To the Graduate Council:

We are submitting herewith a dissertation written by Patrick N. Plyler entitled "Electrophysiologic Correlates of Stop Consonant Perception in Normal and Hearing-Impaired Listeners." We have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Speech and Hearing Science.

A.K. Ananthanarayan, Ph.D., Major Professor

Mark Hedrick, Ph.D., Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of The Graduate School

ELECTROPHYSIOLOGIC CORRELATES OF STOP CONSONANT PERCEPTION IN NORMAL AND HEARING-IMPAIRED LISTENERS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Patrick N. Plyler
August 1998

Copyright Patrick Norton Plyler, 1998
All rights reserved

DEDICATION

This dissertation is dedicated to my parents, James and Rachael Plyler. As a child, you taught me the value of patience, kindness, and honesty. As a teen, you instilled in me a good work ethic and the need for perseverance. In my college years, you taught me the importance of responsibility and accountability by giving me space to grow and by allowing me to learn from my mistakes. As a young adult, you've shown me the importance of standing behind my convictions without convicting those who disagree. It is from of all of your teachings that this document is possible. Therefore, I dedicate this dissertation to you, my parents, for teaching me the lessons of life.

ACKNOWLEDGMENTS

The author would like to take this opportunity to express appreciation to the many people who have made the completion of this project possible. To Dr. Ravi Krishnan and Dr. Mark Hedrick, thank you for working hand-in-hand during the course of this study. Each of you demonstrated a great deal of patience in getting me up to speed in areas of audiology which were not my strengths. You worked well together and never placed more importance on one aspect of the study than the other. I look forward to working with each of you in the future.

I would also like to thank my committee members, Dr. Jim Thelin, Dr. Samuel Burchfield, and Dr. Mary Sue Younger. To Dr. Thelin and Dr. Burchfield, thanks for your daily encouragement and constant guidance. Knowing I could come to you not only as advisors but as friends was and is very comforting. To Dr. Younger, many thanks for your statistical consultations and prompt responses to questions. I feel fortunate that you have played such a large role in my doctoral program, both as a professor and as a committee member.

Special thanks is extended to all of the volunteers who participated in this project. I would also like to personally thank Mrs. Linda Reed whose help with subject scheduling was a gift from above. Once again, your willingness to take on this extra job during my time of need was a true reflection of your Volunteer Spirit. I greatly appreciate the unselfish contribution of your time.

Last but certainly not least, thank you Erin Wendover for always being there for a sounding board and for encouragement. You had confidence in me when I was unsure

and you believed in me when I had doubts. Thank you for being my friend.

ABSTRACT

The effects of varying the onset frequency of the second formant and the overall intensity level of synthetic speech on the identification of stop consonant place of articulation and the neural representation of second formant transitions were compared in normal and hearing-impaired listeners. Comparisons of the phonetic boundary and slope of group identification functions for each stop consonant were made at each intensity level. Phonetic boundaries for normal and hearing-impaired listeners were relatively similar for each stop consonant at each intensity level. Slopes of the /d/ and /g/ identification functions for normal and hearing-impaired listeners were similar at each intensity level, however, slopes of the identification of /b/ functions were significantly steeper for the normal group at each intensity level except 72 dB SPL.

Comparisons of stop consonant identification at each intensity level were also made for the stimuli which corresponded to the highest percentage identification of each stop consonant. Results indicated significantly lower identification of /b/ and /g/ for hearing-impaired listeners at each intensity level and significantly lower identification of /d/ for hearing-impaired listeners at higher intensity levels.

The neural representation of the stimuli which corresponded to the highest percentage identification of each stop consonant was evaluated for each subject using the frequency-following response. Results demonstrated that the frequency-following response was sufficiently dynamic to encode time-varying formant transitions in normal listeners and in some listeners with hearing-impairment. Furthermore, comparisons of identification and neural representation results indicated that the identification of stop

consonant place of articulation was directly related to the neural encoding of the time-variant second formant transition reflected in the human frequency-following response.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE **PAGE**

LIST OF FIGURES

FIGURES **PAGE**

FIGURES

PAGE

CHAPTER I

Introduction

Listeners with normal hearing use the second formant transition as an acoustic cue for the identification of stop consonant place of articulation (Liberman, Delattre, Cooper, and Gerstman, 1954; Kewley-Port, 1981). Listeners with a mild to moderate sensorineural hearing loss have difficulty utilizing the second formant transition for the identification of stop consonants (Dorman, Marton, Hannley, and Lindholm, 1985). However, listeners with hearing-impairment do not have difficulty perceiving stop consonants when the short-term spectral cues embedded in the first 40 ms of the waveform are presented at an audible level (Van Tasell, Hagen, Koblas, and Penner, 1982). These results suggest that errors in the perception of stop consonants may be due to the inaudibility of the second formant transition.

Research investigating the neural representation of speech in the auditory nerve fibers of animals has implicated neural "phase locking" as the basis for the encoding of speech (Miller and Sachs, 1983; Miller and Sachs, 1984; Voight, Sachs, and Young, 1982; Young and Sachs, 1979). Specifically, the spectra of vowels are well represented, across an array of nerve fibers, in the distribution of activity locked to the individual harmonics of the stimulus (Young et al., 1979). Subsequent studies have indicated that this temporal-place scheme not only encodes steady-state vowels, but also whispered vowels, voice pitch, and time-varying stop consonant-vowel syllables (Miller et al., 1983, 1984; Voight et al., 1982). These results suggest that phase-locking plays an important role in the neural encoding of the essential features of speech.

The scalp recorded human frequency-following response (FFR) reflects sustained neural activity in the rostral brainstem which is phase-locked to the individual cycles of the stimulus waveform (Moushegian, Rupert, and Stillman, 1973). The FFR preserves information regarding the spectral features of steady-state back vowels (Ananthanarayan, 1997a, 1997b) and time-varying speech-like sounds (Parkinson, 1997). It is unclear, however, if the neural encoding of time varying features of complex speech sounds, such as the second formant transition, is preserved in the phase-locked neural activity reflected in the FFR.

While the identification of stop consonants in normal and hearing-impaired listeners has been extensively studied, the effects of hearing-impairment on the neural encoding of time varying speech features which serve to cue stop consonant identification remains unclear. Thus, the primary purpose of this study is to evaluate the relationship between the FFR representation of the second formant transition and stop consonant identification in normal hearing and hearing-impaired listeners.

CHAPTER II

Review of Literature

Stop Consonant Perception with Normal Hearing

The production of speech results in prominent resonances known as formants. Formants are characterized in terms of their duration, bandwidth, amplitude and frequency. The formant characteristics of each stop consonant are unique to the stop consonant being produced and serve as acoustic cues to their perception. Factors such as stop consonant place and manner of articulation as well as the degree of voicing interact to create a large number of potential vocal tract shapes and, consequently, a large number of possible perceptual cues (Borden et al., 1990).

Manner Cues

During the production of a stop consonant, the vocal tract is completely occluded then opened, resulting in a momentary cessation of airflow followed by a brief burst of airflow. The momentary cessation of airflow is perceived as silence in voiceless stops and as attenuation in voiced stops, while the subsequent release of the stopped air is perceived as a transient burst of noise. Each of these perceptions, the relative silence and the noise burst, serve as acoustic cues for stop manner of articulation.

Stop manner of articulation has also been shown to be cued by rapid formant transitions. Liberman, Delattre, Gerstman and Cooper (1956) determined that stop consonants could be synthesized by initiating a vowel with a formant transition of 40 milliseconds (ms) or less. However, extending the formant transition to 50 ms or to 150 ms changed the perception to a semivowel glide or to a vowel sequence respectively.

Thus, formant transitions of 40 ms or less serve as acoustic cues for stop manner of articulation.

Voicing Cues

The duration of vocal tract occlusion serves as an acoustic cue for the distinction of stop consonant voicing class. For example, increasing the closure duration of the stop in the word "rabid" by more than 70 ms changes the perception to the word "rapid." In addition, Liberman, Delattre, and Cooper (1958) evaluated the perception of voicing in stop consonants by delaying the onset of the first formant transition in 10 ms steps over a series of successive stimuli. Listeners reported hearing voiced stops for delays less than 30 ms and voiceless stops for delays greater than 30 ms. Listeners also reported that the addition of noise to the upper formants of the delayed stimuli produced a stronger impression of voicelessness than did the delay alone.

Lisker and Abramson (1964) investigated the effects of voice onset time on the ability to distinguish voicing class for initial stop consonants. Results indicated that the voice onset times required for listeners to report hearing a voiceless stop were greater for stops with velar place of articulation than for stops with labial or alveolar place of articulation. Voice onset times were also greater for voiced stops with velar place of articulation than for voiced stops with labial or alveolar place of articulation.

Place Cues

The identification of stop consonant place of articulation can be cued by the frequency position of the noise burst in relation to a vowel and by the second formant transition serve as cues to the stop place of articulation. Liberman, Delattre and Cooper

(1952) studied the effects of burst frequency and vowel pattern on the perception of stop consonants. Subjects listened to seven different two-formant synthetic vowels preceded by bursts of noise with varying center frequencies. Results indicated that high frequency bursts preceding the vowels were perceived as /t/, low frequency bursts preceding the vowels were perceived as /p/, and bursts slightly above the frequency of the second formant of the following vowel were perceived as /k/.

Stevens and Blumstein (1978) observed that the gross shape of the short-term spectrum at consonant release uniquely specified consonant place, regardless of vowel context. The spectral shape at release was related to the acoustic properties of the consonant burst, as well as to the formant frequencies, but not to the later formant transitions into the following vowel. Blumstein and Stevens (1980) later demonstrated that stop consonant place of articulation could be accurately identified using synthetic stop-vowel stimuli consisting only of a noise burst plus 10 ms of voicing. These results suggested that the burst and the attributes of the sound at voicing onset produce a global acoustic property that cues place of articulation for voiced stops.

Liberman et al. (1954) evaluated the role of the second formant transition in the perception of stop consonant place of articulation. A series of two-formant CV (V = /a/) syllables initiated only by a formant transition (no bursts) were synthesized with the first formant transition and the steady state formant frequencies held constant throughout the duration of the stimuli. The slope of the second formant transition varied systematically from sharply rising to sharply falling. Results indicated that normal hearing listeners identified all of the stops cued by rising second formant transitions as labial (/p,b/).

However, the stops cued by falling second formant transitions were identified as alveolar (/t,d/) when the slopes fell slightly and as velar (/k,g/) when the slopes fell sharply. These results suggested that the second formant transition serves as a cue for stop consonant place of articulation.

The perceptual effects of altering the second formant transition reported by Liberman et al. (1954) were limited to the context of the vowel /a/. When stop consonants were combined with vowels other than /a/, the resulting second formant transition pattern differed for each CV combination. For example, when stops were synthesized before a vowel with a high frequency second formant, the second formant transition leading from the labials rose much more steeply than with the /a/. If the synthetic vowel had a low frequency second formant, the second formant transition leading from the velars fell more sharply than with the /a/. The disparity in second formant transition patterns led to the theory of an acoustic locus for each stop place of articulation (Delattre, Liberman, and Cooper, 1955).

The locus theory was based on the premise that when a stop consonant occlusion is released, the shape of the vocal tract is associated with particular formant resonances. These formant resonances change as the vocal tract shape approximates the shape necessary for the following vowel. However, since the place of articulation during the occlusion of a given stop is independent of the following vowel, a relationship should exist between consonant-vowel combinations and the starting frequency of the second formant transition. The locus theory worked well for the alveolar stop /d/, however, identifying a particular locus for the velar stops /g,k/ was more difficult due to the fact

these consonants are not as restricted to a single place of articulation during occlusion (Kewley-Port, 1983).

Kewley-Port (1981) argued that dynamic changes in spectral shape which occur during consonant release and the subsequent articulatory movement towards the following vowel cue stop place of articulation. To substantiate this theory, data was obtained for three sets of stop-vowel syllables. Set one consisted of the initial 20-40 ms of nine natural syllables, set two consisted of the same syllables synthesized to preserve the dynamic spectral features of the natural stimuli, and set three was synthesized to preserve only the static onset spectral characteristics during the 25.6 ms following burst onset. Results indicated that normal hearing subjects identified the dynamic stimuli more accurately than the static stimuli. Kewley-Port concluded that the static spectral information was an insufficient cue for stop place, and that the place cue occurred in the dynamic spectral properties of the initial portion of the stop-vowel syllable waveform.

Burst amplitude and the relative spectral change from burst onset to voicing onset have also been suggested as cues for stop consonant identification. Ohde & Stevens (1983) determined that stop consonants synthesized with relatively low burst amplitudes were identified as /p/. However, increasing the amplitude of the burst resulted in reduced /p/ identification and increased /t/ identification. Lahiri, Gewirth, & Blumstein (1984) concluded that stop consonants could be correctly classified by comparing the ratio of change in the high frequencies (3500 Hz) to the ratio of change in the low frequencies (1500 Hz) over the time interval from stop release to voice onset. This dynamic criterion described a temporal change in spectral tilt.

Summary

In summary, there are several acoustic cues that normal hearing listeners use to determine the manner, voicing class, and place of articulation of stop consonants. The presence of silence, a release-burst, and relatively rapid formant transitions can all serve as cues to the stop manner of articulation. To recognize the voiced-voiceless contrast, listeners use temporal cues such as the first formant transition delay and the duration of the silence marking the stop closure. The acoustic cues for place of articulation of stop consonants are the frequency position of the burst in relation to a vowel, the second formant transition, and the dynamic spectral properties of the initial portion of the stop-vowel syllable waveform.

Stop Consonant Perception with Hearing-Impairment

Overview

Research investigating the effects of sensorineural hearing loss on the perception of various speech features has indicated that listeners with sensorineural hearing loss perceive vowels better than consonants, word initial consonants better than word final consonants, and consonant voicing better than consonant place (Boothroyd, 1984).

Voicing Cues

Studies evaluating the effects of silence duration on the perception of postphonemic stop consonant voicing in normal and hearing-impaired listeners indicate that hearing-impaired listeners need longer silence durations (10-12 ms) to perceive voiceless stop consonants than do normal hearing listeners (Price and Simon, 1984; Dorman et al., 1985).

In contrast, Cazals and Palis (1991) demonstrated that some listeners with mild to moderate hearing-impairment do not require extended silence durations to perceive voicing distinctions of intervocalic stop consonants. Cazals et al. further concluded that voicing distinction errors result from forward masking induced by low frequency energy in the fundamental frequency of the preceding vowel. Therefore, hearing-impaired listeners who demonstrate delayed recovery from forward masking have more difficulty perceiving the voicing contrasts of stop consonants.

Temporal Cues

Listeners with sensorineural hearing-impairment have been found, generally, to have poorer-than-normal acuity for temporal events resulting in larger temporal and gap difference limens (42 vs 18 ms, 78 vs 51 ms) as well as larger gap detection thresholds (12 vs 7 ms) (Tyler, Summerfield, Wood, and Fernandez, 1982). However, temporal acuity differences between normal and hearing-impaired listeners were found to be relatively small when compared to the magnitude of differences in acoustic segment duration which cue phonetic contrasts. For example, the elevation in gap difference limen reported by Tyler et al. was 27 ms, however, a silent interval of 70 ms is required to signal the presence of a stop consonant in a cluster.

Dorman et al. (1985) evaluated the effects of varying the duration of formant transitions on the phonetic identification abilities by young, normal hearing listeners and elderly listeners with a mild to moderate sloping hearing impairment. Results of the identification functions revealed similar phonetic boundaries and slopes for each of the three groups, indicating that the elderly hearing-impaired listeners performed normally on

tasks where phonetic identification was cued by stimulus duration. These results suggest that, although mild to moderately hearing-impaired listeners have poorer than normal temporal acuity, the magnitude of the disability is such that phonetic identification may not be compromised (Dorman et al., 1985).

Place Cues

Studies of phoneme perception by subjects with moderate to severe hearing loss have shown that the phonetic features most likely to be perceived erroneously are those related to place of consonant articulation (Owens and Schubert, 1968, 1972; Byers, 1973; Sher and Owens, 1974; Reed, 1975). Dorman et al. (1985) studied the effects of varying the onset frequency of the second formant transition on the perception of stop consonant place of articulation in normal and hearing-impaired listeners. Results of the identification functions revealed significantly different phonetic boundaries and slopes for each group, indicating that the hearing-impaired listeners perform abnormally on tasks where the identification of stop place of articulation was cued by the onset frequency of the second formant transition.

Schum and Collins (1990) determined that normal hearing listeners extract more phonemic information related to consonant place from the earlier occurring portions of speech waveforms than do hearing-impaired listeners. Van Tasell et al. (1982) investigated the use of acoustic information at syllable onset for the identification of the place feature of synthetic voiced stop consonants in normal and hearing-impaired listeners. Two synthetic speech continua, one full length and one truncated which contained a burst portion plus two glottal pulses, served as the stimuli. Both normal and

hearing-impaired subjects were able to identify the full-length synthetic stimuli accurately, although performance of the hearing-impaired subjects was slightly more variable. For the two-pulse stimuli, performance of the hearing-impaired subjects did not differ from that of the normal group.

Van Tasell et al. (1982) suggested that auditory distortions which occur with sensorineural hearing loss such as impaired frequency resolution, loudness recruitment and the absence of suppression, may have little effect on the perception of speech by moderately hearing-impaired subjects. The researchers further postulated that if acoustic cues are audible, hearing-impaired listeners extract sufficient phonemic information for accurate feature identification in spite of signal distortions, and that the differentiation of phonemes may involve only relatively gross auditory discrimination of certain attributes of the complex speech signal. In contrast, Hedrick (1995) noted that greater perceptual weight was given to the burst relative amplitude cue than to formant transitions by hearing-impaired listeners in labeling voiceless stops. Furthermore, Hedrick and Jesteadt (1996) suggested that recruitment may alter the perceptual weights given to acoustic cues by hearing-impaired listeners.

Summary

In summary, several studies have investigated the effects of sensorineural hearing-impairment on the perception of the various features of speech. Hearing-impaired listeners perform normally on tasks where phonetic identification is cued by stimulus duration. However, hearing-impaired listeners who demonstrate delayed recovery from forward masking have more difficulty distinguishing voicing contrasts than listeners with

normal recovery from forward masking. In addition, listeners with hearing-impairment have more difficulty identifying place of consonant articulation, however, this difficulty is reduced when the cues necessary for accurate feature identification are audible.

Neural Representation of Speech

Single Unit Data

Results from animal experiments in which responses to steady-state synthetic vowels were recorded from large populations of single unit auditory nerve fibers revealed two different schemes of neural representation of vowels. In the first scheme, formant peaks in the stimulus produced peaks in the discharge rate, across an array of nerve fibers, at frequencies corresponding to the formant frequencies. This scheme was termed the "rate-place" representation, and, while the "rate-place" scheme provided a good representation for formant frequencies at stimulus levels below 50 dB SPL, it was unstable and failed for stimulus levels greater than 50 dB SPL due to rate saturation and two-tone suppression (Sachs and Young, 1979).

In the second scheme, it was found that the spectra of vowels were represented in the distribution of activity phase locked, across an array of fibers, to the individual harmonics of the stimulus. This scheme was termed the "temporal-place" or "phase-locking" representation and, unlike the "rate-place" scheme, the "phase-locking" scheme was shown to be extremely robust and exhibited a wide dynamic range of intensity (Young et al., 1979). Subsequent studies demonstrated that the "phase-locking" scheme not only accounted for the encoding of steady-state vowels, but also for the encoding of whispered vowels (Voight et al., 1982) and of time-varying consonant-vowel syllables

(Young et al., 1979; Miller et al., 1983).

Results of these animal studies suggested that phase-locking plays a critical role in the encoding of the essential features of steady-state and time variant speech sounds to the central auditory system. In humans, recordings which reflect sustained neural activity phase-locked to the individual cycles of a stimulus, termed the frequency following response (FFR) (Moushegian et al., 1973; Worden and Marsh, 1978), provide a non-invasive window to evaluate neural encoding of both steady-state and time-variant complex sounds in humans.

The Frequency-Following Response

Neural Generator(s) of the FFR

The FFR is an auditory evoked potential that reflects sustained neural activity phase-locked to the individual cycles of the stimulus waveform (Worden et al., 1968). Moushegian et al. (1973) maintained the FFR is of neural origin because ; (I) it does not have the frequency range of the cochlear microphonic response; (ii) FFR is absent in the presence of white masking noise; and, (iii) latency of the FFR varies inversely with stimulus frequency. Smith, Marsh, and Brown (1975) determined that the onset latency of the FFR recorded from the scalp and from the inferior colliculus brainstem nuclei was similar. In addition, the FFR has been recorded bilaterally from the inferior colliculus, the medial superior olive, and from the scalp. Cryogenic cooling of the inferior colliculus resulted in reduced amplitudes of FFR recordings at both inferior colliculi and the scalp while FFR recordings from the medial superior olive remained at control amplitudes. Therefore, the primary generator of the FFR seems to be the inferior colliculus (Smith et

al., 1975).

Response Characteristics

The scalp-recorded FFR is composed of several components which include a far field recorded cochlear microphonic potential and a low frequency FFR of neural origin (Stillman, Crow and Moushegian, 1978). Moller and Jho (1989) isolated two components of the FFR. One component, equal to the stimulus sinusoid, was isolated by subtracting the responses to tones of opposite polarity. The other component, a slow component with more waveform variability, was isolated by adding the responses to tones of opposite polarity. Batra, Kuwada, and Maher (1986) used continuous tones to isolate two types of FFRs. The latency of the first type of FFR (1 ms) was consistent with cochlear site of origin and was believed to be the cochlear microphonic. The latency of the second type of FFR (8.2 ms, $\pm$ 1 ms) was consistent with mid-brain site of origin, perhaps the inferior colliculus.

Response characteristics of the FFR are dependent on the frequency and intensity of the stimulus. Glaser, Suter, Dasheiff, and Goldberg (1976) observed FFR to stimuli ranging from 70 to 1500 Hz. FFRs to 250 and 500 Hz stimuli were observed at lower levels and with greater amplitudes than FFRs to higher stimulus frequencies. In addition, latency estimates of the FFR have been placed from 5.4 to 7 msec (Gerken, Moushegian, Stillman and Rupert, 1975), while increasing the stimulus intensity reduces FFR latency and increases FFR amplitude. Marsh, Brown and Smith (1975) found a high degree of consistency both within and across subjects in latency, phase, and waveform of averaged FFRs.

Representation of Vowels in FFR

The FFR has been utilized to evaluate the neural representation of synthetic steady-state vowels (Ananthanarayan, 1998a). Results of this study indicated that the spectrum of FFRs to several different back vowels (/a, u, ɔ/) revealed peaks that corresponded to the first two formant frequencies of the vowel stimuli. These formant peaks were robust and identifiable over a range of intensities. Results of this study also indicated that FFRs to these vowel stimuli contained pitch relevant information.

FFR to Complex Stimuli

Galbraith and Brown (1990) reported FFRs to harmonically complex stimuli were characterized by a clear periodicity at the missing fundamental frequency (MF). Galbraith (1994) further revealed that MF frequencies were not coded directly in the peripheral neural response as shown by horizontal channel recordings. However, the MF was present in vertical channel recordings, indicating registration of the MF in the brainstem. Greenberg, Marsh, Brown, and Smith (1987) reported that FFRs to the missing fundamental frequency are spectrally similar to those evoked by pure tones equal in frequency to the missing fundamental. These results indicated that the FFR does indeed reflect neural activity pertinent to the processing of low pitch.

Chertoff and Hecox (1990) introduced an electrophysiological technique which assessed intermodulation distortion in the auditory system. Spectral analysis of the FFR to two tone signals revealed intermodulated distortion products at the difference tone and cubic difference tones which were not stimulus artifact and were definitely neural in origin. Pandya (1996) further evaluated the response characteristics of the FFR distortion

product at 2F1-F2 in normal hearing subjects. He concluded that the magnitude of the F1, F2, and 2F1-F2 components decreased with increasing frequency and that the slope of the 2F1-F2 growth function decreased with increases in the distortion product frequency. In addition, the slope of some of the F2 and 2F1-F2 components were essentially parallel, suggesting a generation site consistent with the F2 place on the cochlea. Level differences generating the maximal distortion product component were also found to be dependent on the level of F1, while the contribution of the brainstem to the 2F1-F2 FFR distortion product was minimal.

The FFR has also been utilized to evaluate neural representation of time-varying stimuli in normal hearing subjects (Ananthanarayan & Parkinson, 1998b). Tone bursts were swept linearly from 400-600 Hz, 600-400 Hz, 1200-1400 Hz, and 1400-1200 Hz in order to approximate the formant transitions of speech. Each tonal sweep was 80 ms in duration and was presented at four intensity levels: 60, 70, 80, and 90 dB SPL.

Results demonstrated that amplitude decreased through the duration of rising sweeps, however, for falling sweeps, amplitude increased through the duration of the 600-400 Hz sweep and decreased through the duration of the 1400-1200 Hz sweep. These results suggested that the phase-locking ability of neurons decreases with increasing frequency. Thus, the ability of neurons to phase-lock to decreasing frequency is dependent on the frequency region stimulated. Results also indicated that the change of frequency in the stimulus and the rate of that change were well preserved in the FFR. These results indicate that the human FFR is sufficiently dynamic for the encoding of time-varying low frequency stimuli (Ananthanarayan et al., 1998b).

Rationale

Studies of phoneme perception by subjects with moderate to severe hearing loss have shown that the phonetic features most likely to be perceived erroneously are those related to place of consonant articulation. For example, Dorman et. al (1985) demonstrated that varying the duration of the second formant transition did not compromise the perception of the /bæ/ to /wæ/ continuum in hearing-impaired listeners, however, varying the onset frequency of the second formant transition did compromise the hearing-impaired's perception of the /ba/ to /da/ continuum. In addition, Van Tasell et al. (1982) reported that hearing-impaired listeners demonstrate normal stop consonant perception when the short-term spectral cues embedded in the first 40 ms of the waveform are audible. Given this, the identification of stop consonant place of articulation should be normal in the hearing-impaired when the necessary cues are audible to the listener.

Auditory nerve single-unit population studies have demonstrated that neural phase-locking is the primary basis for the neural representation of both steady-state and time-variant complex speech-like sounds. Additionally, the scalp recorded human frequency-following response, reflecting phase-locked neural activity in the rostral brainstem, has been shown to preserve information about certain spectral features of steady-state back vowels and time-varying speech-like sounds . Given this, it is reasonable to postulate that the phase-locked neural activity underlying the FFR generation is sufficiently dynamic to encode time varying features of complex speech sounds such as formant transitions. Therefore, information regarding the neural encoding

of the acoustic cues used to perceive stop consonant place of articulation, such as the second formant transition, should be preserved in the phase-locked neural activity reflected in the FFR. The primary purpose of this study is to evaluate the relationship between the FFR representation of the second formant transition and stop consonant identification in normal hearing and hearing-impaired listeners. Specifically, (i) what are the effects of intensity level on the identification of stop consonant place of articulation in listeners with normal and impaired hearing (ii) is the human frequency-following response sufficiently dynamic to encode time-varying speech, and (iii) does a relationship exist between stop consonant identification and the neural representation of the second formant transition reflected in the human frequency-following response.

CHAPTER III

Methods

Subjects

Thirty two human subjects ranging in age from 20 to 67 years participated in this experiment. The subjects were equally divided into a normal hearing group and a hearing-impaired group. The criteria for inclusion in the normal hearing group included: (i) hearing sensitivity of 20 dB HL (ANSI S3.6-1989) or better for frequencies from 250 to 8000 Hz at the right ear; (ii) normal appearance of ear canal and pinna; and (iii) normal tympanograms bilaterally. The inclusion criteria for the hearing-impaired group differed from that of the normal hearing group in one aspect: (i) hearing sensitivity between 20 dB and 60 dB HL for frequencies from 250 to 2000 Hz at the right ear (Appendix A). There were two experimental sessions for each subject and all subjects were unpaid volunteers.

Stimuli

Stimuli for all experiments were generated via Klatt's cascade/serial formant synthesizer (Klatt, 1980) at a sampling rate of 10 kHz. The stop consonant-vowel syllable /bɑ/ was synthesized with four formants. The first formant began at 150 Hz and reached steady state at 750 Hz. The second formant began at 900 Hz and reached steady state at 1250 Hz. The total duration of each syllable was 100 ms while duration of the first and second formant transitions were 40 ms. The third and fourth formants were 2400 Hz and 3300 Hz throughout the duration of the syllable. Fourteen additional stimuli were created by varying the onset frequency of the second formant from 900 to 2300 Hz in 100 Hz steps in order to create a /bɑ/, /dɑ/, to /gɑ/ perceptual continuum (Appendix

B). For all the stimuli, the first formant was synthesized at ± 3 dB relative to the second formant, all transitions were linear, and the fundamental frequency fell from 130 to 100 Hz over the duration of the syllable.

Experiment I: Speech Perception

Recording System

Stop consonant identification ability was evaluated using the 15 stimuli and four intensity levels. The stop-consonant stimuli were digitally synthesized and controlled by a signal generation system (Tucker-Davis, System II) interfaced to a microcomputer (Compaq 2000, 166 MHZ). Digital signal generation, including control of parameters such as overall intensity and programming of the research protocol, was accomplished by interactive signal generation and control software (CSRE Version 4.5). The digital stimuli were routed from the microcomputer to the digital-to-analog converter (TDT-DD1). The output of the D/A converter was routed to a low pass filter (TDT-PF1, 4.9 kHz low-pass) then to a separate programmable attenuator (TDT-PA4) then to the headphone buffer (TDT-HB) before being delivered to a magnetically shielded insert earphone (Etymotic, ER-3A) located in a sound treated examination room (Industrial Acoustics). Sound levels are expressed as the sound pressure level measured in a 2cc cavity using a one inch condenser microphone coupled to a sound level meter (Quest Model OB-145).

Subjects were seated in the examination room and were instructed to view a computer monitor through the examination room window. Following each stimulus presentation, subjects indicated which stop consonant they perceived by selecting the

appropriate symbol displayed on the monitor (B, D, or G) via the computer mouse. Each stimulus was presented at random at each level over 20 trials for a total of 1200 responses. Generation of random stimulus orders and online data collection were performed using a commercially-available identification program (CSRE, Version 4.5).

Response Evaluation

Subject responses were imported into Microsoft Excel (Version 5.0) in order to sort the data by onset frequency and level. Psychometric functions displaying the percentage identification as a function of onset frequency were generated for each stop consonant at each level for each subject and group.

Experimental Protocol

Speech perception data was collected during the first experimental session. Subjects first read and signed the informed consent form containing all pertinent information regarding the experiments and their participation in them (Appendix C). Pure-tone air-conduction and bone-conduction thresholds, tympanometry and otoscopy were performed at the beginning of the experimental session. The remainder of the first session was devoted to data collection. Data was collected for the normal hearing group first in order to establish perceptual norms.

Perceptual data was obtained under the following experimental conditions:

1) Trial data was obtained in order to familiarize the subject with the task. The parameters of the stimuli in the trial were:

Onset frequencies: 0.9, 1.0, 1.5, 1.6, 2.2, 2.3 kHz
Level: 82 dB SPL
Randomizations: 5

2) Upon completion of the trial, subjects proceeded to the experimental protocol. Parameters of the stimuli in the experimental protocol were:

Onset frequencies: 0.9, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3 kHz
Levels: 92, 82, 72, 62 dB SPL
Randomizations: 20

3) Upon completion of the experimental protocol, psychometric functions were generated.

Experiment II: Electrophysiology

Recording System

Frequency-following responses were recorded using the stop consonant-vowel stimuli with onset frequencies of 0.9, 1.6, and 2.3 kHz. These stimuli were the same as those used in Experiment I. Control of parameters such as overall intensity and programming of the research protocol was accomplished by interactive signal generation and control software (TDT-Siggen) interfaced to a microcomputer (Zenith, 486DX, 66 MHZ). Pulses were delivered from the timing generator module (TDT-TG6) which controlled the timing of the stimuli and provided a synchronous trigger pulse for the onset of digitization (AD-1). The stimulus generation module (TDT-DA1) was triggered by the pulses delivered from the timing generator module. The output of the stimulus generation module was then routed to a programmable attenuator (TDT-PA4) before being delivered to a magnetically shielded insert earphone (Etymotic, ER-3A) located in the sound treated examination room (Industrial Acoustics).

Subjects reclined comfortably in the examination room after being instructed to sit quietly and remain still during the recording session. FFRs were recorded differentially

between scalp electrodes placed on the midline of the forehead at the hairline (FpZ) and the 7th cervical vertibra (C7). Another electrode placed on the lower forehead served as the common ground. The interelectrode impedances were maintained below 3000 ohm. The EEG inputs were amplified by 200,000 and band-pass filtered from 100 to 3000 Hz (6 dB roll-off). Each response waveform represented an average of 1500 stimulus presentations over a 110 ms analysis window using a sampling rate of 20 kHz. Sound levels were expressed in the same manner as mentioned previously. Data acquisition was accomplished by the signal averaging and data analysis software (Biosig) of the TDT, System II.

Response Evaluation

Neural phase-locking to envelope periodicity dominates the FFR. In order to identify neural activity phase-locked to the spectral components of the stimulus, FFRs obtained using condensation polarity were subtracted from FFRs obtained using rarefaction polarity. A joint time-frequency analysis was then performed on the differential FFR waveform. This analysis yielded: (i) a three dimensional representation of how the FFR response changes in frequency and amplitude as a function of time; (ii) power spectrum of each instant in time of the FFR waveform; and (iii) magnitude and frequency of the FFR spectral peak as a function of time. Joint time-frequency analyses were performed for each onset frequency at each level for each subject.

Experimental Protocol

FFR data was collected during the second experimental session. As before, waveforms were collected for the normal hearing group first in order to establish

normative response data. FFRs were obtained under the following experimental conditions:

1) Stimuli were presented to the same ear in which perceptual data was obtained

2) Stimuli with onset frequencies of 0.9, 1.6, and 2.3 kHz were utilized.

3) Stimuli were presented at 92, 82, and 72 dB SPL.

4) Stimuli were presented at a rate of 5/sec using both rarefaction and condensation polarity.

CHAPTER IV

Results

Experiment I: Speech Perception

Slope and Phonetic Boundary

Each subject was presented 20 randomizations of the 15 stimuli at four intensity levels totaling 1200 stimulus presentations. Following each stimulus presentation, subjects indicated which stop consonant they perceived by selecting the appropriate symbol (B, D, or G) displayed on the computer monitor via the computer mouse. Group percentage identification functions generated for each stop consonant at each intensity level were analyzed for slope and for the location of the phonetic boundary.

Differences in slopes for the normal and impaired subjects were tested using SAS by fitting a multiple linear regression model with an indicator variable and an interaction term. If the coefficient on the interaction term did not differ significantly from zero, indicating no significant difference in slopes, the interaction term was dropped and a multiple regression model with only an indicator variable was fitted. If the coefficient on the indicator variable did not differ significantly from zero, meaning that the heights (intercepts) of the lines did not differ, then the indicator variable was dropped and a single simple linear regression relating percent correct responses to onset frequency for both normals and hearing-impaired listeners was fitted to the data.

Percentage /b/ Identification

Normal and hearing-impaired mean percentage /b/ identification functions for each intensity level are shown in Figure 1. Slopes of the percentage /b/ identification

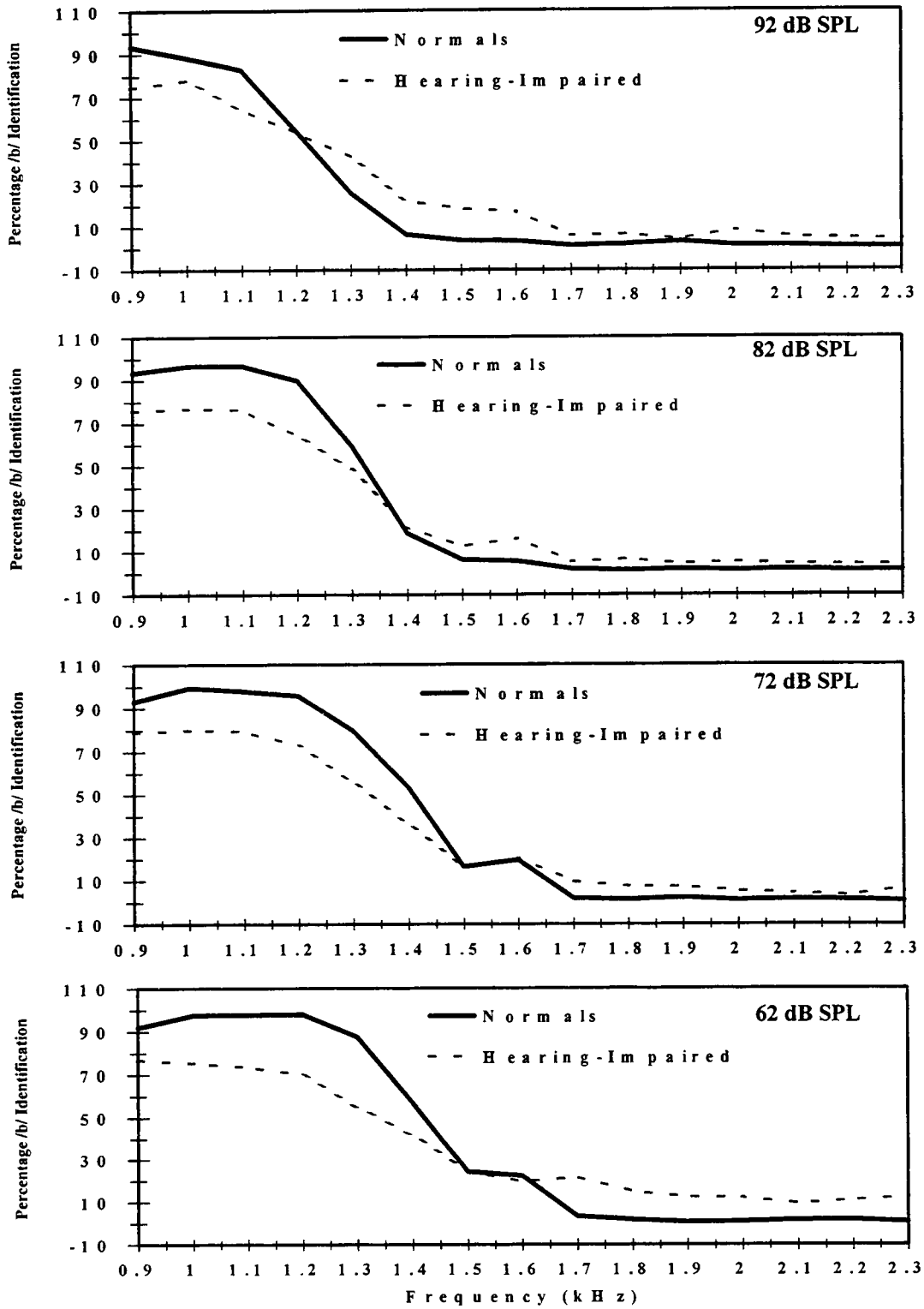


Figure 1: Mean percentage /b/ identification functions at each intensity level.

functions were significantly different between the two groups of listeners at 92, 82, and 62 dB SPL and were similar for each group at 72 dB SPL (Table 1).

The phonetic boundaries for each group at each intensity level are displayed in Table 2. In general, the phonetic boundaries for the percentage /b/ identification functions indicated relatively similar boundaries for each group at each intensity level. For each group, the phonetic boundary decreased in frequency with each increase in stimulus intensity. In addition, the difference in phonetic boundary between the two groups of listeners also decreased with each increase in stimulus intensity level.

Percentage /d/ Identification

Normal and hearing-impaired mean percentage /d/ identification functions for each intensity level are displayed in Figure 2. Slope and phonetic boundary measures were analyzed for data less than and greater than 1.6 kHz. Slopes of the percentage /d/ identification functions were not significantly different between the two groups of listeners at any intensity level with the exception of d > 1.6 kHz at 92 dB SPL (Table 1).

The phonetic boundaries for the percentage /d/ identification functions indicated relatively similar boundaries for each group at each intensity level (Table 2). For the /d/ < 1.6 kHz data, the only difference in phonetic boundary was at 92 dB SPL (0.13 kHz). The only phonetic boundary difference for the /d/ > 1.6 kHz data was at 82 dB SPL (0.146 kHz).

Percentage /g/ Identification

Normal and hearing-impaired mean percentage /g/ identification functions for each intensity level are shown in Figure 3. Slopes of the percentage /g/ identification

Table 1: Summary of mean slopes for each condition

Level	t-value	p>/t/ (normal vs hearing-impaired)
	Percentage /b/ identification functions	
92	-3.017	0.0033
82	-2.677	0.0088
72	-0.507	0.6132
62	-2.929	0.0041
	Percentage /d/ < 1.6 kHz identification functions	
92	0.84	0.4021
82	1.249	0.2139
72	1.768	0.0795
62	2.048	0.0427
	Percentage /d/ > 1.6 kHz identification functions	
92	-2.116	0.0359
82	-0.285	0.7761
72	1.519	0.1314
62	1.079	0.2828
	Percentage /g/ identification functions	
92	1.606	0.1095
82	0.104	0.9171
72	-0.038	0.9701
62	1.015	0.3119

Table 2: Summary of mean phonetic boundaries in Hertz

Level	Normals	Hearing-Impaired	Difference (N-I)
	Percentage /b/ identification funcitons		
92	1216	1233	-17
82	1313	1275	38
72	1416	1335	81
62	1433	1325	108
	Percentage /d/ < 1600 Hz identification functions		
92	1289	1419	-130
82	1385	1385	0
72	1492	1492	0
62	1501	1501	0
	Percentage /d/ > 1600 Hz identification functions		
92	1758	1758	0
82	1866	1720	0
72	1726	1726	0
62	1773	1773	0
	Percentage /g/ identification functions		
92	1756	1923	-167
82	1833	1833	0
72	1795	1795	0
62	1852	1852	0

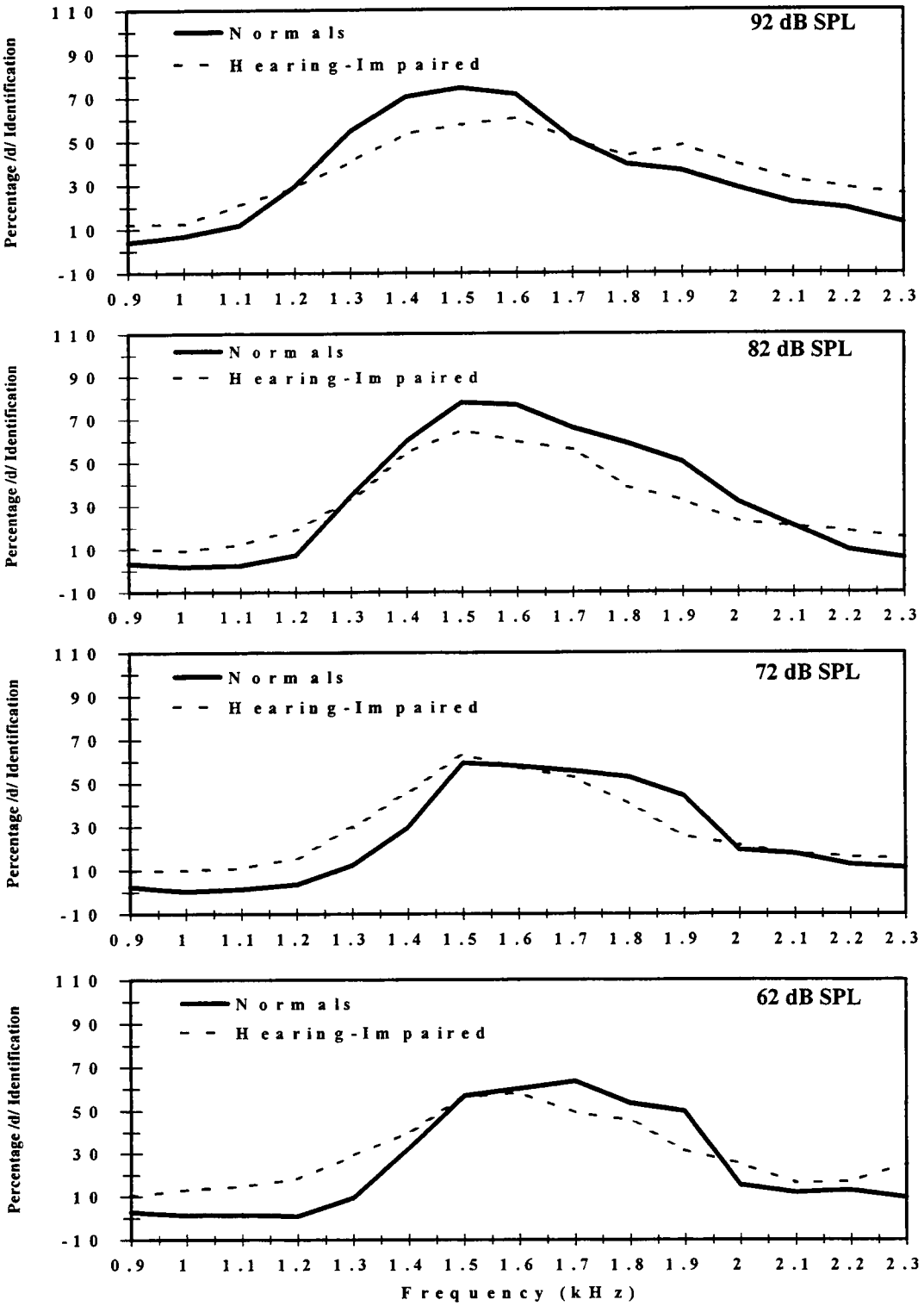


Figure 2: Mean percentage /d/ identification functions at each intensity level.

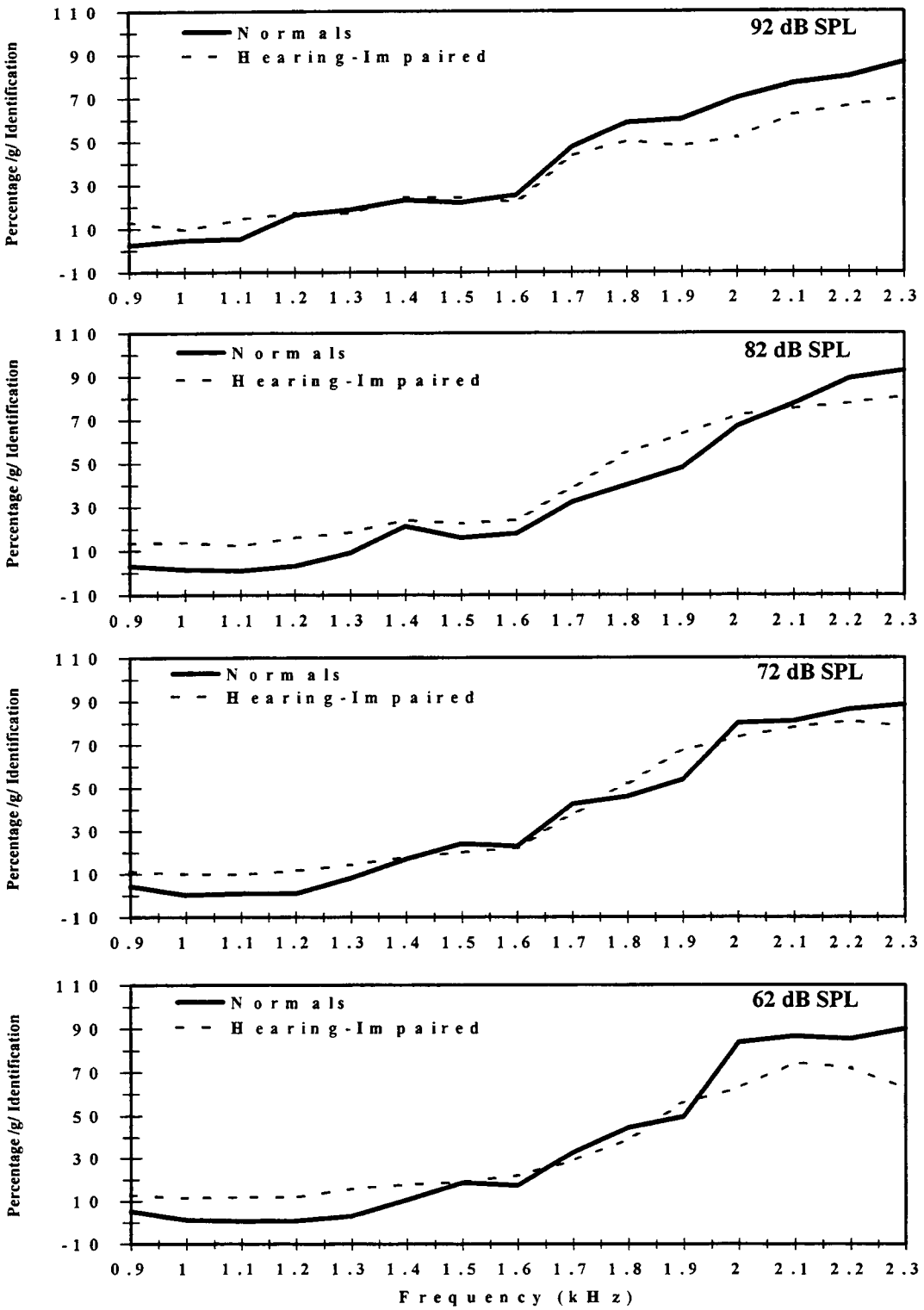


Figure 3: Mean percentage /g/ identification functions at each intensity level.

functions were not found to be significantly different between the two groups of listeners at any intensity level (Table 1).

The phonetic boundaries for the percentage /g/ identification functions indicated relatively similar boundaries for each group at each intensity level, however, a 0.167 kHz phonetic boundary difference was noted at 92 dB SPL (Table 2).

Onset Frequency

Results of the identification data were analyzed across group and intensity level to determine which second formant onset frequency corresponded to the highest percentage identification for each stop consonant. Results indicated that stimuli with rising second formant transitions resulted in the greatest percentage of /b/ responses (0.9-1.0 kHz), stimuli with gradually falling second formant transitions resulted in the greatest percentage of /d/ responses (1.5-1.6 kHz), and stimuli with sharply falling second formant transitions resulted in the greatest percentage of /g/ responses (2.2-2.3 kHz).

Normal and hearing-impaired mean percentage /b, d, g/ identification of the 0.9, 1.6, and 2.3 kHz onset stimuli at each intensity level are displayed in Figure 4. In order to determine if group identification was similar at each level for each onset frequency, a between subjects analysis of variance with repeated measures was completed on the percentage /b, d, g/ identification of the 0.9, 1.6, and 2.3 kHz onset stimuli at each intensity level. Contrasts were written to compare group means within each intensity level for the three frequency conditions.

The analysis revealed that the mean percent /b/ identification for the 0.9 kHz onset stimulus was significantly higher for the normal group at each intensity level (Table 3).

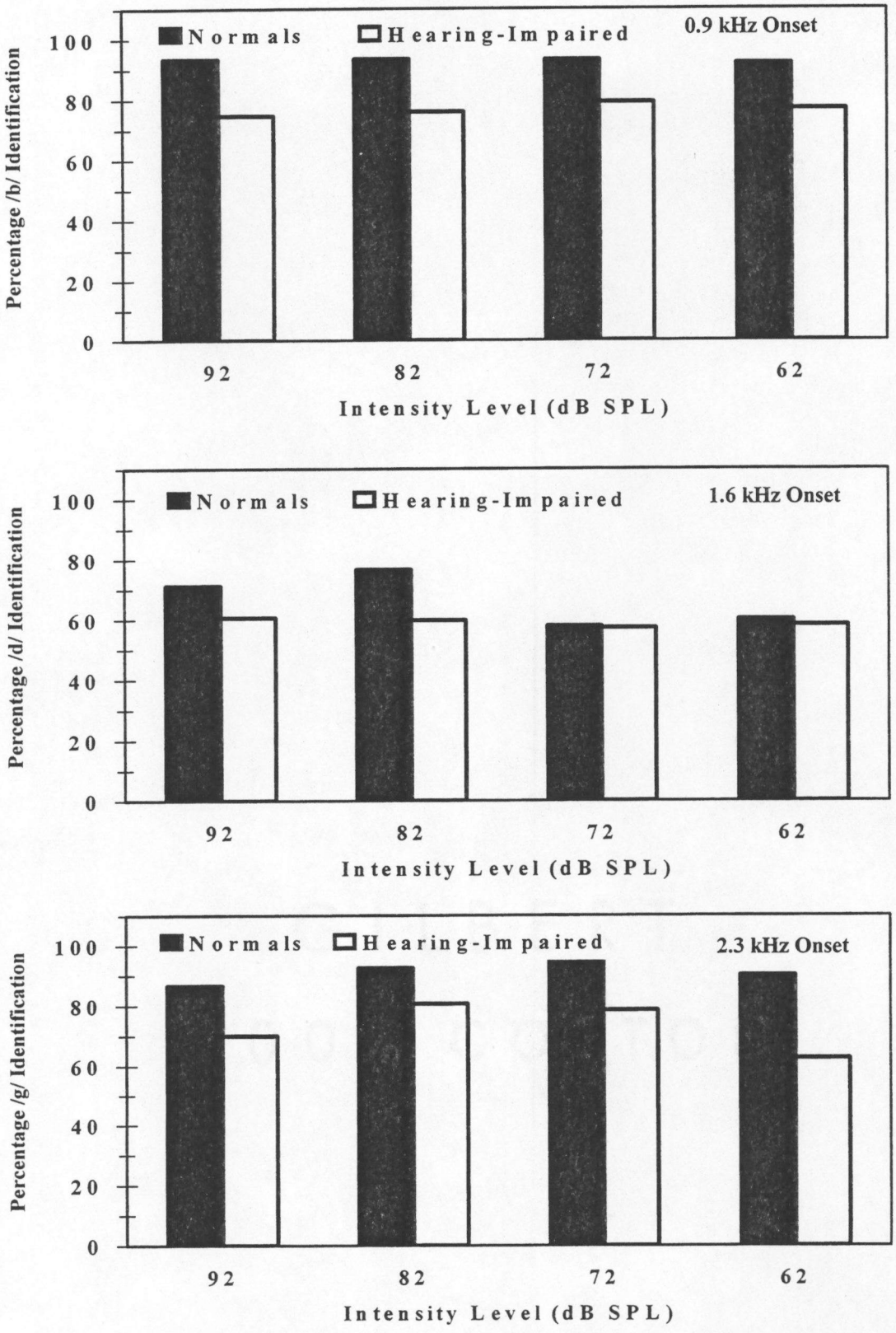


Figure 4: Mean percentage /b, d, g/ identification for the 0.9, 1.6, 2.3 kHz onset stimuli.

Table 3: Summary of mean identification percentages

Level	Normals	Hearing-Impaired	Difference	F-value	p>F
		0.9 kHz onset stimulus			
92	93.437	74.687	18.75	7.32	0.0111
82	93.437	75.937	17.50	6.38	0.0171
72	93.125	79.062	14.06	4.12	0.0514
62	91.875	76.875	15.00	4.68	0.0385
		1.6 kHz onset stimulus			
92	71.25	60.625	10.62	2.46	0 .1274
82	76.562	59.687	16.87	6.20	0.0185
72	57.812	57.187	0.63	0.01	0.9271
62	60.00	58.125	1.87	0.08	0.7839
		2.3 kHz onset stimulus			
92	86.562	70.00	16.56	7.84	0.0088
82	92.5	80.625	11.87	4.03	0.0537
72	94.687	78.437	16.25	7.55	0.0101
62	90	62.5	27.50	21.62	0.0001

The analysis also revealed that the mean percent /d/ identification for the 1.6 kHz onset stimulus was significantly greater for the normal group at 82 dB SPL, however, the mean percent /d/ identification was not significantly different at 62, 72, or 92 dB SPL (Table 3). For the 2.3 kHz onset stimulus, the analysis revealed that the mean percent /g/ identification was significantly greater for the normal group than the hearing-impaired group at each intensity level (Table 3).

Summary of Speech Perception Results

1. Slopes of the percentage /b/ identification functions were steeper for the normal group at 92, 82, and 62 dB SPL but were similar at 72 dB SPL.

2. Slopes of the percentage /d, g/ identification functions were similar for each group at each intensity level.

3. Phonetic boundaries of the /b, d, g/ identification functions were relatively similar for each group at each intensity level.

4. For both groups, stimuli with rising second formant transitions resulted in the highest percentage of /b/ identifications, stimuli with gradually falling second formant transitions resulted in the highest percentage of /d/ identifications, stimuli with sharply falling second formant transitions resulted in the highest percentage of /g/ identifications at each intensity level.

5. Normal hearing listeners identified the 0.9 kHz second formant onset stimulus as /b/ and the 2.3 kHz second formant onset stimulus as /g/ a significantly higher percentage of the time than the hearing-impaired listeners at each intensity level.

6. Each group identified the 1.6 kHz second formant onset stimulus as /d/ a similar percentage of the time at all levels except 82 dB SPL where normal listeners identified /d/ significantly more often.

Experiment II: Electrophysiology

Frequency-following responses were recorded using the 0.9, 1.6, and 2.3 kHz onset stimuli at 92, 82, and 72 dB SPL. A joint time-frequency analysis was performed

on each FFR obtained under each experimental condition in order to evaluate the FFRs ability to represent the time-variant second formant transition. Figure 5 displays a joint time-frequency analysis of an FFR obtained from a normal hearing subject using the 2.3 kHz onset stimulus presented at 92 dB SPL. The bottom panel in Figure 5 displays the differential FFR while the top panel displays the frequency components of the phase-locked activity as a function of time. Inspection of Figure 5 clearly reveals a systematic shift in the frequency of phase-locked activity over time corresponding to the temporal change in frequency of the second formant transition.

Instant spectra were obtained from the joint time-frequency analyses at 10 ms intervals for each subject under each experimental condition. The initial spectral measurement was made at 10 ms post stimulus onset and the final spectral measurement was made at 60 ms post stimulus onset. For each time interval, spectral data corresponding to the frequency region comprising the second formant transition were analyzed for response peak amplitude and frequency. Histograms displaying the response peak amplitude as a function of frequency were generated for each subject at each time interval.

0.9 kHz Onset Stimulus: /ba/

The second formant transition for the 0.9 kHz onset stimulus ranged from 0.9 kHz to 1.25 kHz for a duration of 40 ms. Spectral data less than 0.8 kHz and greater than 1.347 kHz were not included in the analyses. Peak amplitudes and frequencies were determined from the remaining spectral data for each subject at each of the six time intervals. The spectral data from these analyses are presented for the three intensity levels

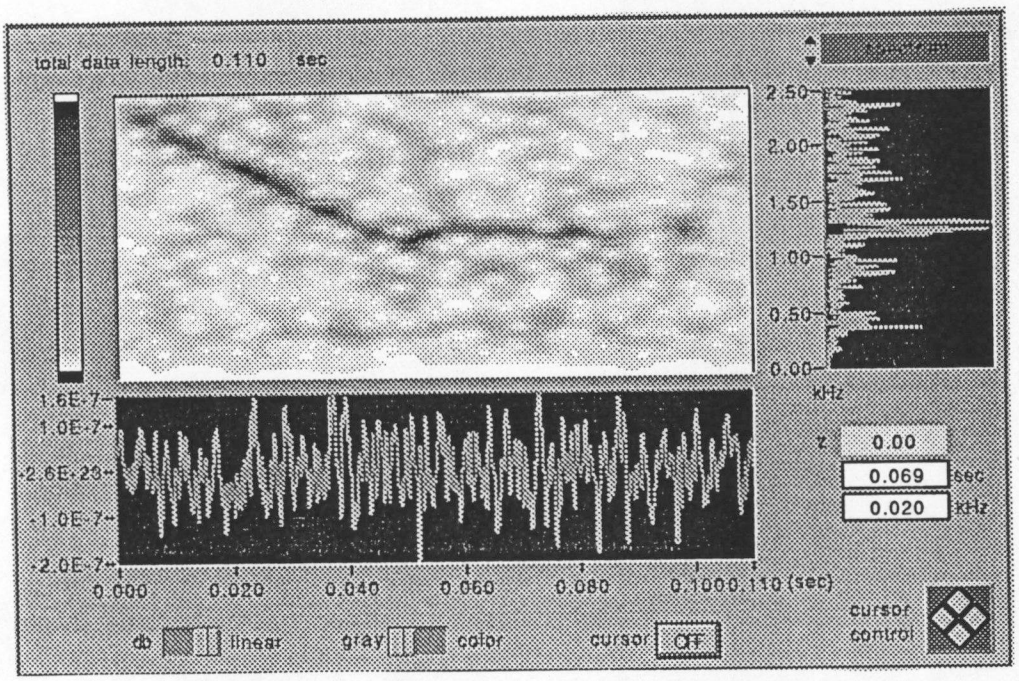


Figure 5: Example joint time-frequency analysis of an FFR obtained for the 2.3 kHz onset stimulus at 92 dB SPL.

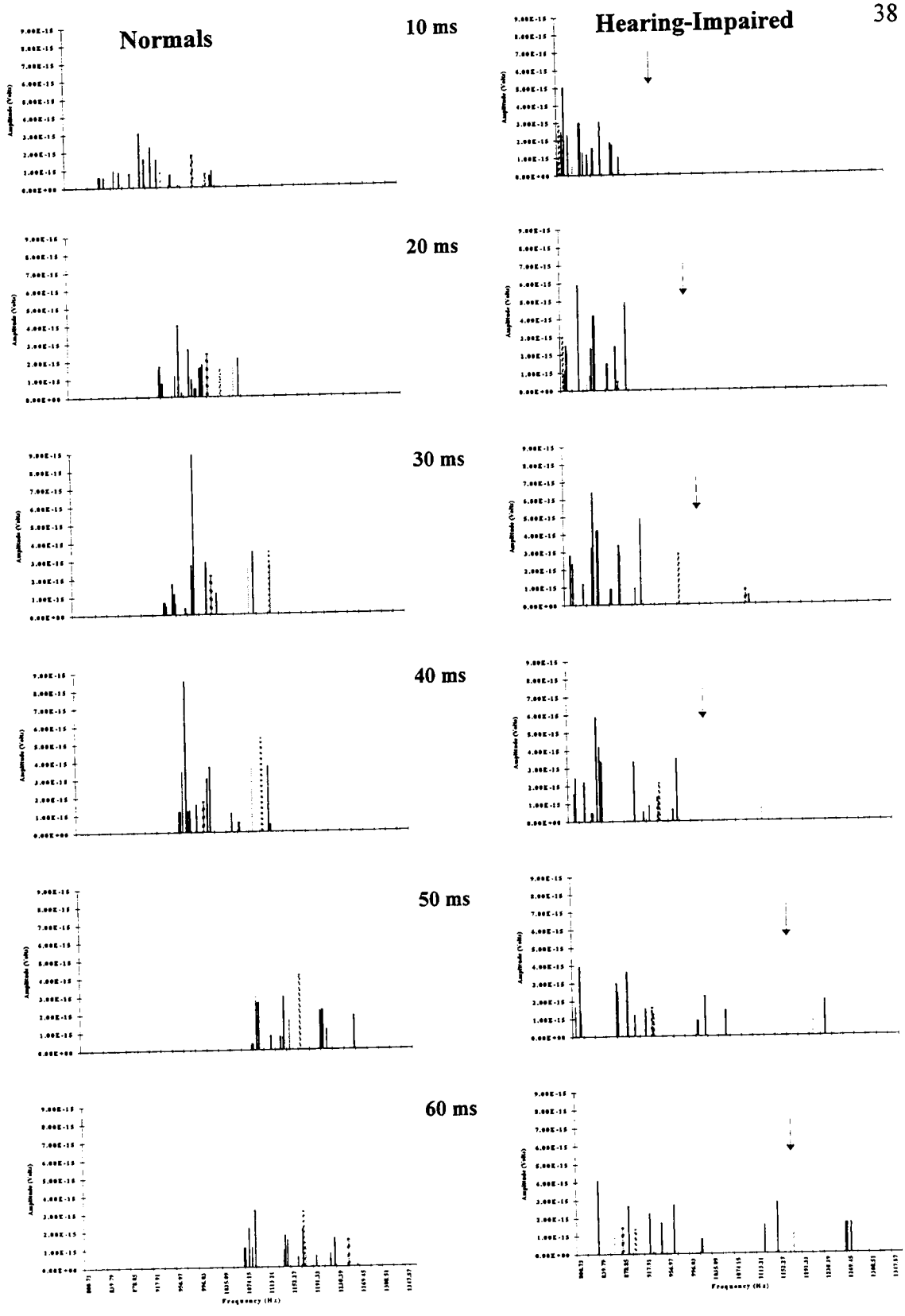


Figure 6: FFR spectral data for the 0.9 kHz onset stimulus at 92 dB SPL.

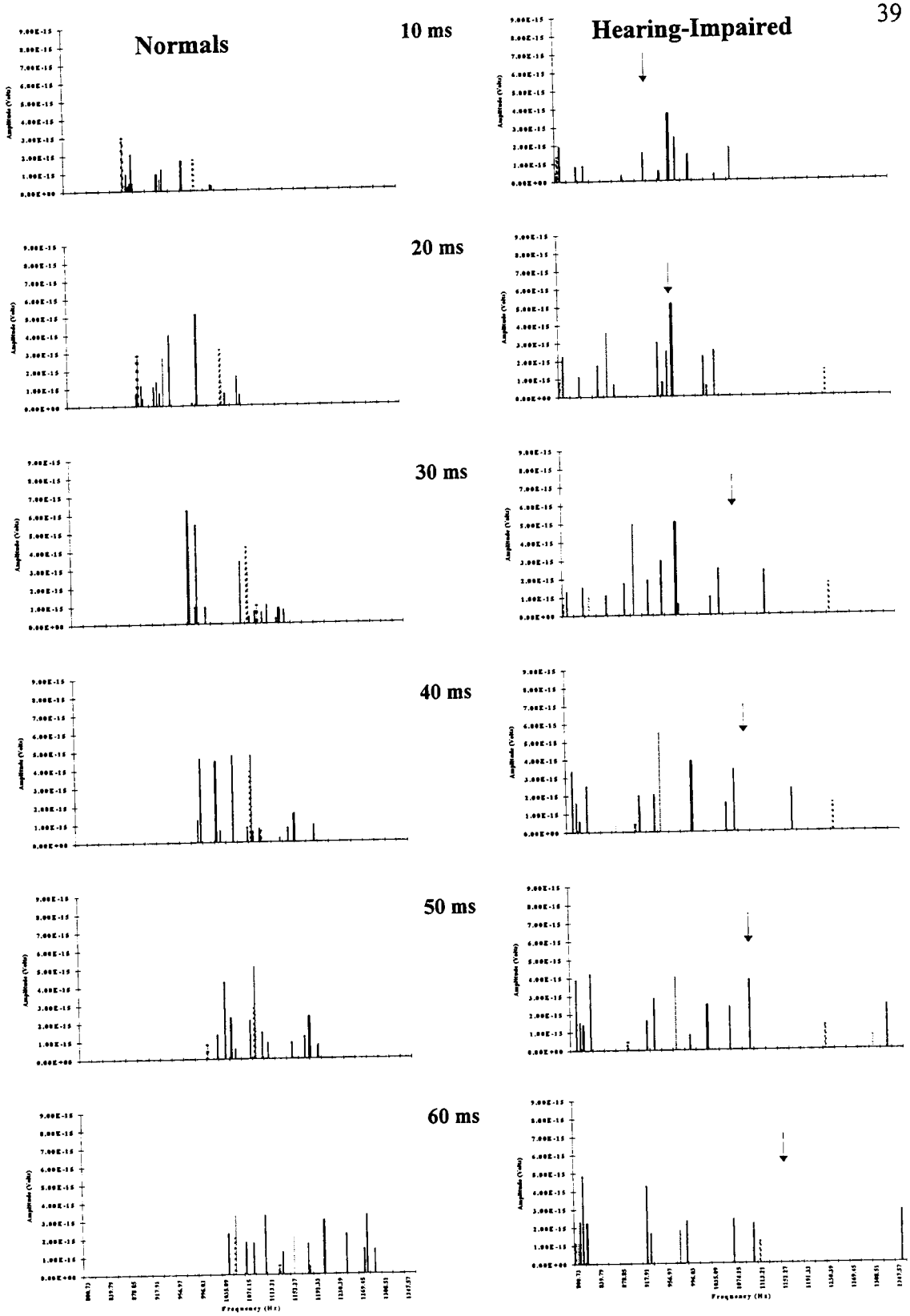


Figure 7: FFR spectral data for the 0.9 kHz onset stimulus at 82 dB SPL.

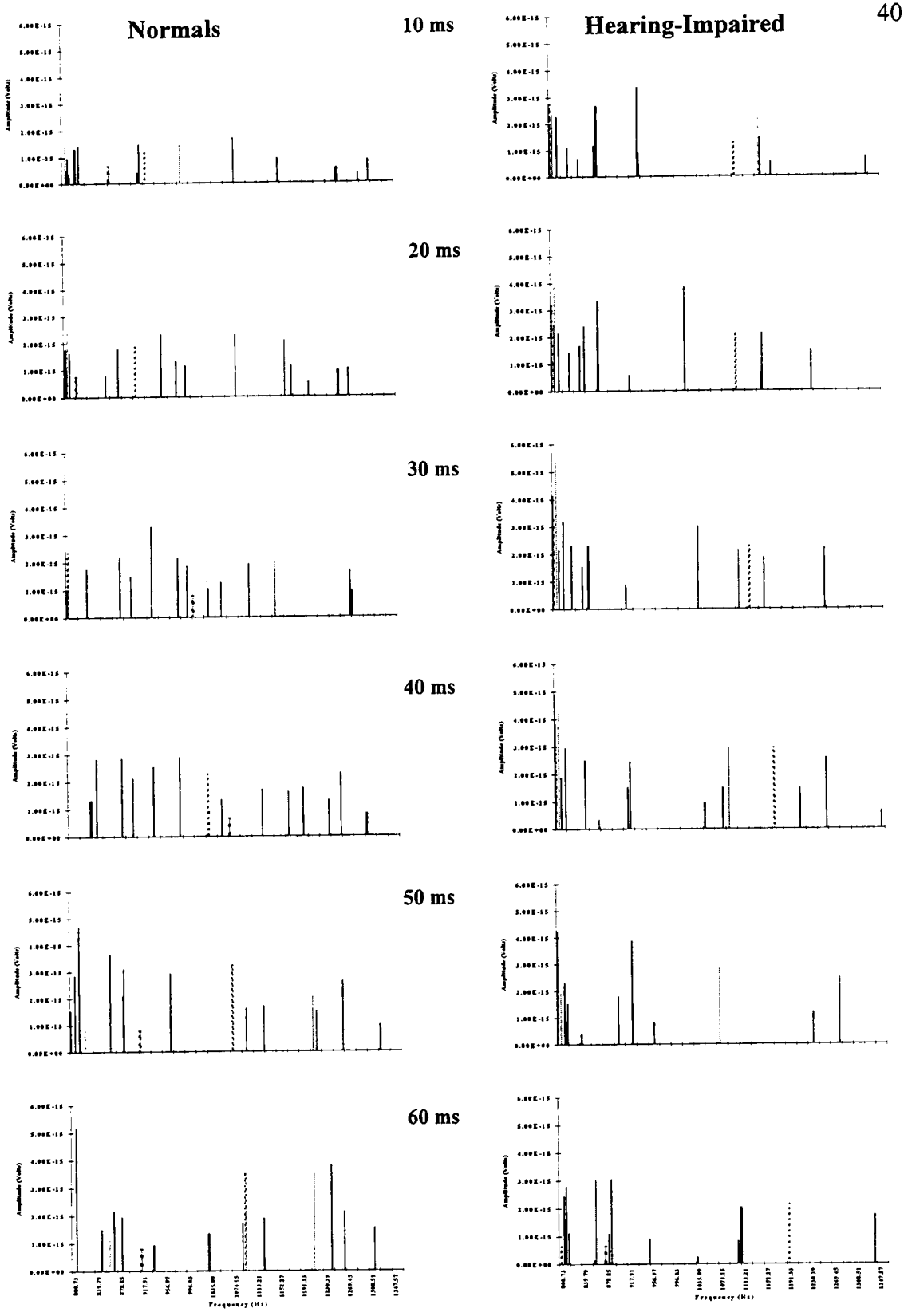


Figure 8: FFR spectral data for the 0.9 kHz onset stimulus at 72 dB SPL.

in Figures 6, 7, and 8.

Figure 6 displays the spectral data from the 92 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group clearly shows that the FFR spectral peaks shift progressively upwards (i.e. to higher frequencies) as the second formant transition rises from 0.9-1.25 kHz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. However, note that some peaks in the hearing-impaired data do approach the mean peak frequency of the normal group (indicated by the arrow).

Figure 7 displays the spectral data from the 82 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group clearly shows that the FFR spectral peaks shift progressively upwards (i.e. to higher frequencies) as the second formant transition rises from 0.9-1.250 kHz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. Once again, note that some peaks in the hearing-impaired data do approach the mean peak frequency of the normal group.

Figure 8 displays the spectral data from the 72 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time reveals no systematic spectral peak shifting for either group across the 0.8-1.347 kHz range.

The mean peak amplitude for each time interval during the second formant transition is plotted in Figure 9 for the 92 and 82 dB SPL data from the normal group.

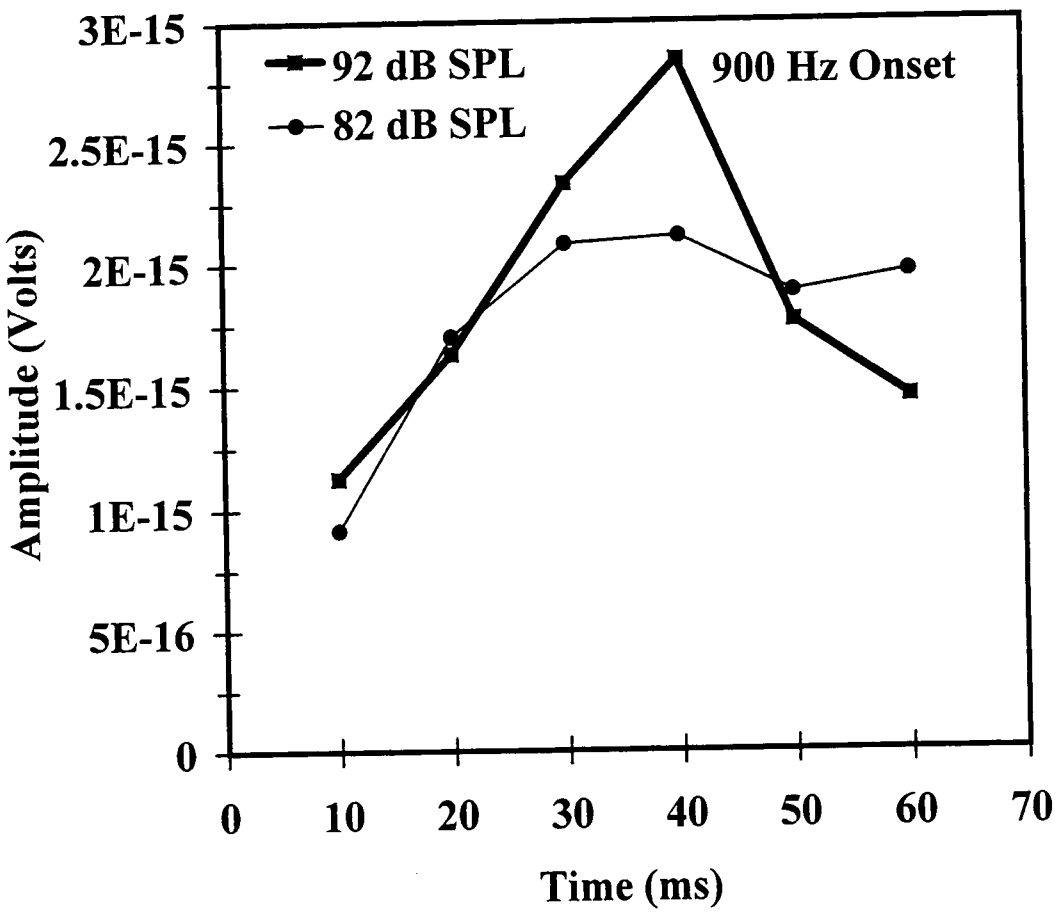


Figure 9: Mean FFR peak amplitudes for the 0.9 kHz onset stimulus at 92 and 82 dB SPL.

Recall that no systematic responses were detected for the normal group at 72 dB SPL or for the hearing-impaired group at any intensity level. In general, this figure indicates that the amplitude of the response increases during the first 40 ms at which point the response amplitude begins to decrease.

1.6 kHz Onset Stimulus: /da/

The second formant transition for the 1.6 kHz onset stimulus ranged from 1.6 kHz to 1.25 kHz for a duration of 40 ms. Spectral data for each time interval less than 1.152 kHz and greater than 1.718 kHz were not included in the analyses. Peak amplitudes and frequencies were determined from the remaining spectral data for each subject at each of the six time intervals. The spectral data from these analyses are presented for the three intensity levels in Figures 10, 11, and 12.

Figure 10 displays the spectral data for the 92 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group clearly shows that the FFR spectral peaks shift progressively downwards (i.e. to lower frequencies) as the second formant transition falls from 1.6-1.25 kHz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. However, note that some peaks in the hearing-impaired data do approach the mean peak frequency of the normal group (indicated by the arrow).

Figure 11 displays the spectral data from the 82 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group clearly shows that the FFR spectral peaks shift progressively downwards (i.e. to lower

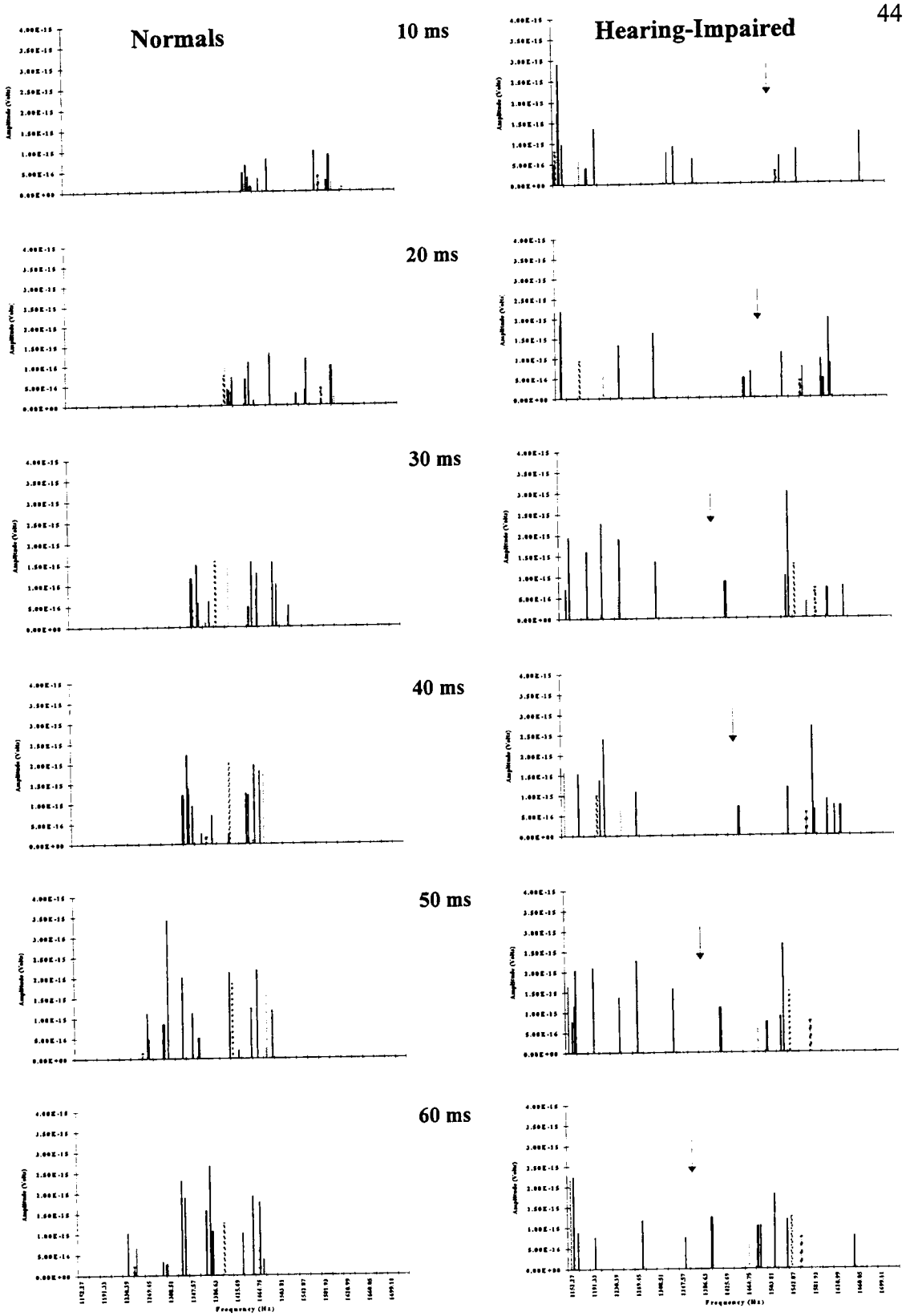


Figure 10: FFR spectral data for the 1.6 kHz onset stimulus at 92 dB SPL.

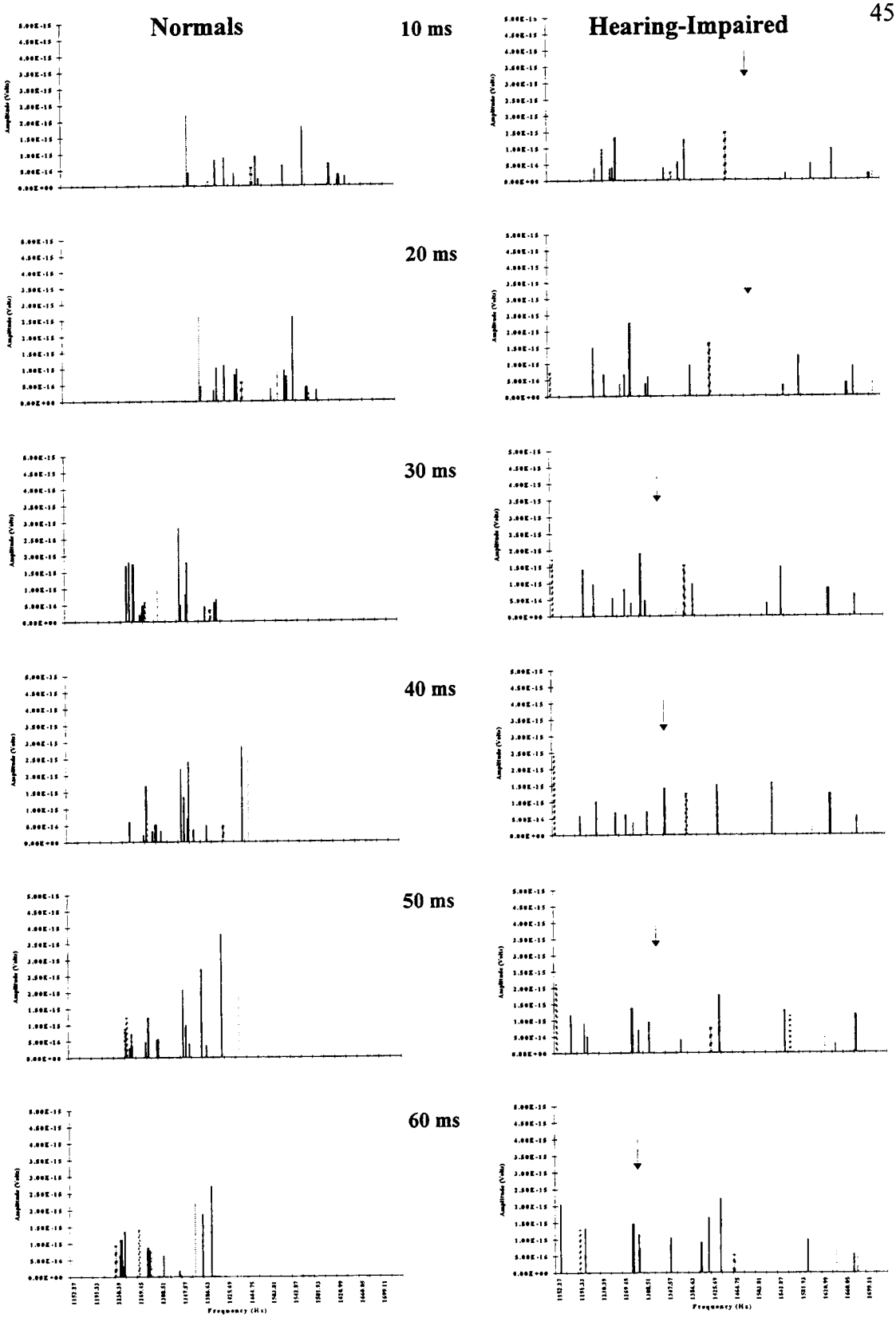


Figure 11: FFR spectral data for the 1.6 kHz onset stimulus at 82 dB SPL.

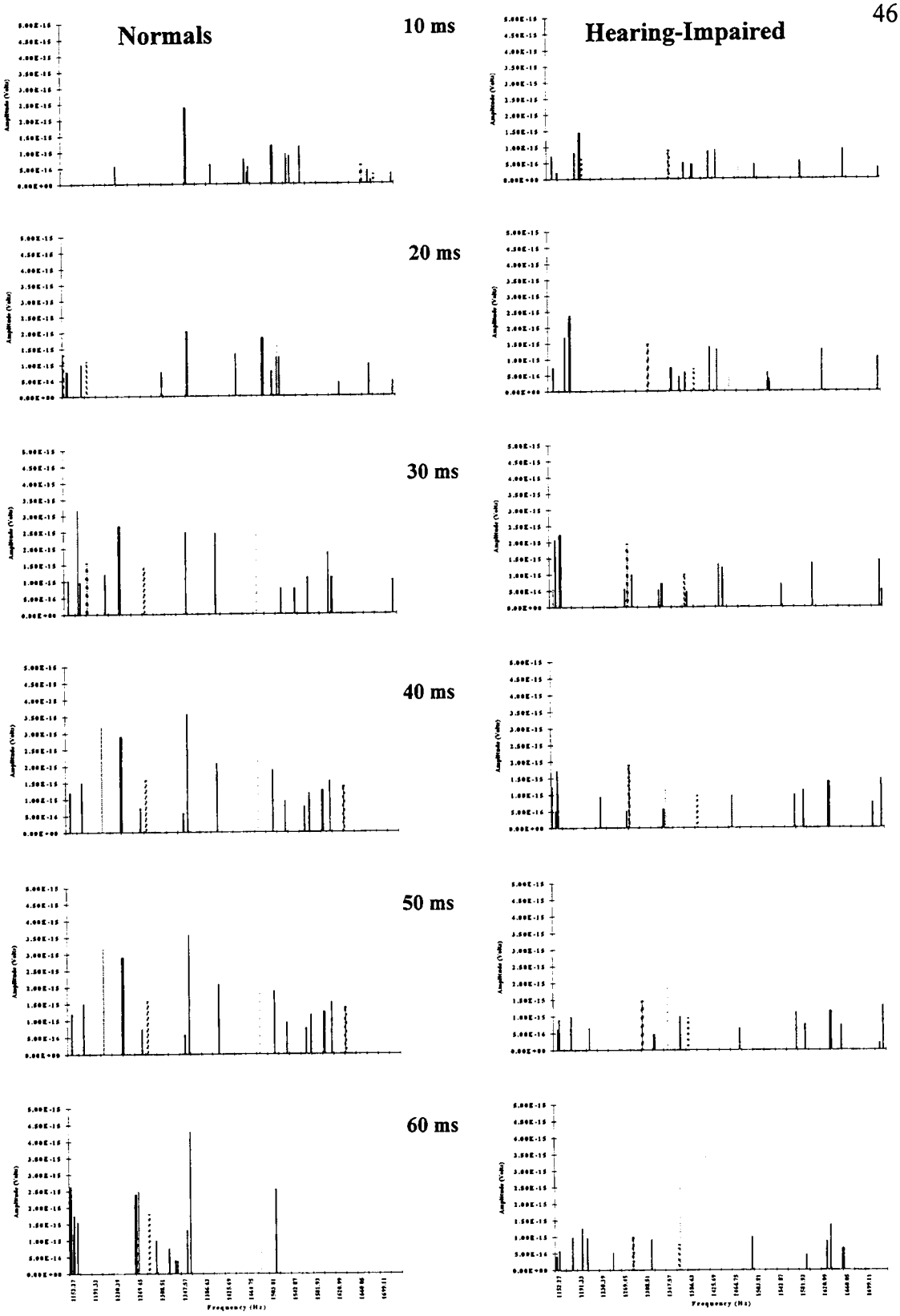


Figure 12: FFR spectral data for the 1.6 kHz onset stimulus at 72 dB SPL.

frequencies) as the second formant transition falls from 1.6-1.250 kHz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. Once again, note that some peaks in the hearing-impaired data do approach the mean response peak frequency of the normal group.

Figure 12 displays the spectral data from the 72 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time reveals no systematic spectral peak shifting for either group across the 1.152-1.718 kHz range.

The mean peak amplitude for each time interval during the second formant transition is plotted in Figure 13 for the 92 and 82 dB SPL data from the normal group. Recall that no systematic responses were detected for the normal group at 72 dB SPL or for the hearing-impaired group at any intensity level. In general, this figure indicates that the amplitude of the response increases as a function of time for each intensity level.

2.3 kHz Onset Stimulus: /ga/

The second formant transition for the 2.3 kHz onset stimulus ranged from 2.3 kHz to 1.25 kHz for a duration of 40 ms. Spectral data for each time interval less than 1.152 kHz and greater than 2.499 kHz were not included in the analysis. Peak amplitudes and frequencies were determined from the remaining spectral data for each subject at each of the six time intervals. The spectral data from these analyses are presented for the three intensity levels in Figures 14, 15, and 16.

Figure 14 displays the spectral data from the 92 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group

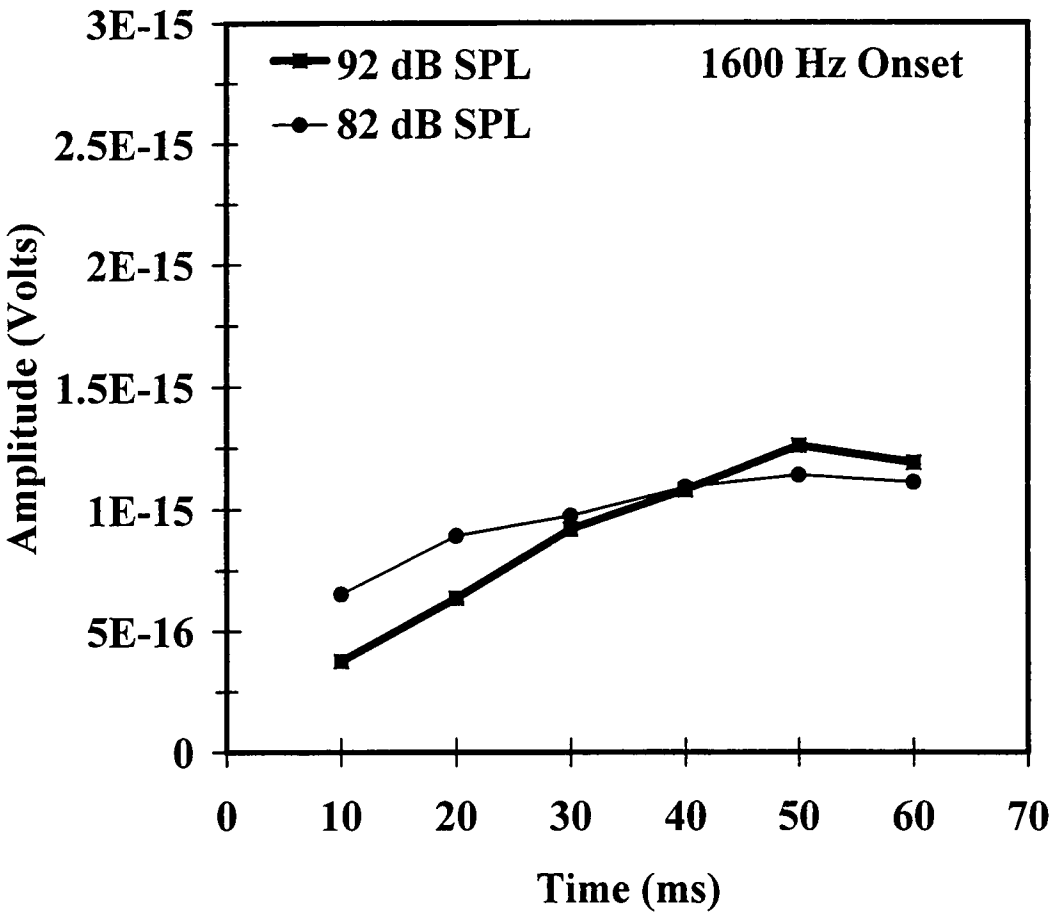


Figure 13: Mean FFR peak amplitudes for the 1.6 kHz onset stimulus at 92 and 82 dB SPL.

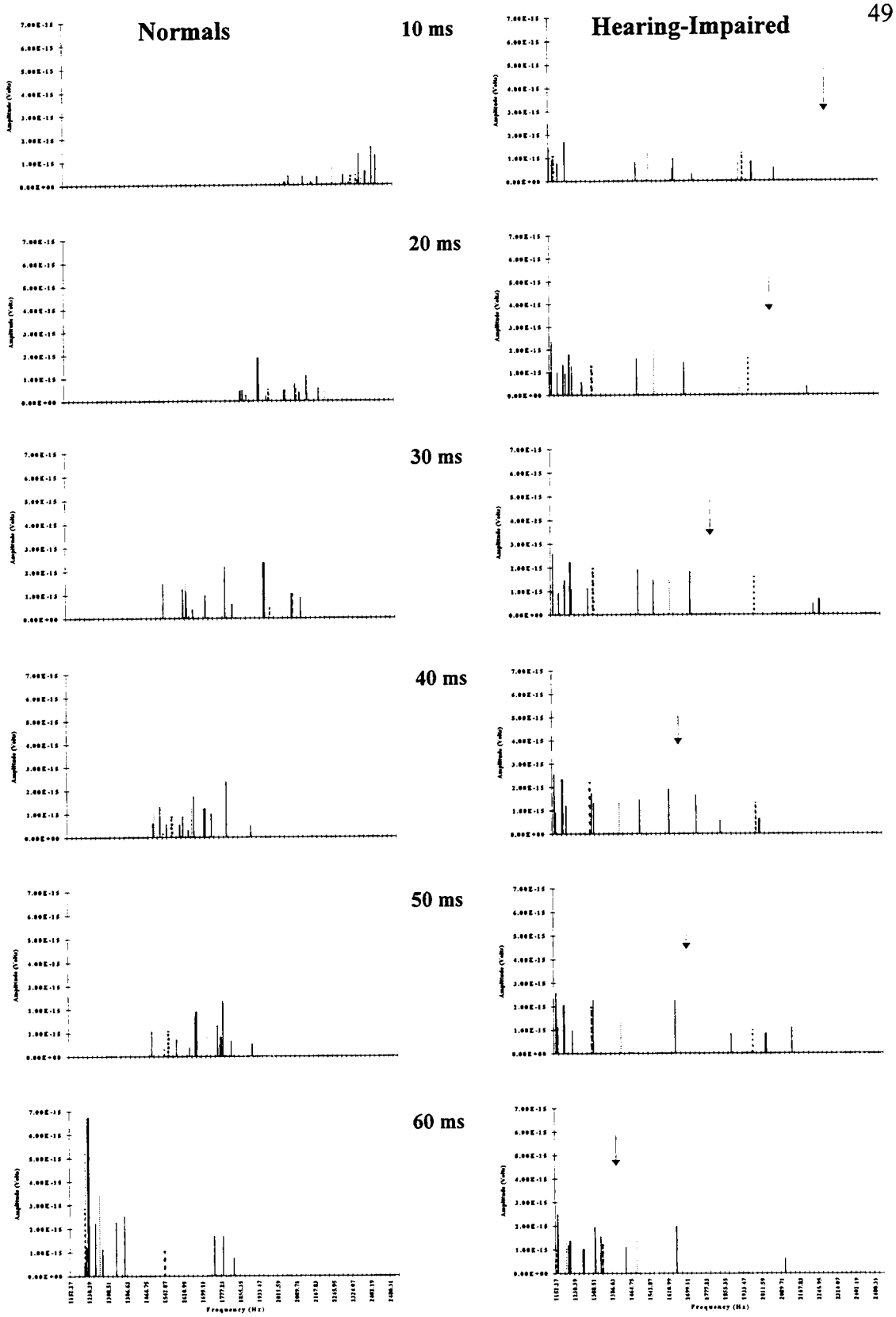


Figure 14: FFR spectral data for the 2.3 kHz onset stimulus at 92 dB SPL.

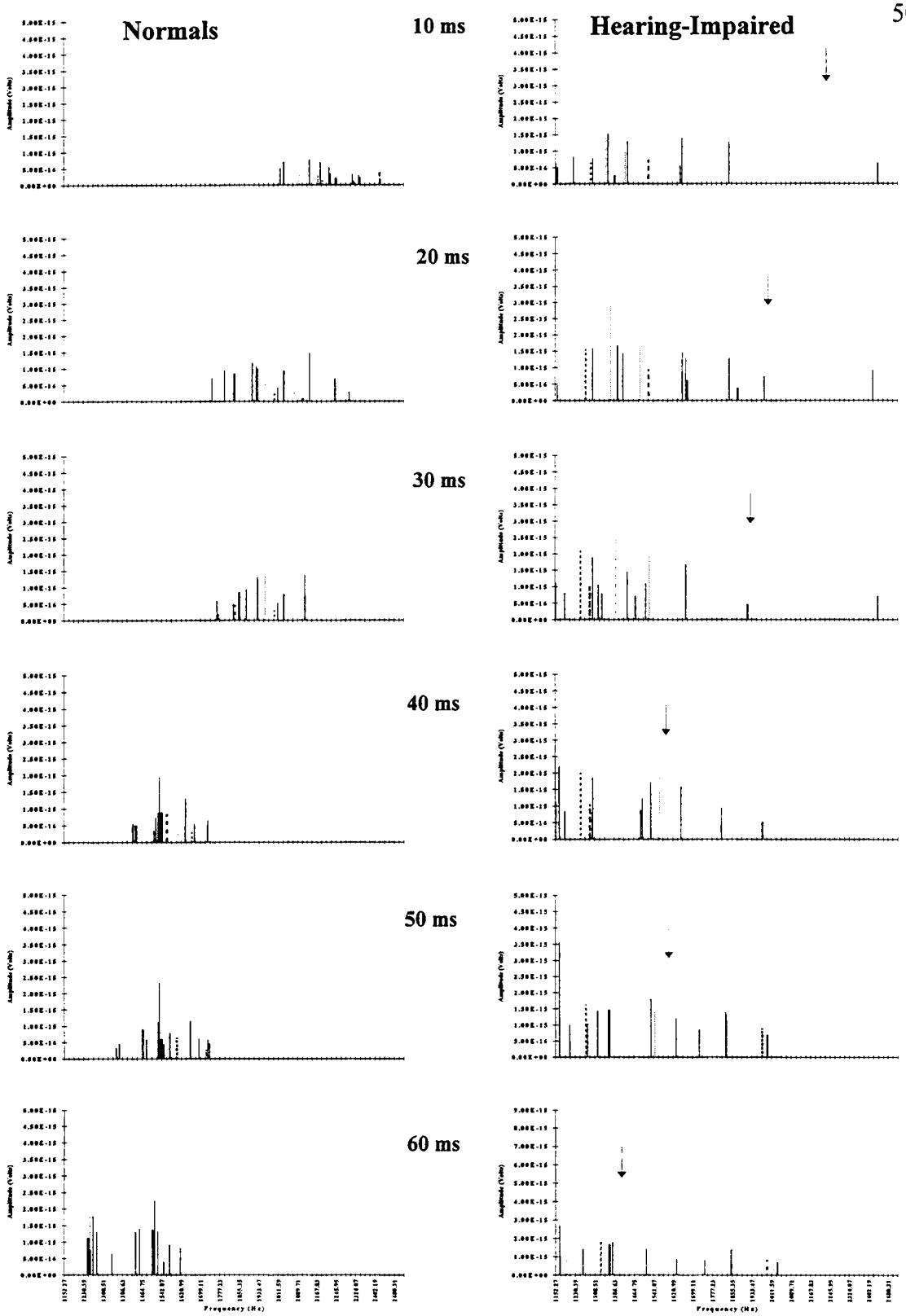


Figure 15: FFR spectral data for the 2.3 kHz onset stimulus at 82 dB SPL.

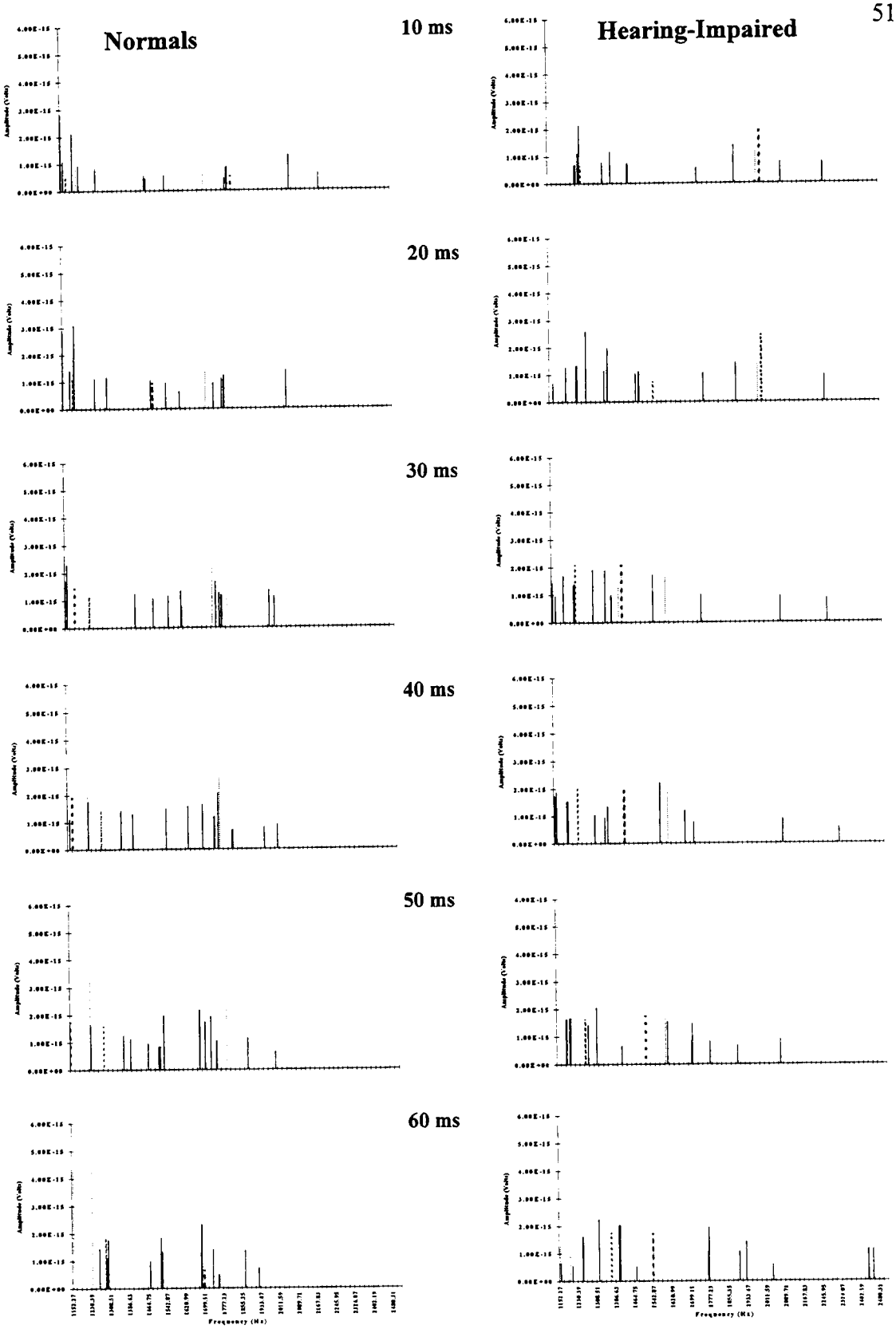


Figure 16: FFR spectral data for the 2.3 kHz onset stimulus at 72 dB SPL.

clearly shows that the FFR spectral peaks shift progressively downwards (i.e. to lower frequencies) as the second formant transition falls from 2.3-1.25 Hz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. However, note that some peaks in the hearing-impaired data do approach the mean peak frequency of the normal group (indicated by the arrow).

Figure 15 displays the spectral data from the 82 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time for the normal group clearly shows that the FFR spectral peaks shift progressively downward (i.e. to lower frequencies) as the second formant transition falls from 2.3-1.25 Hz over a 60 ms duration. In contrast, inspection of the spectral data for the hearing-impaired group does not reveal a systematic shift in the FFR spectral peaks over the 60 ms duration. Once again, note that some peaks in the hearing-impaired data approach the mean peak frequency of the normal group.

Figure 16 displays the spectral data from the 72 dB SPL stimulus for each group at each time interval. Inspection of the spectral data across time reveals no systematic spectral peak shifting for either group across the 1.152-2.499 kHz range.

The mean peak amplitude for each time interval during the second formant transition is plotted in Figure 17 for the 92 and 82 dB SPL data from the normal group. Recall that no systematic responses were detected for the normal group at 72 dB SPL or for the hearing-impaired group at any intensity level. In general, this figure indicates that the amplitude of the response increases as a function of time for each intensity level.

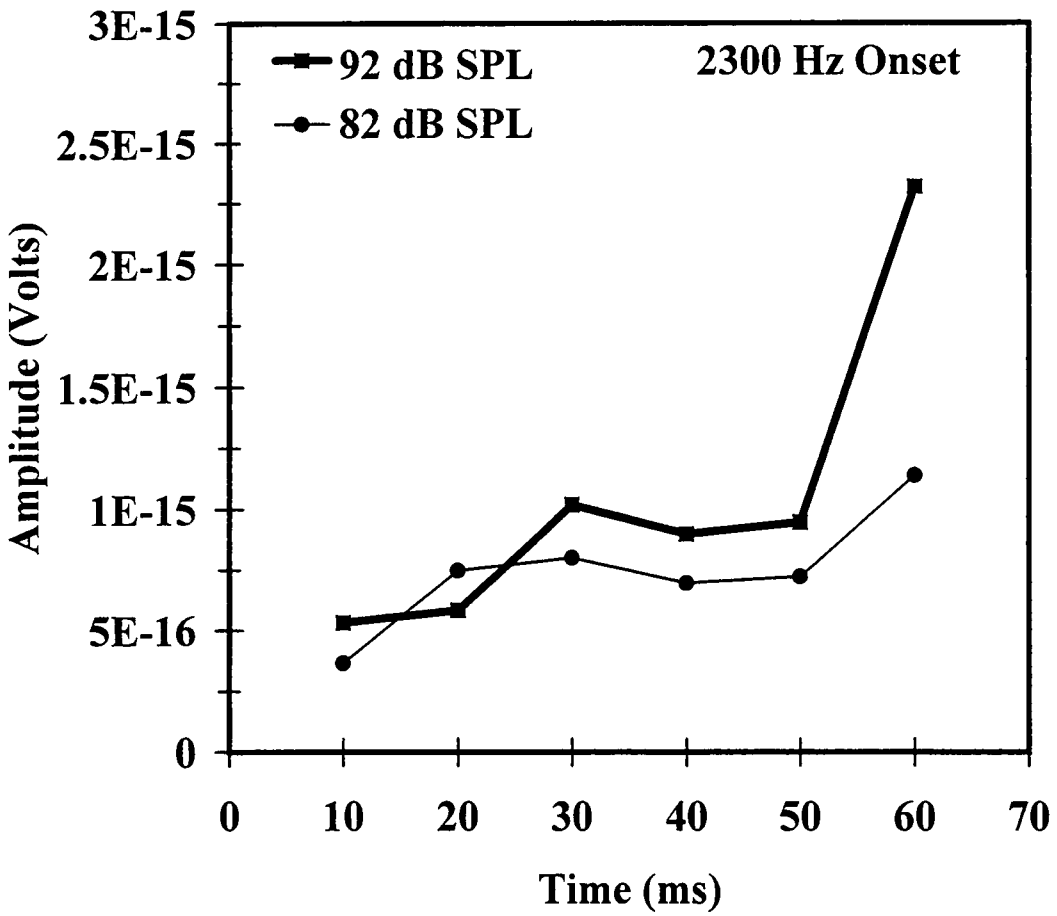


Figure 17: Mean FFR peak amplitudes for the 2.3 kHz onset stimulus at 92 and 82 dB SPL.

Summary of Electrophysiology Results

1. FFR spectral peaks shift with the second formant transition of the 0.9, 1.6, and 2.3 kHz onset stimuli at 92 and 82 dB SPL in normal hearing listeners.

2. FFR spectral peaks do not shift with the second formant transition of the 0.9, 1.6, and 2.3 kHz onset stimuli at 72 dB SPL in normal hearing listeners.

3. FFR spectral peaks do not shift with the second formant transition of the 0.9, 1.6, and 2.3 kHz onset stimuli at 92, 82, or 72 dB SPL in hearing-impaired listeners.

4. For normal hearing listeners, FFR amplitude increases as a function of time with the exception of responses to the 0.9 kHz onset stimulus, for which amplitude increases during the first 40 ms at which point amplitude begins to decrease.

5. For normal hearing listeners, FFR amplitude increases as stimulus intensity increases with the exception of responses to the 1.6 kHz onset stimulus, for which amplitude was relatively similar for each intensity level.

6. FFR response data for individual hearing-impaired listeners approached normal values for various conditions.

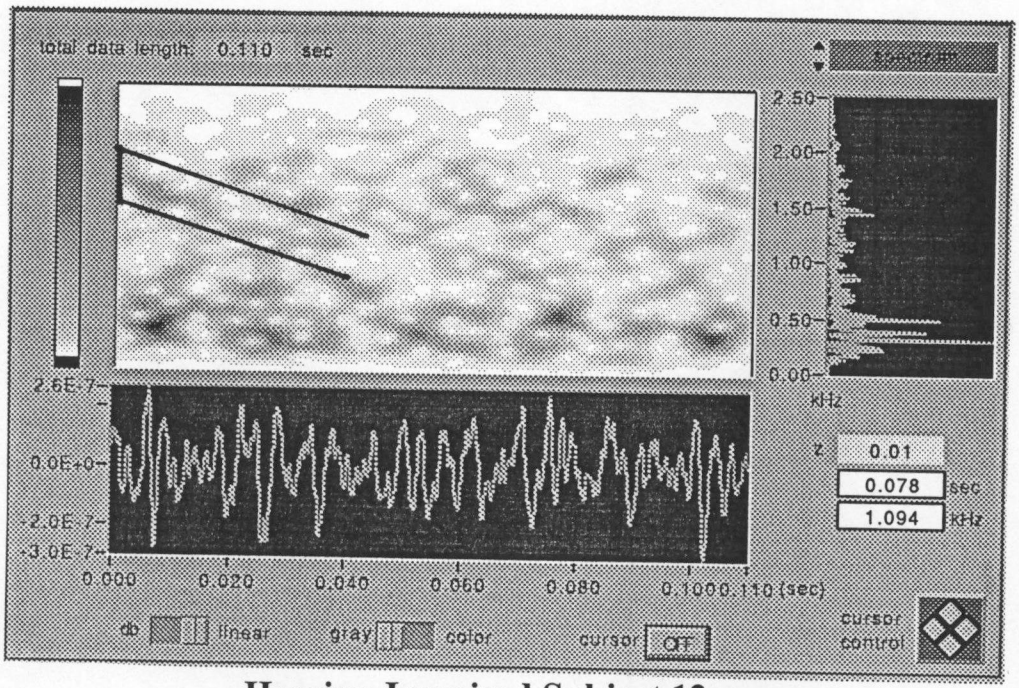
Relationship Between Encoding and Identification

Results of the present study indicate that hearing-impaired listeners do not identify stop consonant place of articulation as accurately as normal listeners for stimuli with second formant onset frequencies of 0.9, 1.6 and 2.3 kHz presented at 82 and 92 dB SPL. In addition, the FFR of most hearing-impaired listeners revealed degraded encoding of the time-variant second formant transitions of the above mentioned stimuli. These results suggest that accurate identification of stop consonant place of articulation is dependant on the auditory systems ability to phase-lock to the time-variant second formant transition.

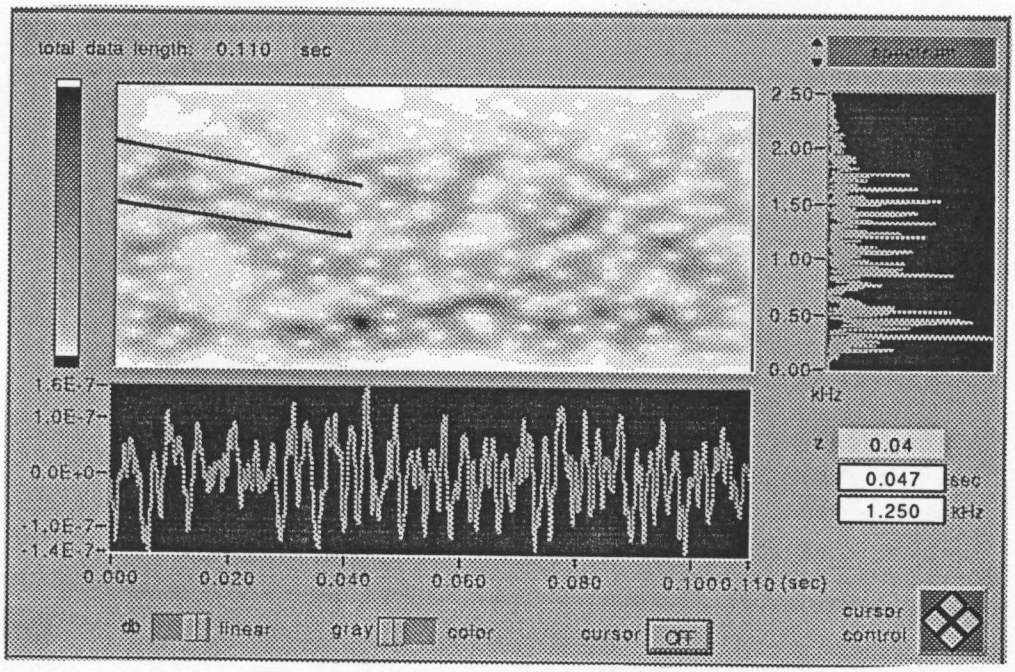
To further investigate the relationship between identification and encoding, consider the electrophysiologic and identification data from hearing-impaired subjects 12, 14, 16, and 5. Hearing-impaired subjects 12 and 14 demonstrated neural activity phase-

locked to the second formant transition of the 1.6 kHz onset stimulus presented at 92 dB SPL (Figure 18). Evaluation of their identification data revealed that the 1.6 kHz onset stimulus presented at 92 dB SPL was identified as /d/ 30-35% more often than the hearing-impaired group and 19-24% more often than the normal group (Figure 20). The FFR of hearing-impaired subject 16 demonstrated neural activity phase-locked to the second formant transition of the 2.3 kHz onset stimuli presented at 92 dB SPL (Figure 19). Evaluation of this listener's identification data revealed that the 2.3 kHz onset stimulus presented at 92 dB SPL was identified as /g/ 30 % more often than the hearing-impaired group and 13% more often than the normal group (Figure 20).

The FFR of hearing-impaired subject 5 to the 1.6 kHz onset stimulus presented at 92 dB SPL is displayed in Figure 21. Inspection of Figure 21 indicated that, although neural activity phase-locked to the first formant transition shifted in frequency over time, neural activity phase-locked to the second formant transition did not shift in frequency over time. Instead, the frequency of the phase-locked neural activity remained relatively constant throughout the duration of the second formant transition. Furthermore, evaluation of this hearing-impaired subject 5's identification data revealed that the 1.6 kHz onset stimulus presented at 92 dB SPL was identified as /d/ 16% more often by the hearing-impaired group and 26% more often by the normal group (Figure 22). These results suggest that listeners who demonstrated neural activity phase-locked to the second formant transition identified stop consonant place of articulation more accurately than listeners who demonstrated a degradation in phase-locking to the second formant transition.



Hearing-Impaired Subject 12



Hearing-Impaired Subject 14

Figure 18: Joint time-frequency analyses of FFRs obtained from hearing-impaired subjects 12 and 14 to the 1.6 kHz onset stimulus at 92 dB SPL.

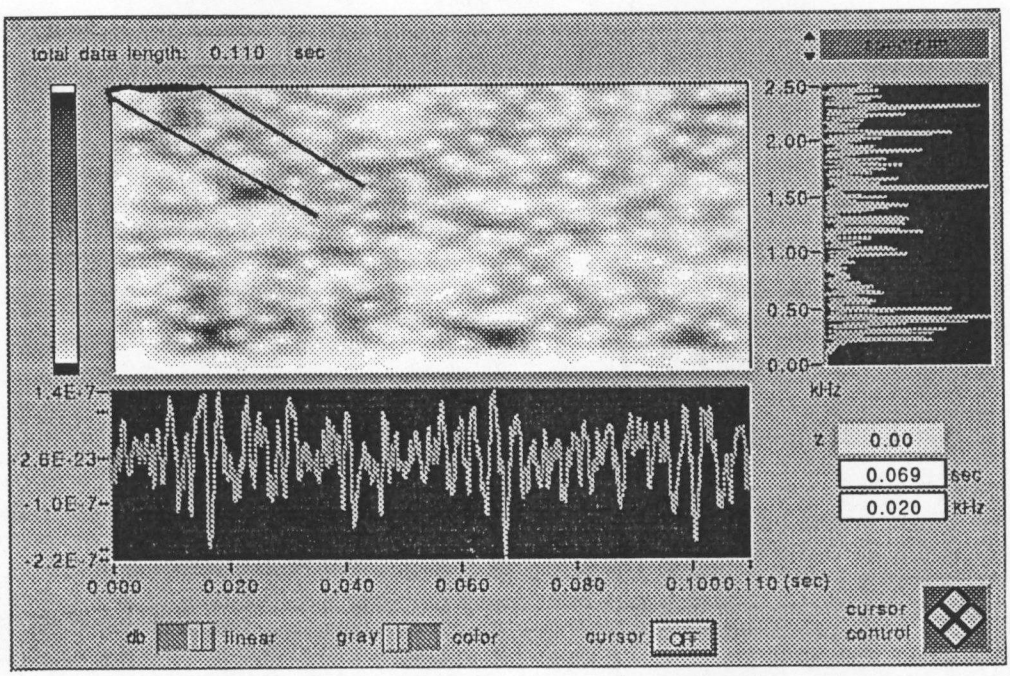
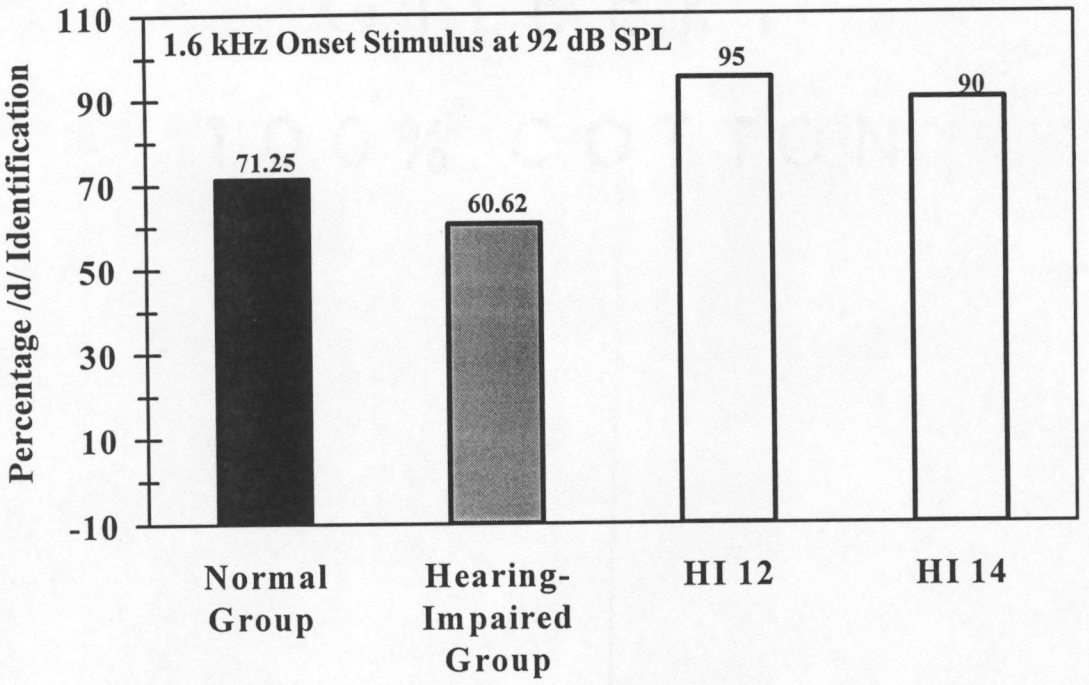


Figure 19: Joint time-frequency analysis of the FFR obtained from hearing-impaired subject 16 to the 2.3 kHz onset stimulus at 92 dB SPL.



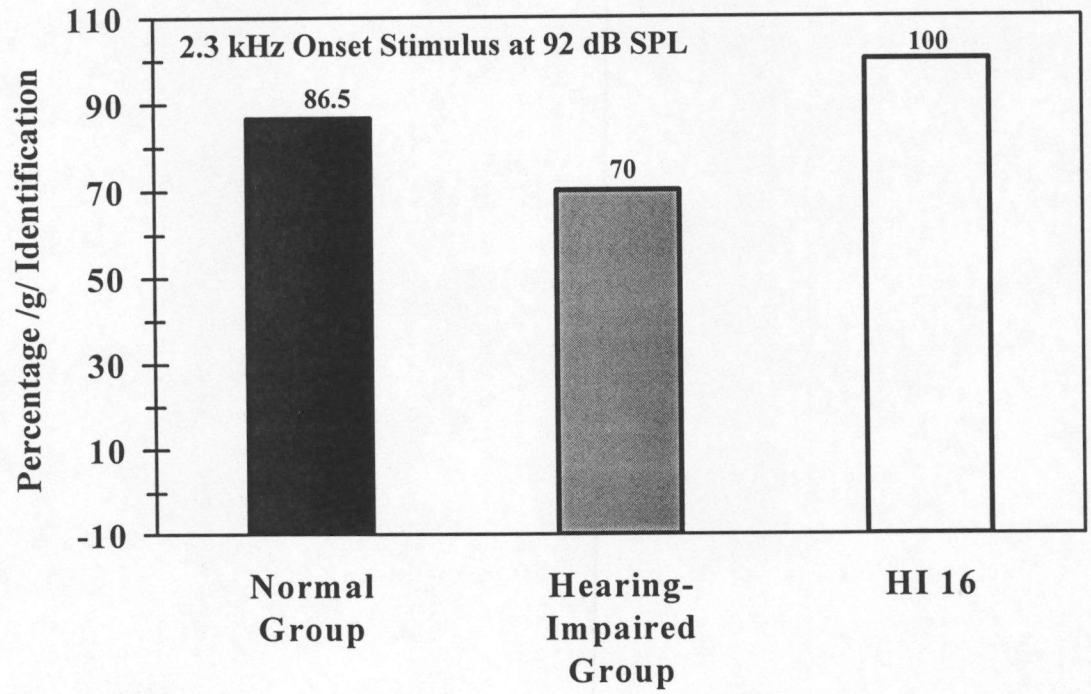


Figure 20: Psychoacoustic data for hearing-impaired subjects 12, 14 and 16.

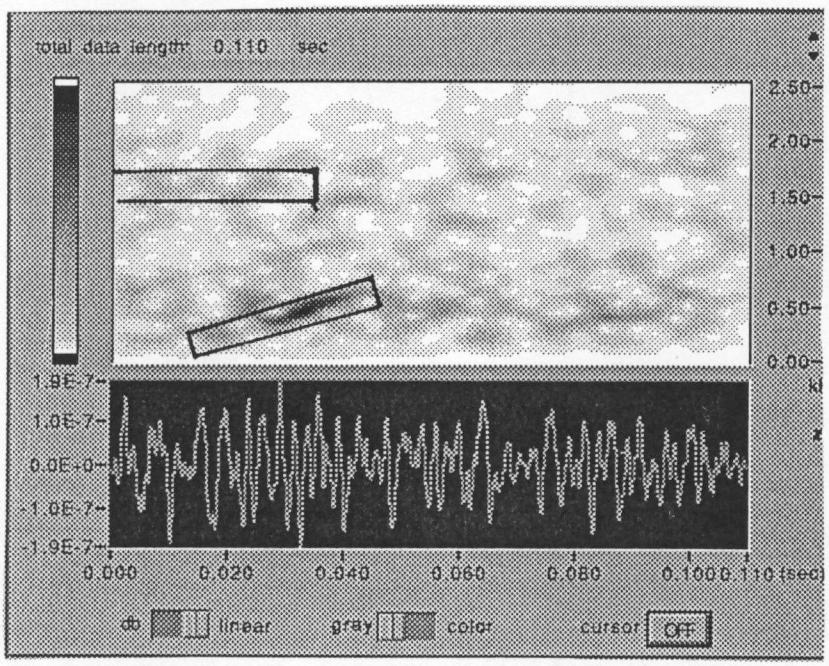


Figure 21: Joint time-frequency analysis of the FFR obtained from hearing-impaired subject 5 to the 1.6 kHz onset stimulus at 92 dB SPL.

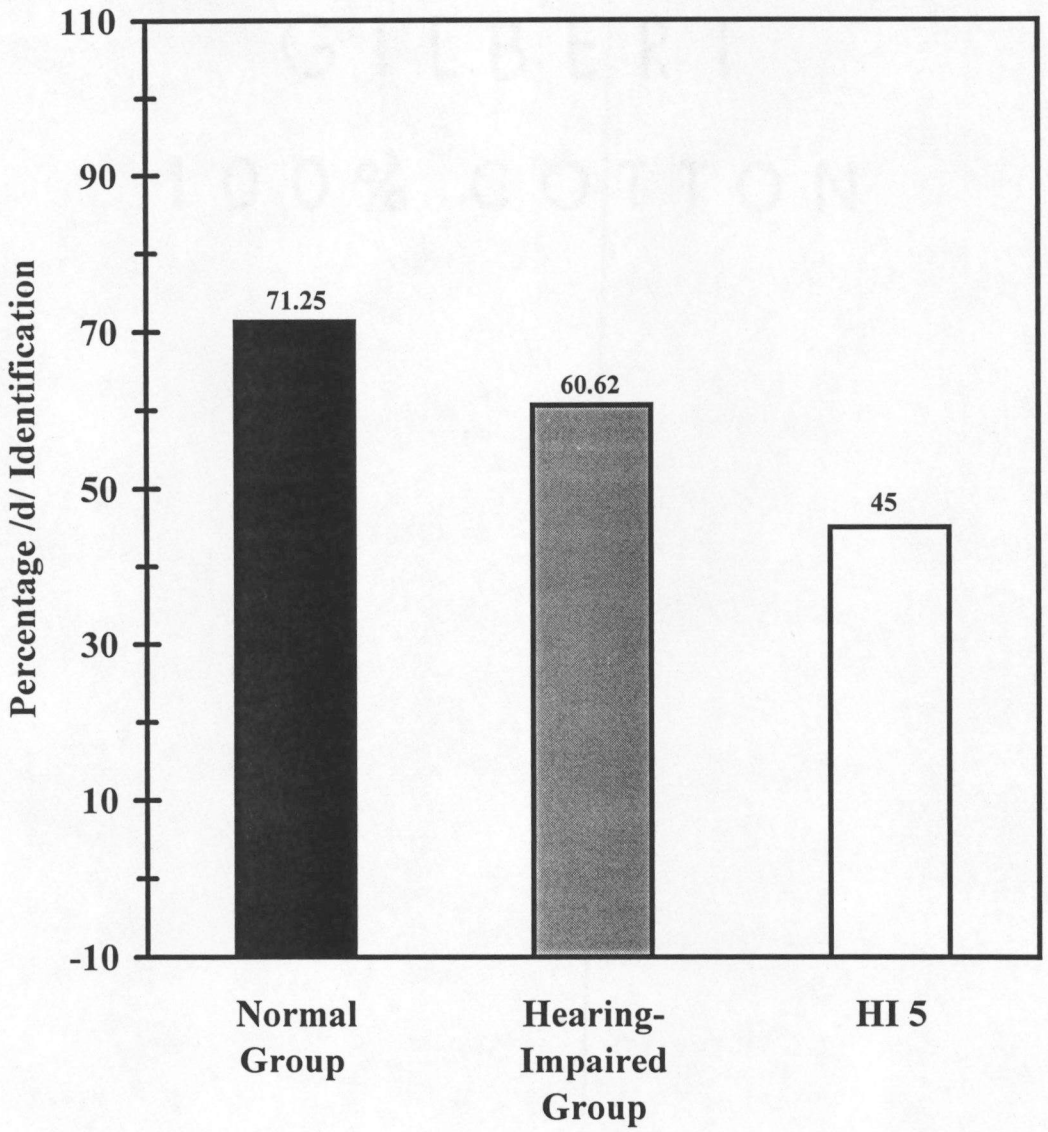


Figure 22: Identification data for hearing-impaired subject 5.

CHAPTER V

Discussion

Experiment I: Speech Perception

Slope and Phonetic Boundary

The present study evaluated the effects of varying the onset frequency of the second formant and the overall intensity level of the synthetic speech stimuli on the identification of stop consonant place of articulation in normal and hearing-impaired listeners. Slopes of the hearing-impaired listener's identification functions were similar to those of the normal group for the identification of /d/ and /g/ but were more shallow for the identification of /b/. For the identification of /d/ and /g/, the inclusion of the third and fourth formants in the present study provided hearing-impaired listeners with the relatively static spectral information which was omitted in the two-formant synthesis used by Dorman et al. (1985). Slopes of the identification functions for /d/ and /g/ which were steeper for normal hearing listeners than for hearing-impaired listeners for the two-formant synthesis (Dorman et al., 1985) were similar when the static spectral information provided by the third and fourth formants was available to the hearing-impaired listeners. The results of the present study suggest the static spectral information provided by the third and fourth formants had a greater effect on the identification of /d/ and /g/ in hearing-impaired listeners than in normal hearing listeners.

For the identification of /b/, the harmonics of the first formant and the fundamental frequency could have induced an upward spread of energy along the basilar membrane which effectively masked the low frequency portion of the rising second

formant transitions (0.9-1.1 kHz). Greater masking of rising second formant transitions would have occurred in the hearing-impaired listeners than in the normal hearing listeners because of the broader auditory filters and reduced frequency selectivity associated with cochlear hearing loss. Therefore, the acoustic cue for the identification of /b/ would have been less audible to hearing-impaired listeners than to normal hearing listeners. The reduced audibility of rising second formant transitions in hearing-impaired listeners could explain the shallower slopes seen in the percentage /b/ identification functions.

Phonetic boundaries of the /b, d, g/ identification functions were also determined for each group at each intensity level. The differences in the phonetic boundaries of each group at each intensity level were relatively small. These results suggest that the frequency at which the phonetic boundary occurs for given intensity level is minimally effected by sensorineural hearing-impairment. However, the phonetic boundaries were effected by changing the intensity level of the stimuli. The phonetic boundaries tended to decrease in frequency as the stimuli increased in intensity level. Although stimulus effects were seen for each group, the effects were typically greater for the normal hearing listeners.

Onset Frequency

Identification data were analyzed across group and intensity level to determine which second formant onset frequency corresponded to the highest percentage identification for each stop consonant. Results indicated that stimuli with rising second formant transitions resulted in the greatest percentage /b/ responses (0.9-1.0 kHz), stimuli with gradually falling second formant transitions resulted in the greatest percentage /d/

responses (1.5-1.6 kHz), and stimuli with sharply falling second formant transitions resulted in the greatest percentage /g/ responses (2.2-2.3 kHz). These results were expected and were consistent with the findings reported by Liberman et al. (1954).

Van Tasell et al. (1982) reported that stimuli presented at lower intensity levels could result in substantially poorer hearing-impaired identification ability because necessary spectral regions of the stimuli would be below listeners' thresholds. Van Tassel et al. (1982) also suggested the presence of hearing loss did not appear to affect listeners ability to identify stimuli given that the necessary spectral information was presented at suprathreshold levels.

In the present study, comparisons of the mean percentage /b/ and /g/ identification for the 0.9 and 2.3 kHz onset stimuli indicated reduced identification in the hearing-impaired listeners at each intensity level. For the 1.6 kHz onset stimulus, comparisons of the mean percentage /d/ identification at 62 and 72 dB SPL were similar for each group of listeners. However, hearing-impaired listeners identification of the 1.6 kHz onset stimulus was reduced at 82 dB SPL and approached abnormality at 92 dB SPL ($p = .12$). Stated differently, the presence of hearing loss did not reduce the hearing-impaired listeners' ability to identify the 1.6 kHz stimulus as /d/ at lower intensity levels but did reduce the identification ability of stimuli presented at higher intensity levels. The presence of hearing loss also reduced the hearing-impaired groups' ability identify the 0.9 kHz stimulus as /b/ and the 2.3 kHz stimulus as /g/ regardless of the intensity level.

The results outlined above are not consistent with the findings reported by Van Tasell et al. (1982). Increasing the overall intensity level of the stimulus did not improve

the identification of stop consonant place of articulation in the hearing-impaired listeners. This was evidenced by the hearing-impaireds' relatively small improvement in stop consonant identification as a function of stimulus intensity level (Figure 5). In fact, the only condition that resulted in similar identification of stop consonant place of articulation for each group were those in which the acoustic cue for /d/ was presented at lower intensity levels. For these conditions, percentage /d/ identification in the normal listeners decreased and approached the identification percentage of the hearing-impaired listeners. This result was consistent with that of Dorman and Dougherty (1981) who also found a reduction in percentage /d/ identification as stimulus intensity levels were reduced.

The hearing-impaired listeners lack of systematic identification improvement with increasing stimulus intensity level may be related to the manner in which the stimuli were amplified. With each increase in overall stimulus intensity, the relative difference in sound pressure level between the first and second formants remained constant. Therefore, increasing the overall level of the stimulus would not have improved the audibility of the second formant relative to that of the first formant. Stated differently, the lack of systematic identification improvement in the hearing-impaired listeners may be a result of increasing the intensity level of the entire stimulus rather than increasing the level of the second formant alone.

Another possible explanation for the lack of systematic identification improvement with increasing stimulus intensity level could be that listeners with sensorineural hearing loss have inherent noise in their perceptual system (Jesteadt, Neff,

Humes, & Leek, 1997). Such an inherent noise would have further reduced the audibility of the second formant transition in the hearing-impaired listeners thereby making the discrimination of /b/, /d/, and /g/ more difficult.

In summary, results of the speech perception experiment indicate hearing-impaired listeners categorize /d/ and /g/ as well as normal hearing listeners but have more difficulty categorizing /b/ than normal hearing listeners. Results also indicate that each group used the same onset frequencies as their primary cues for the identification of /b, d, g/, however, hearing-impaired listeners did not utilize these primary cues as well as normal hearing listeners. In addition, the identification of /b, d, g/ in the hearing-impaired listeners did not systematically increase with each increase in stimulus intensity level.

Experiment II: Electrophysiology

Encoding of Time-Variant Speech

Frequency-following responses were recorded using the 0.9, 1.6, and 2.3 kHz onset stimuli at 92, 82, and 72 dB SPL. Results from the present study indicate that the human FFR is sufficiently dynamic to encode time varying formant transitions under certain conditions. The response characteristics of the FFR to synthetic speech stimuli will be discussed in terms of intensity encoding and frequency encoding.

Intensity Encoding: Amplitude of the FFR

The response amplitude of an evoked potential is directly proportional to the number of neural elements activated in a synchronous fashion. A decrease in intensity results in fewer neural elements phase-locked to the stimulus, therefore, the overall

amplitude of the sustained response decreases as well. In the present study, FFR amplitude varied with intensity for all stimuli with the exception of the 1.6 kHz stimulus in which the response amplitude remained relatively stable. In addition, decreasing the intensity level of the stimulus resulted in an increase in response variability. Examination of the responses from the normal group revealed an increase in response variability as the stimulus intensity level decreased from 82 to 72 dB SPL. It is possible the FFR threshold for some of the normal hearing listeners was between 82 and 72 dB SPL for these synthetic speech stimuli. Stimuli presented below the FFR threshold of these listeners would result in degraded phase locking thereby increasing the response variability of the group. Stated differently, the histograms plotted at each 10 ms interval may reveal neural activity phase locked to the stimulus for listeners whose FFR threshold was below 72 dB SPL and random neural activity for listeners whose FFR threshold was above 72 dB SPL. This same logic may serve as a possible explanation for the response variability seen in the hearing-impaired group for the higher level stimuli. At the 92 dB SPL level, it is possible the stimuli may have been below the FFR threshold of some hearing-impaired listeners thereby resulting in an increase in response variability. These results suggest response variability increases as the presentation level of the stimulus approaches the FFR threshold of the listener.

Previous studies have shown that FFR amplitude decreases as a function of frequency over a range of about 70-1800 Hz (Glaser et al., 1976; Gardi, Merzenich, and McKean, 1979; Rickman and Chertoff, 1991; Pandya, 1996). Parkinson further suggested that FFR amplitude varies inversely with frequency for time-variant stimuli

(1997). The results of the present study obtained using the 1.6 and 2.3 kHz onset stimuli were consistent with previous findings in that FFR amplitude increased as the frequency components in the time-variant second formant transition decreased. Since the amplitude of the FFR is an index of the degree of phase locking in the neural elements that generate the FFR, the increase in amplitude over time (frequency decreasing) could be taken to mean the ability of neurons to phase lock increases with decreasing frequency. Results obtained using the 0.9 kHz onset stimulus did not completely support this theory. For the 0.9 kHz stimulus, FFR amplitude increased as a function of frequency for the first 40 ms of the response at which point the amplitude began to decrease as a function of frequency. One possible explanation for the unexpected increase in response amplitude during the first 40 ms of the response could be that the upward spread of energy from the fundamental frequency and first formant contributed to the FFR amplitude. Amplitude levels began to decrease when the frequency components of the stimulus reached frequencies less effected by this upward spread of energy. It is also possible that the amplitude of the response began to decrease when the frequency components of the 0.9 kHz second formant transition increased to frequencies in which phase locking began to suffer. In this sense, the frequency components corresponding to the 40 ms time interval may have served as a "phase locking kneepoint."

Frequency Encoding: Second Formant Transition

The ability of the human FFR to represent single frequency tones (Moushegian et al., 1973; Gerken et al., Glaser et al., 1976; Sohmer and Pratt, 1977; Ananthanarayan and Durrant, 1992), multiple frequency steady-state tones (Galbraith et al., 1990; Galbraith,

1994; Greenberg et al., 1987; Pandya, 1996), steady state synthetic vowels (Ananthanarayan, 1994) and time-variant tonal glides (Parkinson, 1997) has been established. The results of this study have demonstrated that the human FFR reflecting sustained phase-locked neural activity in a population of neural elements in the rostral brainstem does indeed encode time-varying speech events such as formant transitions.

For the normal group, results demonstrated a systematic shift in the frequency peak over time corresponding to the temporal change in frequency of the second formant transition for each stimuli presented at 92 and 82 dB SPL. These results were consistent with recordings made from neurons at the auditory nerve level which demonstrated the ability to encode time-varying stop consonant-vowel stimuli (Miller et al., 1983). Systematic shifting of the spectral peaks over time was not clearly demonstrated for any stimuli presented at 72 dB SPL. It is reasonable to postulate that the encoding of formant transitions involves a progressive shift in the neural population phase locked to the stimulus along a tonotopically organized frequency map of the cochlea in the anatomical generator site of the FFR. For the hearing-impaired group, results did not demonstrate a systematic shift in the frequency peak over time for any stimuli at any level. One possible explanation for the lack of systematic peak shifting in the FFR of the hearing-impaired may be that a reduction in hearing sensitivity may result in degraded phase locking. This explanation is supported by the fact the FFR of the normal hearing group demonstrated more variable phase locking when the stimuli were at the lower presentation level. It is also possible that the broad dispersion of peaks in the hearing-impaired FFR are a reflection of the wider critical bands and reduced frequency selectivity associated with

cochlear hearing loss. This is supported by the fact phase locked activity in the FFR of the hearing-impaired is not limited to a narrow range of frequencies as it is in the normal listeners. In either case, cochlear hearing-impairment appears to disrupt the temporal distribution of neural activity phase locked to the second formant transition.

Although the hearing-impaired group did not demonstrate systematic phase-locked activity, systematic phase-locking to time varying speech was evident in some hearing-impaired individuals. The mean peak response frequency of the normal group was determined for each condition at each time interval by averaging the 16 frequency peaks in each normal group histogram. The arrow present on each hearing-impaired histogram represents the mean peak frequency of the normal group for that condition at that time interval (Figures 6, 7, 10, 11, 14, and 15). Examination of the hearing-impaired histograms revealed that systematic shifting of neural activity phase-locked to the second formant transition did occur in some hearing-impaired listeners. Joint-time-frequency analyses of FFRs obtained from hearing-impaired subjects 12 and 14 clearly demonstrate a systematic shift in frequency over time corresponding to the temporal change in frequency of the 1.6 kHz onset stimuli (Figure 18). In addition, the joint-time-frequency analysis of an FFR obtained from hearing-impaired subject 16 also revealed neural activity phase locked to the time-variant second formant transition of the 2.3 kHz onset stimuli (Figure 19). These results suggest that, although cochlear hearing-impairment appears to disrupt the temporal distribution of neural activity phase locked to the second formant transition in the hearing-impaired group, systematic phase-locking to time varying speech was evident in some hearing-impaired individuals.

Relationship Between Encoding and Identification

The results of the present study suggest that accurate identification of stop consonant place of articulation is dependant on the auditory systems ability to encode the time-variant second formant transition. For example, hearing-impaired listeners which demonstrated neural phase locking to the temporal changes of the second formant also demonstrated above average stop consonant identification. However, hearing-impaired listeners that did not demonstrate a systematic shifting of phase locked activity typically demonstrated reduced stop consonant identification. In addition, listeners with sloping, moderate to moderately severe hearing-impairments typically exhibited greater degradation in neural activity phase locked to the second formant transition than listeners with relatively flat, mild to moderate hearing impairments. The active processing of the cochlea which enhances frequency selectivity in normal hearing listeners is less efficient in listeners with moderate to moderately severe hearing-impairment than in listeners with mild to moderate hearing-impairment. The results of the present study suggest the active processing of the cochlea may play an important role in the auditory systems ability to encode time-variant stimuli. These results also suggest the FFR may serve as a useful tool in the evaluation of algorithms used in various hearing aid and cochlear implant signal processing schemes.

Conclusions

The effects of varying the onset frequency of the second formant and the overall intensity level of synthetic speech stimuli on the identification of stop consonant place of articulation were compared in normal and hearing-impaired listeners. These comparisons

included the slope and phonetic boundary of the identification functions of each stop consonant at each intensity level. Comparisons of stop consonant identification at each intensity level were also made for the three onset frequencies which corresponded to the highest percentage identification.

Group data revealed relatively similar phonetic boundaries for each stop consonant at each intensity level. Group data also revealed similar slopes for the /d/ and /g/ identification functions at each level however significantly different slopes were found between the groups for the identification of /b/. The results of this study indicate hearing-impaired listeners use relatively static information in the third and fourth formants to improve the identification of velar and alveolar stop consonants. However, hearing-impaired listeners have more difficulty identifying labial stop consonants due to their increased susceptibility to the upward spread of masking.

Comparisons of stop consonant identification were made for the three onset frequencies corresponding to the highest percentage identification of each stop. These results revealed reduced hearing-impaired identification of labial and velar stop consonants at all intensity levels and reduced hearing-impaired identification of alveolar stop consonants at higher intensity levels. Results also indicate that increasing the intensity level of the entire stimulus does not improve hearing-impaired listeners' identification of stop consonant place of articulation.

FFRs were recorded using the three stimuli corresponding to the highest percentage stop consonant identification at 92, 82, and 72 dB SPL. Results demonstrated that the FFR is sufficiently dynamic to encode time-varying formant transitions in normal

listeners at 92 and 82 dB SPL. Results also demonstrated that, although systematic phase-locking to the second formant transition did not occur in the hearing-impaired listeners as a group, systematic phase-locking to the second formant transition did occur at higher intensity levels in some hearing-impaired individuals. Comparisons of perceptual data and electrophysiologic data indicated that the identification of stop consonant place of articulation may be directly related to the neural representation of the second formant transition at the rostral brainstem. Furthermore, the neural representation of the second formant transition appears to become more degraded as the degree and slope of hearing-impairment increases.

Future Research

Future research should focus on developing a better understanding of the neural mechanism responsible for the encoding of time-variant stimuli. Developing a better understanding of the neural mechanism responsible for the encoding of time-variant stimuli may provide insight on issues ranging from hearing aid candidacy to the diagnosis and treatment of central auditory processing disorder. Future studies should investigate:

(i) the effects of relative formant amplitude levels on stop consonant representation and identification in hearing-impaired listeners.

(ii) the relationship between neural encoding and speech identification in persons with various degrees of hearing-impairment.

(iii) the effects of different signal processing algorithms on the neural encoding of speech.

(iv) the relationship between the active processing of the cochlea and the neural encoding of speech.

REFERENCES

American National Standards Institute. (1989). Specification for Audiometers, ANSI S3.6-1989. American National Standards Institute, New York.

Ananthanarayan, A.K., (1998a). Human frequency-following responses to two-tone approximations of steady state vowels. Journal of Audiology and NeuroOtology. In Press.

Ananthanarayan, A.K., (1998b). Human frequency-following response to tonal sweeps. Journal of Audiology and NeuroOtology. In Press.

Ananthanarayan, A.K., & Durrant, J.D. (1992). The frequency-following response and the onset response: evaluation of frequency specificity using a forward masking paradigm. Ear and Hearing, 13, (4), 228-232.

Batra, R., Kuwada, S., & Maher, V.L. (1986). The frequency-following response to continuous tones in humans. Hearing Research, 21, (2), 167-177.

Blumstein, S.E., & Stevens, K.N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. Journal of the Acoustical Society of America, 66, 1001-1017.

Blumstein, S.E., & Stevens, K.N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. Journal of the Acoustical Society of America, 67, 648-662.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. Journal of Speech & Hearing Research, 27, 134-144.

Borden, Harris, & Raphael. (1990). Speech Perception. In Borden, Harris, & Raphael. The Speech Primer.

Byers, V. (1973). Initial consonant intelligibility by hearing impaired children. Journal of Speech & Hearing Research, 16, 48-55.

Cazals, Y., & Palis, L. (1991). Effect of silence duration in intervocalic velar plosive on voicing perception for normal and hearing-impaired subjects. Journal of the Acoustical Society of America, 89, 2916-2921.

Chertoff, M.E. & Hecox, K.E. (1990). Auditory nonlinearities measured with auditory-evoked potentials. Journal of the Acoustical Society of America, 87, 1248-1254.

Delattre, P.C., Liberman, A.M., & Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. Journal of the Acoustical Society of America, 24, 597-606.

Delattre, P.C., Liberman, A.M., Cooper, F.S., & Gerstman, L.J. (1952). An experimental study of the acoustic determinants of vowel color: Observations on one and two formant vowels synthesized from spectrographic patterns. Word, 8, 195-210.

Dorman, M.F., & Dougherty, K. (1981). Shifts in phonetic identification with changes in signal presentation level. Journal of the Acoustical Society of America, 69, 1439-1440.

Dorman, M.F., Marton, K., Hannley, M.T., & Lindholm, J.M. (1985). Phonetic identification by elderly normal and hearing-impaired listeners. Journal of the Acoustical Society of America, 77, 664-670.

Galbraith, G.C. (1994). Two-channel brain-stem frequency-following response to pure tone and missing fundamental stimuli. Electroencephalography and Clinical Neurophysiology, 92, (4), 321-330.

Galbraith, G.C, & Brown, W.S. (1990). Cross-correlation and latency compensation analysis to click-evoked and frequency-following brain-stem responses in man. Electroencephalography and Clinical Neurophysiology, 77, (4), 295-308.

Gardi, J., Merzenich, M., & McKean, C. (1979). Origins of the scalp-recorded freqeuncy-following-response in the cat. Audiology, 18, 353-381.

Gerken, Moushegian, Stillman, & Rupert. (1975). Human frequency-following response to monaural and binaural stimuli. Electroencephalography and Clinical Research, 38, 379-386.

Glaser, Suter, Dasheiff, & Goldberg. (1976). The human frequency-following response: Its behavior during continuous tone and tone burst stimulation. Electroencephalography and Clinical Neurophysiology, 40, 25-32.

Greenberg, S., Marsh, J.T., Brown, W.S., & Smith, J.C. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hearing Research, 25, (2-3), 91-114.

Jesteadt, W., Neff, D.L., Humes, L., & Leek, M.R. (1997). Modeling hearing loss as an additional source of masking. In Jesteadt. Modeling Sensorineural Hearing Loss.

Kewley-Port, D. (1981). Representations of spectral change as cues to place of articulation in stop consonants. Unpublished doctoral dissertation: Indiana University, Bloomington, IN.

Kewley-Port, D. (1983a). Time-varying features as correlates of place of articulation in stop consonants. Journal of the Acoustical Society of America, 73, 322-335.

Kewley-Port, D. (1983b). Measurement of formant transitions in naturally produced stop-consonant syllables. Journal of the Acoustical Society of America, 72, 379-389.

Klatt, D.H. (1980). Software for a cascade/parallel formant synthesizer. Journal of the Acoustical Society of America, 67, 971-995.

Lahiri, A., Gewirth, L., & Blumstein, S.E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: evidence from a cross-language study. Journal of the Acoustical Society of America, 76, 391-404.

Liberman, A.M., Delattre, P.C., & Cooper, F.S. (1952). The role of selected stimulus-variables in the perception of the unvoiced stops. American Journal of Psychology, LXV, 497-516.

Liberman, A.M., Delattre, P.C., & Cooper, F.S. (1958). Some rules for the distinction between voiced and voiceless stops in initial position. Language & Speech, 1 153-167.

Liberman, A.M., Delattre, P.C., Cooper, F.S., & Gerstman, L.J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. Psychology Monagr., 68, 1-13.

Liberman, A.M., Delattre, P.C., Gerstman, L.J., & Cooper, F.S. (1956). Tempo of frequency change as a cue for distinguishing classes of speech sounds. Journal of Experimental Psychology, 52, 127-137.

Lisker, L., & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. Word, 20, 384-422.

Marsh, J.T., Brown, W.S., & Smith, J.C. (1975). Far-field recorded frequency-following responses: correlates of ow pitch auditory perception in humans. Electroencephalography and Clinical Neurophysiology, 38, 113-119.

Miller, M.I., & Sachs, M.B. (1983). Representation of stop consonants in the discharge patterns of auditory-nerve fibers. Journal of Acoustical Society of America, 74, 502-517.

Miller, M.I., & Sachs, M.B. (1984). Representation of voice pitch in discharge patterns of auditory-nerve fibers. Journal of the Acoustical Society of America, 74, 502-517.

Moller, A.R., & Jho, H.D. (1989). Response from the exposed intracranial human auditory nerve to low-frequency tones: basic characteristics. Hearing Research, 38, (1-2), 163-175.

Moushegian, G., Rupert, A.L., & Stillman, R.D. (1973). Scalp-recorded early responses in man to frequencies in the scalp range. Electroencephalography and Clinical Neurophysiology, 35, 665-667.

Ohde, R.N., & Stevens, K.N. (1983). Effect of burst amplitude on the perception of the stop consonant place of articulation. Journal of the Acoustical Society of America, 74, 706-714.

Owens, E., Benedict, M., & Schubert, E.D. (1972). Consonant phonemic errors associated with pure-tone configurations and certain kinds of hearing impairment. Journal of Speech & Hearing Research, 15, 308-322.

Owens, E., & Schubert, E.D. (1968). The development of constant items for speech discrimination testing. Journal of Speech & Hearing Research, 11, 656-667.

Pandya, P.K. (1996). Evaluation of the human frequency following response distortion product at 2F1-F2. Unpublished Thesis: University of Tennessee, Knoxville, Tennessee.

Parkinson, J.D. (1997). Human frequency-following response to tonal sweeps in normal hearing subjects. Unpublished Thesis: University of Tennessee, Knoxville, Tennessee.

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

Price, P.J., & Simon, H.J. (1984). Perception of temporal differences in speech by normal hearing adults: Effects of age and intensity. Journal of Acoustical Society of America, 76, 405-410.

Reed, C. (1975). Identification and discrimination of vowel-consonant syllables in listeners with sensorineural hearing loss. Journal of Speech & Hearing Research, 18, 773-794.

Rickman, M.D., & Chertoff, M.E. (1991). Electrophysiological evidence of nonlinear distortion products of two-tone stimuli. Journal of the Acoustical Society of America, 89, 2818-2826.

Sachs, M.B., & Young, E.D. (1979). Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. Journal of the Acoustical Society of America, 66, 470-479.

Schum, D.J., & Collins, M.J. (1990). The time course of acoustic/phonemic cue integration in the sensorineurally hearing-impaired listener. Journal of the Acoustical Society of America, 87, 2716-2728.

Sher, A.E., & Owens, E. (1974). Consonant confusions associated with hearing loss above 2000 Hz. Journal of Speech & Hearing Research, 17, 669-681.

Smith, J.C., Marsh, J.T., & Brown, W.S. (1975). Far-Field recorded frequency-following-responses: Evidence for the locus of brainstem sources. EEG Clinical Neurophysiology, 39, 465-472.

Sohmer, H., & Pratt, H. (1977). Identification and seperation of acoustic frequency-following responses (FFRs) in man. Electroencephalography and Clinical Neurophysiology, 42, 492-500.

Stevens, K.N. (1975). The potential role of property detectors in the perception of consonants. Auditory Analysis and Perception of Speech, ed. G. Fant, & M.A.A. Tatham (Academic New York), 303-330.

Stevens, K.N., & Blumstein, S.E. (1978). Invariant cues for place of articulation in stop consonants. Journal of the Acoustical Society of America, 64, 1358-1368.

Stillman, R.D., Crow, G., & Moushegian, G. (1978). Components of the frequency-following response in man. Electroencephalography and Clinical Neurophysiology, 44, (4), 438-446.

Tyler, R., Summerfield, A.Q., Wood, E., & Fernandes, M. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. Journal of the Acoustical Society of America, 73, 740-752.

Worden, F.G., & Marsh, J.T. (1978). Frequency-following (microphonic-like) neural responses evoked by sound. Electroencephalography and Clinical Neurophysiology, 25, 42-52.

Van Tasell, D.J., Hagen, L.T., Koblas, L.L., & Penner, S.G. (1982). Perception of short-term spectral cues for stop consonant place by normal and hearing-impaired subjects. Journal of the Acoustical Society of America, 72 (6), 1771-1780.

Verbrugge, R.R., Strange, W., Shankweiler, D.P., & Edman, T.R. (1976). What information enables a listener to map a talker's vowel space? Journal of the Acoustical Society of America, 27, 761-764.

Voight, M.F., Sachs, M.B., & Young, E.D. (1982). Representation of whispered vowels in discharge patterns of auditory-nerve fibers. Hearing Research, 8, 49-58.

Young, E.D., & Sachs, M.B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. Journal of the Acoustical Society of America, 66, (5) 1381-1403.

APPENDICES

APPENDIX A

AUDIOMETRIC DATA

Test Frequency Hertz	Test Ear Threshold in dB HL for Normal Subjects 1-8							
	1	2	3	4	5	6	7	8
125	5	5	10	0	0	0	5	-5
250	5	5	10	0	10	0	5	-5
500	0	0	5	-5	10	0	0	-5
750	5	5	5	-5	10	0	0	-5
1000	5	5	0	5	5	0	0	-5
1500	5	0	0	0	5	-5	0	-5
2000	5	-5	0	-5	5	10	5	0
3000	0	-5	-5	-5	5	0	5	0
4000	-5	-5	0	-5	5	10	5	0
6000	-5	0	15	0	5	5	5	0
8000	-5	5	10	0	0	5	5	15

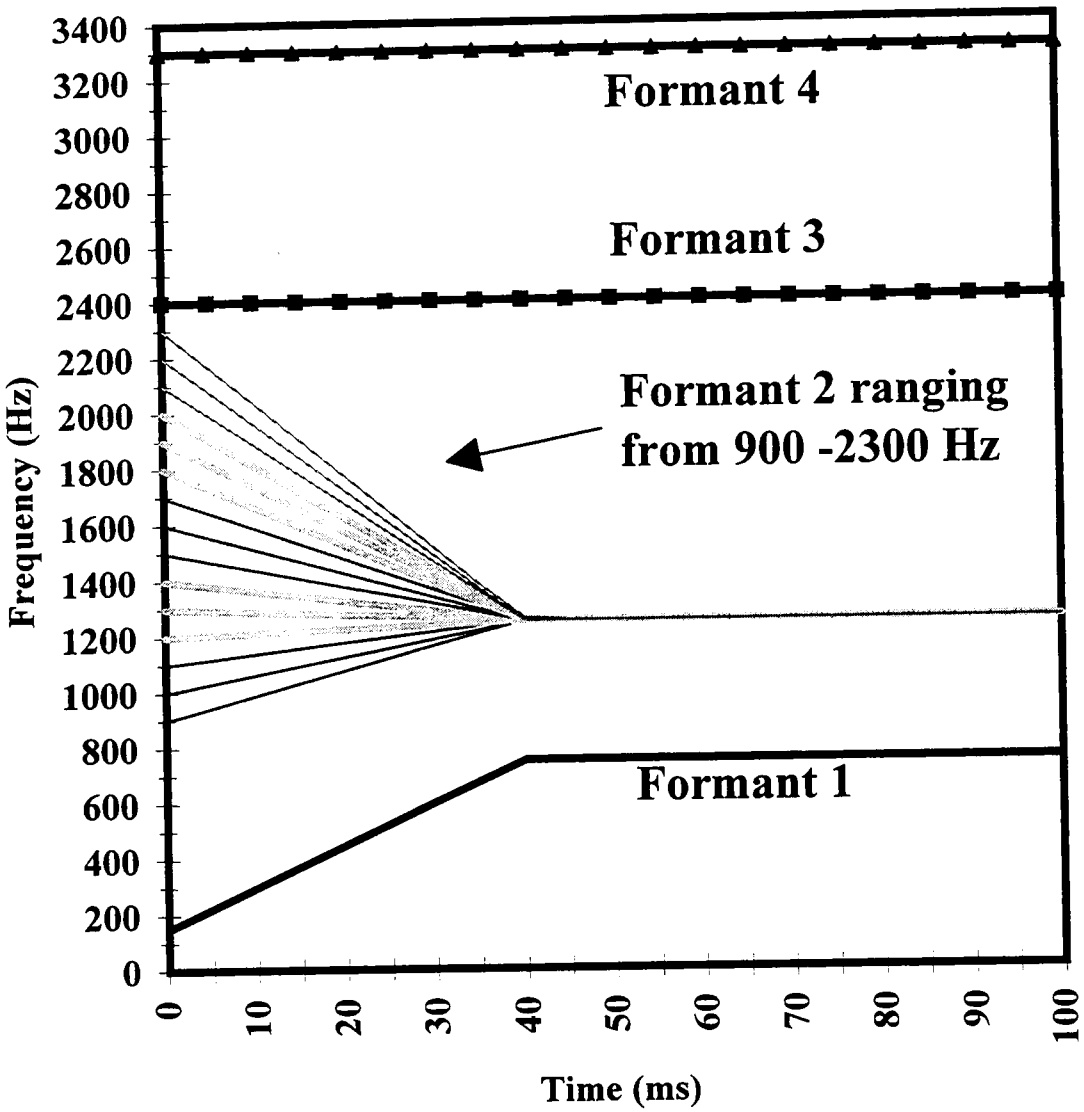
Test Frequency Hertz	Test Ear Threshold in dB HL for Normal Subjects 9-16							
	9	10	11	12	13	14	15	16
125	5	0	0	-5	0	5	5	5
250	5	0	0	-5	0	5	0	0
500	5	5	0	0	0	5	0	5
750	5	5	0	0	0	5	5	0
1000	0	0	0	-5	0	5	5	0
1500	-5	0	-5	-5	0	5	0	0
2000	-5	-5	-5	-5	5	10	5	5
3000	-5	0	-5	-5	0	5	0	0
4000	0	0	-5	-5	-5	0	10	5
6000	5	0	0	-5	0	5	5	0
8000	5	0	5	-5	10	5	5	5

Test Frequency	Test Ear Threshold in dB HL for Hearing-Impaired Subjects 1-8							
Hertz	1	2	3	4	5	6	7	8
125	25	25	50	20	65	25	30	35
250	25	25	50	20	65	25	30	35
500	35	25	55	25	55	40	45	40
750	45	30	55	25	50	35	50	40
1000	55	30	60	30	45	35	60	40
1500	55	35	60	35	50	35	65	35
2000	60	50	60	35	40	40	65	35
3000	65	60	65	40	60	55	65	45
4000	60	65	60	45	65	60	60	50
6000	60	65	60	55	70	65	90	55
8000	65	65	65	70	75	70	90	55

Test Frequency	Test Ear Threshold in dB HL for Hearing-Impaired Subjects 9-16							
Hertz	9	10	11	12	13	14	15	16
125	25	20	20	35	15	40	20	20
250	25	20	20	35	15	40	20	20
500	25	30	30	30	15	40	30	25
750	30	40	40	30	15	45	30	30
1000	35	45	50	30	35	45	35	40
1500	40	50	55	40	40	50	40	45
2000	50	50	45	40	40	50	40	45
3000	55	45	55	40	45	45	55	40
4000	60	40	55	45	50	55	60	45
6000	55	65	70	45	55	60	55	45
8000	50	60	70	65	55	60	55	45

APPENDIX B

SCHEMATIC OF STIMULI


Formant 4
Formant 3
Formant 2 ranging
from 900 -2300 Hz
Formant 1
Frequency (Hz)
Time (ms)
3400
3200
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
10
20
30
40
50
60
70
80
90
100

APPENDIX C

SUBJECT CONSENT FORM

INFORMED CONSENT DOCUMENT

PROJECT DESCRIPTION

You are being asked to participate in an experimental study aimed at evaluating how the central auditory system of normal hearing and hearing-impaired individuals responds to variations of the speech sounds /b/ and /d/, the results of which will be included in a dissertation by Patrick N. Plyler. The responses recorded from the central auditory system in this project are called frequency-following responses (FFRs). Should you agree to participate, the test session will not exceed three consecutive hours.

Before the study, you will be given a hearing test (tympanograms and routine audiometry) to make sure your hearing is within the normal (30 dB HL or better) or mild to moderate (30 dB HL to 55 dB HL) range of hearing. Then you will be seated in sound treated booth where you will listen to 160 speech sounds through earphones. After each speech sound, you will be asked to judge if the sound you heard was /b/ or /d/.

Following the speech testing, FFRs will be recorded by placing electrodes, which are small metal discs, on your forehead and the back of your neck. The electrodes are held in place by electrode paste and plastic tape after the surface of the skin has been prepared with a mild abrasive solution. During the test session, you will be asked to simply remain still and relax in a reclining chair while you listen to the same series of speech signals through headphones. Should you need to excuse yourself at any time, for any reason, during the testing, you will have a button to signal your request for attention. You will be given a ten minute break after the qualification testing and after the first aspect of the experimental procedure.

The above test and procedures are routine clinical procedures which offer no known health risks. Regardless, every effort will be made to minimize any potential risk or discomfort. Although you may not benefit directly from this study, you will have contributed to our understanding of the way the central auditory system encodes variations of speech sounds.

CONFIDENTIALITY

Any information obtained about you for this study will be kept strictly confidential. Confidentiality of your data as well as any identifying material will be protected by limiting access to only the Principal and Co-Principal Investigators. Your identity will not be revealed in any description or publication of this research. This consent form will be stored for three years following completion of the study at a secured location at the University of Tennessee.

RIGHT TO REFUSE OR TO END PARTICIPATION

You are free to refuse to participate in this study or to withdraw at any time without having it affect your care in any of the University of Tennessee Clinics and without any

loss of benefits to which you might otherwise be entitled.

VOLUNTARY CONSENT

I have read the above information, and the Principal Investigator has explained the procedures and answered all my questions concerning this project. It has been explained to me that my identity will not be revealed in any description or publication of this research. Therefore, I consent to such publication for scientific purposes.

Any questions I have concerning my rights as a research subject will be answered by Patrick N. Plyler, the Principal Investigator, or by Dr. A.K. Ananthanarayan, the Co-Principal Investigator, at 570 South Stadium Hall [telephone: (423) 974-1811].

My signature below means that I have had all my questions answered and freely agree to participate in the experimental study.

Signature of Subject: _____ Date _____

INVESTIGATOR'S CERTIFICATION

I certify that I have explained to the above individual the nature and purpose, the potential benefits, and the possible risks associated with participating in this experimental study. I have answered any questions that have been raised, and have witnessed the above signature.

Signature of Investigator: _____ Date _____

VITA

Patrick N. Plyler was born in Statesville, North Carolina on July 12, 1969. He received his Bachelor of Science in Speech-Language-Auditory Pathology from East Carolina University in May of 1992 and his Master of Arts in Audiology from the University of Tennessee in December of 1993. Mr. Plyler entered the Doctoral program in Speech and Hearing Science at the University of Tennessee upon completion of his Clinical FellowshipYear at the Bill Wilkerson Center in Nashville, Tennessee. Upon completion of his Doctorate, Mr. Plyler will join the faculty of the Department of Communication Sciences & Disorders at Louisiana State University in Baton Rouge, Louisiana as an Assistant Professor.