	378.9632	38.05924	0.000165
A x L	3	38.1053	30	67.57927	0.563861	0.643056
M x L	3	46.45095	30	79.51569	0.584173	0.63004
A x C	9	1242.196	90	470.5587	2.639832	0.009267
M x C	9	1326.864	90	323.2039	4.105347	0.00019
L x C	27	121.9436	270	73.33104	1.662919	0.023738
AxMxL	3	72.35149	30	79.51569	0.909902	0.447894
AxMxC	9	377.036	90	323.2039	1.166558	0.326031
AxLxC	27	81.44819	270	73.33104	1.110692	0.326381
MxLxC	27	76.69572	270	76.62672	1.000901	0.467654
AxMxLxC	27	74.19026	270	76.62672	0.968204	0.513689

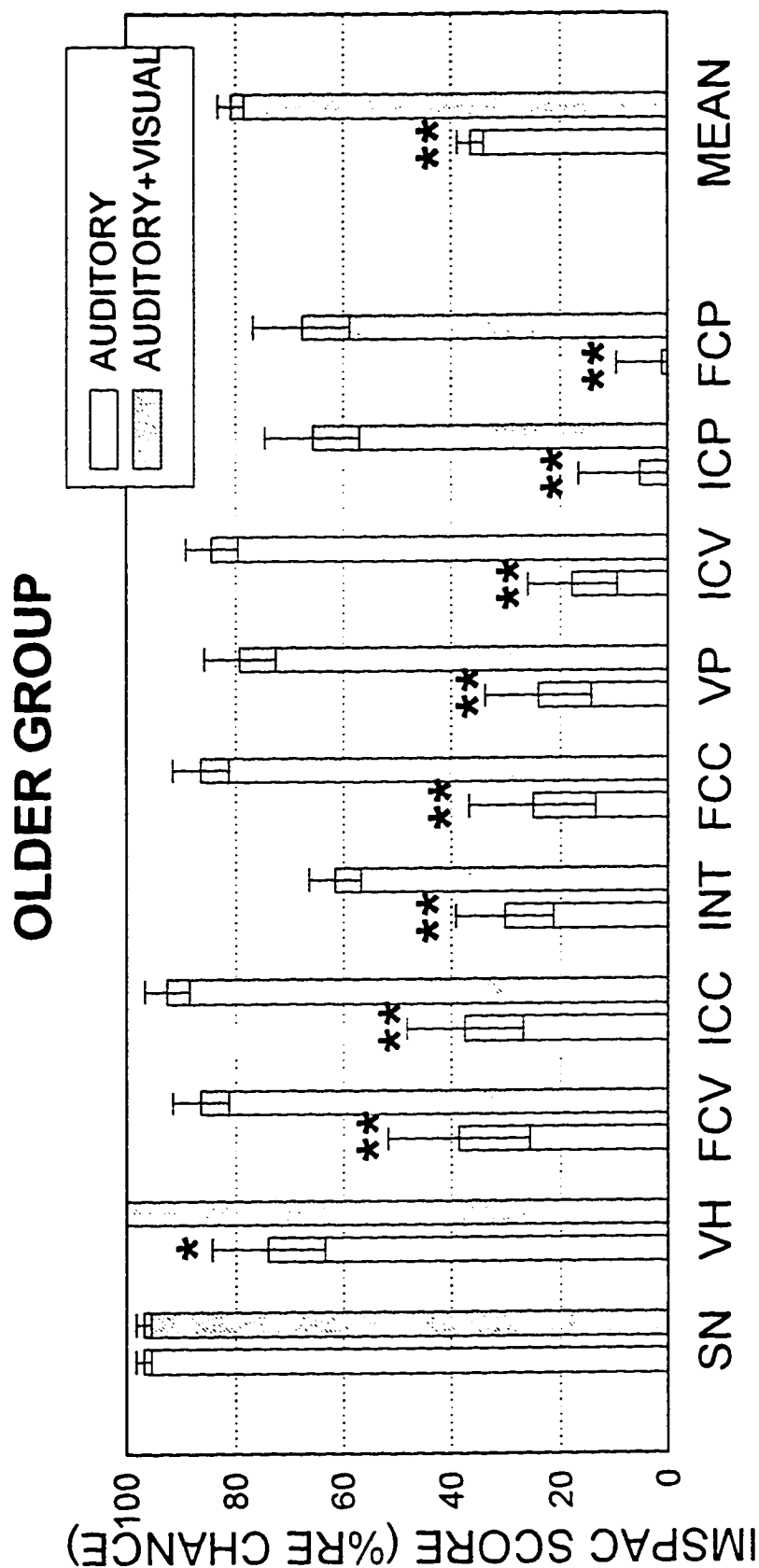
*Significant effects are highlighted

YOUNGER GROUP



SPEECH-PATTERN CONTRASTS

Figure 5.2. The auditory and auditory-plus-visual IMSPAC contrasts profile for the younger group. The contrasts have been arranged in descending order of performance based on the auditory scores. The error bars show one standard error of the mean for each bar. Double asterisks show significance at the 1% level. The key to the contrast abbreviations will be found in Appendix D.



SPEECH-PATTERN CONTRASTS

Figure 5.3. The auditory and auditory-plus-visual IMSPAC contrasts profile for the older group. The contrasts have been arranged in descending order of performance based on the auditory scores. The error bars show one standard error of the mean for each bar. Single and double asterisks show significance at the 5% and 1% level, respectively. The key to the contrast abbreviations will be found in Appendix D.

To identify significant differences in performance between the auditory-only and the auditory-visual input modalities, a post hoc analysis with planned comparison was applied to the data of the younger subjects. Results confirmed that the younger group took advantage of visual cues only to differentiate between vowel place and final consonant place of articulation contrasts. Significant differences between input modalities are shown by the double asterisks in Figure 5.2.

In contrast, it is clear that the older group benefited greatly from the addition of visual input (Figure 5.3). In fact, a post hoc analysis with planned comparison applied to the older group data revealed that performance on nine of the ten contrasts was significantly better on the auditory-visual condition than on the auditory-only condition. These findings are illustrated by the asterisks in Figure 5.3. The absence of a significant difference in performance for the syllable number (SN) contrast is attributable to a ceiling effect for both input modalities.

It is appropriate to ask whether the individual contrast data support the earlier conclusions regarding differences in performance between older and younger

subjects. Figure 5.4 shows the auditory-only contrast profiles for the two age groups. The overall pattern is similar for the two age groups. A post hoc analysis with planned comparison found, however, that the younger group scored significantly higher than the older group on the initial consonant place of articulation contrast.

Figure 5.5 shows the auditory-visual contrast profiles for the two age groups. Overall, the older group outperformed the younger group. A post hoc analysis with planned comparison, however, revealed that the older group was performing significantly better on intonation, initial consonant voicing, final consonant voicing, initial consonant continuance, and final consonant continuance contrasts.

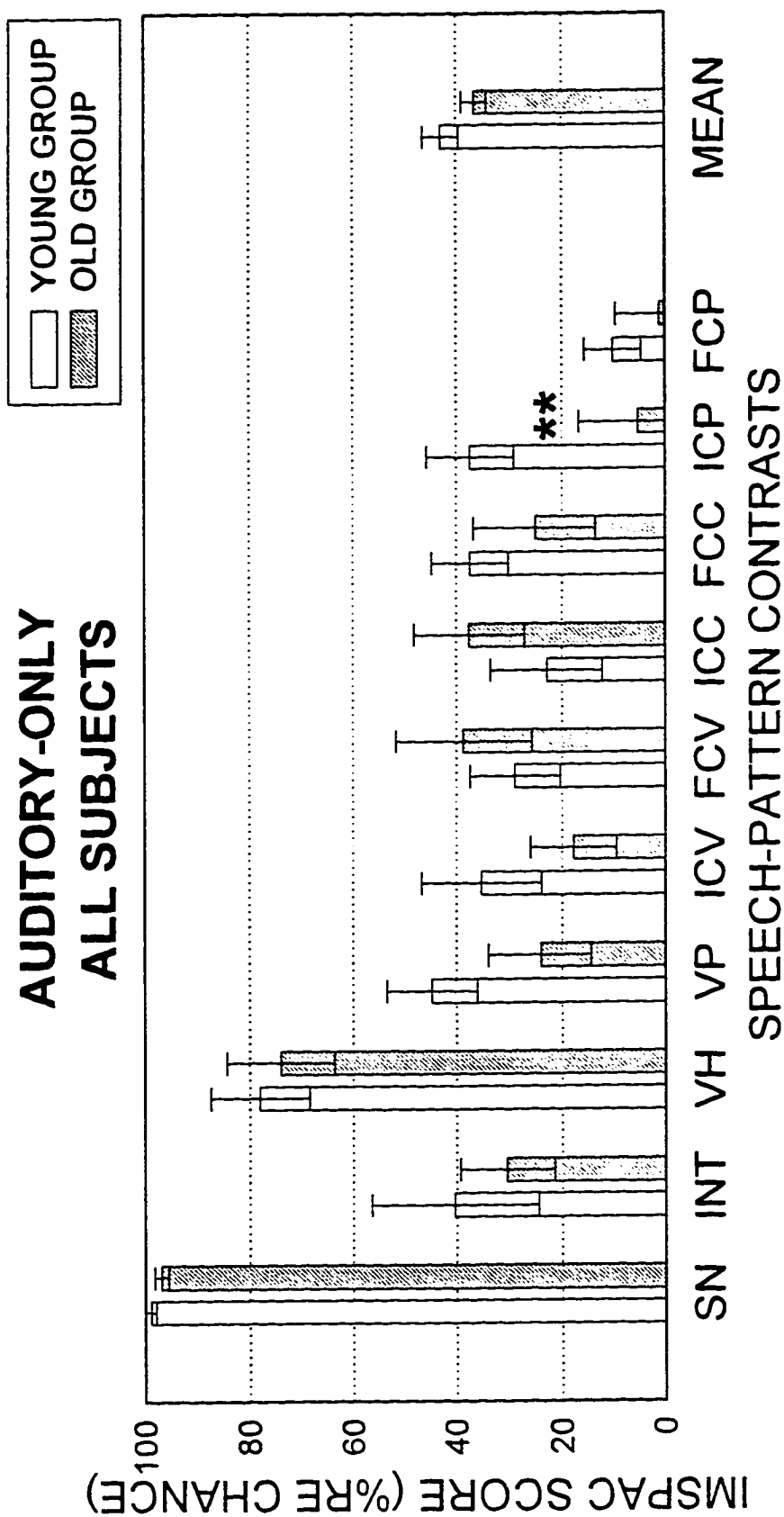


Figure 5.4. Group mean IMSPAC auditory scores as a function of speech contrasts and age group. The error bars show one standard error of the mean for each bar. Double asterisks show significance at the 1% level. The key to the contrast abbreviations will be found in Appendix D.

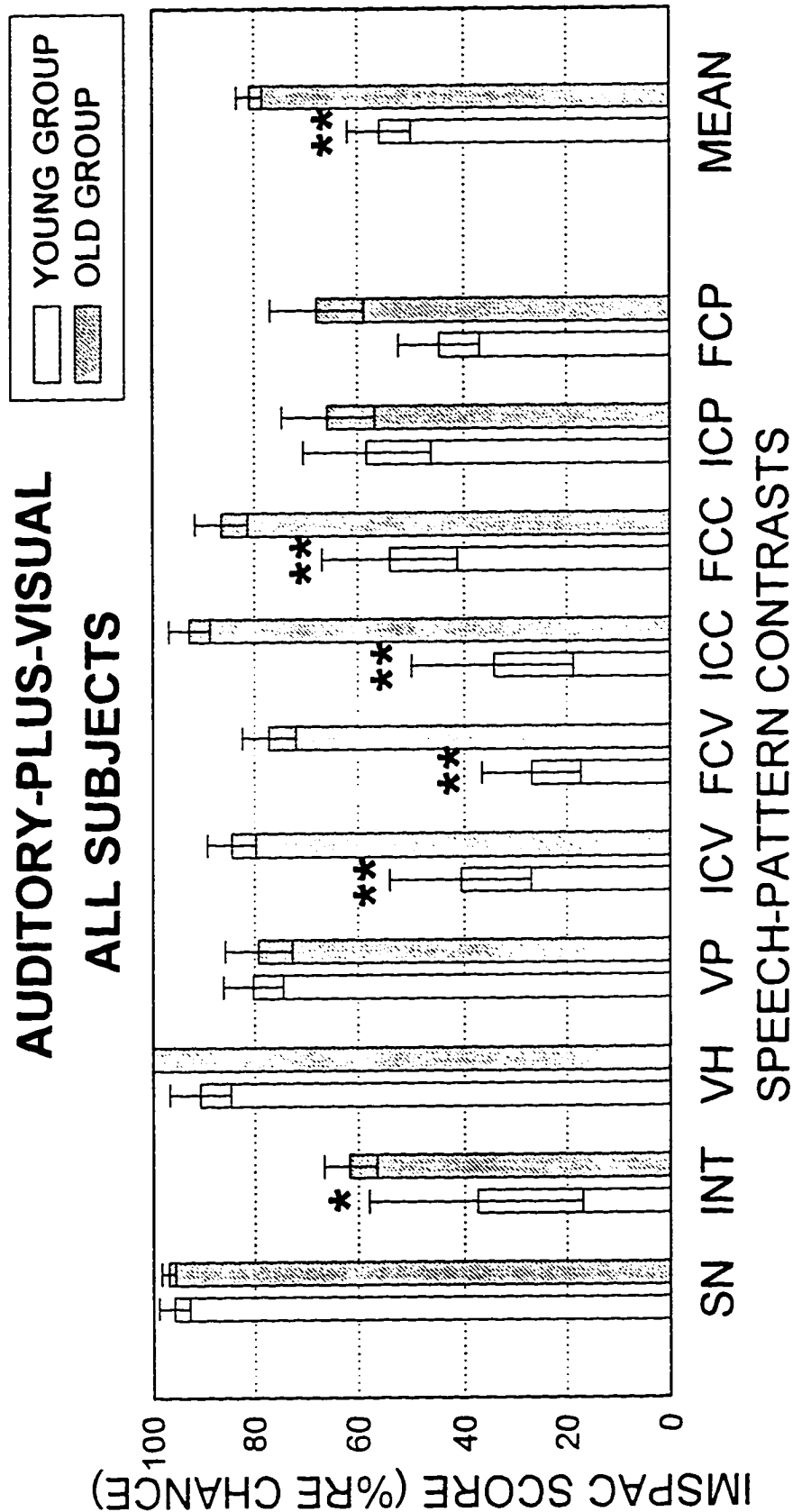


Figure 5.5. Group mean IMSPAC auditory-plus-visual scores as a function of speech contrasts and age group. The error bars show one standard error of the mean for each bar. Single and double asterisks show significance at the 5% and 1% level, respectively. The key to the contrast abbreviations will be found in Appendix D.

Analysis of the open-set listening task

A second question of the study was whether there is an age effect for the auditory-only condition when an open-set evaluation procedure is used. To address this question listeners' transcriptions of subjects' utterances, for both input modalities, were scored as percent phonemes recognized. In other words, each phoneme was scored as either correct or incorrect. For example, if the listener wrote down "shoo" when the target for imitation presented to the child was "choo", then the syllable was scored incorrect for the consonant and correct for the vowel. No attempt was made to penalize distortions within the same phoneme category. For each subject, scores within each input modality were averaged across the four listeners. This averaged score represents one data point for one child. The phoneme recognition scores obtained by each listener for each of the 12 subjects can be found in Appendix H. Table 5.8 shows the mean open-set scores (collapsed across listeners) for each of the 12 hearing-impaired subjects. Group means and standard errors for each condition are also shown in the table and illustrated in Figure 5.6.

Table 5.8. Percent recognition of phonemes in syllables produced by imitation, by 12 hearing-impaired subjects as a function of age group and input modality. Group means and standard errors for each condition are also shown.

AGE GROUP	SUBJECT NUMBER	OPEN-SET AUDITION (%)	OPEN-SET AUDITION+ VISION (%)
YOUNGER	1	50.39	62.11
YOUNGER	2	18.36	34.76
YOUNGER	3	17.19	38.28
YOUNGER	4	37.11	44.95
YOUNGER	5	41.02	53.91
YOUNGER	6	49.61	55.86
MEAN		35.61	48.31
SE		6.01	4.37
OLDER	7	16.41	51.56
OLDER	8	24.22	63.28
OLDER	9	29.69	80.08
OLDER	10	35.16	67.19
OLDER	11	35.16	82.43
OLDER	12	22.27	60.94
MEAN		27.15	67.6
SE		3.07	4.82

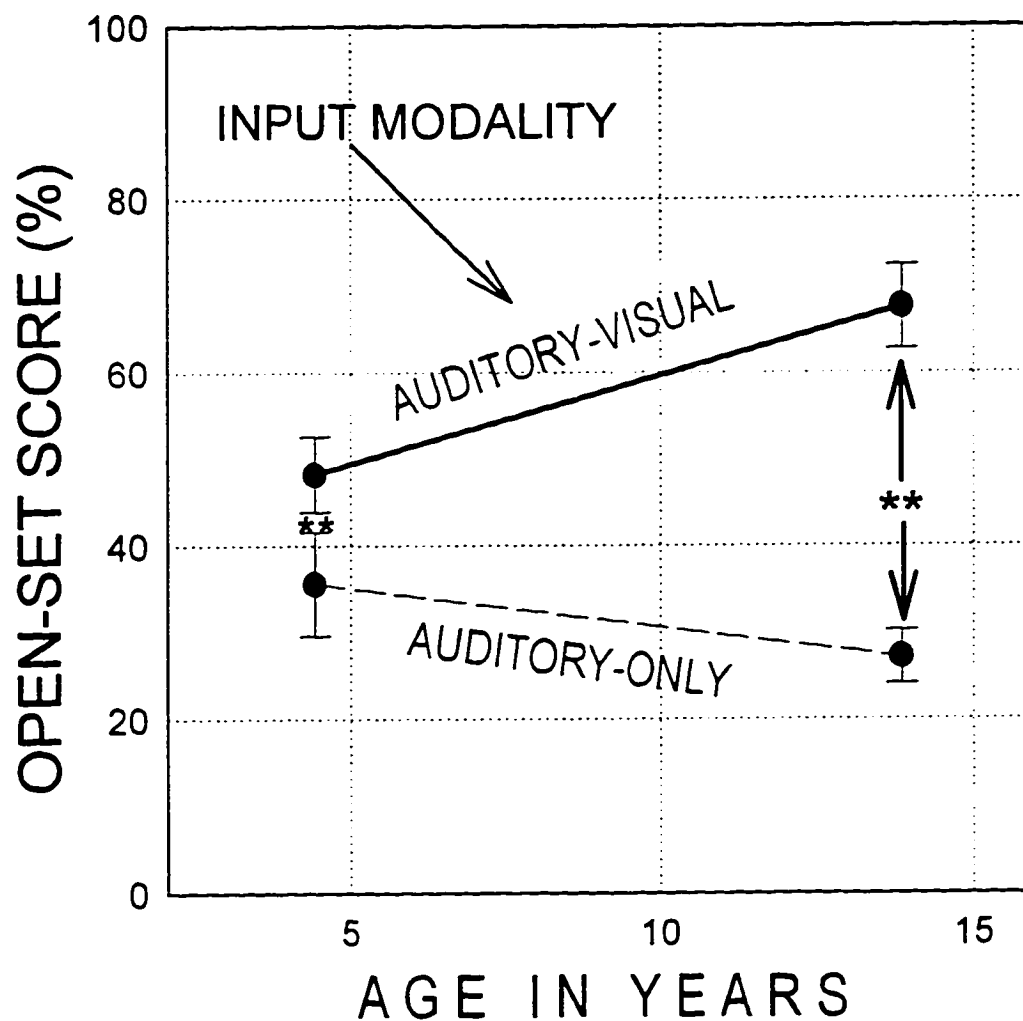


Figure 5.6. Group mean (± 1 std. error) open-set scores for 6 younger and 6 older children. Heavy lines and double asterisks show significance at the 1% level.

A two-factor analysis of variance was performed on the arcsine transformed open-set data. The results are shown in Table 5.9.

Table 5.9. Analysis of variance: effects of AGE (A) and INPUT MODALITY (M) on the Open-set scores of the 12 hearing-impaired subjects.

Effect	df	MS	df	MS	F	p-level
Effect	Effect	Effect	Error	Error		
Age	1	79.563	10	15.2363	0.6904	0.425411
Modality	1	1932.828	10	10.4677	84.6462	0.000000*
AxM	1	506.19	10	10.4677	48.3572	0.000039

*Significant effects are highlighted

The important finding of this analysis is that there was a highly significant interaction between age and input modality ($F(1,10)=48.36$, $p<0.01$). It will be seen, from an examination of Figure 5.6, that performance of the older subjects on the auditory-visual condition was very different from that of the younger subjects. It is not clear, however, whether differences in performance on the auditory-only condition also contributed to the interaction. To investigate further whether there was an age effect on

the auditory-only condition, a post-hoc analysis with the Scheffe test was applied to the means of the arcsine transformed data. The results, shown in Table 5.10, strongly suggest that the age by input modality interaction was due solely to the auditory-visual condition. This finding further supports the original observation of the absence of an age effect on the auditory-only condition.

In addition to the presence of an age effect for the auditory-visual scores, and its absence for the auditory-only scores, there was a significant input modality effect for both age groups. These findings are also shown in Table 5.10 and illustrated in Figure 5.6. These findings are consistent with the findings reported for both the original and new IMSPAC data.

Table 5.10. Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data.

Age Group	Input Modality	(1) 40.22649	(2) 48.98965	(3) 34.68294	(4) 61.81619
YOUNGER	Audition (1)		0.006926*	0.08567	0.000004
YOUNGER	Audition (2) +Vision	0.006926*		0.000166	0.00041
OLDER	Audition (3)	0.08567	0.000166		0.000001
OLDER	Audition (4) +Vision	0.000004	0.00041	0.000001	

*Significant effects are highlighted

It will be recalled that the third research question was whether the magnitude of the input modality effect on the performance of younger subjects will remain essentially unchanged when an open-set evaluation procedure is used.

Table 5.11 shows the composite open-set scores and the new IMSPAC closed-set scores for the 12 hearing-impaired subjects. Note that syllable-number and intonation contrasts were omitted from the composite score of the closed-set procedure to make it comparable to the open-set segmental scores. Thus, the IMSPAC

closed-set composite scores shown in Table 5.11 are the average of only the eight segmental speech contrasts.

Table 5.11. Open-set scores and new segmental closed-set scores for 12 hearing-impaired subjects as a function of age group and input modality. The closed-set data are based on the average of eight segmental speech contrasts.

AGE GROUP	SUBJECT NUMBER	LISTENING TASK AND CONDITION			
		Open-set Audition (%)	Open-set Audition+ Vision (%)	New Segmental Closed- set Audition (%)	New Segmental Closed- set Audition+ Vision (%)
YOUNGER	1	50.39	62.11	71.71	86.31
YOUNGER	2	18.36	34.76	62.35	70.21
YOUNGER	3	17.19	38.28	67.83	81.63
YOUNGER	4	37.11	44.95	69.0	72.17
YOUNGER	5	41.02	53.91	72.17	81.62
YOUNGER	6	49.61	55.86	67.07	69.02
MEAN		35.61	48.31	68.36	76.83
SE		6.01	4.37	2.93	5.91
OLDER	7	16.41	51.56	59.37	86.34
OLDER	8	24.22	63.28	62.9	90.64
OLDER	9	29.69	80.08	64.07	94.15
OLDER	10	35.16	67.19	62.51	89.85
OLDER	11	35.16	82.43	67.97	95.32
OLDER	12	22.27	60.94	66.81	88.67
MEAN		27.15	67.60	63.94	90.83
SE		3.07	4.82	2.54	2.76

Figure 5.7 shows the relationship between mean segmental speech perception scores and age for the two input modalities and the two evaluation procedures. The upper panel shows new IMSPAC segmental closed-set scores, corrected for guessing. The lower panel shows open-set scores.

A three-factor repeated-measures analysis of variance was performed on the arcsine-transformed open-set and closed-set segmental IMSPAC data (corrected for the effect of guessing). The results are shown in Table 5.12.

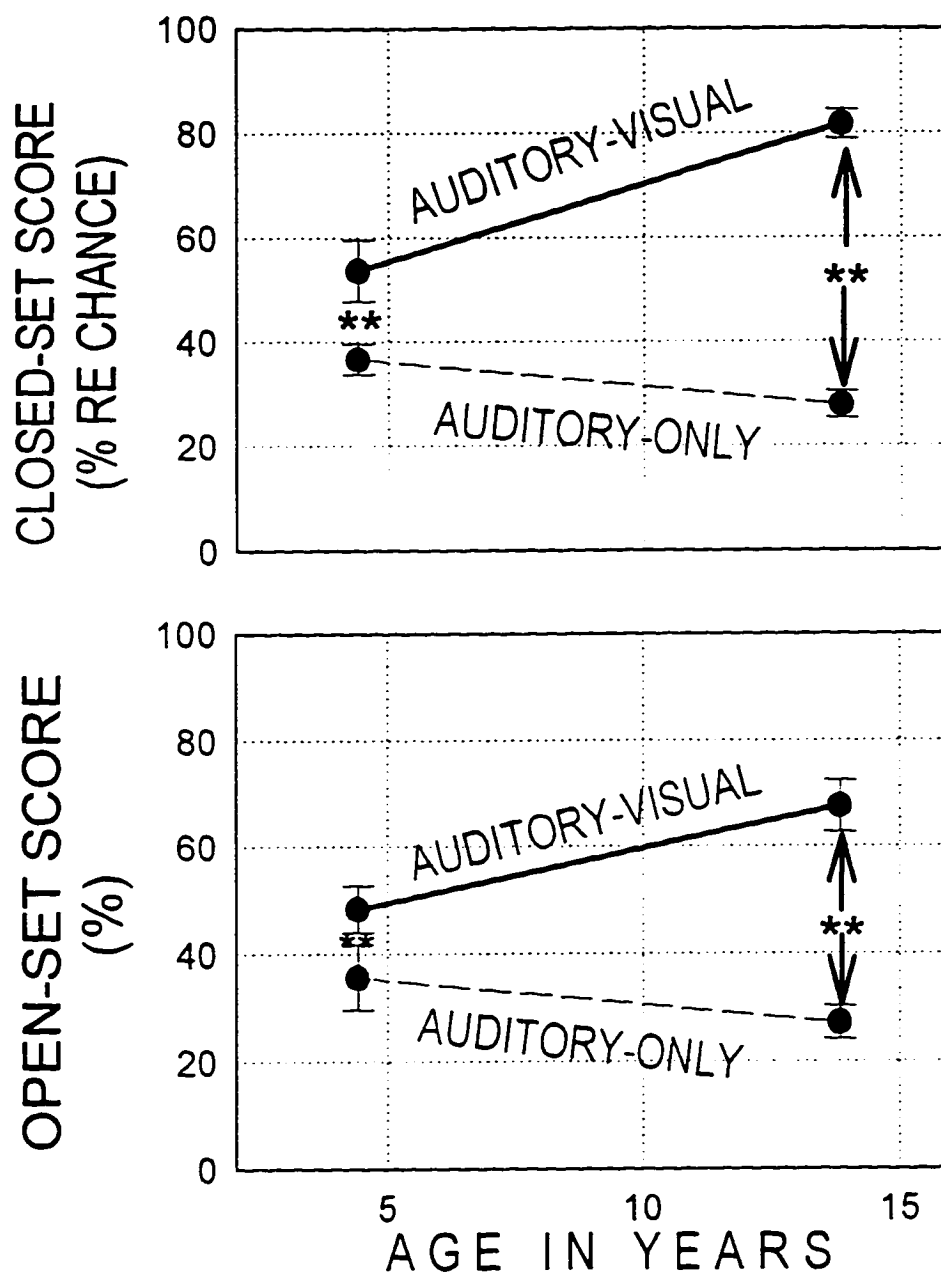


Figure 5.7. Group mean (± 1 std. error) new IMSPAC and open-set scores for 6 younger and 6 older children. Heavy lines and double asterisks show significance at the 1% level. The composite closed-set scores are based on the mean of eight contrasts (averaged across four listeners).

Table 5.12. Analysis of variance: effects of AGE (A), TEST (OPEN VS. NEW CLOSED-SET IMSPAC) (T), and INPUT MODALITY (M) on the composite scores of the 12 hearing-impaired subjects.

Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
Age	1	528.207	10	195.9871	2.6951	0.131689
Test	1	153.152	10	56.7989	2.6964	0.131608
Modality	1	8512.206	10	32.8683	258.9795	0.000000*
AxT	1	104.462	10	56.7989	1.8392	0.204879
AxM	1	2292.071	10	32.8683	69.7351	0.000008*
TxM	1	914.143	10	12.5114	73.0648	0.000007*
AxTxM	1	253.131	10	12.5114	20.232	0.001147*

*Significant effects are highlighted

The important findings of this analysis are that there was no significant main effect of evaluation procedure (open-set versus closed-set) ($F(1,10)=2.69$, $p=0.131$), and no interaction between age group and evaluation procedure ($F(1,10)=1.84$, $p=0.205$). Figure 5.8 shows the performance of the younger group for both

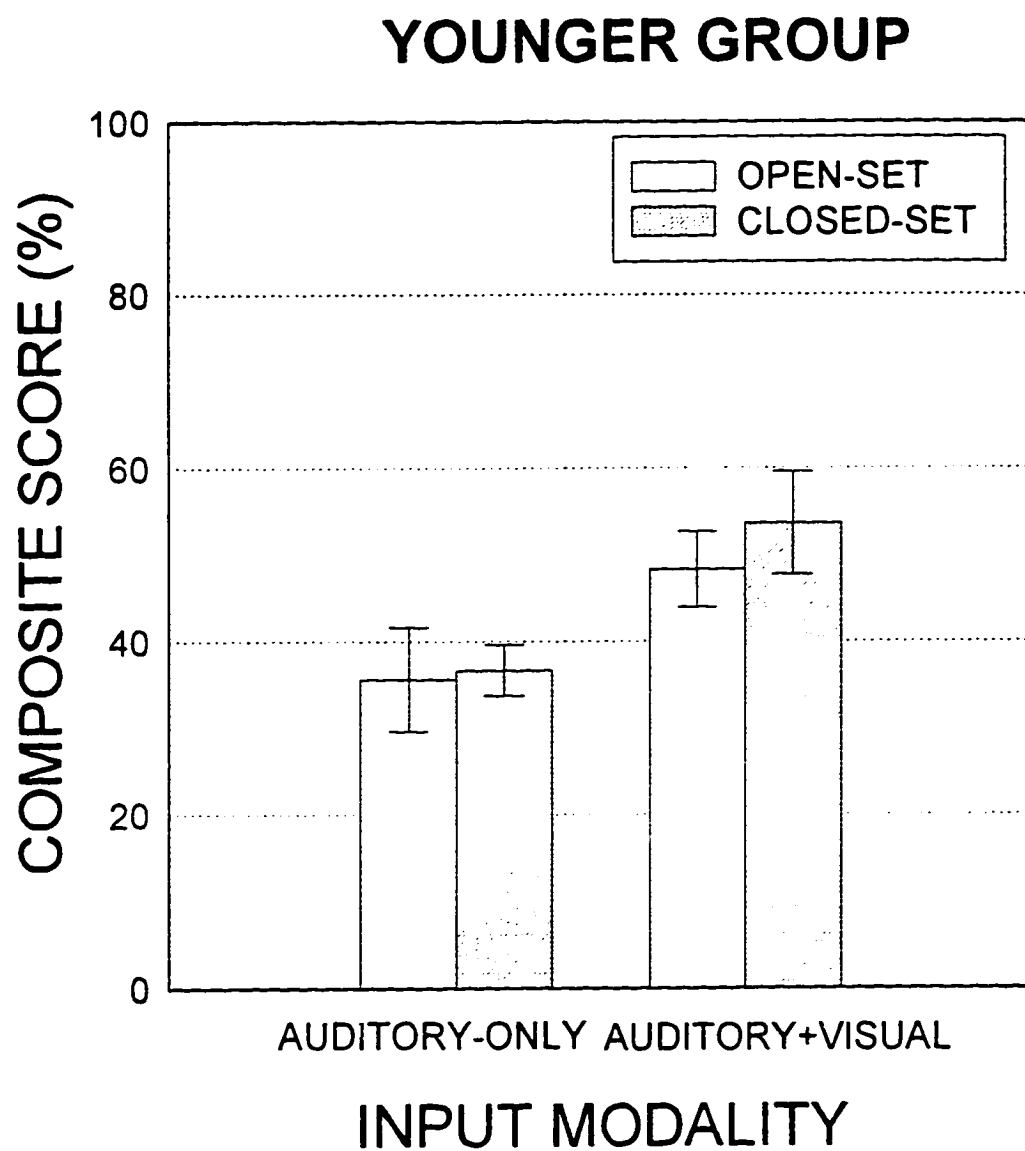


Figure 5.8. Mean new closed-set IMSPAC and open-set scores for the younger group as a function of input modality. The open-set scores are expressed in percent phoneme recognition. The closed-set scores (mean of eight segmental contrasts) are expressed in percent correct (re chance).

evaluation procedures as a function of input modality. It will be seen that the effect of input modality on performance remains consistent across evaluation procedures. A post hoc analysis with the Scheffe test showed that the differences in performance for the younger group, when measured with the open-set and the closed-set evaluation procedures, are non-significant ($p > 0.05$) for both input modalities (see Table 5.13).

Table 5.13. Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data and the new closed-set data (mean of eight segmental speech contrasts corrected for guessing).

Age	Evaluation Input	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Group	procedure Modality	40.22649	48.90445	36.71333	53.66167	34.68294	61.81619	27.885	81.66
YOUNGER	OPEN-SET Audition (1)	0.084679	0.86717	0.005367*	0.45493	0.00011	0.009885	0.000000	
YOUNGER	OPEN-SET Audition (2) 0.084679		0.010767	0.621713	0.003504	0.007171	0.000139	0.000002	
	+Vision								
YOUNGER	CLOSED-SET Audition (3) 0.86717	0.010767		0.000873	0.991773	0.000028	0.077491	0.000000	
YOUNGER	CLOSED-SET Audition (4) 0.005367	0.621713	0.000873		0.000338	0.115033	0.000022	0.00001	
	+Vision								
OLDER	OPEN-SET Audition (5) 0.45493	0.003504	0.991773	0.000338		0.000014	0.245967	0.000000	
OLDER	OPEN-SET Audition (6) 0.00011	0.007171	0.000028	0.115033	0.000014		0.000002	0.00023	
	+Vision								
OLDER	CLOSED-SET Audition (7) 0.009885	0.000139	0.077491	0.000022	0.245967	0.000002		0.000000	
OLDER	CLOSED-SET Audition (8) 0.000000	0.000002	0.000000	0.00001	0.000000	0.00023			
	+Vision								

*Significant effects are highlighted

The post hoc analysis also revealed that, similarly to the younger group, the difference in performance for the older group across evaluation procedures was not significant for the auditory-only condition. This finding is illustrated in Figure 5.9. In contrast to the finding for the younger group, however, a significant difference was found for the older group between the open-set and the closed-set procedures for the auditory-visual condition. Figure 5.10 shows the closed-set and open-set mean scores for the older group as a function of input modality. It will be seen that the mean closed-set score for the auditory-visual condition is higher than the open-set score.

To examine further the relationship between the closed-set and open-set evaluation procedures, a Pearson product moment correlation technique was used. Figures 5.11 and 5.12 are scatter plots of the subject's open-set scores as a function of their closed-set scores, for the auditory-only and the auditory-visual input modalities, respectively. Each data point represents a single hearing-impaired

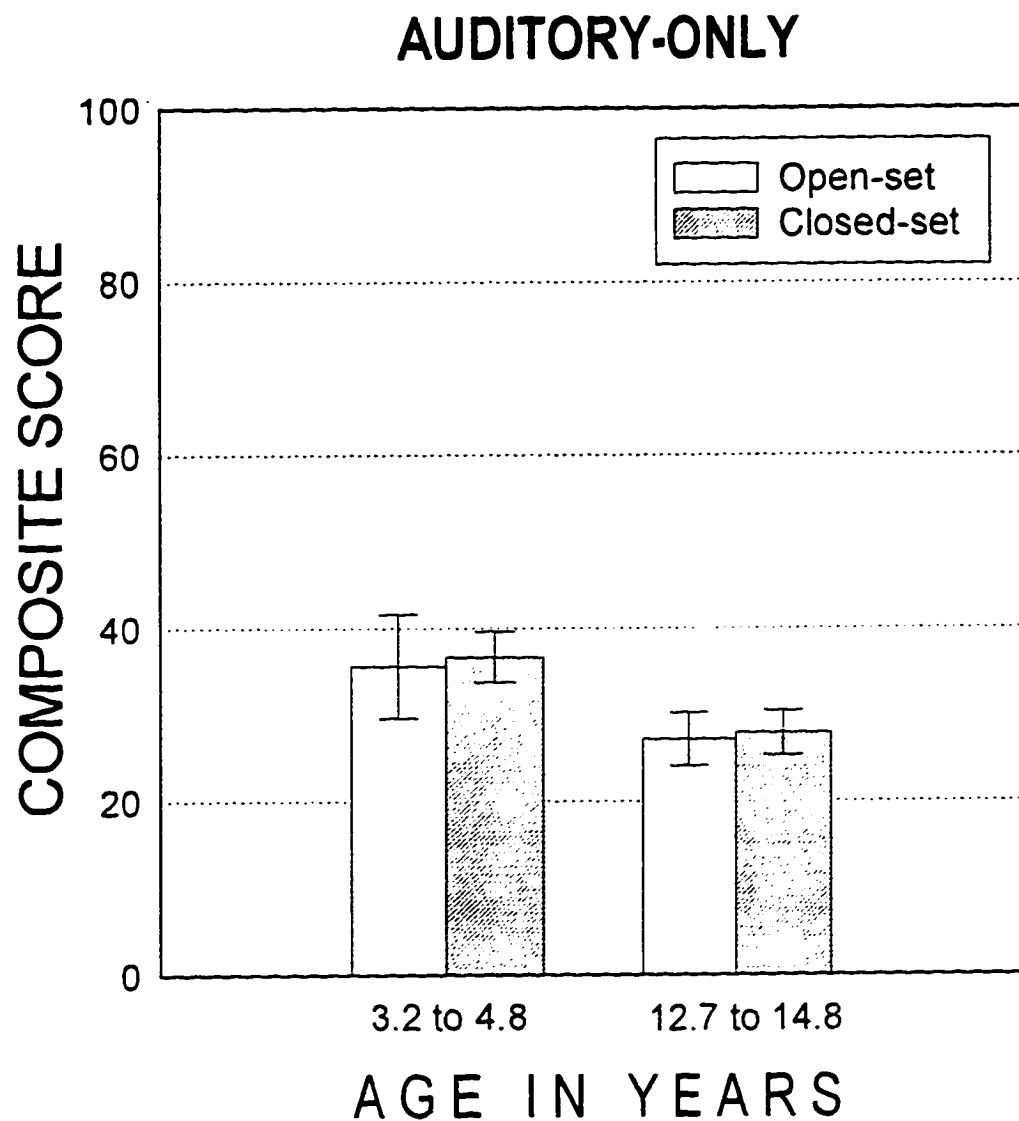


Figure 5.9. Group mean new IMSPAC and open-set scores for the auditory only condition. The composite open-set scores are expressed in percent phoneme recognition. The composite closed-set scores (mean of eight contrasts) are expressed in percent correct (re chance).

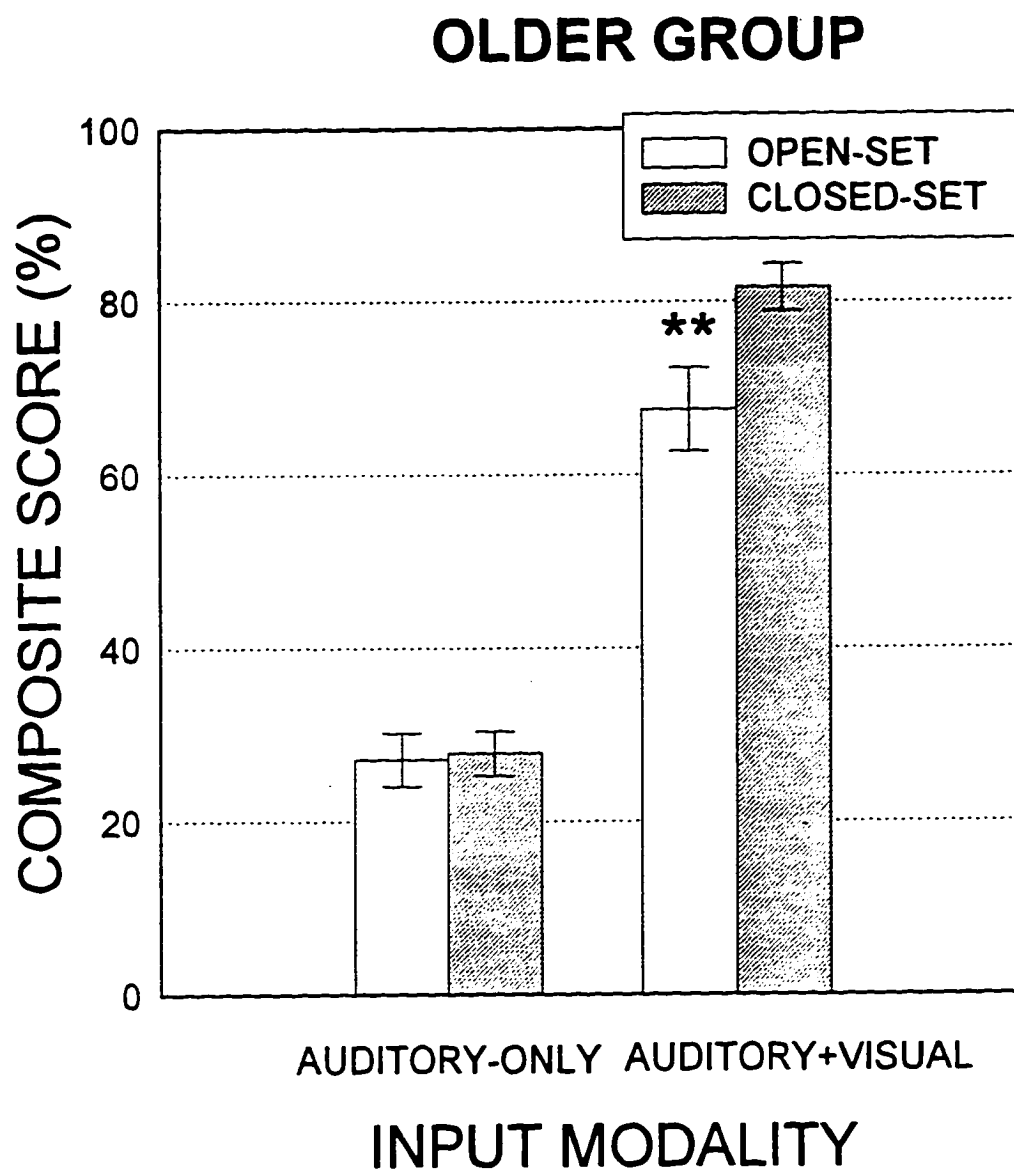


Figure 5.10. Mean new closed-set IMSPAC and open-set scores for the older group as a function of input modality. The open-set scores are expressed in percent phoneme recognition. The closed-set scores (mean of eight segmental contrasts) are expressed in percent correct (re chance). The double asterisks show significance at the 1% level.

subject. Across input modalities, scores for both evaluation procedures were closely related. The Pearson product moment correlation technique produced significant coefficients of correlation ($r(11) = 0.65$, $p < 0.05$ and $r(11) = 0.79$, $p < 0.05$ for the auditory-only and the auditory-visual conditions, respectively). The regression functions also show relative equivalency for the two sets of auditory-only scores, but higher scores for the auditory-visual closed-set scores than the auditory-visual open-set scores. The two scatter plots suggest that the difference between input modalities may be a function of the magnitude of the scores rather than the modality, per se (i.e., differences are higher for higher scores). Again, this issue will be addressed in more detail in the discussion chapter.

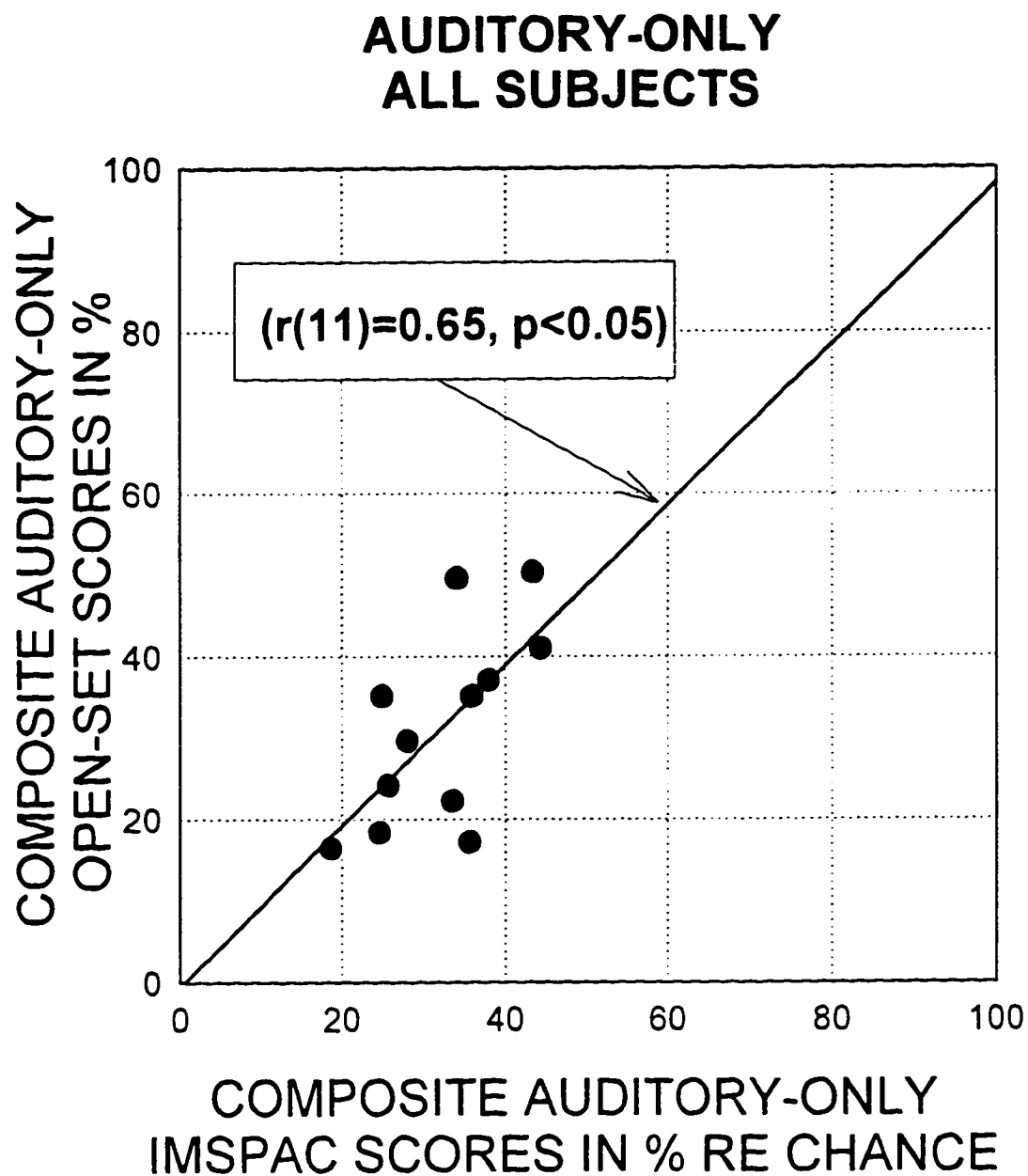


Figure 5.11. Open-set scores as a function of closed-set scores for the auditory-only condition. Each data point represents a single subject. The composite closed-set scores are based on the mean of eight segmental contrasts. The solid line shows linear regression function.

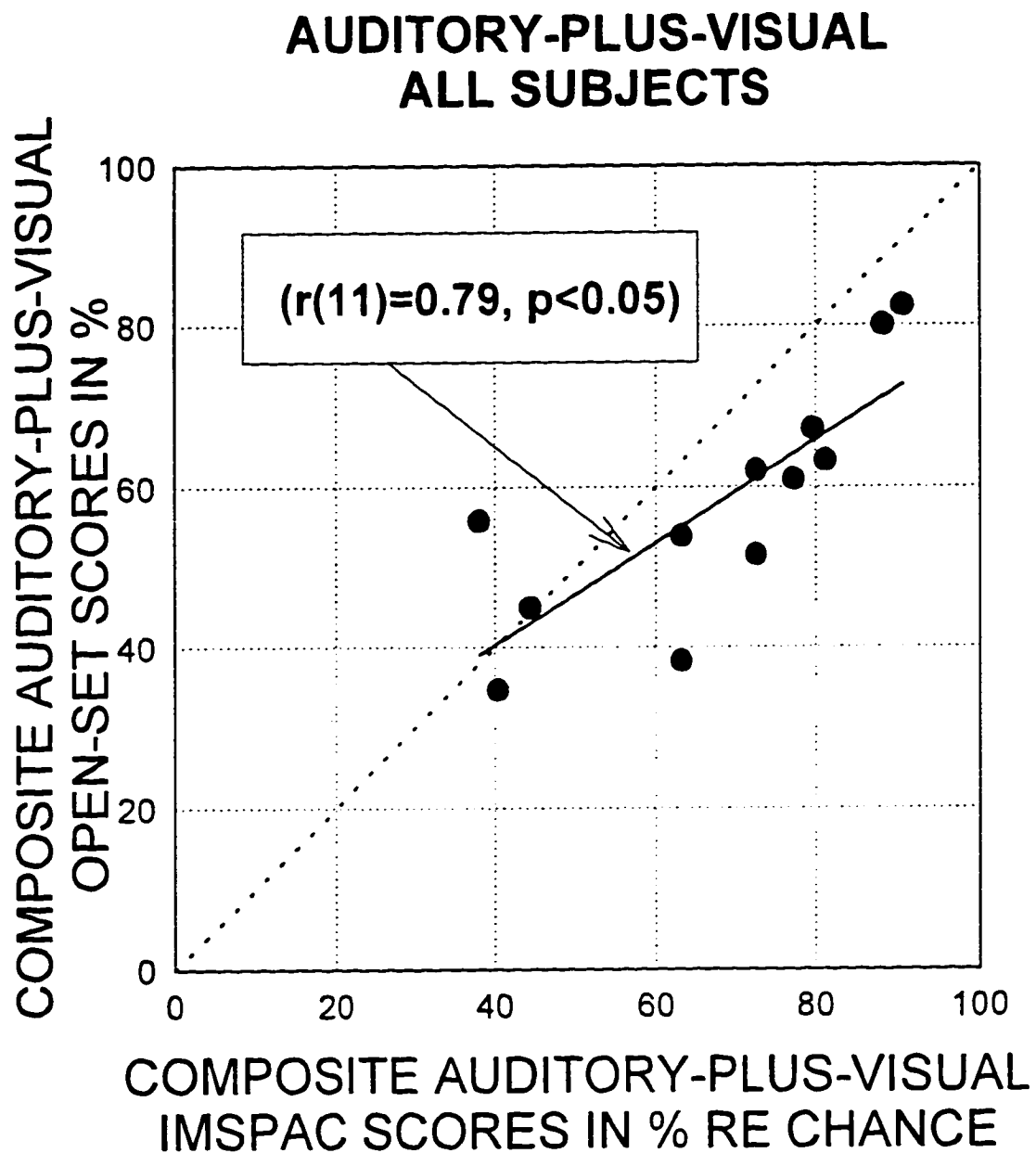


Figure 5.12. Open-set scores as a function of closed-set scores for the auditory-plus-visual condition. Each data point represents a single subject. The composite closed-set scores are based on the mean of eight segmental contrasts. The solid line shows linear regression function. The dotted line shows the diagonal.

Listener differences

The results of the analysis of variance of the IMSPAC data presented in Table 5.7 revealed a significant main effect of listener ($F(3,30)=4.47$, $p=0.0104$) and an interaction between listener and speech contrast ($F(27,270)=1.663$, $p=0.024$). Figure 5.13 shows the mean IMSPAC scores for both input modalities collapsed across 12 subjects as a function of listener. In general, listeners were similar in performance and observable differences were minimal. In fact the important findings of the analysis were that there was no significant interaction between listener and input modality ($F(3,30)=0.584$, $p=0.63$), nor among listener, input modality and speech contrast ($F(27,270)=1.00$, $p=0.467$). Figure 5.14 illustrates the listeners' variability, collapsed across the 12 subjects, as a function of individual speech contrasts. The upper panel illustrates performance for the auditory-visual input modality and the lower panel for the auditory-alone. The overall pattern of performance on individual contrasts is similar for the four listeners across modalities. This last observation provides justification for the averaging of scores across listeners.

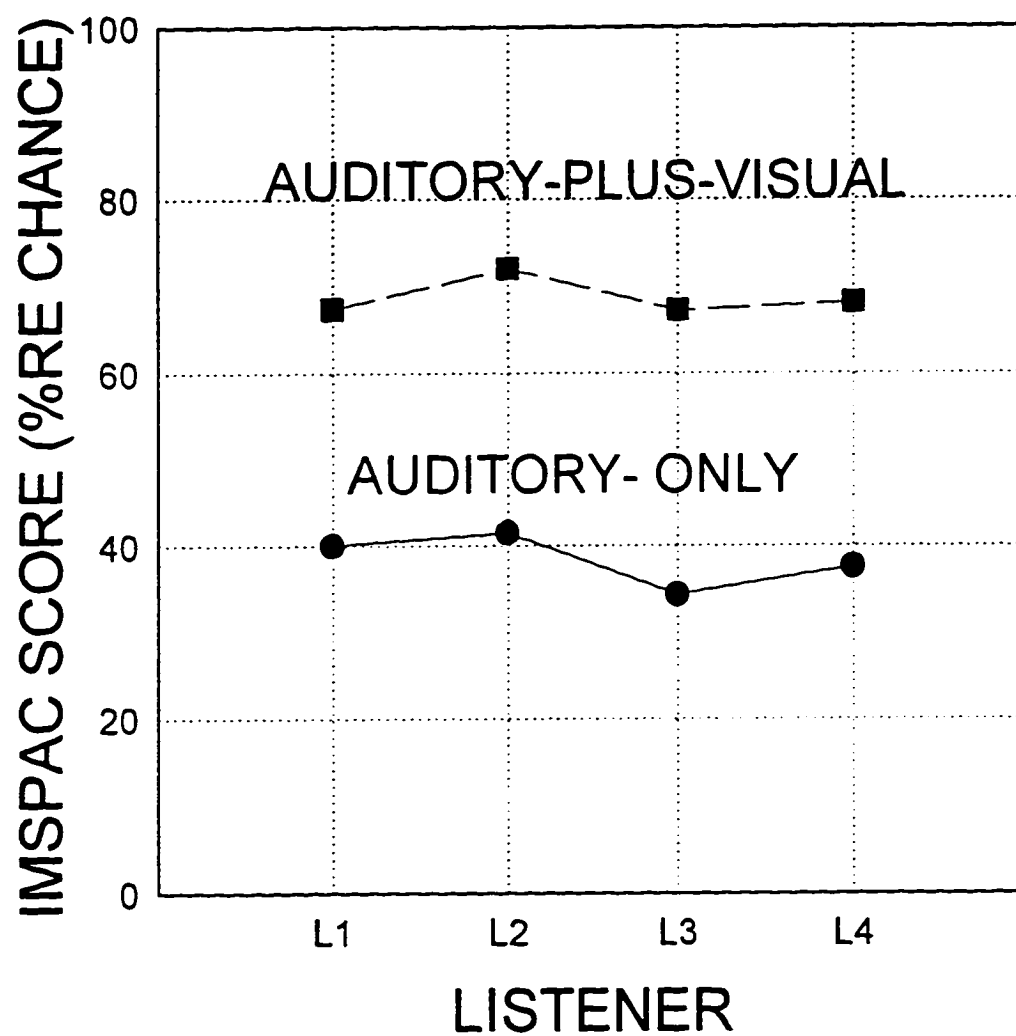
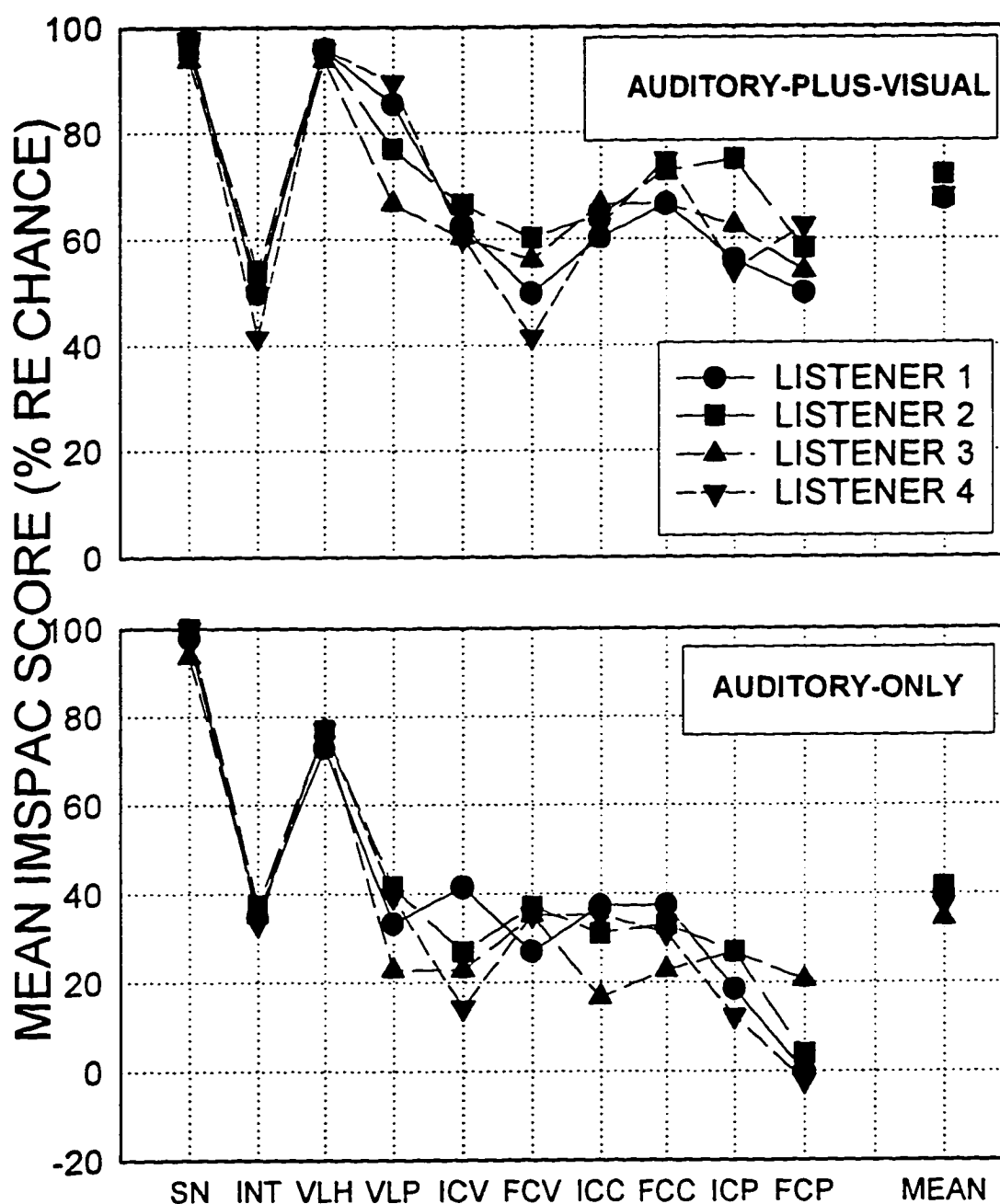


Figure 5.13. Auditory-plus-visual and auditory-only closed-set IMSPAC scores collapsed across the 12 hearing-impaired subjects as a function of listener.

LISTENERS' PERFORMANCES



SPEECH CONTRASTS

Figure 5.14. Auditory-plus-visual and auditory-only IMSPAC scores, as a function of listener and speech contrast. Scores are collapsed across the 12 subjects.

The effect of listener differences was also examined with respect to the open-set scores. A three-factor analysis of variance was performed on the arcsine transformed open-set scores of the four listeners for the two age groups. The results, shown in Table 5.14, indicate that there is a main effect of listener ($F(3,30)=20.6$, $p < 0.01$). A significant although somewhat weak interaction was found between listener and input modality ($F(3,30)=2.95$, $p = 0.048$).

Table 5.14. Analysis of variance: effects of AGE (A), INPUT MODALITY (M), LISTENER (L) on the open-set scores of the 12 hearing-impaired subjects.

Effect	df	MS	df	MS	F	p-level
	Effect	Effect	Error	Error		
Age	1	392.7799	10	485.1342	0.809631	0.389383
Modality	1	8049.711	10	53.88285	149.3928	0.000000*
Listener	3	219.9984	30	10.68001	20.59908	0.000000
A x M	1	2190.369	10	53.88285	40.65057	0.000081
A x L	3	10.85508	30	10.68001	1.016392	0.399191
M x L	3	25.72367	30	8.714765	2.951734	0.048469
AxMxL	3	2.606328	30	8.714765	0.29907	0.825777

*Significant effects are highlighted

Figure 5.15 shows the open-set mean scores for both input modalities collapsed across the 12 subjects as a function of listener. The figure suggests that the first (L1) and second (L2) listeners tended to obtain somewhat higher scores than the other two listeners. A post-hoc analysis with the Scheffe test revealed a significant difference between listeners, but only for the auditory-visual condition. Scores obtained by the first (L1) and the second (L2) listeners were, indeed, significantly different from scores of the third (L3) and fourth (L4) listeners. The results are shown in Table 5.15. Possible explanations of listener differences will be considered in the discussion chapter.

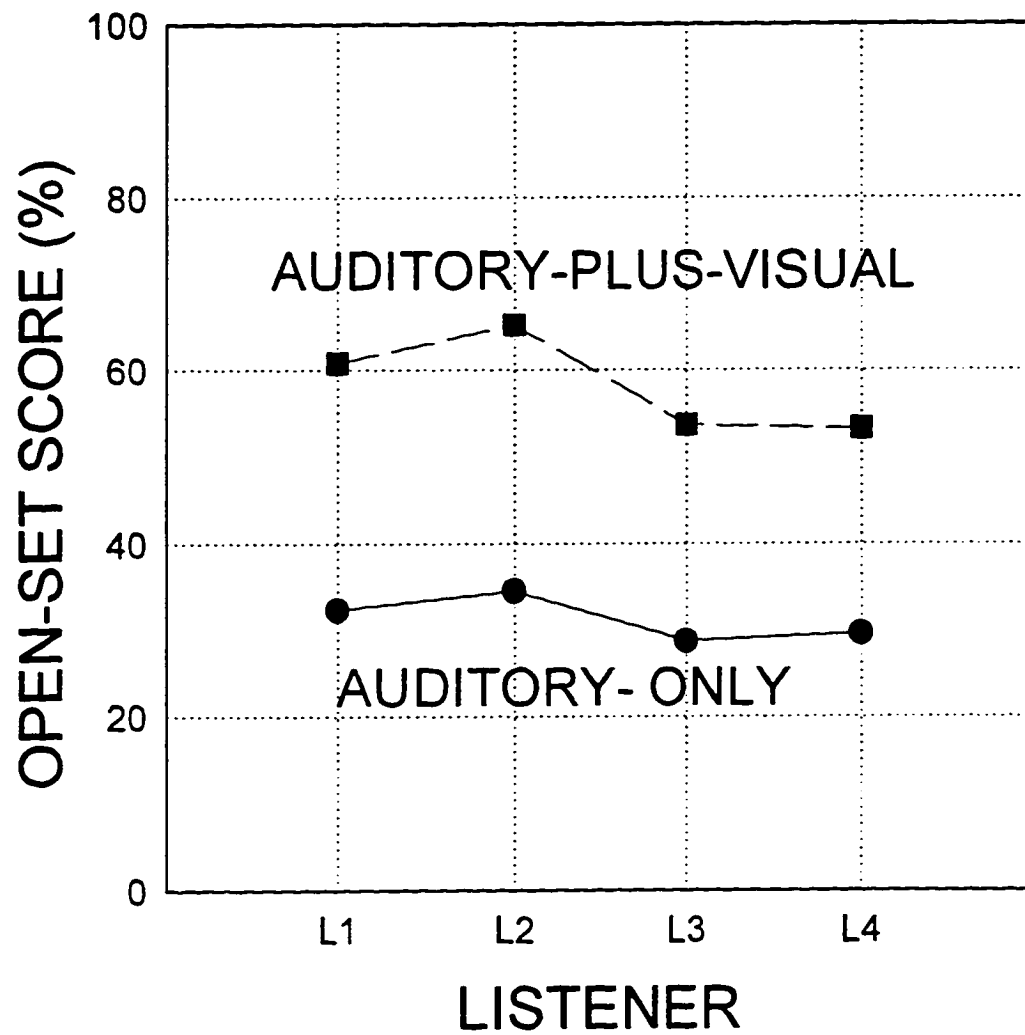


Figure 5.15. Auditory-plus-visual and auditory-only open-set scores collapsed across the 12 hearing-impaired subjects as a function of listener.

Table 5.15. Results of a post-hoc analysis with the Scheffe test applied to the means of the arcsine transforms of the open-set data.

Input	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Modality Listener	38.26826	39.75973	35.37082	36.04542	57.41826	60.57756	52.41333	52.29131
AUDITION (L1)	(1)	0.978068	0.574041	0.837094				
AUDITION (L2)	(2)	0.978068	0.105653	0.259335				
AUDITION (L3)	(3)	0.574041	0.105653	0.999852				
AUDITION (L4)	(4)	0.837094	0.259335	0.999852				
AUDITION+ (L1) VISION	(5)				0.462677	0.040163*	0.032721	
AUDITION+ (L2) VISION	(6)				0.462677	0.000098	0.000077	
AUDITION+ (L3) VISION	(7)				0.040163	0.000098		1.000000
AUDITION+ (L4) VISION	(8)				0.032721	0.000077	1.000000	

*Significant effects are highlighted

Test-retest reliability

To establish a within-listener test-retest reliability for both evaluation procedures, eight recordings, from two, randomly selected, children in each age group, were re-administered to all four listeners. Tables 5.16 and 5.17 show test-retest data for the four hearing-impaired subjects for the closed-set and open-set evaluation procedures, respectively. Composite scores, averaged across the four listeners, are shown for the auditory-only and auditory-visual input modalities. Differences between test and retest scores are also shown.

To test for the presence of a learning effect from one testing period to another, a two-factor analysis of variance was performed on the data for both evaluation procedures. The results of the analysis of variance of the closed-set IMSPAC data, presented in Table 5.18, revealed a non-significant main effect of test replication. In contrast, a small but significant main effect of test replication ($F(1,3)=13.16$, $p=0.036$) was evident for the open-set data (Table 5.19). From examination of the open-set scores presented in Table 5.17, however, it is clear that any differences that exist are very minimal.

To establish test-retest reliability of repeated presentation of the test for a single subject, the error term in the analysis of variance, presented in Tables 5.18 and 5.19, was used to estimate the standard deviations for both evaluation procedures. The resulting standard deviations in percentage points were ± 3.7 for a single closed-set score, and ± 0.3 for a single open-set score.

Table 5.16. Test and retest composite IMSPAC scores, in percent, for the auditory-only and the auditory-visual input modalities for four of the hearing-impaired subjects. Differences between test and retest scores are also shown.

AGE GROUP	SUBJECT NUMBER	Test 1		Test 2		Test 1		Test 2		Difference Audition (%)	Difference Audition+ Vision (%)
		Closed- Set Audition (%)	Closed- Set Audition (%)	Closed- Set Audition (%)	Closed- Set Audition (%)	Closed- Set Audition+ Vision (%)	Closed- Set Audition+ Vision (%)	Closed- Set Audition+ Vision (%)	Closed- Set Audition+ Vision (%)		
YOUNGER	1	76.91	77.39	77.39	77.39	87.16	86.5	86.5	86.5	-0.66	0.66
YOUNGER	5	74.59	71.81	71.81	71.81	84.98	83.39	83.39	83.39	1.59	1.59
OLDER	10	67.20	72.76	72.76	72.76	89.07	87.76	87.76	87.76	1.31	1.31
OLDER	11	70.32	72.73	72.73	72.73	94.07	93.75	93.75	93.75	0.32	0.32

Table 5.17. Test and retest open-set scores, in percent, for the auditory-only and the auditory-visual input modalities for four of the hearing-impaired subjects. Differences between test and retest scores are also shown.

AGE GROUP	SUBJECT NUMBER	Test 1		Test 2		Test 1		Test 2		Difference Audition+ Vision (%)
		Open-Set Audition (%)	Open-Set Audition (%)	Open-Set Audition (%)	Difference Audition (%)	Open-Set Audition + Vision (%)	Open-Set Audition + Vision (%)			
YOUNGER	1	50.39	50.40	50.40	-0.01	62.11	63.70	63.70	-1.59	
YOUNGER	5	41.02	41.00	41.00	0.02	53.91	55.50	55.50	-1.59	
OLDER	10	35.16	35.50	35.50	-0.34	67.19	67.60	67.60	-0.41	
OLDER	11	35.16	34.80	34.80	0.36	82.43	83.20	83.20	-0.77	

Table 5.18. Analysis of variance: effects of TEST REPLICATION (R), and INPUT MODALITY (M) on the closed-set scores of four hearing-impaired subjects.

Effect	df	MS	df	MS	F	p-level
	Effect	Effect	Error	Error		
Replication	1	0.801	3	13.5863	0.05896	0.823805
Modality	1	3780.405	3	146.9450	25.72666	0.014799
RxM	1	22.801	3	11.5216	1.97894	0.254199

*Significant effects are highlighted

Table 5.19. Analysis of variance: effects of TEST REPLICATION (R), and INPUT MODALITY (M) on the open-set scores of four hearing-impaired subjects.

Effect	df	MS	df	MS	F	p-level
	Effect	Effect	Error	Error		
Replication	1	1.172	3	0.0890	13.16297	0.036042
Modality	1	2814.568	3	281.9456	9.98266	0.050889
RxM	1	1.205	3	0.1294	9.31154	0.055363

*Significant effects are highlighted

Summary of results

1. The pattern of results from the original IMSPAC evaluation of 12 subjects was unchanged when:

(a) The same listeners were used for closed-set evaluation procedure of all subjects and recordings.

(b) Performance was evaluated by open-set recognition rather than closed-set identification.

2. This pattern of results showed:

(a) A significant effect of age on imitation of combined auditory-visual input. On average the older subjects' score was about 1.5 times the score of the younger subjects.

(b) No significant effect of age on imitation of auditory-only input.

(c) A small but significant effect of input modality in the younger subjects. For these subjects, the auditory-visual scores were about 1.5 times the auditory-only scores.

(d) A large and significant effect of input modality in the older subjects. For these subjects, the auditory-visual scores were 2.25 times the auditory-only scores.

3. Individual speech contrast analysis showed that the modality effect for the younger subjects was limited to vowel place and final consonant place of articulation contrasts.

4. In the older subjects, however, modality effects were present for all speech contrasts except syllable number (the mean scores for which were close to 100%).

5. Contrast analysis showed, further, that any age effect for auditory-only input was limited to initial consonant place of articulation. This effect was, however, in favor of the younger subjects, whose mean score for this contrast was about twice that of the older subjects.

6. For auditory-visual input, significant age effects were found for five of the ten contrasts: intonation, initial and final consonant voicing, and initial and final consonant continuance.

7. The open- and closed-set scores for auditory-only input were highly correlated ($r=0.65$) and not significantly different.

8. The open- and closed-set scores for auditory-visual input, however, were correlated ($r=0.79$) but not necessarily equal. The closed-set auditory-visual scores tended to be around 10 percentage points higher when evaluated by the open-set task.

9. There were small but significant differences among the four listeners, but only for the open-set evaluation procedure and only for the auditory-visual input. These differences, however, did not influence the basic pattern of findings regarding the effects of age, modality, and evaluation procedure.

10. Repeated evaluation of the same recordings, using both closed-set and open-set evaluation procedures, revealed little or no learning effect for the listeners.

11. The estimated standard deviation of repeated evaluation of the same recording by the same group of four listeners was 3.7 and 0.3 percentage points for the closed-set and open-set evaluation procedures, respectively.

The principal conclusion, drawn from the previous data and supported by the new data, is that the ability to imitate speech from auditory-only input does not change significantly between ages 4 and 14. Furthermore, the fact that there is a highly significant and large effect of age on the imitation of auditory-visual input indicates that the negative finding for auditory-only input is not simply a reflection of low power of either the test instruments or the design. Note, further, that the only evidence of a significant age effect for auditory-only input is in the opposite direction from that expected. Namely, the older group performed less well than the younger group on the initial consonant place contrast. Some possible implications of this finding will be discussed later.

Chapter 6

Acoustic analysis of vowels

The result obtained so far, that the ability to imitate speech from auditory input does not change with age, is consistent with the conclusion that neither maturation nor improved production influenced the auditory imitation performance. Hence, this finding is in keeping with the conclusion that auditory imitation reflects peripheral auditory speech perception capacity.

However, the performance measures thus far reported are limited to behavioral measures. These measures addressed imitation in terms of its ability to convey both contrastive and categorical information. It remains possible, however, that there are some subtle contrastive acoustic differences between the imitations of older and younger subjects that are not revealed in the behavioral measures. Thus, it remains to be determined whether there is an age-related change in ability to imitate auditory input that is revealed when imitations are evaluated in terms of their acoustic properties.

The purpose of this section is to examine the acoustic characteristics of the hearing-impaired

subjects' productions of vowels. The general goal is to assess the subjects' ability to make appropriate phonemic contrasts as measured by means of acoustic analysis.

The specific goals are:

1) To establish age-sensitive vowel templates in F_1/F_2 frequency space.

2) To measure acoustic performance in terms of discriminability in F_1/F_2 frequency space, given auditory-visual and auditory-only input modalities.

3) To examine the relationship between mean acoustic performance and behavioral results on the closed-set listening task described in an earlier chapter.

The experimental questions to be answered are as follows:

1. Are there age-related changes in the ability to imitate phonemically contrastive auditory input that are revealed when the imitations are evaluated in terms of their acoustic properties?

2. Are the behavioral results predictable from acoustic measures of performance?

An acoustic analysis of vowels was performed for all subjects on all recordings. The restriction of this

study to vowels arose for two practical reasons. First, vowels are believed to convey information mainly by their formant patterns and formant frequencies, whereas consonants convey information by numerous acoustic cues and therefore require a variety of possible measures. Second, design of the IMSPAC results in more tokens of each vowel used than tokens of each consonant used.

Method

Stimuli

A total of 730 CV and VC utterances were acoustically analyzed for the frequency values of the first and second formants (F_1 and F_2). The four vowels included in the IMSPAC test were measured, namely: /i/ as in the word 'beet', /u/ as in 'boot', /E/ as in 'bed', and /c/ as in 'bought'. The symbols used for representation of the vowels are one-character representations of phonemes used in DECTalk (Klatt, 1987). This representation is nearly identical to the ARPAbet code (Shoup, 1980). Measurements of F_1 and F_2 were taken at a central steady-state location of the vowel.

Instrumentation

The Computerized Speech Lab (CSL), Model 4300B by Kay Elemetrics, was used for the acoustic analysis.

Procedure

The 12-bit data files of subjects' utterances were converted to 16-bit files for compatibility with the 16-bit sound board of the CSL system. The general protocol for acoustic measurements was as follows (with some variations based on specific needs):

- 1) In the CSL program, the screen was divided into two windows (see Figure 6.1). The upper window displayed the waveform of the utterance (pressure as a function of time) and the lower window a wide-band spectrogram (frequency/amplitude as a function of time). The frequency range on the wide-band spectrogram (300 Hz bandwidth) was from 0 to 4000 Hz.

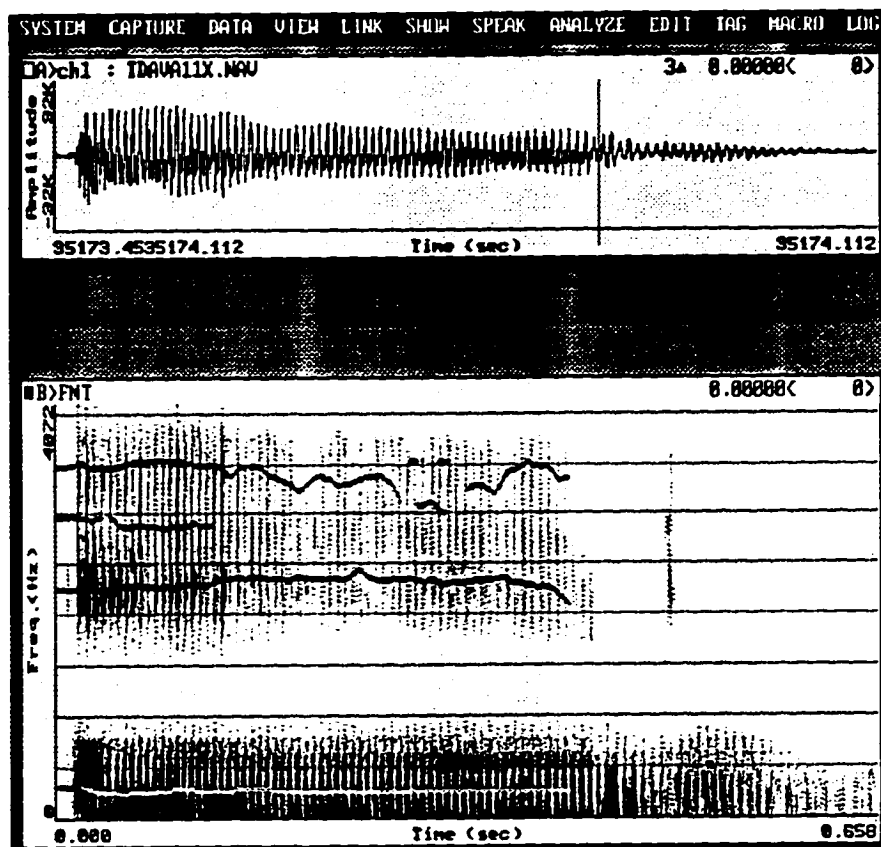


Figure 6.1. Waveform (top panel) and wide-band spectrogram (bottom panel) for the utterance "eed", produced by a hearing-impaired child in the auditory-visual condition. Also seen is an LPC formant history superimposed on the spectrogram.

2) Time cursors were placed on the waveform display, defining the boundaries of the syllable.

3) A wide-band spectrographic display was obtained.

4) Locations of F1 and F2 were visually determined at a steady-state portion of the vowel and measurements of the formant frequency values were taken.

5) When the formants of the vowels were not clearly defined, two procedures were used. The first was a **linear predictive coding (LPC) formant history**. The tracing of the formant history was superimposed on the wide-band spectrogram, as can be seen in Figure 6.1. The second procedure was a **linear predictive coding (LPC) spectrum** (see window "C" in Figure 6.2). The combination of LPC and wide-band spectra was used for approximately 25% of the vowels.

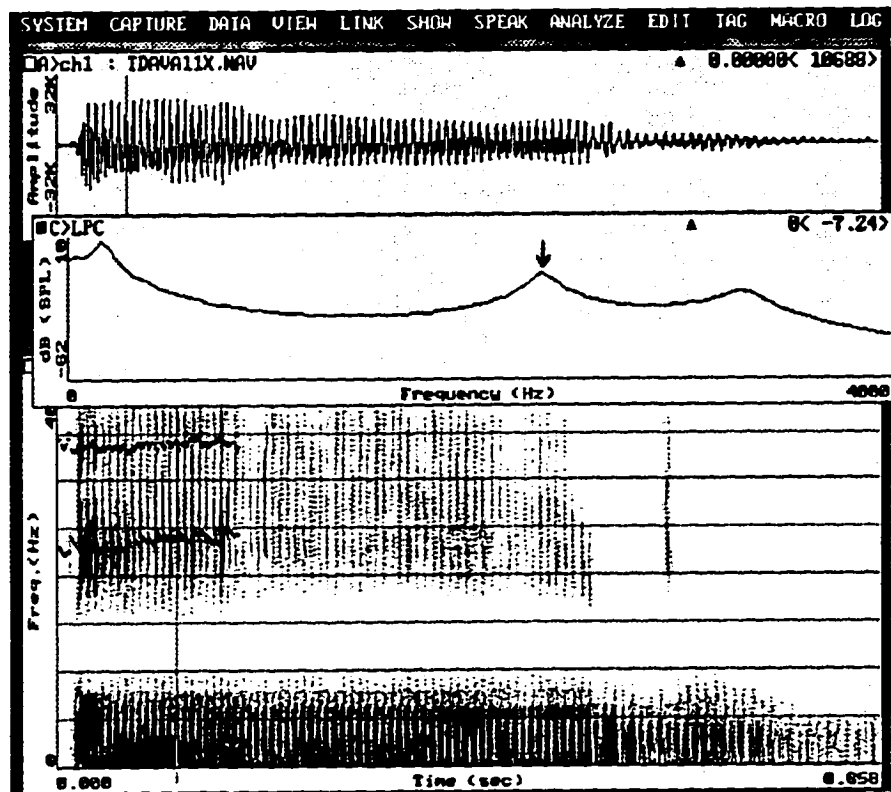


Figure 6.2. The same spectrogram and waveform as in Figure 6.1., but with an LPC spectrum displayed in window "C". The arrow points to the peak of F_2 .

Results

Analysis of vowels correctly identified in the open-set task

The first question to be addressed was "What are the F_1 and F_2 values of vowels that were correctly identified by the four experienced listeners in the open-set task?" The answer to this question provides information on the extent to which the speech of the hearing-impaired children approximates that of age-matched normally hearing children available in the published literature. It also provides templates, specific to this sample of children, against which to evaluate the complete body of F_1/F_2 data including both correct and incorrect imitations. Table 6.1 shows the number of vowels that were correctly recognized by all four listeners in the auditory-visual condition for each of the 12 hearing-impaired subjects, together with their mean values for the frequencies of the first (F_1) and second (F_2) formants.

Table 6.1. Mean values for the frequencies of the first (F1) and second (F2) formants and number of vowels that were correctly identified by all four listeners in the auditory-visual condition on the open-set listening task for each of the 12 hearing-impaired subjects.

Age Group	Sub No.	/u/			/i/			/c/			/E/		
		No.of tokens	n	Mean F1	Mean F2	No.of tokens	n	Mean F1	Mean F2	No.of tokens	n	Mean F1	Mean F2
Younger	H21	10	3	513.3	1195.0	9	7	474.7	2729.9	9	2	753.5	1282.0
	H44		2	571.5	1671.5		2	554.0	2876.0		0		
	H46		1	537.0	1750.0		0				0		
	H47		2	545.5	1316.5		3	375.0	3159.0		0		
	H66		4	445.0	2298.0		7	489.6	3128.4		0		
	H70		10	474.2	1120.5		3	496.3	2893.0		6	783.0	1240.5
	Mean	6	10	22 514.2	1558.3	9	22	477.6	2957.0	9	8	768.0	1261.0
Older	H23		10			9	4	268.0	2650.0	9	0		
	H78		9	386.5	1663.1		3	369.0	2177.0		0		
	H80		5	356.6	1399.4		7	289.1	2333.9		1	623.0	1230.0
	H81		7	260.0	1700.1		3	265.3	2228.7		2	692.5	1351.0
	H84		10	286.9	1278.3		8	304.9	2169.7		4	671.0	1190.5
	H106		7	269.3	1744.4		4	237.7	2126.2		0		
	Mean	6	10	38 311.9	1557.1	9	29	289.0	2280.9	9	7	662.2	1257.2

No. of tokens = number of presented models.

n = number of imitations judged correct by the listeners

Figure 6.3 displays the mean values of F_2 versus F_1 for the 'template' vowels /i/, /u/, /E/, and /c/ produced by the hearing-impaired subjects in the auditory-visual condition. The left panel shows formant values for the younger subjects and the right panel shows values for the older subjects. Each symbol represents, for a single vowel, the mean frequency values for one child. The circles show the group mean for each vowel. As mentioned earlier, only vowels correctly identified by all four listeners are included. Thus, not all six subjects are accounted for in the present analysis. For example, only two out of the six younger subjects produced correct tokens of the vowel /E/ (left panel). On the other hand, five out of the six older subjects produced tokens of the same vowel that were correctly perceived by all listeners.

It will be seen from Figure 6.3 that the frequency separation of group means along the F_2 space is greater for the younger group (left panel) than for the older group (right panel), whereas the frequency separation along the F_1 space is greater for the older group than for the younger group. It will also be seen that, overall, the values of F_1 and F_2 for /i/ and /E/ are higher for the younger

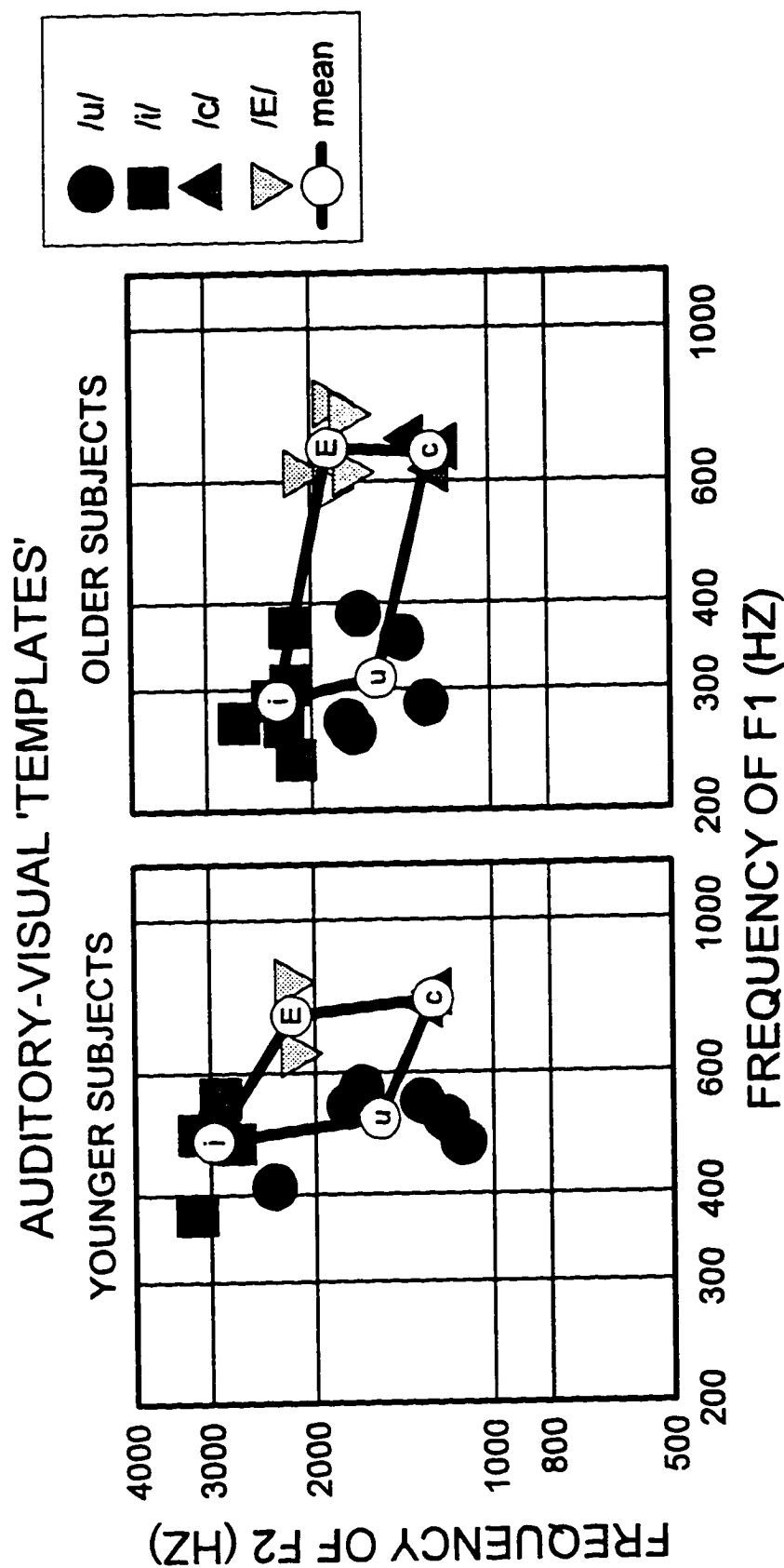


Figure 6.3. Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-visual condition. The left panel shows the formant values for the younger subjects and the right panel for the older subjects. Only vowels that were correctly identified by all four listeners in the open-set listening task are included. The means of these vowels will serve as an age-sensitive templates.

subjects. Kent and Read (1992) suggested that as children grow, their formant frequencies decrease because of overall growth of facial structures, particularly lengthening and change in shape of the vocal tract.

It is appropriate to ask how these 'template' values compare with published data from normally hearing talkers. Peterson and Barney (1952) in their classic study measured the formant frequencies of 10 vowels embedded in monosyllabic words in a /hVd/ format spoken by 15 children, 28 women, and 33 men. No information is given about the age or gender of the child talkers. Peterson and Barney also included in their study an identification task in which a group of 70 listeners were asked to write down the monosyllabic words spoken by subjects. Listening results were not reported separately for children, women and men talkers. Furthermore, Peterson and Barney did not include the frequency values for correctly identified vowels. Therefore, their data reported here include frequency values for all vowel productions. Figure 6.4 gives a graphic summary of their data for the same four vowels that were examined in the present study (i.e., /i/, /E/, /u/,

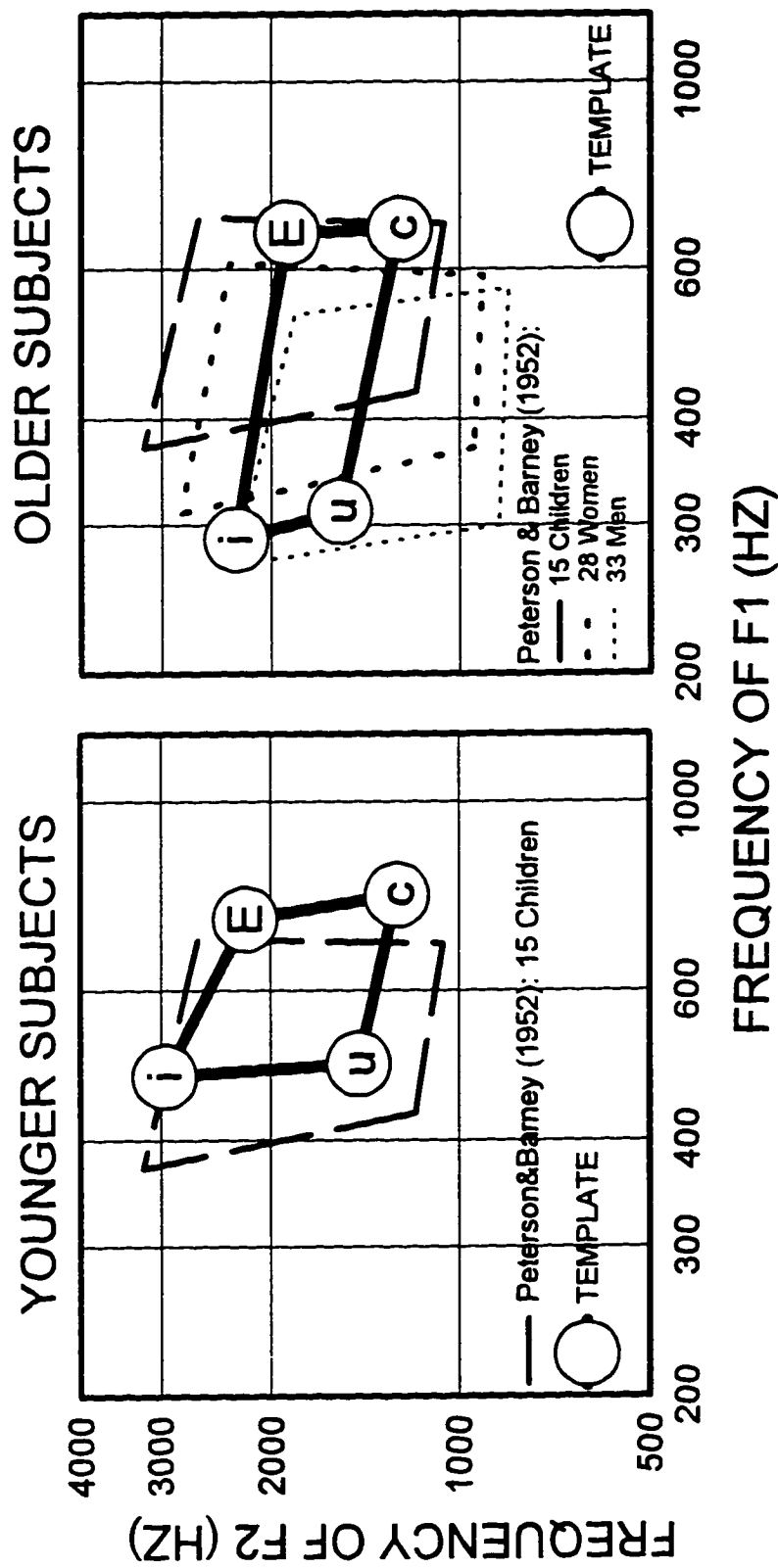


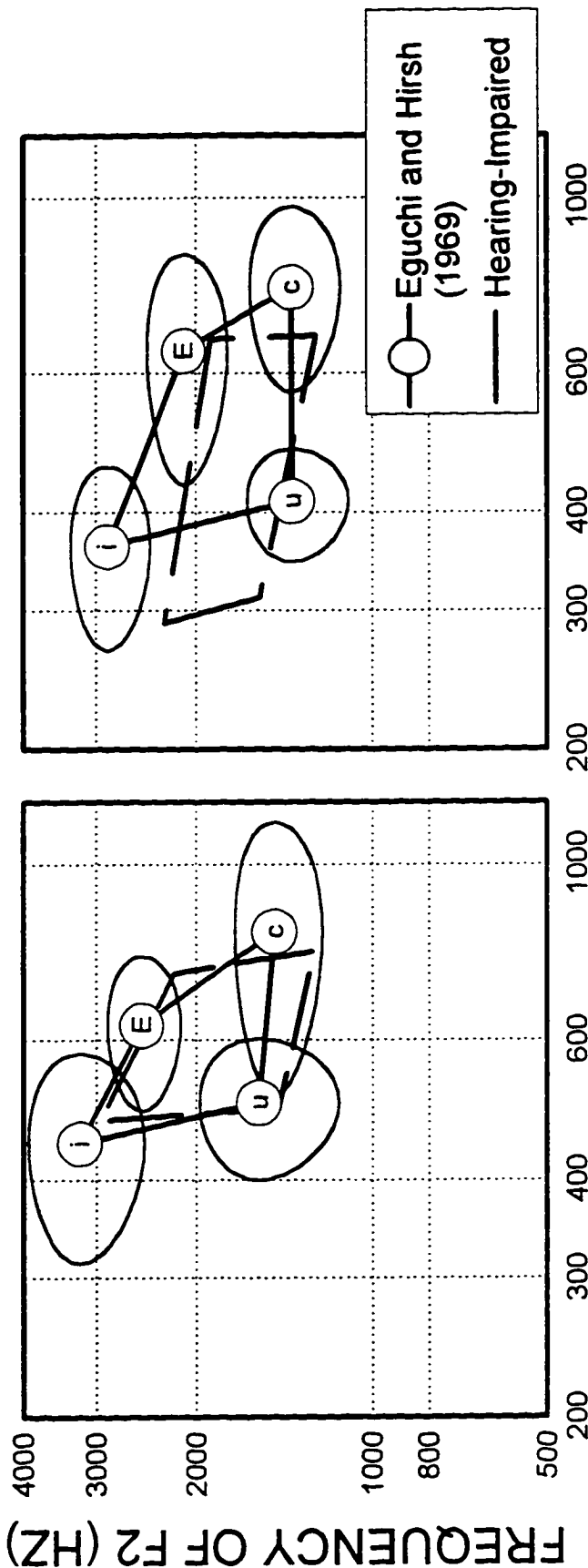
Figure 6.4. Second versus first formant frequency for the vowel templates produced by the 12 hearing-impaired subjects, and for the Peterson and Barney (1952) data for children, women, and men.

and /c/). The figure also displays the vowel templates for the 12 hearing-impaired subjects taken from Figure 6.3. The left panel shows the formant values for the younger hearing-impaired subjects together with the data reported by Peterson and Barney for their child talkers. The right panel shows the formant values for the older hearing-impaired subjects together with the Peterson and Barney data for their men, women, and children groups. It will be seen that the mean formant values and the formant separations for vowels produced by the younger hearing-impaired subjects are similar to these of the normal hearing children. In contrast, the second formant frequency separation for the older hearing-impaired children is significantly reduced compared to that for normal hearing subjects. However, since Peterson and Barney did not specify the ages of the children included in their study, they may not be appropriate for comparison with of the sample of hearing-impaired children included in the present study. A later study by Eguchi and Hirsh (1969) provides formant frequencies as well as between-subject standard deviations for six vowels spoken by 70 children divided into age groups from 3 to 13. The vowels were extracted from sentences that were either

read by children over seven years of age, or imitated by the younger children. Similarly to Peterson and Barney, Eguchi and Hirsh also included in their study an identification task, in which a group of 20 normally hearing adult listeners identified a sample of vowels taken from the recordings of several age groups. No acoustic measurements are reported, however, for vowels that were correctly identified.

For the purpose of comparing the hearing-impaired templates with age-appropriate normal hearing templates, formant frequency data were extracted from Eguchi and Hirsh's study for children ages 3 to 5 ($n=15$), and 12 to 13 ($n=20$). Figure 6.5 shows means $\pm$ 2 standard deviations for the normal hearing children for the same four vowels examined in the present study. The left panel shows data for 15 younger children aged 3 to 5, and the right panel shows data for 20 older children aged 12 to 13. Also shown are the vowel templates for younger and older hearing-impaired subjects taken from Figure 6.3. It will be seen that, similarly to Peterson and Barney's data, the mean formant values and the formant separations for vowels produced correctly by the younger hearing-impaired subjects are similar to those of the normal hearing children. In contrast to

OLDER SUBJECTS



FREQUENCY OF F1 (HZ)

Figure 6.5. Second versus first formant frequency for four vowels produced by 30 hearing children. Data was taken from Eguchi and Hirsh (1969). The left panel shows the formant values for children aged 3 to 5 years. The right panel shows the formant values for children aged 12 to 13 years. Broken lines show vowel templates for the older and younger hearing-impaired subjects taken from Figure 6.3. The width and the height of each ovoid shape are ± 2 standard deviations between subjects, pooled for three groups of five children in each, for the F1 and F2 values, respectively. The circles are the mean for each vowel.

the younger subjects, it will be seen that there is an overall decrease in mean formant values for the older hearing-impaired children relative to the older normal hearing children, with a greater drop of the second formant than the first formant. Particularly, the F_2 value for /i/ is significantly lower compared to the normal hearing children. Because of that drop, the range of F_2 variation for the /i/ and /u/ contrast is reduced.

Figures 6.6 and 6.7 illustrate the relationship between age-representative vowel templates obtained from Eguchi and Hirsh's data to the distributions of correctly identified vowels in the auditory-visual imitations of the hearing-impaired children. Figure 6.6 illustrates individual data of the six younger hearing-impaired subjects together with means $\pm$ 2 standard deviations for normal hearing 3- to 5-year-olds. Figure 6.7 illustrates individual data of the six older hearing-impaired subjects together with means $\pm$ 2 standard deviations for normal hearing 12- to 13-year-olds. It will be seen from these figures that the formant values of correctly perceived vowels produced by the younger hearing-impaired children are close to the formant values of normal hearing children of similar age. In contrast

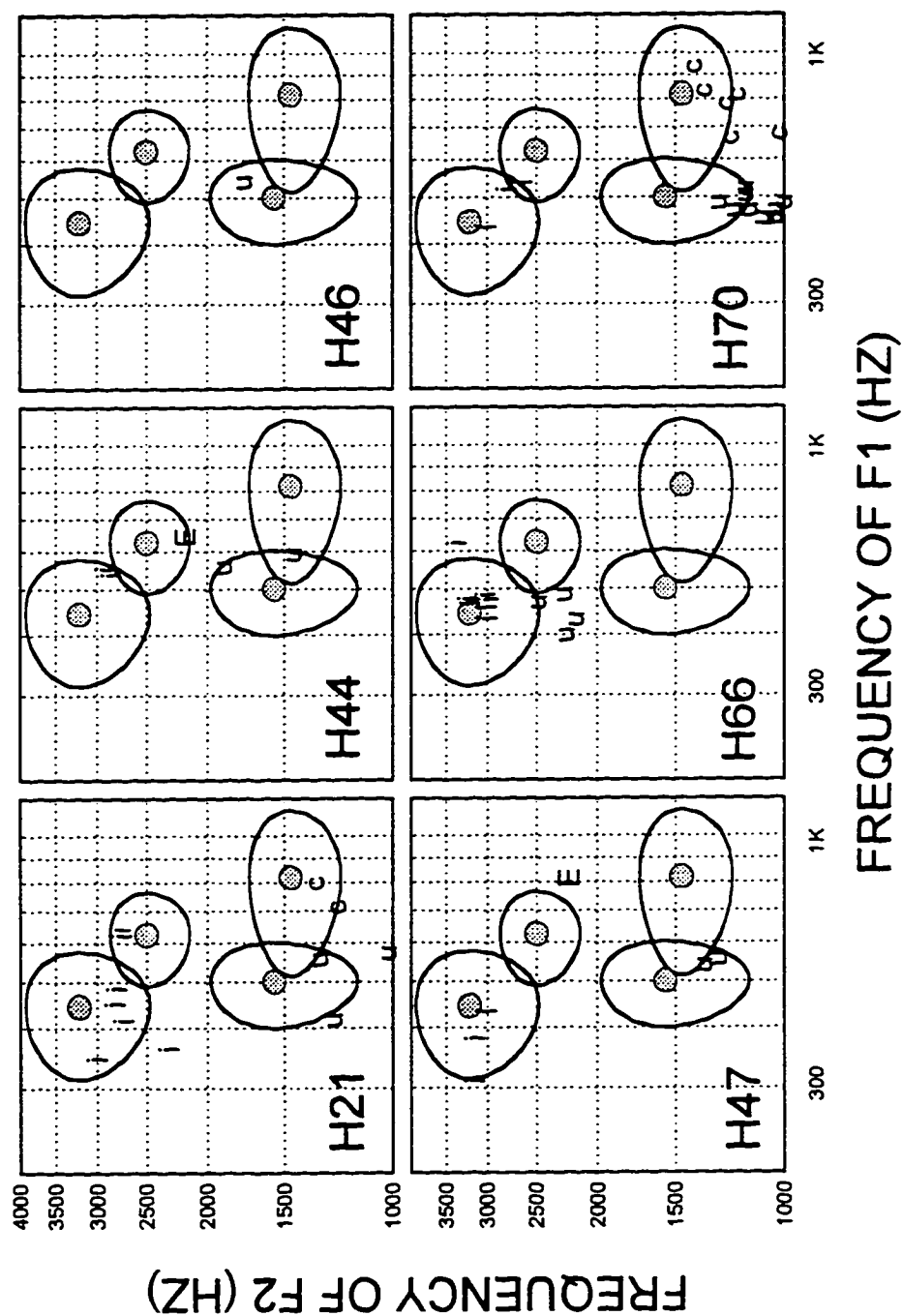
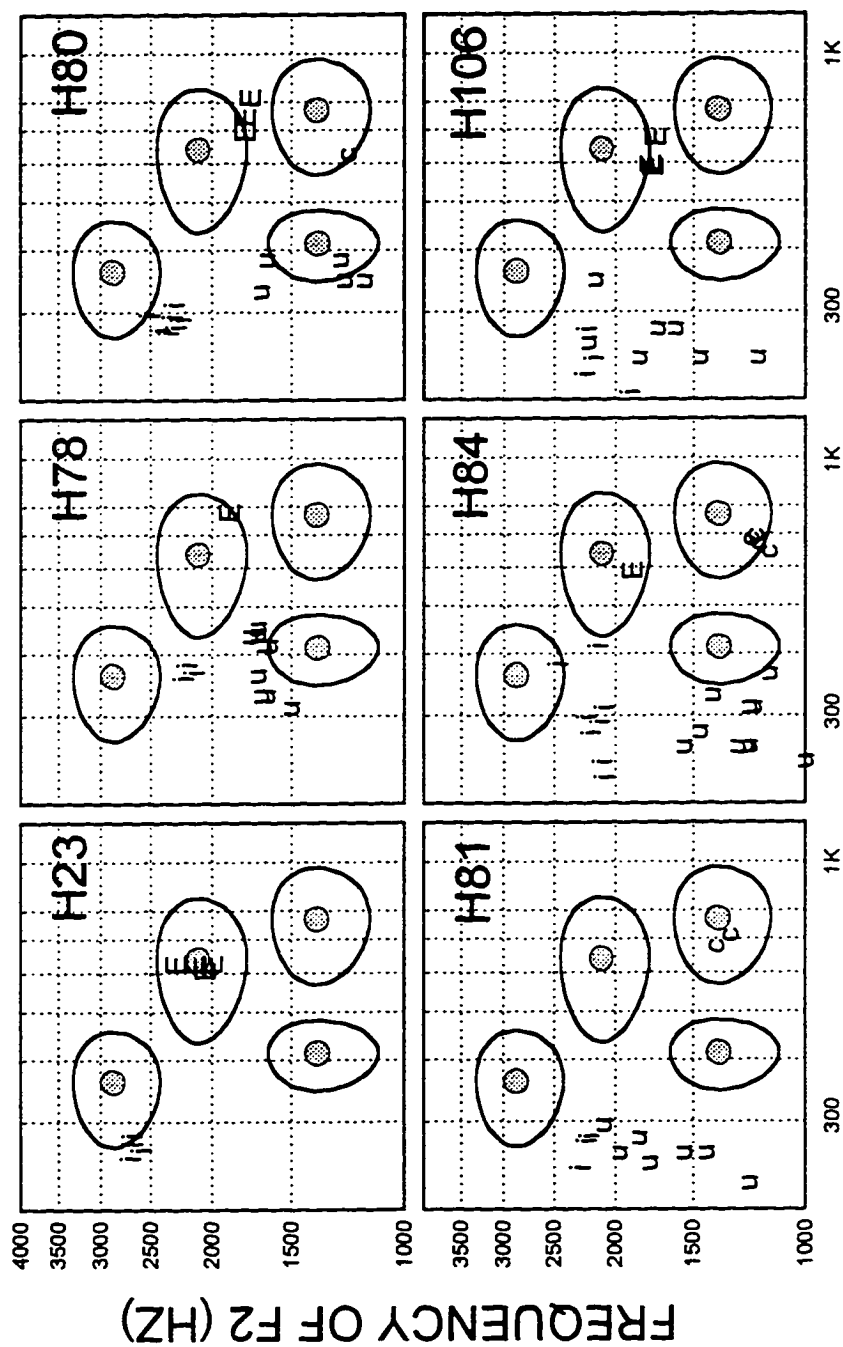


Figure 6.6. Distributions of first and second formant frequencies of correctly identified vowels in the auditory-visual imitations of 6 younger hearing-impaired children. The ovoid shapes show the means $\pm$ 2 standard deviations for normal hearing 3 to 5 year olds, taken from Figure 6.5.



FREQUENCY OF F1 (HZ)

Figure 6.7. Distributions of first and second formant frequencies of correctly identified vowels in the auditory-visual imitations of 6 older hearing-impaired children. The ovoid shapes show the means $\pm$ 2 standard deviations for normal hearing 12 to 13 year olds, taken from Figure 6.6.

the mean formant values, mainly for /i/ and /u/, of the older hearing-impaired children are different from those of the older normal hearing children.

It is possible that the older children taken from Eguchi and Hirsh's study are, again, not comparable in age to the older hearing-impaired group in the present study. Recall that the age of the older hearing-impaired children ranged from 12.7 to 14.8 years, compared with 12 to 13 years for the normal hearing children included in Eguchi and Hirsh's study. It is possible that more dramatic differences associated with puberty occur within the age range of the hearing-impaired group.

A more logical explanation is that the differences in mean formant values become more pronounced with age as a result of bad production habits, with a greater reliance on visual cues due to the hearing impairment. It seems, for example, that in producing /i/, hearing-impaired subjects are appropriately spreading the lips (visible information) but are reluctant to front the tongue (invisible information), resulting in lower F_2 values.

The above findings justify using the hearing-impaired children to provide their own templates

rather than deriving templates from the published literature. In fact, the open-set listening data provide exactly the necessary information re the F_1/F_2 characteristics of vowels that are "correct" for these children. The difference between using self-generated normative data and published data is particularly important for the older hearing-impaired subjects.

Analysis of all auditory-visual productions of vowels

In this analysis, all vowels produced under the auditory-visual condition by the 12 hearing-impaired subjects were included. In other words, both correct and incorrect productions, as perceived by the four listeners on the open-set task, were included. Here, the purpose was to assess the accuracy of production as reflected by the open-set listening task. Table 6.2 shows the mean frequency values and standard deviations for the first and second formants for both younger and older subjects.

Table 6.2. Mean frequency values and standard deviations (SD) for the first (F₁) and second (F₂) formants for the auditory-visual condition for both younger and older subjects.

AGE GROUP	VOWEL	NUMBER OF VOWELS	MEAN OF F ₁ (Hz)	S.D. (Hz)	MEAN OF F ₂ (Hz)	S.D. (Hz)
YOUNGER	/u/	56	549.8	92.6	1597.9	512.6
	/i/	54	539.2	106.7	2875.9	352.5
	/c/	52	732.1	179.6	1585.9	365.2
	/E/	24	995.2	283.9	2161.8	377.7
OLDER	/u/	50	339.5	78.4	1661.9	357.8
	/i/	54	317.4	57.3	2204.3	214.9
	/c/	55	707.8	80.3	1457.4	167.4
	/E/	24	687.0	93.4	1748.1	206.7

Figure 6.8 shows the values of F₂ versus F₁ for the four vowels /i/, /u/, /E/, and /c/ produced by the hearing-impaired subjects. The left panel shows the formant values for the younger subjects and the right panel for the older subjects. As indicated above, all vowels, both correct and incorrect productions, were included. Note that the symbol for each vowel on the graph is based on the model presented to the child rather than his production (unlike earlier figures in which symbols denoted

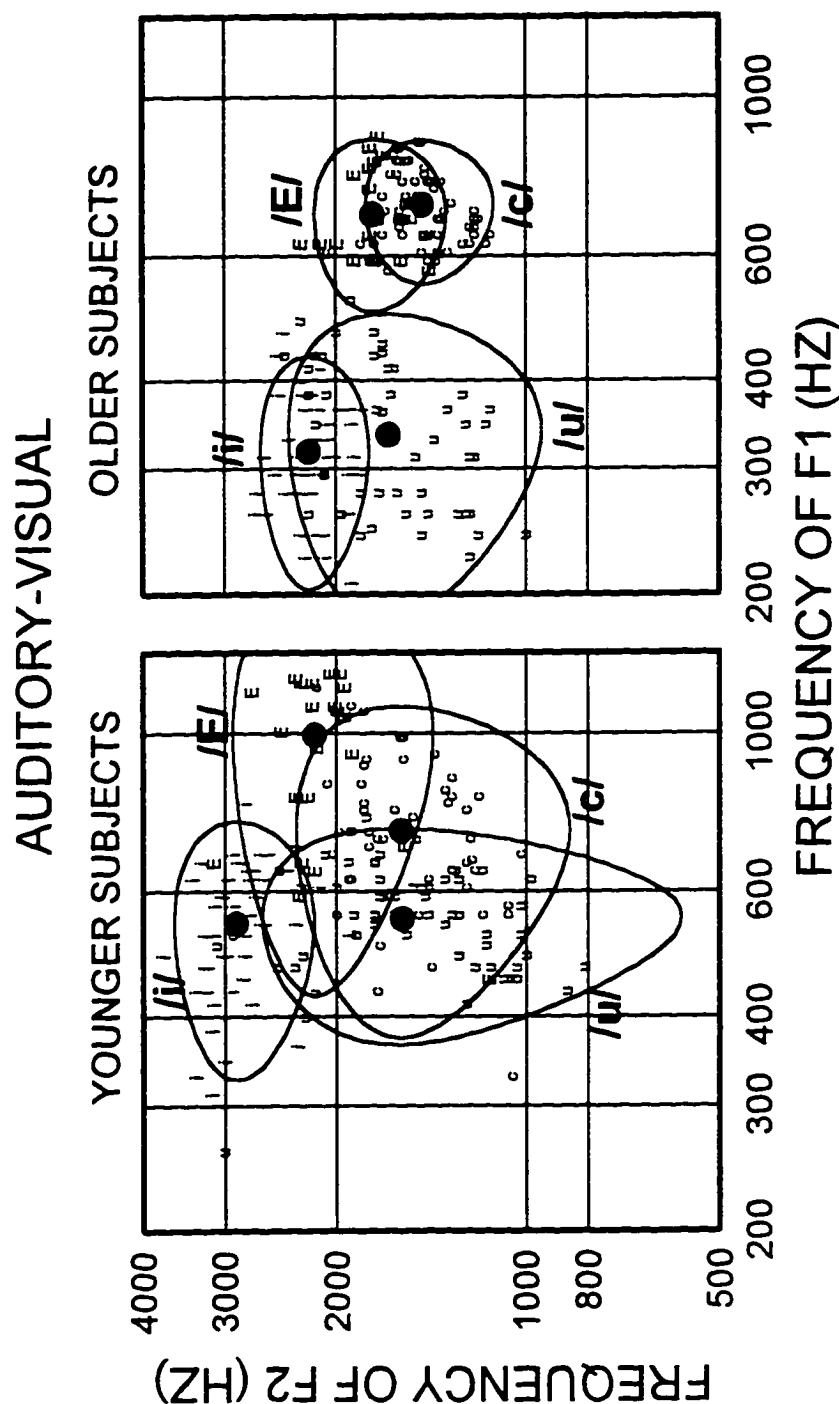


Figure 6.8. Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-visual condition. The left panel shows the formant values for the young subjects and the right panel for the older subjects.

The ovoid shapes are drawn to encompass approximately 90% of the tokens for each vowel.

The width of each shape is ± 2 standard deviations of the F1 values. The height is ± 2 standard deviations of the F2 values. The shapes are actually ellipses on linear frequency scales but are somewhat distorted here because of the use of logarithmic frequency scales. The black circles are the mean for each vowel.

correctly perceived productions). Thus, all vowels produced in response to the model /c/ will be marked as /c/. The ovoid shapes show the means ± 2 standard deviations of the first and second formant values derived from all data. Also shown are the means for each vowel. It can be seen that the distributions for individual vowels are greater, and show more overlap, for the younger subjects than for the older subjects. This observation suggests less discriminable productions by the younger subjects than by the older subjects.

Analysis of all auditory-only production of vowels

All vowels produced under the auditory-only condition by the 12 hearing-impaired subjects were included in the present analysis. The purpose was to assess the discriminability of the different vowels as well as to measure the consistency in their production. Table 6.3 shows the mean frequency values and standard deviations for the first and second formants for both the younger and older subjects.

Table 6.3. Mean frequency values and standard deviations (SD) for the first (F1) and second (F2) formants for the auditory-only condition for both younger and older subjects.

AGE GROUP	VOWEL	NUMBER OF VOWELS	MEAN OF F ₁ (Hz)	S.D.	MEAN OF F ₂ (Hz)	S.D.
YOUNGER	/u/	58	537.6	96.7	1699.1	527.9
	/i/	48	501.6	101.6	2719.9	343.0
	/c/	51	695.7	129.0	1478.8	299.9
	/E/	23	893.8	175.9	1849.0	253.5
OLDER	/u/	55	444.4	152.5	1874.7	344.1
	/i/	50	382.3	85.4	2072.9	268.8
	/c/	52	684.3	104.7	1428.3	159.6
	/E/	24	705.9	115.8	1632.7	219.0

Figure 6.9 shows the values of F₂ versus F₁ for the vowels /i/, /u/, /E/, and /c/ produced by the hearing-impaired subjects in the auditory-only condition. Also shown are the means $\pm$ 2 standard deviations of the first and second formant values for each vowel. The left panel shows the formant values for the younger subjects and the right panel for the older subjects. It is clear from the figure that both groups encountered some degree of difficulty in the auditory-only condition. There is confusion in production that

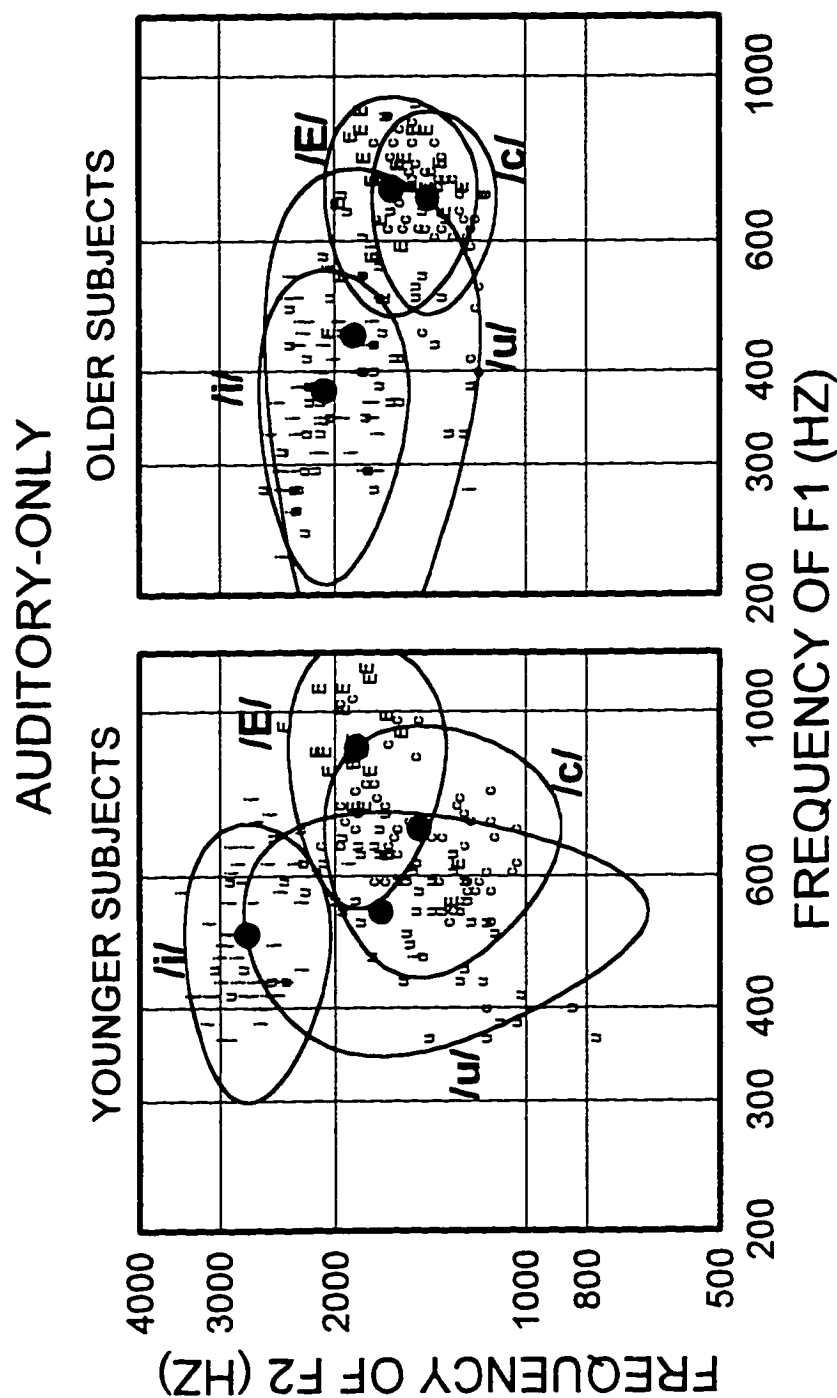


Figure 6.9. Second versus first formant frequency for four vowels produced by 12 hearing-impaired children in the auditory-only condition. The left panel shows the formant values for the young subjects and the right panel for the older subjects. The ovoid shapes are drawn to encompass approximately 90% of the tokens for each vowel. The width of each shape is ± 2 standard deviations of the F1 values. The height is ± 2 standard deviations of the F2 values. The shapes are actually ellipses on linear frequency scales but are somewhat distorted here because of the use of logarithmic frequency scales. The black circles are the mean for each vowel.

is evident from vowel overlap and unclear vowel boundaries.

Prediction of closed-set performance

To give a measure of discriminability of vowels on the basis of F_1 and F_2 and to predict scores for contrasts of vowel height and vowel place, a d' analysis (Green and Swets, 1974) was used. These predicted scores will be compared with the overall pattern of performance as reflected on the forced-choice closed-set listening task.

The d' analysis is derived from signal detection theory. As applied here, three important assumptions are made: 1) That the probability density distribution of the formant values for a given target vowel is 'normal' (i.e., Gaussian). 2) That the variances of these distributions for each pair of vowels, along a single formant dimension, are the same. 3) That a listener judgment is based on a single dimension (F_1 or F_2) and criteria that are invariant across judgments.

The value of d' , for a pair of vowels, is given by the ratio between the separation of their mean formant values and the standard deviation of the probability density distribution as shown in Figure 6.10. In the present analysis, the standard deviation was estimated

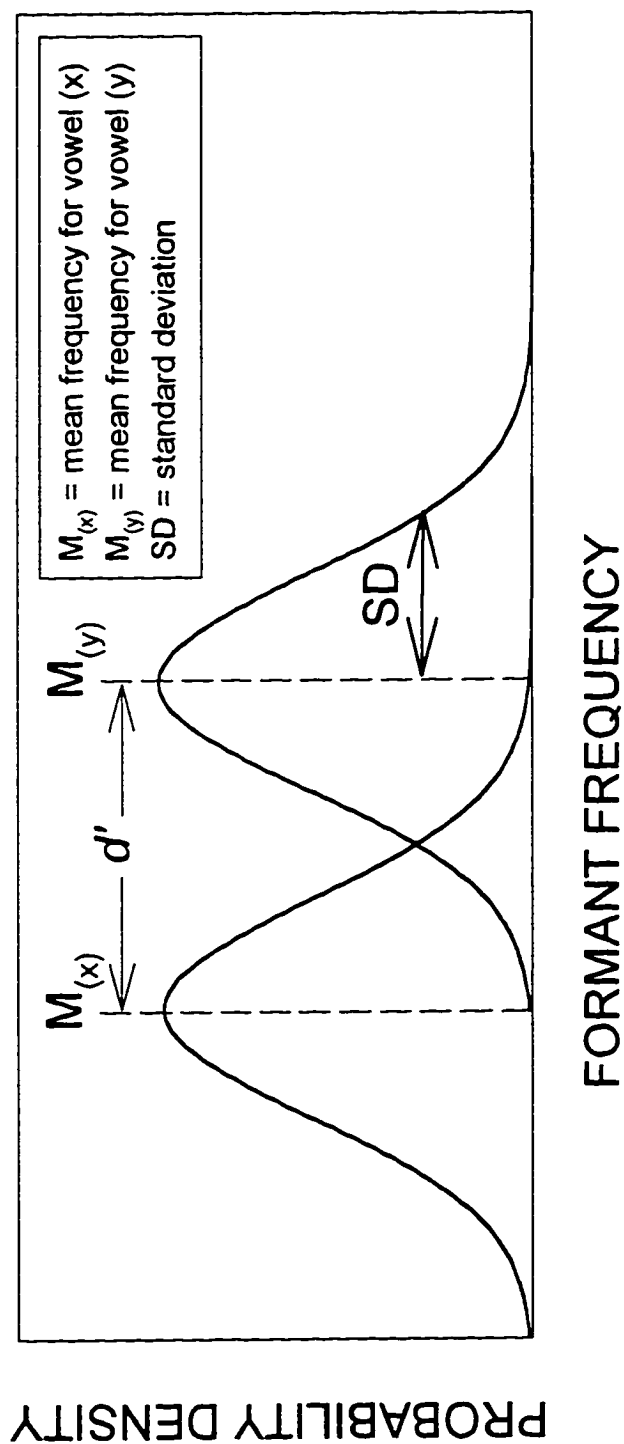


Figure 6.10. The value of d'' for a pair of vowels is determined by the ratio between the separations of their mean formant values and the standard deviation of the probability density distribution (see text).

by pooling the measured standard deviations for the formant frequencies of the two vowels:

$$SD = \sqrt{(\sigma_{(x)}^2 * (N_{(x)} - 1) + \sigma_{(y)}^2 * (N_{(y)} - 1)) / (N_{(x)} + N_{(y)} - 1)} \dots\dots\dots (2)$$

where: SD = pooled standard deviation

$\sigma_{(x)}$ = standard deviation for vowel (x)

$\sigma_{(y)}$ = standard deviation for vowel (y)

$N_{(x)}$ = total number of vowel (x)

$N_{(y)}$ = total number of vowel (y)

Thus,

$$d' = (M_{(x)} - M_{(y)}) / SD \dots\dots\dots (3)$$

where: $M_{(x)}$ = mean frequency for vowel (x)

$M_{(y)}$ = mean frequency for vowel (y)

Values of d' were calculated for differences between the means of F_1 and F_2 for the following vowels:

1. F_1 for:

/u/ and /c/

/i/ and /E/

2. F_2 for:

/u/ and /i/

/c/ and /E/

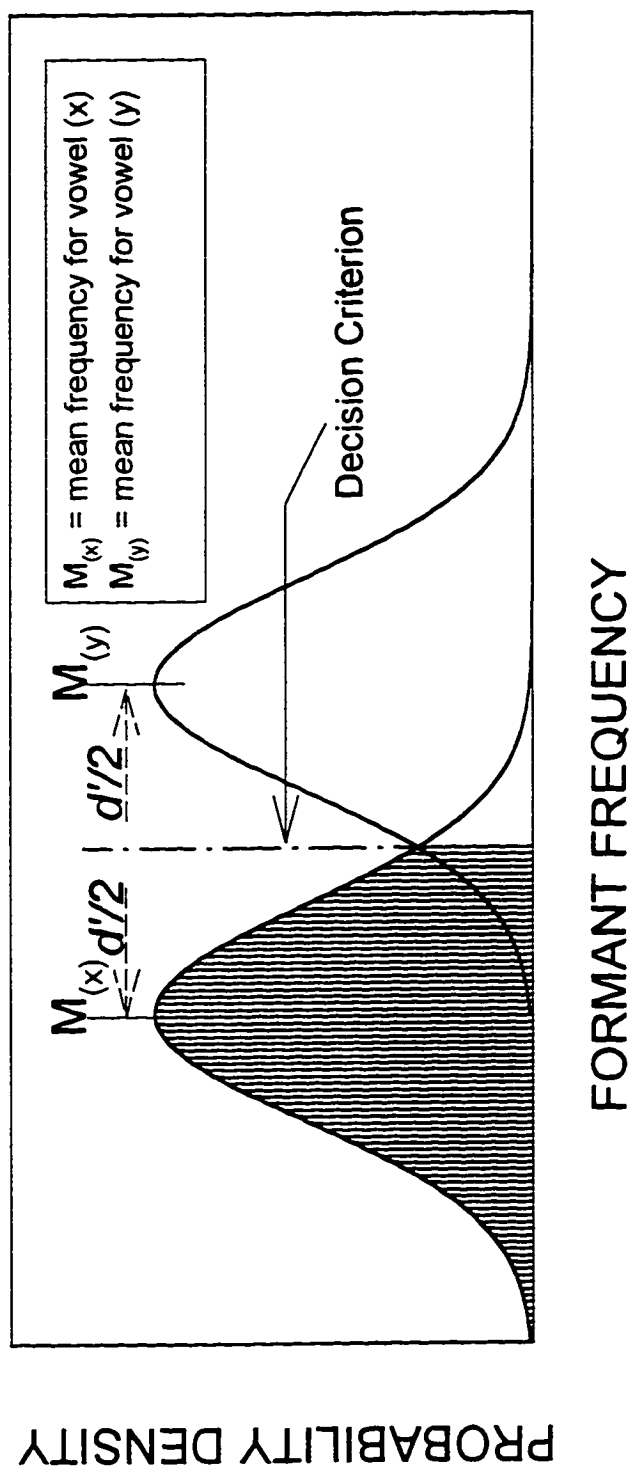


Figure 6.11. Assume that a listener identifies any formant whose frequency is above the decision criterion as belonging to vowel (y) and any formant whose frequency is below the decision criterion as belonging to vowel (x). Assume also, that the decision criterion is mid-way between the two means. Then, the probability of a correct response is given by the shaded area to the left of the decision criterion. This value can be determined from z tables.

Tables 6.4 and 6.5 show the values of d' for F_1 and F_2 for the auditory-visual and auditory-only conditions, respectively, for the younger and older groups. Also shown are the computed predicted scores in percent correct (re chance) for vowel height and vowel place.

To obtain the predicted scores, the area under the normal probability density distribution, to the left of a value $d'/2$ from the mean, was determined. This area, obtained from standard z tables, is the probability that a subject who places his decision criterion midway between the means for the two distributions will correctly identify a token (see Figure 6.11). Note that, for $d' = 0$, this value is 50% and represents the results of random guessing in a binary choice task. The data in the last two columns of tables 6.4 and 6.5 have been corrected for guessing, as were the behavioral data reported in previous chapters.

Table 6.4. *d'* values for frequencies of the second (F2) and first (F1) formants for the auditory-visual condition for both younger and older subjects. Also shown are predicted scores (% correct re chance) for Vowel Height and Vowel Place. These predicted scores are derived from the *d'* values (see text).

Age Group	x vs. y	F	Mean x	SDx	Nx	Mean y	SDy	Ny	Differ-ence pooled	SD	<i>d'</i>	Predicted (d'/2)	Predicted Mean* p(t)	Grand Mean
													Chance	
Younger	u/i	F2	1597.9	512.6	56	2875.9	352.5	54	1278.0	439.3	2.91	92.77	PLACE	
	c/E	F2	1585.9	365.2	52	2161.8	377.7	24	575.9	366.7	1.57	78.38	85.58	71.15
	u/c	F1	549.8	92.6	56	732.1	179.6	52	182.3	140.6	1.30	74.22	HEIGHT	
	i/E	F1	539.2	106.7	54	995.2	283.9	24	456.0	178.6	2.55	89.88	82.05	64.10 67.63
Older	u/i	F2	1661.9	357.8	50	2204.3	214.9	54	542.4	291.0	1.86	82.38	PLACE	
	c/E	F2	1457.4	167.4	55	1748.1	206.7	24	290.7	178.9	1.63	79.25	80.82	61.63
	u/c	F1	339.5	78.4	50	707.8	80.3	55	368.3	79.0	4.66	99.01	HEIGHT	
	i/E	F1	317.4	57.3	54	687.0	93.4	24	369.6	69.8	5.30	99.60	99.31	98.61 80.12

* Mean of the p values derived for the two vowel contrasts.

Table 6.5. d' values for frequencies of the second (F2) and first (F1) formants for the auditory-only condition for both younger and older subjects. Also shown are predicted scores (% correct re chance) for Vowel Height and Vowel Place. These predicted scores are derived from the d' values (see text).

Age Group	x vs. y	F	Mean x	SDx	Nx	Mean y	SDy	Ny	Differ-ence pooled	SD	d'	Predicted (d'/2)	Predicted Mean* p(%)	Grand Mean
													Chance	
Younger	u/i	F2	1699.1	527.9	58	2719.9	343.0	48	1020.8	451.6	2.26	87.08	PLACE	
	c/E	F2	1478.8	299.9	51	1849.0	253.5	23	370.2	284.6	1.30	74.22	80.65	61.30
	u/c	F1	537.6	96.7	58	695.7	129.0	51	158.1	112.4	1.41	75.96	HEIGHT	
	i/E	F1	501.6	101.6	48	893.8	175.9	23	392.2	129.1	3.04	93.57	84.77	69.53 65.42
Older	u/i	F2	1874.7	344.1	55	2072.9	268.8	50	198.2	309.1	0.64	62.55	PLACE	
	c/E	F2	1428.3	159.6	52	1632.7	219.0	24	204.4	179.0	1.14	71.57	67.06	34.12
	u/c	F1	444.4	152.5	55	684.3	104.7	52	239.9	130.9	1.83	81.99	HEIGHT	
	i/E	F1	382.3	85.4	50	705.9	115.8	24	323.6	95.5	3.39	95.50	88.75	77.49 55.81

* Mean of the p values derived for the two vowel contrasts.

Predicted scores for vowel height and vowel place for both input conditions are illustrated in Figure 6.12 together with the obtained IMSPAC closed-set scores (data taken from Tables 5.5 and 5.6) in the same format used to illustrate results for the behavioral data reported in Chapter 5. In the Figure, predicted and obtained vowel scores, corrected for the effect of guessing, are presented as a function of age for the auditory-only and the auditory-visual input modalities. Figure 6.13 presents the same data. Here, predicted and obtained IMSPAC closed-set scores for vowel height and vowel place are plotted together for each input condition. As can be seen, the similarities between predicted and obtained vowel scores are striking. The important finding is that even with the predicted vowel scores there is still no evidence of an improvement of auditory imitation with age. In fact, if anything, there is evidence of a slight decrease with age. This occurs mainly because of a more discriminable production along the F_2 dimension by the younger group (see Table 6.5).

Predicted scores for vowel height and vowel place were averaged together to yield one vowel score for each age group for both input conditions. The results, shown in Figure 6.14, are once again illustrated in the

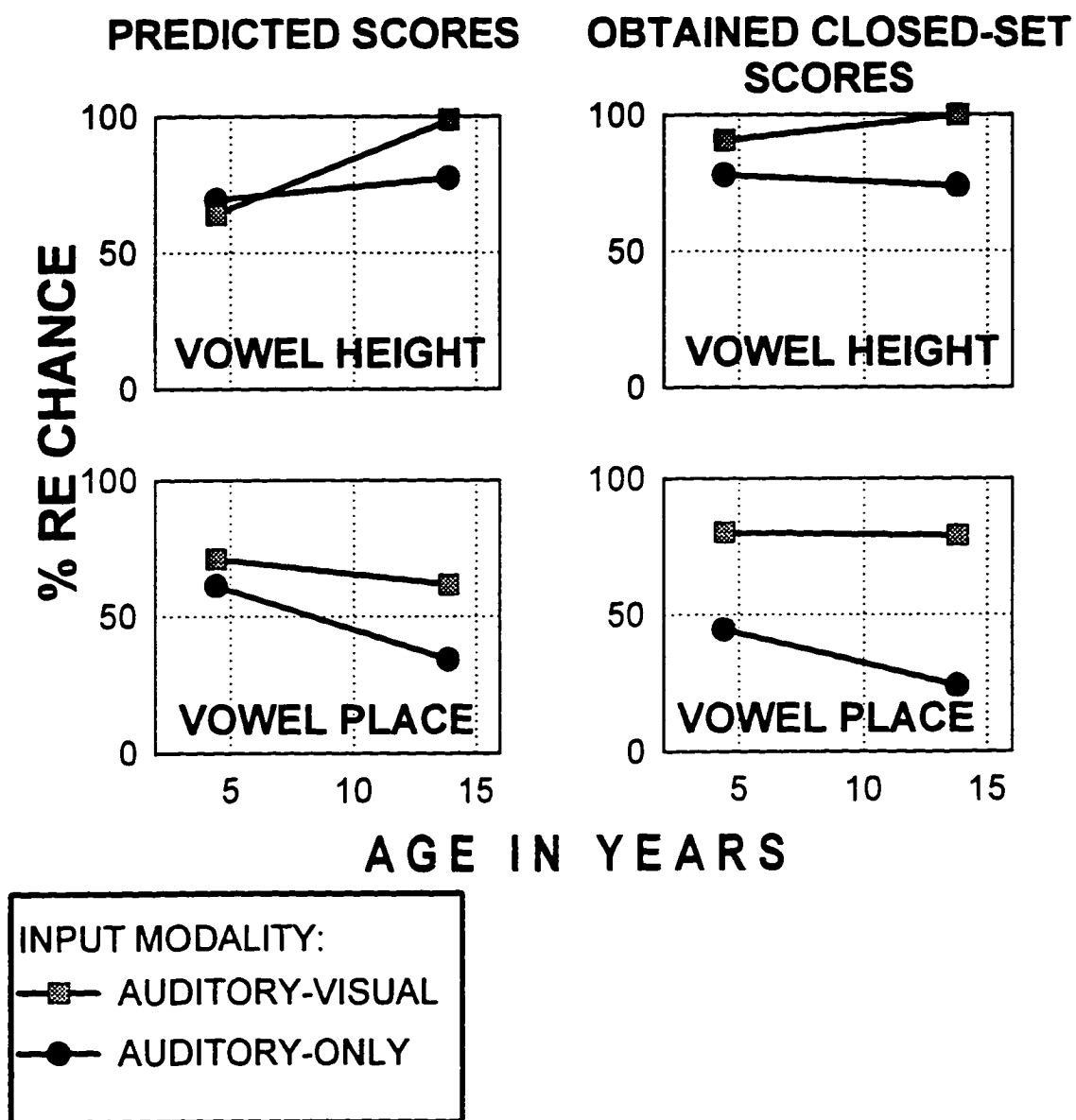


Figure 6.12. Obtained closed-set scores (% re chance) and predicted vowel scores (% re chance) for the younger and older subjects. The predicted scores were derived from d' analysis.

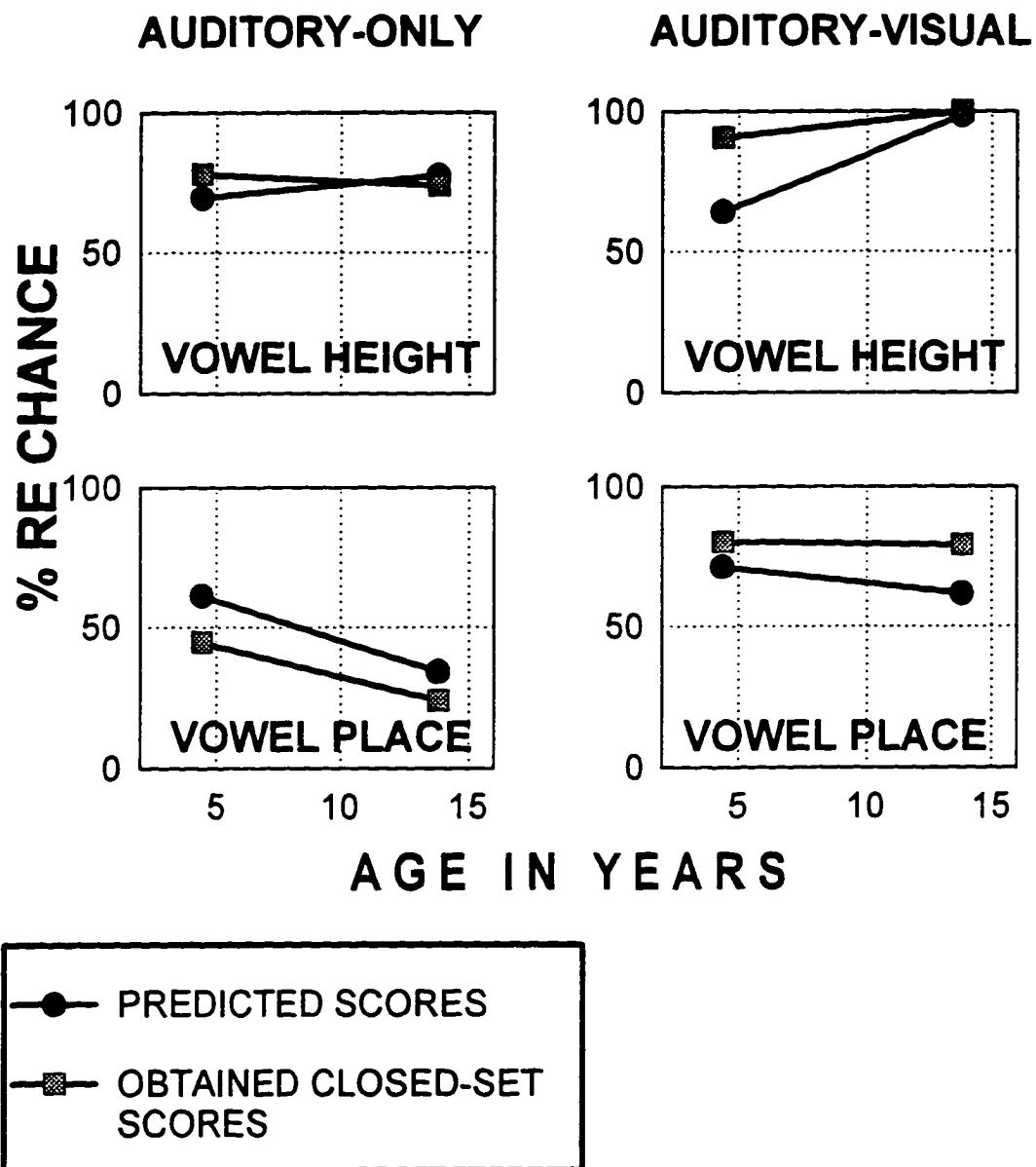


Figure 6.13. Obtained closed-set scores (% re chance) and predicted vowel scores (% re chance) as a function of age for auditory-only and auditory-visual modalities. The predicted scores were derived from d' analysis.

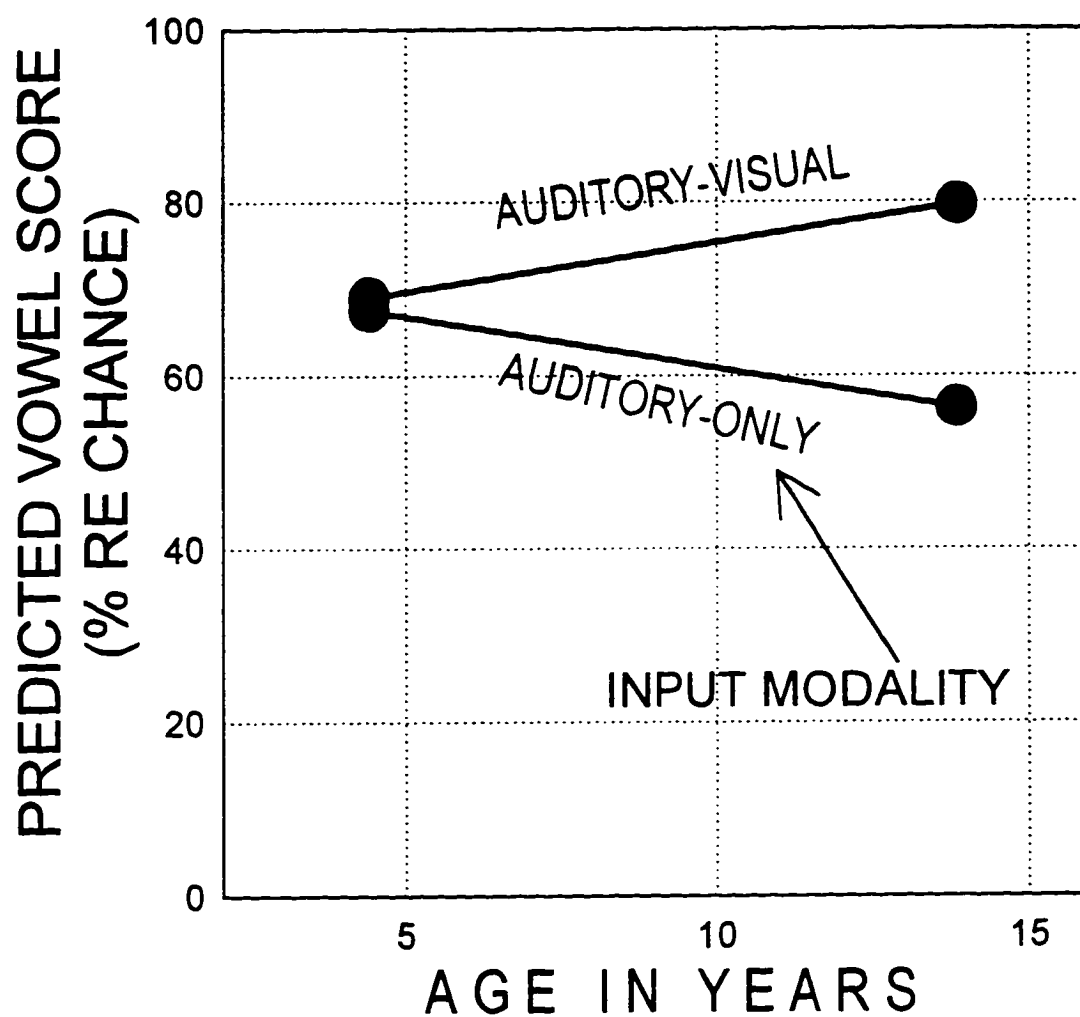


Figure 6.14. Combined predicted vowel scores for the younger and older subjects. Scores were derived from d' analysis.

same format used to illustrate results for the closed-set and open-set listening tasks (see Figures 5.1, 5.6, and 5.7). Overall there is no improvement (there is actually a drop) on the auditory-only condition with age. In the auditory-visual condition, there is an obvious increase in performance in favor of the older group. Across input modalities the younger subjects are performing at the same level, whereas there is a sharp increase in performance for the older subjects with the addition of visual input. This pattern of performance is consistent with the findings reported earlier for both the closed-set and the open-set behavioral data.

Summary of Results

1. The vowel templates derived from the production of the younger hearing-impaired subjects were similar to those derived from the data reported by both Peterson and Barney (1952) and Eguchi and Hirsh (1969) for their normal hearing children.

2. The vowel templates derived from the older hearing-impaired subjects were not comparable to those of either Peterson and Barney or Eguchi and Hirsh. In particular, the second formant frequency separation was significantly reduced compared to that for any group of normal hearing subjects.

3. When acoustic measures of F1 and F2 were used to assess vowel discriminability, the following results were obtained:

(a) There was strong evidence of an improvement of auditory-visual imitation with age. On average the older subjects' mean predicted closed-set composite vowel score based on d' analysis was about 13 percentage points higher than the score of the younger subjects.

(b) There was no evidence of an improvement of auditory imitation with age. In fact, if anything,

there was evidence of a decrement with age. The better performance by the younger group was attributable to more accurate production along the F_2 dimension (vowel place) (see Table 6.5).

(c) There was a strong effect of input modality in the older subjects. For these subjects the predicted closed-set auditory-visual scores were about 1.5 times the auditory-only scores.

(d) There was no effect of input modality in the younger subjects.

Discussion of Acoustic Findings

The vowel templates for normal hearing speakers taken from Peterson and Barney and Eguchi and Hirsh were based on the acoustic analysis of all productions of vowels regardless of their correctness. In contrast, the vowel templates for the hearing-impaired speakers included in the present study are based only on vowels correctly identified by four listeners. Although Peterson and Barney as well as Eguchi and Hirsh included in their studies an identification task, formant values for correctly identified vowels were not reported. Peterson and Barney, however, did report that when they plotted

formant values only for vowels that were correctly identified by all listeners, considerable scatter and overlap were still observed within and across vowels. Peterson and Barney, somewhat puzzled by their findings, concluded first, that the correlation is poor between formant values and phonetic (perceptual) values, and second, that the complex acoustic pattern of speech requires consideration of more than one or two factors.

Peterson and Barney's conclusions further justify using the hearing-impaired subjects to provide their own templates rather than going to the published literature. First, Peterson and Barney's children's data are not necessarily representative of the sample included in the present study. Second, as indicated earlier, a simple location in the F1/F2 space is not an adequate specification of a vowel's essential characteristics. And third, it seems that what is important is how vowels are perceived rather than whether they meet some acoustic criterion.

The acoustic measure of vowels generally supports the behavioral results, at least as reflected by the predicted vowel scores (Figures 6.12, 6.13, and 6.14). The F₂ advantage demonstrated by the younger children is consistent with their significantly higher score

than the older group's on the initial consonant place of articulation (ICP) contrast in the auditory-only condition (Figure 5.4). Similarly, the absence of acoustic evidence of an input modality effect in the younger subjects is supported by the behavioral findings of only a small effect of input modality for the younger subjects.

The principal conclusion that can be drawn from the foregoing results is that when subjects' vowel productions were evaluated in terms of F_1 and F_2 and the discriminability of production, there is still no evidence that the imitation of auditory-only input improves with age.

Chapter 7

Discussion

It will be recalled that the purpose of the present study was to investigate the effects of age and evaluation method on measured speech imitation performance in aided profoundly hearing-impaired children. The study reevaluated existing recordings of younger and older hearing-impaired children under conditions in which the subject and listener variables were more tightly controlled and with various evaluation methods. Specifically, recorded imitations of representative matched samples of younger and older prelingually, profoundly deaf children were evaluated by three methods: (a) behaviorally, using a closed-set identification task, (b) behaviorally, using an open-set recognition task, and (c) acoustically, in terms of vowel formant frequencies. Performance on the three measures was examined in relation to age (3-5 versus 13-15) and input modality (auditory-visual versus auditory-only).

The results of the study for both behavioral and acoustic measures were similar and are in keeping with the original observations of a significant effect of age on imitation of combined auditory-visual input, but the

absence of a significant effect of age on imitation of auditory-only input. These results do not support the suggestion that the earlier findings were the result of an insensitive test or inherent design issues. The use of the same group of listeners did not reduce error variance for the new IMSPAC scores. Although there was a tendency for new IMSPAC scores to be higher than earlier IMSPAC scores, differences were small. This result does not support the conclusion that the earlier findings were a function of the use of different listeners for different subjects.

When the data were analyzed in term of individual speech contrast perception, a significant age effect for auditory-only input was evident for initial consonant place of articulation. The effect, however, was in favor of the younger subjects. In fact, a closer examination of the data (see Figure 5.4 in the results chapter) reveals higher scores, although not statistically different, for the younger group on some of the other contrasts in the auditory-only condition. Examination of the audiometric characteristics of the younger and older subjects (Table 4.2) revealed a 3 dB difference in favor of the younger group (though not statistically significant). In fact the younger group had flatter hearing loss with somewhat better high frequencies

hearing. This alone can explain the tendency for the younger group to perform better than the older group. However, in the larger study (Boothroyd et al., 1996) the effect of pure tone average amounted only to 1.4% per dB. Thus a 3 dB difference would account only for about 4% change in composite score, whereas the difference for the composite auditory-only IMSPAC score found here was 8% (Figure 5.1).

It is possible, however, that the younger children actually had even better hearing than indicated by their obtained pure tone thresholds, for two reasons. First, during pure tone audiometry testing, young children may be responding behaviorally only to the stimuli above their true threshold levels. This argument has long been supported by clinicians. Boothroyd (1993), for example, stated that "Young deaf children often give behavioral thresholds that are 5 to 10 dB above their true thresholds until they have had adequate listening experience" (p. 9).

A second reason to believe that younger children may present an elevated threshold is that they are more prone to have otitis media. An undetected conductive component may result in an overestimation of the sensorineural component.

It is also possible that, by chance, the younger group had better peripheral auditory capacity which would not be reflected by their pure tone thresholds. The limitations of the pure tone average as a predictor of speech perception ability have been noted by several researchers.

Contrast analysis showed, further, what was already apparent from the composite score: that for auditory-visual input, the older children took greater advantage of the addition of visual input than the younger children. Surprisingly, however, the older children also performed better on some of the contrasts for which no visual cues are believed to have been available. For example, the older group obtained significantly higher scores than the younger group on voicing contrasts and Fo cues (see Figure 5.5 in the results chapter). Recall that one purpose of the auditory-visual testing is to demonstrate that production is not a limiting factor for the auditory imitation task. Thus, the child is given every possible input. He is encouraged to listen, speechread, and read the printed syllables if he/she can read. It is appropriate to assume that the older subjects have established better reading skills than the younger subjects. Therefore, it seems that the auditory-visual advantage demonstrated by the older group

reflects a combination of superior reading skills as well as the effects of experience and formal training on phonological and motor speech skills.

One of the most important factors that influences perceptual judgments of speech production is the listener's prior experience in listening to deaf speech. Numerous studies have shown that experienced listeners are significantly better than inexperienced listeners at recognizing utterances produced by hearing-impaired talkers (Markides, 1970; Monsen, 1978, 1983; McGarr, 1983; Boothroyd, 1985b). A solution to the problem of listener experience was offered by the use of forced-choice procedures (Monsen, 1981; Boothroyd, 1985b). It was demonstrated that forced-choice procedures provide measures of speech production that are virtually independent of listener experience (Boothroyd, 1985b). A four-alternative forced-choice format was employed in IMSPAC testing for the present study, and differences among the four listeners were examined. Out of the four listeners, two were experienced (L1 and L2) and two inexperienced (L3 and L4) in listening to deaf talkers. Results of the closed-set forced-choice task revealed no significant differences among the four listeners. The results of this study support previous findings that the

use of a forced-choice procedure to assess the speech production performance of the hearing-impaired provides a score that is not influenced by listeners' experience.

In contrast, however, small but significant differences were found among the four listeners on the open-set evaluation procedure for the auditory-visual input. For this condition the experienced listeners (L1 and L2) gave somewhat higher scores than did the inexperienced listeners (see Figure 5.15 in the results chapter). This result supports the conclusion that the effect of listener experience is task-related and is consistent with findings reported in the literature. Boothroyd's data (1985b), for example, showed that the listener effect was bigger on an open-set task than on a closed-set task. Nevertheless, the small differences between the listeners, in the present study, did not influence the basic pattern of findings regarding the effects of age, modality, and evaluation procedure.

Evaluation of the relationship between scores obtained with the closed-set and open-set tasks revealed a high correlation between the two procedures. Similar findings were reported by other investigators (e.g., Monsen, 1981; Boothroyd, 1985). The similarity of the scores for closed-set and open-set testing (as opposed to the correlation between them) was unexpected, given

that the decision process on the listener's part, is markedly different for each of the listening tasks. The closed-set procedure is a forced-choice task of phonetic contrast perception. The listener is provided with phonetic context and must make his/her decision on the basis of the characteristics of a single utterance. On each trial of IMSPAC, the decision is made along only two phonetic dimensions. For example, for the four alternatives "faw," "vaw," "saw," and "zaw," the listener must decide whether the child's utterance conveys information about either or both phonetic contrasts—initial consonant place of articulation and initial consonant voicing. The drawback of the closed-set format is that there is no way to determine which type of error has been made. In fact, the listener is forced to choose one of the alternatives even if none matches his/her perceptual judgment. For that reason measured scores are corrected for the effect of guessing. The open-set task involves recognition of phonemes in nonsense syllables. Here, a panel of listeners write down what they think the child has said. No context is provided and the listener decision is based on categorical rather than contrastive perception. The main advantage of this procedure is that it allows

monitoring of distortions within the same phoneme category. The main disadvantage, as discussed earlier, is the bias introduced by the listener's prior experience in listening to deaf speech.

In conclusion, the results of the present investigation further support the use of a closed-set procedure to assess the speech production abilities of hearing-impaired children, as it appears to provide a score that is context-free, and that is not influenced by listeners' past experience with deaf speech.

To establish the critical difference that is needed to show a significant age effect on the closed-set and open-set procedures, a two-factor repeated-measures analysis of variance was performed, separately, on the arcsine transform open-set and closed-set IMSPAC data (see Appendix I). The error term in each analysis of variance was used to estimate the critical differences for both evaluation procedures. The resulting critical differences in percentage points for a group mean of six were 4.7 for the closed-set scores, and 6.3 for the open-set score. Therefore, a difference between the older and younger groups of 4.7 and 6.3 percentage points, would be required to reveal a significant age

effect for the closed-set and open-set listening tasks, respectively.

The consistent result obtained with the behavioral measures, that the ability to imitate speech from auditory input does not change with age, is in keeping with the conclusion that auditory imitation reflects peripheral auditory speech perception capacity. The behavioral measures addressed imitation in terms of its ability to convey both contrastive and categorical information. It remains possible, however, that the acoustic productions of the older children are better in some way than those of the younger children, even if their effectiveness did not change significantly. To explore the possibility of an age-related change in the ability to imitate auditory input, which was not revealed by the behavioral measures, an acoustic analysis was performed on the children's vowel productions, in terms of F_1 and F_2 and the discriminability of their production. The pattern of the acoustic measure of vowels generally supports the behavioral results, at least as reflected by the predicted vowel scores. Specifically, there was no evidence of an improvement of auditory-only imitation with age.

Some words of caution are appropriate at this point concerning the conclusions drawn from the acoustic analysis. First, it is acknowledged that a simple location in the F_1/F_2 space is not an adequate specification of a vowel's essential characteristics. It is possible that such a simple representation is not sufficient for the acoustic description of vowels. Kent and Read (1992) suggested that a full account of the acoustic cues for vowels perception requires consideration of additional factors such as duration, fundamental frequency, and formant bandwidth and amplitude. Second, the predicted scores derived from d' analysis are based on several questionable assumptions as discussed in the previous chapter. The first is that the probability density distribution of the formant values for a given target vowel is 'normal'. The second is that the variances of these distributions for each pair of vowels, along a single formant dimension, are the same. And the last is that a listener judgment is based on a single dimension (either F_1 or F_2) and invariant criteria. Nevertheless, the analysis offers an excellent prediction of the general pattern of performance observed on the behavioral measures.

When the vowel templates derived from the production of the hearing-impaired subjects were

compared with published data from normally hearing talkers some similarities were found with the younger hearing-impaired group but not with the older group. As indicated earlier, the younger hearing-impaired subjects had in fact somewhat better hearing than the older subjects. It is also possible that the older normal hearing subjects are not representative of the older hearing-impaired subjects because of age differences. Another possibility is that the dissimilarities reflect poor judgments by listeners in the present study. For example, examination of Figure 6.7 in Chapter 6 reveals that for subjects number H81 and H106, the frequency values of F2 for some tokens of the vowel /u/ are very high. These frequency values are significantly deviant from the mean frequency values of the same vowel produced by normal hearing talkers as well as hearing-impaired fellows (e.g., H78 and H80). In fact, an interesting observation made by the advisor of the present work was that these vowel tokens sounded more like the French /y/ than English productions. Nevertheless, all of these vowels included in the older subject template were correctly identified by all four listeners.

A threat to the external validity of the study is introduced by the small sample of children included in each age group. These groups, however, were carefully matched with respect to various independent variables such as educational placement, age at onset of deafness, and degree of hearing loss (in spite of the 3 dB advantage of the younger group). Thus, to the extent that these variables affect performance on the test, the composition of the groups reduces the error variance for inter-group comparison and increases the power of the analysis. An encouraging sign in favor of this argument is the presence of a change over time in the auditory-visual condition. Ideally, a repeated-measures design would be preferable, in which change over time within subjects is one of the dependent variables.

In summary, the study investigated the effects of age and evaluation method on measured auditory speech perception performance in profoundly hearing-impaired children. The results continue to suggest that the auditory imitation of syllables does not improve between ages four and fourteen.

It will be recalled that the reason for going to such lengths to establish the presence or absence of age-related improvements of auditory-only imitations was

because the absence of an age effect carries two important implications. The first is that production measures can be used to assess peripheral auditory speech perception capacity. The second is that this capacity is determined only by the status of the peripheral system, together with the characteristics of sensory assistance, and not by maturation, experience, and training. If either of these implications were untrue, the imitation of an auditory input would be expected to improve with age.

A common criticism of an imitative approach to the assessment of auditory perception is that the results will reflect the limited production abilities of the subjects rather than their perception abilities. If this were the case, however, age-related improvements of production would be accompanied by improvements of auditory imitation. The fact that this was not observed argues against the criticism. There is no evidence that the auditory-only performance of the younger children was being limited by production difficulties because their auditory-only score was the same as that of the older children. In fact, the similarity of the auditory-only and the auditory-visual scores of the younger children argues for an alternative interpretation--namely that the production abilities of the younger

children are being limited by their peripheral auditory speech perception capacity. This is in contrast to the older subjects whose production abilities have improved beyond the limits set by their peripheral auditory speech perception capacity.

The data indicate that the younger group was taking less advantage of visual cues than the older group. In fact they were using visual cues to improve perception of only vowel place and final consonant place of articulation contrasts. These results support the notion that in young hearing-impaired children, speech production performance is driven primarily by peripheral auditory speech perception capacity. In other words, the young hearing-impaired child imitates what he/she hears without taking advantage of other available sensory inputs, such as visual cues.

In contrast to the younger group, it is clear that the older group was benefiting greatly from the addition of visual input. In fact, modality effects were present for all speech contrasts except syllable number. The auditory-visual performance of the older group presumably reflects the influence of maturation, training and experience on the development of speech production skills. Because the older subjects were both more mature and better talkers, the absence of an

auditory-only change with age was consistent with the conclusion that neither maturation nor improved production influenced the auditory imitation performance.

As already emphasized throughout this work peripheral auditory speech perception capacity is a hypothetical construct that cannot be measured directly. Performance, however, can be operationally defined and measured. We have devised a test that ought to depend on peripheral auditory capacity but is clearly influenced by other factors such as production skills and phonological knowledge. Therefore, we expect performance to improve over time, approaching some asymptote that is determined by peripheral auditory capacity. If we can be sure that the asymptote has been reached, then we have established an indirect measure of peripheral auditory speech perception capacity as defined here. The empirical data obtained in the present study provide convincing evidence that asymptotic performance has been reached by the age of four years, i.e., that auditory performance is not likely to improve significantly, at least not between the ages of four and fourteen. Therefore, we can conclude that IMSPAC provides a reasonably valid measure of speech perception capacity in profoundly aided hearing-impaired children aged four

years or more. However, since IMSPAC requires speech production, its application is limited to children who have developed a certain level of competence in production.

The idea that peripheral auditory speech perception capacity might be fixed, rather than responsive to maturation, experience, and training is potentially controversial for two reasons. First, many educators and clinicians invest considerably amounts of time and effort in auditory training and have formal and informal evidence to support improved performance over time. Second, there is converging animal and human evidence to suggest plasticity of the auditory system at the level of the brain stem and cortex. It should be emphasized, however, that the present data, and the conclusions drawn from them, are deliberately restricted to the more peripheral parts of the auditory system. These data say nothing about the benefits of training and experience in terms of higher level auditory and linguistic functions. It is important to distinguish between the generation of sensory evidence by the peripheral auditory system and the use of that evidence for higher level perception. The second part of this process should clearly benefit from training.

In summary, the absence of age-related improvements in the auditory-only imitation of syllables carries two implications: (a) Peripheral auditory speech perception capacity depends only on the status of the auditory periphery and the characteristics of a hearing aid and is not influenced by maturation, experience, and training, (b) This capacity can be inferred from measures of imitative performance.

These are the same implications that were drawn from original IMSPAC data, and the results of the present study continue to support their validity.

Clinical Implications And Future Research

The results of this study have several clinical implications. One implication arises from the findings that maturation, experience and training do not affect the peripheral auditory speech perception capacity of aided profoundly deaf children as measured by means of imitations of auditory input, but do affect production abilities as reflected by imitations of auditory-visual input. These findings support the concept of a multi-sensory approach to habilitation management of the pediatric hearing-impaired population. In other words,

the findings support the potential benefits of speechreading as a supplement to audition. If there is a ceiling on the auditory sensory evidence that can be made available by virtue of the child's physical characteristics and those of the sensory assistance, then it makes sense that any method of increasing total sensory evidence should have positive benefits. In fact, the auditory-visual performance of the older subjects support that conclusion. There is no evidence, here, to suggest that the simultaneous use of hearing and vision would, in any sense, reduce the amount of information being provided by the auditory system and sensory aid.

These findings have serious implications for planning an appropriate habilitation program for the hearing-impaired child. Knowledge of the hearing-impaired child's auditory speech perception capacity at an early age will assure the appropriate selection of sensory aids, educational placement, and intervention programs.

The findings also raise some concerns regarding the appropriateness of some habilitative approaches employed with profoundly hearing-impaired children. One is Auditory-Verbal Therapy. There is much support from auditory-verbal therapists for an auditory-based approach regardless of the degree of hearing loss (e.g.,

Ling, 1976; Ling and Milne, 1981; Estabrooks, 1994). Advocates have implied that peripheral auditory speech perception capacity can be enhanced with appropriate auditory-verbal training. No empirical data, however, have ever been provided in support of this argument. The data in the present study, of course, do not suggest that there are hearing-impaired children who cannot benefit from an auditory-verbal approach. Furthermore, the results do not imply a reduced need for listening training. In fact, the assumption of this study was that appropriate listening training had already occurred with these subjects and that it had continued during their educational experience. Moreover, the data stress the need to insure that peripheral auditory speech perception capacity is optimized by proper fitting of sensory device.

The subjects included in the data were limited to the profoundly deaf. It is suggested that the effect of age on auditory performance be tested in the future with children in different hearing-status groups and with different sensory aids (hearing aids or cochlear implants). It is further suggested that testing be done with hearing-impaired children attending different educational/habilitation programs.

Early and accurate quantification of auditory capacity is particularly crucial when determining whether a child is a suitable candidate for a cochlear implant. A cochlear implant is designed to stimulate the remaining auditory nerve fibers directly, bypassing the damaged inner ear structures. During surgery, an electrode array is inserted into the Scala Tympani of the cochlea. This procedure results in irreversible destruction of all residual hearing a child might have. Since cochlear implants became a viable option for prosthetic management, the main target population has become the very young, prelingually and preverbally deaf children. Therefore, there is a need for a valid, reliable, and sensitive speech perception test. IMSPAC, though it seems to satisfy these requirements, is applicable only to children who have developed competent production skills. In a continuing effort to meet the need for appropriate pediatric tests to provide valid information about speech perception capacity in very young children with no speech production experience, we are currently exploring the use of a video game that probes contrast perception (Boothroyd et al., 1992). In this approach, the child is required only to have the concept of object permanence and to be able to point. Another option under investigation is the measurement of

electrophysiological responses to phonetic changes (Boothroyd, 1991b).

The applicability of IMSPAC as a clinical diagnostic tool is also limited by the lengthiness and complexity of the editing and listening stages of the evaluation procedure. An on-line version of IMSPAC, where a single listener responds to the child's imitations of recorded stimulus models, is currently under investigation (Boothroyd, 1991b). The effect of number of listeners on perceptual score, however, remains to be determined.

Caution is called for in interpreting the results because of the small sample of hearing-impaired children included in each age group. It is strongly suggested that the effect of age on measured speech imitation performance be examined with a larger sample of subjects, preferably using multiple single subject design.

Valuable information would also be provided by a longitudinal study in which a group of hearing-impaired children is tested periodically from the age of three. In the early years, children could be tested with IMSPAC. Later a battery of speech perception tests with varying levels of contextual influence (e.g., phoneme

recognition in CVC words, CVC word recognition in isolation, word recognition in sentences) could be included and the relationship among scores could be explored. In addition, the relationship between speech perception and speech production should be further examined.

APPENDIX - A

Instructions for listeners in the open-set procedure.

YOU WILL BE LISTENING TO A SERIES OF IMITATIONS OF
NONSENSE WORDS SPOKEN BY HEARING-IMPAIRED CHILDREN.

EACH NONSENSE WORD CONSISTS OF TWO SEGMENTS OF
EITHER: CONSONANT-VOWEL OR VOWEL-CONSONANT.

PLEASE WRITE DOWN WHATEVER YOU THINK THE CHILD
WAS TRYING TO IMITATE. IF YOU WOULD LIKE TO LISTEN TO
AN ITEM AGAIN BEFORE YOU RESPONSE, PUSH THE LETTER 'A'
ON THE KEYBOARD.

REMEMBER THESE ARE ENGLISH-SPEAKING CHILDREN.

GUESS IF YOU ARE NOT SURE.

APPENDIX - B
OPEN SET WORD RECOGNITION

LISTENER'S NAME: _____

DATE: _____

CHILD'S ID. NUMBER: _____

ID #: _____

CONDITION: _____ FORM: _____

PLEASE PRINT**ITEM**

1	_____
2	_____
3	_____
4	_____
5	_____
6	_____
7	_____
8	_____
9	_____
10	_____
11	_____
12	_____
13	_____
14	_____
15	_____
16	_____
17	_____
18	_____
19	_____
20	_____
21	_____
22	_____
23	_____
24	_____
25	_____
26	_____
27	_____
28	_____
29	_____
30	_____
31	_____
32	_____
33	_____
34	_____
35	_____
36	_____
37	_____
38	_____
39	_____
40	_____

APPENDIX - C

Instructions for listeners in the closed-set procedure.

YOU WILL BE LISTENING TO A SERIES OF IMITATIONS OF NONSENSE WORDS SPOKEN BY HEARING-IMPAIRED CHILDREN. FOR EACH WORD YOU HEAR, YOU WILL SEE FOUR (4) RESPONSE ALTERNATIVES ON THE MONITOR. PLEASE PUSH THE NUMBER ON THE KEYBOARD CORRESPONDING TO THE WORD YOU HEARD FOR EACH ITEM. IF YOU WOULD LIKE TO LISTEN TO AN ITEM AGAIN BEFORE YOU RESPONSE, PUSH THE LETTER 'A' ON THE KEYBOARD.

IT IS POSSIBLE THAT SOME ITEMS MAY NOT SOUND LIKE THE ALTERNATIVES DISPLAYED ON THE MONITOR. ALWAYS GUESS IF YOU ARE UNSURE. IT IS IMPORTANT THAT YOU RESPONSE FOR EACH ITEM.

HINTS: UPPERCASE "TH" IS VOICED AS IN THE WORD
 "THESE".
 LOWERCASE "th" IS UNVOICED AS IN THE WORD
 "THINK".
 AH? HAS A RISING INTONATION AS IN A QUESTION.
 AH! HAS A FALLING INTONATION AS IN A
 STATEMENT.

APPENDIX - D

Phonologically significant contrasts represented in the IMSPAC test.

Contrast	Abbreviation
Syllable Number	SYL
Intonation	INT
Vowel Height	VH
Vowel Place	VP
Initial Consonant Voicing	ICV
Final Consonant Voicing	FCV
Initial Consonant Continuance	ICC
Final Consonant Continuance	FCC
Initial Consonant Place	ICP
Final Consonant Place	FCP

Appendix - E

An example for a closed-set scoring form for one child. The listeners' responses were analyzed for each subject separately as the proportion of correct responses for each contrast for the auditory and the auditory-visual input modalities. For each listener, a composite score, the mean of ten contrasts, was computed. The Composite scores averaged across the four listeners represents one data point for each subject.

LISTENING TASK:CLOSED-SET

CHILD'S NAME:TD
TEST: IMSPAC

CONDITION: AUDITORY-ONLY

LISTENER	SPEECH CONTRAST*										MEAN OF 10 CONTRASTS
	SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	
AH (L1)	87.5	75.0	75.0	75.0	75.0	50.0	75.0	75.0	37.5	50.0	67.50
RW (L2)	100.0	75.0	87.5	75.0	62.5	50.0	87.5	87.5	25.0	37.5	68.75
EG (L3)	100.0	75.0	87.5	75.0	37.5	50.0	62.5	75.0	50.0	50.0	66.25
TC (L4)	100.0	75.0	75.0	75.0	37.5	50.0	87.5	62.5	62.5	37.5	66.25
MEAN	96.9	75.0	81.3	75.0	53.1	50.0	78.1	75.0	43.8	43.8	67.19

Average of 4 composite scores=
1 data point per child

*The key to the contrasts' abbreviations can be found in Appendix D.

APPENDIX - F

CLOSED-SET SCORES:
 Listeners' individual and composite speech contrast scores for the 12 hearing-impaired subjects for the auditory-only and the auditory-visual input modalities.

LISTENER: AH (L1)
 CONDITION: AUDITORY-ONLY

SUBJECT	AGE GROUP	SPEECH CONTRAST*											
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN	
H21	YOUNGER	100.0	100.0	100.0	87.5	75.0	50.0	75.0	62.0	75.0	37.5	76.20	
H44	YOUNGER	100.0	50.0	75.0	62.0	62.0	37.5	75.0	75.0	75.0	37.5	64.90	
H46	YOUNGER	100.0	50.0	87.5	75.0	87.5	75.0	75.0	87.5	62.0	62.0	76.15	
H47	YOUNGER	100.0	62.0	87.5	75.0	100.0	75.0	62.0	75.0	62.0	62.0	76.05	
H66	YOUNGER	100.0	62.0	87.5	75.0	75.0	75.0	62.0	87.5	62.0	50.0	73.60	
H70	YOUNGER	100.0	87.5	100.0	75.0	62.0	62.0	50.0	50.0	87.5	50.0	72.40	
H23	OLDER	100.0	62.5	37.5	50.0	75.0	50.0	50.0	87.5	62.5	50.0	62.50	
H78	OLDER	100.0	75.0	100.0	37.5	62.5	62.5	87.5	62.5	37.5	37.5	66.25	
H80	OLDER	100.0	87.5	100.0	50.0	50.0	87.5	50.0	37.5	25.0	62.5	65.00	
H81	OLDER	87.5	75.0	75.0	75.0	75.0	50.0	75.0	75.0	37.5	50.0	67.50	
H84	OLDER	100.0	50.0	87.5	75.0	75.0	62.5	87.5	50.0	87.5	62.5	73.75	
H106	OLDER	100.0	50.0	100.0	62.5	50.0	75.0	75.0	75.0	37.5	37.5	66.25	

* The key for the speech contrasts can be found in Appendix D.

LISTENER: AH (L1)
CONDITION: AUDITORY-VISUAL

SUBJECT	AGE GROUP	SPEECH CONTRAST											
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN	
H21	YOUNGER	100.0	75.0	100.0	100.0	87.5	62.0	87.5	75.0	87.5	87.5	86.20	
H44	YOUNGER	100.0	25.0	100.0	87.5	75.0	75.0	37.5	62.0	62.0	62.0	68.60	
H46	YOUNGER	100.0	62.0	75.0	87.5	75.0	62.0	87.5	100.0	87.5	75.0	81.15	
H47	YOUNGER	100.0	75.0	100.0	87.5	50.0	50.0	50.0	62.0	87.5	50.0	71.20	
H66	YOUNGER	100.0	100.0	100.0	87.5	37.5	75.0	62.0	87.5	87.5	87.5	82.45	
H70	YOUNGER	100.0	87.5	100.0	100.0	75.0	50.0	50.0	50.0	50.0	62.0	72.45	
H23	OLDER	100.0	75.0	100.0	87.5	100.0	87.5	100.0	87.5	62.5	50.0	85.00	
H78	OLDER	87.5	87.5	100.0	100.0	100.0	87.5	100.0	100.0	62.5	100.0	92.50	
H80	OLDER	100.0	87.5	100.0	100.0	87.5	100.0	100.0	100.0	87.5	100.0	96.25	
H81	OLDER	100.0	62.5	100.0	75.0	100.0	87.5	100.0	87.5	87.5	87.5	88.75	
H84	OLDER	100.0	87.5	100.0	100.0	100.0	75.0	100.0	100.0	87.5	87.5	93.75	
H106	OLDER	100.0	75.0	100.0	100.0	87.5	87.5	87.5	87.5	87.5	50.0	86.25	

LISTENER: RW (L2)
CONDITION: AUDITORY-ONLY

SUBJECT	AGE GROUP	SPEECH CONTRAST												
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN		
H21	YOUNGER	100.0	87.5	100.0	100.0	75.0	62.0	75.0	62.0	75.0	62.0	79.9		
H44	YOUNGER	100.0	50.0	62.0	62.0	50.0	62.0	75.0	62.0	75.0	62.0	66.0		
H46	YOUNGER	100.0	62.0	87.5	87.5	37.5	62.0	62.0	75.0	50.0	62.0	68.6		
H47	YOUNGER	100.0	62.0	87.5	50.0	87.5	62.0	50.0	62.0	62.0	50.0	67.3		
H66	YOUNGER	100.0	75.0	100.0	75.0	75.0	75.0	50.0	87.5	75.0	50.0	76.3		
H70	YOUNGER	100.0	100.0	100.0	75.0	62.0	62.0	50.0	75.0	87.5	50.0	76.2		
H23	OLDER	100.0	50.0	75.0	50.0	62.5	62.5	62.5	75.0	62.5	50.0	65.0		
H78	OLDER	100.0	37.5	100.0	50.0	37.5	87.5	87.5	62.5	37.5	37.5	63.8		
H80	OLDER	100.0	75.0	87.5	75.0	75.0	87.5	50.0	50.0	50.0	62.5	71.3		
H81	OLDER	100.0	75.0	87.5	75.0	62.5	50.0	87.5	87.5	25.0	37.5	68.8		
H84	OLDER	100.0	75.0	75.0	75.0	75.0	62.5	62.5	37.5	87.5	50.0	70.0		
H106	OLDER	100.0	75.0	100.0	75.0	62.5	87.5	75.0	62.5	75.0	50.0	76.3		

LISTENER: RW (L2)
CONDITION: AUDITORY-VISUAL

SUBJECT	AGE GROUP	SPEECH CONTRAST												
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN		
H21	YOUNGER	100.0	75.0	100.0	87.5	100.0	62.0	100.0	75.0	100.0	87.5	88.7		
H44	YOUNGER	75.0	25.0	100.0	87.5	75.0	75.0	62.0	75.0	62.0	37.5	67.4		
H46	YOUNGER	100.0	50.0	75.0	75.0	87.5	75.0	75.0	100.0	87.5	75.0	80.0		
H47	YOUNGER	100.0	62.0	100.0	75.0	62.0	62.0	50.0	75.0	87.5	62.0	73.6		
H66	YOUNGER	100.0	100.0	100.0	100.0	50.0	75.0	75.0	100.0	100.0	75.0	87.5		
H70	YOUNGER	100.0	100.0	100.0	100.0	62.0	37.5	50.0	50.0	75.0	75.0	75.0		
H23	OLDER	100.0	87.5	100.0	87.5	87.5	100.0	87.5	87.5	87.5	75.0	90.0		
H78	OLDER	100.0	87.5	100.0	100.0	87.5	100.0	100.0	100.0	87.5	100.0	96.3		
H80	OLDER	100.0	87.5	100.0	87.5	87.5	100.0	100.0	100.0	87.5	100.0	95.0		
H81	OLDER	100.0	75.0	100.0	75.0	100.0	87.5	100.0	87.5	87.5	100.0	91.3		
H84	OLDER	100.0	87.5	100.0	87.5	100.0	87.5	100.0	100.0	100.0	87.5	95.0		
H106	OLDER	100.0	87.5	100.0	100.0	100.0	100.0	87.5	87.5	87.5	75.0	92.5		

LISTENER: EG (L3)
CONDITION: AUDITORY-ONLY

SUBJECT	AGE GROUP	SPEECH CONTRAST											
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN	
H21	YOUNGER	100.0	87.5	100.0	75.0	75.0	62.0	87.5	37.5	62.0	62.0	74.9	
H44	YOUNGER	87.5	50.0	62.0	50.0	50.0	37.5	75.0	75.0	50.0	62.0	59.9	
H46	YOUNGER	100.0	50.0	75.0	75.0	50.0	50.0	75.0	75.0	62.0	87.5	70.0	
H47	YOUNGER	100.0	62.0	100.0	50.0	87.5	62.0	37.5	75.0	75.0	62.0	71.1	
H66	YOUNGER	100.0	62.0	100.0	75.0	75.0	87.5	25.0	62.0	75.0	62.0	72.4	
H70	YOUNGER	100.0	100.0	100.0	62.0	62.0	62.0	50.0	50.0	87.5	50.0	72.4	
H23	OLDER	100.0	62.5	75.0	50.0	50.0	75.0	62.5	75.0	50.0	62.5	66.3	
H78	OLDER	87.5	37.5	87.5	50.0	50.0	87.5	75.0	62.5	62.5	50.0	65.0	
H80	OLDER	100.0	75.0	87.5	62.5	75.0	87.5	37.5	37.5	62.5	62.5	68.8	
H81	OLDER	100.0	75.0	87.5	75.0	37.5	50.0	62.5	75.0	50.0	50.0	66.3	
H84	OLDER	87.5	75.0	87.5	75.0	75.0	75.0	50.0	50.0	62.5	50.0	68.8	
H106	OLDER	100.0	75.0	100.0	37.5	50.0	75.0	62.5	62.5	62.5	62.5	68.8	

LISTENER: EG (L3)
CONDITION: AUDITORY-VISUAL

SUBJECT	AGE GROUP	SPEECH CONTRAST												
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN		
H21	YOUNGER	100.0	87.5	87.5	75.0	100.0	50.0	87.5	75.0	75.0	87.5	82.5		
H44	YOUNGER	100.0	37.5	87.5	75.0	75.0	75.0	87.5	62.0	50.0	62.0	71.2		
H46	YOUNGER	100.0	50.0	87.5	75.0	87.5	75.0	87.5	87.5	87.5	75.0	81.3		
H47	YOUNGER	100.0	37.5	100.0	75.0	75.0	62.0	62.0	75.0	87.5	75.0	74.9		
H66	YOUNGER	100.0	100.0	100.0	100.0	50.0	87.5	50.0	100.0	87.5	62.0	83.7		
H70	YOUNGER	87.5	87.5	100.0	100.0	62.0	50.0	62.0	62.0	87.5	62.0	76.1		
H23	OLDER	87.5	87.5	100.0	100.0	75.0	75.0	87.5	87.5	62.5	75.0	83.8		
H78	OLDER	100.0	87.5	100.0	87.5	87.5	87.5	87.5	87.5	87.5	75.0	88.8		
H80	OLDER	100.0	100.0	100.0	87.5	75.0	100.0	100.0	100.0	75.0	100.0	93.8		
H81	OLDER	100.0	75.0	100.0	75.0	75.0	87.5	100.0	75.0	87.5	87.5	86.3		
H84	OLDER	100.0	87.5	100.0	100.0	100.0	87.5	100.0	100.0	100.0	87.5	96.3		
H106	OLDER	87.5	75.0	100.0	50.0	100.0	100.0	87.5	87.5	87.5	75.0	85.0		

LISTENER: TC (L4)
CONDITION: AUDITORY-ONLY

SUBJECT	AGE GROUP	SPEECH CONTRAST										
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN
H21	YOUNGER	100.0	87.5	87.5	62.0	62.0	75.0	62.0	62.0	75.0	75.0	74.8
H44	YOUNGER	100.0	50.0	87.5	50.0	37.5	62.0	75.0	75.0	75.0	62.0	67.4
H46	YOUNGER	100.0	50.0	62.0	87.5	62.0	50.0	62.0	75.0	37.5	50.0	63.6
H47	YOUNGER	100.0	62.0	87.5	75.0	75.0	100.0	62.0	50.0	62.0	37.5	71.1
H66	YOUNGER	100.0	75.0	100.0	87.5	87.5	75.0	62.0	75.0	62.0	37.5	76.2
H70	YOUNGER	100.0	100.0	100.0	87.5	50.0	62.0	37.5	75.0	75.0	37.5	72.5
H23	OLDER	100.0	50.0	75.0	50.0	50.0	37.5	75.0	75.0	37.5	37.5	58.8
H78	OLDER	100.0	50.0	100.0	50.0	37.5	87.5	100.0	75.0	25.0	25.0	65.0
H80	OLDER	100.0	75.0	100.0	62.5	62.5	87.5	62.5	50.0	37.5	75.0	71.3
H81	OLDER	100.0	75.0	75.0	75.0	37.5	50.0	87.5	62.5	62.5	37.5	66.3
H84	OLDER	100.0	50.0	87.5	75.0	75.0	37.5	62.5	50.0	75.0	75.0	68.8
H106	OLDER	100.0	75.0	100.0	75.0	50.0	87.5	62.5	62.5	50.0	37.5	70.0

LISTENER: TC (L4)
CONDITION: AUDITORY-VISUAL

SUBJECT	AGE GROUP	SPEECH CONTRAST										
		SYL	INT	VLH	VLP	ICV	FCV	ICC	FCC	ICP	FCP	MEAN
H21	YOUNGER	100.0	87.5	100.0	100.0	87.5	75.0	100.0	75.0	100.0	87.5	91.3
H44	YOUNGER	87.5	37.5	87.5	100.0	75.0	62.5	12.5	75.0	50.0	75.0	66.3
H46	YOUNGER	100.0	50.0	87.5	100.0	75.0	50.0	100.0	87.5	75.0	75.0	80.0
H47	YOUNGER	100.0	62.0	100.0	87.5	50.0	62.0	50.0	75.0	75.0	87.5	74.9
H66	YOUNGER	100.0	87.5	100.0	100.0	50.0	75.0	75.0	100.0	100.0	75.0	86.3
H70	YOUNGER	100.0	87.5	100.0	100.0	62.0	37.5	50.0	62.0	50.0	75.0	72.4
H23	OLDER	100.0	75.0	100.0	87.5	87.5	100.0	100.0	87.5	75.0	75.0	88.8
H78	OLDER	100.0	62.5	100.0	100.0	87.5	75.0	100.0	100.0	37.5	75.0	83.8
H80	OLDER	100.0	87.5	100.0	87.5	87.5	100.0	100.0	100.0	75.0	87.5	92.5
H81	OLDER	100.0	75.0	100.0	75.0	100.0	75.0	100.0	100.0	87.5	87.5	90.0
H84	OLDER	100.0	50.0	100.0	100.0	100.0	75.0	100.0	100.0	100.0	87.5	91.3
H106	OLDER	100.0	87.5	100.0	100.0	100.0	62.5	87.5	87.5	100.0	87.5	91.3

APPENDIX - G

A two-factor analysis of variance: effects of AGE (A) and INPUT MODALITY (M) on the original closed-set scores (first table), and the new closed-set scores (second table) of the 12 hearing-impaired subjects.

Effect	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
Age	1	210.9442*	10	7.95697	7.5453	0.020588
Modality	1	778.7400	10	5.50480	41.4657	0.000000
AxM	1	181.4814	10	5.50480	32.9679	0.000187

*Significant effects are highlighted

Effect	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
Age	1	114.8440*	10	2.22159	5.1681	0.046306
Modality	1	813.8639	10	5.89959	37.9525	0.000000
AxM	1	269.5924	10	5.89959	45.6968	0.000050

*Significant effects are highlighted

Analysis is based on normalized arcsine transformed data:

$$X_{tr} = 100 * (\text{ARCSIN}(\text{SQRT}(X/100)) / \text{ARCSIN}(1)) \dots \dots \dots (4)$$

Where: X_{tr} = arcsine transformed score

X = untransformed score in %

Using this transformation a score of 100% remains 100%, a score of 50% remains 50%, and a score of 0% remains 0%.

APPENDIX - H

OPEN-SET SCORES:

Listeners' phoneme recognition scores (in %) for the 12 hearing-impaired subjects, as well as a composite score averaged across the four listeners. The first table shows data for the auditory-only condition, and the second table for the auditory-visual condition.

CONDITION: AUDITORY-ONLY

L I S T E N E R S

SUBJECT	AGE GROUP	AH (L1)	RW (L2)	EG (L3)	TC (L4)	MEAN
H21	YOUNGER	54.69	54.69	46.88	45.31	50.39
H44	YOUNGER	23.44	17.19	15.63	17.19	18.36
H46	YOUNGER	18.75	29.69	9.38	10.94	17.19
H47	YOUNGER	35.94	34.38	39.06	39.1	37.11
H66	YOUNGER	40.63	40.63	40.62	42.19	41.02
H70	YOUNGER	48.44	51.56	50.00	48.44	49.61
H23	OLDER	18.75	17.19	17.19	12.5	16.41
H78	OLDER	23.44	32.81	17.19	23.44	24.22
H80	OLDER	28.12	35.94	26.56	28.13	29.69
H81	OLDER	32.81	39.06	32.81	35.94	35.16
H84	OLDER	35.94	35.94	32.81	35.94	35.16
H106	OLDER	28.13	26.56	17.19	17.19	22.27

CONDITION: AUDITORY-VISUAL

L I S T E N E R S

SUBJECT	AGE GROUP	AH (L1)	RW (L2)	EG (L3)	TC (L4)	MEAN
H21	YOUNGER	67.19	68.75	56.25	56.25	62.11
H44	YOUNGER	42.19	43.75	26.56	26.56	34.76
H46	YOUNGER	37.5	46.87	31.25	37.5	38.28
H47	YOUNGER	48.43	48.43	45.44	37.5	44.95
H66	YOUNGER	56.25	56.25	50	53.13	53.91
H70	YOUNGER	54.69	57.81	57.81	53.13	55.86
H23	OLDER	48.44	62.5	46.88	48.44	51.56
H78	OLDER	70.31	78.13	53.13	51.56	63.28
H80	OLDER	89.06	75.00	78.13	78.13	80.08
H81	OLDER	73.44	76.56	64.06	54.69	67.19
H84	OLDER	81.25	90.63	73.44	84.38	82.43
H106	OLDER	60.94	64.06	60.94	57.81	60.94

APPENDIX - I

A two-factor analysis of variance: effects of AGE (A) and INPUT MODALITY (M) on the Open-set scores (first table), and the closed-set scores (second table) of the 12 hearing-impaired subjects.

Effect	df	MS	df	MS	F	p-level
Effect	Effect	Effect	Error	Error		
Age	1	79.563	10	15.2363	0.6904	0.425411
Modality	1	1932.828	10	10.4677	84.6462	0.000000*
AxM	1	506.19	10	10.4677	48.3572	0.000039
*Significant effects are highlighted						

Effect	df	MS	df	MS	F	p-level
Effect	Effect	Effect	Error	Error		
Age	1	114.8440*	10	2.22159	5.1681	0.046306
Modality	1	813.8639	10	5.89959	37.9525	0.000000
AxM	1	269.5924	10	5.89959	45.6968	0.000050
*Significant effects are highlighted						

Bibliography

- Ary, D., Cheser-Jacob, L., Razavieh, A. (1985). Validity and reliability. In: Introduction to research in education. Holt, Rinehart & Winston. 213-243.
- Binnie, C.A., Danilooff, R.G., & Buckingham, H.W. (1982). Phonetic disintegration in a five-year-old following sudden hearing loss. Journal of Speech and Hearing Disorders. 47:181-189.
- Boothroyd, A. (1969). Distribution of hearing levels in the student population of Clarke School for the Deaf. Sensory Aids Research Project #3. Northampton, MA.
- Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. Journal of Speech and Hearing Research. 27:134-144.
- Boothroyd, A. (1985a). Residual hearing and the problem of carry-over in the speech of the deaf. ASHA Reports. 15:8-14.
- Boothroyd, A. (1985b). Evaluation of speech production of the hearing impaired: Some benefits of forced-choice testing. Journal of Speech and Hearing Research. 28:185-196.
- Boothroyd, A. (1988). Perception of speech pattern contrasts from auditory presentation of voice fundamental frequency. Ear and Hearing. Vol. 9, No.6:313-321.
- Boothroyd, A. (1991a). Assessment of speech perception capacity in profoundly deaf children. American Journal of Otology. 12(suppl):67-72.
- Boothroyd, A. (1991b). Speech perception measures and their role in the evaluation of hearing aid performance in a pediatric population. In: Feigin, J.A. and Stelmachowicz, P.G. (Eds.). Pediatric amplification. Omaha, NE: Boys Town National Research Hospital. 77-91.

Boothroyd, A., Hanin, L., Yeung, E., and Chen, Q. (1992). Video-game for speech perception testing and training of young hearing-impaired children. Proceeding of the Johns Hopkins national search for computing applications to assist persons with disabilities.

Boothroyd, A. (1993). Speech perception, sensorineural hearing loss, and hearing aids. In G. Studebaker and I. Hochberg (Eds.). Acoustical factors affecting hearing aid performance. Allyn and Bacon, Boston. 277-299.

Boothroyd, A. (1994). Speech perception and production in hearing-impaired children. Paper presented at Vanderbilt International Meeting on Pediatric Amplification. Palm Springs, CA.

Boothroyd, A. (1995). Speech perception tests and hearing-impaired children. In: Plant, G. and Spens, K.E. (Eds.). Profound deafness and speech communication. Whurr Publishers Ltd., London. 345-371.

Boothroyd, A., and Eran, O. (1994). Auditory speech perception capacity of child implant users expressed as equivalent hearing loss. Volta Review. 96, No.5:151-167.

Boothroyd, A., Eran, O., and Hanin, L. (1995). Speech perception and production in children with hearing impairment. In: Bess, F.H. (Ed.). Amplification for children with auditory deficits. Bill Wilkerson Center Press. (In press).

Boothroyd, A., and Lambert, L. (1980). Influence of age and hearing loss on speech intelligibility. Sensory Aids Research Project Report No.33. Northhampton, MA: The Clark School for the Deaf.

Boothroyd, A., Springer, N., Smith, L., Schulman, J. (1988). Compression amplification and profound deafness. Journal of Speech and Hearing Research. 33:362-376.

Brownlee, V.A. (1965). Statistical theory and methodology in science and engineering. 2nd Ed. Wiley, New York.

Caplan-Hack, Z., and Erber, N.P. (1982). Auditory, visual, and auditory-visual perception of vowels by hearing-impaired children. Journal of Speech and Hearing Research. 25:100-107.

Cramer, K.D., and Erber, N.P. (1974). A spondee recognition test for young hearing impaired children. Journal of Speech and Hearing Disorders. 39:304-311.

Eguchi, S., and Hirsh, I.J. (1969). Development of speech sounds in children. Acta Oto-Laryngologica. Suppl. 257.

Elliot, L.L., Katz, D.R. (1980). Northwestern University, Children's Perception of Speech (NU-CHIPS). St. Louis:Auditec.

Elkins, E. (1984). Recommendation for speech recognition research. ASHA Reports. 14:77-79.

Eran, O. & Boothroyd, A. (1993). An actuarial approach to cochlear implant candidacy in deaf children. 1: Development of test instruments. Journal of the Acoustical Society of America, 4, 2336.

Erber, N.P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. Journal of Speech and Hearing Research. 15:413-422.

Erber, N.P. (1974). Pure-tone threshold and word recognition abilities of hearing impaired children. Journal of Speech and Hearing research. 17:194-202.

Erber, N.P. (1979). An approach to evaluating auditory speech perception ability. The Volta Review. 81:16-25.

Erber, N.P. (1980). Use of the Auditory Numbers Test to evaluate speech perception abilities of hearing-impaired children.

Journal of Speech and Hearing Disorders.
45:527-532.

Erber, N.P. (1981). Speech perception by hearing impaired children. In: Bess, F.H., Freeman, B.A., and Sinclair, J.S. (Eds.). Amplification in Education. Washington, DC: A.G. Bell Association for the Deaf. 69-88.

Erber, N.P. (1982). Speech Perception. In: Erber, N.P. Auditory Training. Washington, DC: A.G. Bell Association for the Deaf. 11-28.

Erber, N.P. (1982). Screening auditory abilities. In: Erber, N.P. Auditory Training. Washington, DC: A.G. Bell Association for the Deaf. 47-71.

Erber, N.P., and Alencewicz, C.M. (1976). Audiologic evaluation of deaf children. Journal of Speech and Hearing Disorders. 41:256-267.

Estabrooks, W. (Ed.) (1994). Auditory-Verbal Therapy For Parents And Professionals. Washington, DC: A.G. Bell Association for the Deaf.

Geers, A. E. and Moog, J.S. (1992). Speech perception and production skills of students with impaired hearing from oral and total communication education settings. Journal of Speech and Hearing Research. 35:1384-1393.

Geers, A.E., and Tobey, E. (1992). Effects of cochlear implants and tactile aids on the development of speech production skills in children with profound hearing impairment. The Volta Review. Volume 94:135-163.

Godfrey, J., and Millay, K. (1978). Perception of rapid spectral change in speech by listeners with mild and moderate sensorineural hearing loss. Journal of the American Audiology Society. 3:200-208.

Gold, T. (1978). Speech and hearing skills: A comparison between hard-of-hearing and deaf children. Doctoral dissertation, City University of New York.

Green, D.M., and Swets, J.A. (1974). Signal detection theory and psychophysics. Krieger, New York.

Haskins, H.A. (1949). A phonetically balanced test of speech discrimination for children (Kindergarten Phonetically Balanced word lists: K-PB). Unpublished master thesis, Northwestern University, Evanston, IL.

Harris, K.S. (1990). The assessment of speech: What progress? In: Assessment of speech and voice production: Research and clinical applications. NIDCD Monograph. 5-16.

Jensema, C.J., Karchmer, M.A., and Trybus, R.J. (1978). The rated speech intelligibility of hearing-impaired children: Basic relationships and a detailed analysis. Washington, DC: Office of Demographic Studies, Gallaudet College. (Series R, No.6).

Kent, D.R., and Read, C. (1992). The acoustic analysis of speech. Singular publishing group, Inc. San Diego, California.

Klatt, D.H. (1987). Review of text-to-speech conversion for English. Journal of the acoustical Society of America. 82:737-793.

Levitt, H., and Resnick, S.B. (1978). Speech perception by the hearing-impaired: Methods of testing and development of new tests. Scandinavian Audiology. Suppl., 6:107-129.

Levitt, H. (1987). Speech production in deaf children: Longitudinal data over three years. Paper presented at the meeting of A.G. Bell Association for the Deaf, St. Louis.

Lewis, M. (1951). Infant speech: A study of the beginnings of language. London: Routledge and Kegan Paul.

Ling, D. (1976). Speech and the hearing-impaired child: Theory and practice. Washington, DC: A.G. Bell Association for the Deaf.

Ling, D. and Milne, M. (1981). In: Bess, F.H., Freeman, B.A., and Sinclair, J.S. (Eds.). Amplification in Education. Washington, D.C.: A.G. Bell Association for the Deaf.

Markides, A. (1970). The speech of deaf and partially-hearing children with special reference to factors affecting intelligibility. British Journal of Disorders of Communication. 5:126-140.

McGarr, N. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. Language and Speech. 24:255-264.

McGarr, N. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. Journal of Speech and Hearing Research. 26:451-458.

McGarr, N. (1987). Communication skills of hearing-impaired children in schools for the deaf. Language and Communication Skills of Deaf Children, ed. by H. Levitt, N. McGarr, and D. Geffner, ASHA Monograph, 26. Washington, DC: American Speech-Language-Hearing Association.

Metz, D.E., Schiavetti, N., and Sitler, R.W. (1980). Toward an objective description of the dependent and independent variables associated with intelligibility assessments of hearing-impaired adults. In J.D. Subtelny (Ed.). Speech assessment and speech improvement for the hearing-impaired. Washington, DC: A.G. Bell Association for the Deaf.

Montgomery, G. (1967). Analysis of pure-tone audiometric responses in relation to speech development in the profoundly deaf. Journal of the Acoustical Society of America. 41:53-59.

Monsen, R.B. (1978). Toward measuring how well hearing-impaired children speak. Journal of Speech and Hearing Research. 21:197-219.

Monsen, R.B. (1981). A usable test for the speech intelligibility of deaf talkers. Americans Annals of the Deaf. 126:845-852.

Monsen, R.B. (1983). The oral speech intelligibility of hearing-impaired talkers. Journal of Speech and Hearing Disorders. 48:286-296.

Moog, J. S. (1989). The CID Phonetic Inventory. St. Louis: Central Institute for the Deaf.

Moog, J.S., and Geers, A.E. (1990). Early speech perception battery. St. Louis: Central Institute for the Deaf.

Most, T. (1985). Assessment of the perception of intonation by severely and profoundly hearing-impaired children. Unpublished doctoral dissertation. New-York: City University.

Osberger, M. J. (1992). Speech intelligibility in the hearing impaired: Research and clinical implications. In Kent, R. D. (Ed.). Intelligibility in speech disorders. John Benjamins Publishing Company.

Osberger, M. J., Maso, M., Sam, L. K. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. Journal of Speech and Hearing Research. 36:186-203.

Osberger, M.J., Miyamoto, R.T., Zimmerman-Phillips, S., Kemink, J.L., Stroer, B.S., Firszt, J.B., and Novak, M.A. (1991a). Independent evaluation of the speech perception abilities of children with the Nucleus 22-Channel cochlear implant system. Ear and Hearing. 12, No. 4 (Suppl):66s-80s.

Osberger, M.J., Robbins, A.M., Miyamoto, R.T., Berry, S.W., Myres, W.A., Kessler, K., & Pope, M.L. (1991b). Speech perception abilities of children with cochlear implants, tactile aids, or hearing aids. The American Journal of Otology. 12 (Suppl):105-115.

Penrod, J.P. (1994). Speech threshold and word recognition/ discrimination testing. In: Katz, J. (Ed.). Handbook of clinical audiology. (4th Edition). William and Wilkins. Baltimore, Maryland. 147-164.

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

Pickett, J.M., Martin, E.S., Johnson, D., Brandsmith, S., Daniel, Z., Willis, D., & Otis, W. (1972). On patterns of speech features reception by deaf listeners. In Fant, G. (Ed.). International symposium of speech communication ability and profound deafness. Washington, DC: A.G. Bell Association for the Deaf.

Plant, G. (1984). The effects of an acquired profound hearing loss on speech production. British Society of Audiology. 18:39-48.

Resnick, S.B. (1984). Speech recognition testing: Nonconventional testing techniques and application to children. ASHA Reports. 14:65-68.

Ross, M., and Lerman, J. (1970). A picture identification test for hearing-impaired children. Journal of Speech and Hearing Disorders. 13:44-53.

Ruffin-Simon, C. (1983). Relationship between production and perception of speech: Some developmental issues. In: Hochberg, I., Levitt, H., & Osberger, M.J. (Eds.). Speech of the Hearing Impaired: Research, Training, and Personal Preparation. University Park Press, Baltimore. 233-250.

Shoup, J.E. (1980). Phonological aspects of speech recognition. In: Lea, W.A. (Ed.) Trends in Speech Recognition. Prentice-Hall, New York. 125-138

Smith, C.R. (1975). Residual hearing and speech production in deaf children. Journal of Speech and Hearing Research. 18:795-811.

Stark, R.E., and Levitt, H. (1974). Prosodic feature reception and production in deaf children. Paper presented at the 87th meeting of the Acoustical Society of America.

Trammel, J.L., and Owens, S.L. (1980). Los Angeles County. Office of the Los Angeles County Superintendent of Schools, Audiologic Services, and Southwest School for the hearing-impaired. Test of Auditory Comprehension. Foreworks, North Hollywood, CA.

Tyler, R.S. (1993). Speech perception by children. In: Tyler, R.S. (Ed.). Cochlear Implants: Audiological Foundations. Singular Publishing Group. 191-256.

Walden, B.E. (1984). Validity issues in speech recognition testing. ASHA Reports. 14:16-18.

Yeung, E., Chant, G., Redmond, C., and Boothroyd, A. (1991). Stimulus presentation system. Center for Research in Speech and Hearing Sciences. Engineering Support Services. ESS#6.1.

Yeung, E. (1995). Digital waveform editing software using the data translation 2801A data acquisition board (version 2.0). Center for Research in Speech and Hearing Sciences. Engineering Support Services. ESS#22.