Veterans Administration

Journal of Rehabilitation Research and Development Vol. 24 No. 4
Pages 229–238

Auditory filtering and the discrimination of spectral shapes by normal and hearing-impaired subjects*

CHRISTOPHER W. TURNER, Ph.D., LENORE A. HOLTE, M.S., and EVAN RELKIN, Ph.D.
Communication Sciences and Disorders and Institute for Sensory Research, Syracuse University, Syracuse, New York 13244

Abstract—A review of the literature suggests that many hearing-impaired patients suffer from sensory deficits in addition to the reduced audibility of speech signals. Poor frequency resolution, or abnormal spread of masking, is a consistently identified deficit in sensorineural hearing loss. Frequency resolution was measured in individual subjects using the *input filter pattern paradigm*, and the minimum detectable amplitude of a second-formant spectral peak in a spectral-shape discrimination task was also determined for each subject. The two tasks were designed to test the identical frequency regions in each subject. A nearly perfect correlation was found between the degree of frequency resolution as measured by the input filter pattern and performance on the spectral-shape discrimination task. These results suggest that measures of frequency selectivity may offer predictive value as to the degree of impairment that individual hearing-impaired patients may have in perceiving the spectral characteristics of speech, and also lead to suggestions for signal processing strategies to aid these patients.

INTRODUCTION

Many persons with hearing loss of the sensorineural type using currently available hearing aids will testify that their hearing aids do not return speech recognition ability to an entirely normal status. Conventional amplification is, by its nature, primarily capable of restoring the audibility of speech signals to the hearing-impaired patient. A substantial body of evidence indicates however, that when sensorineural hearing loss is modeled as a purely attenuative loss, the consequences of the hearing loss are not completely described for many patients.

If the speech recognition abilities of hearing-impaired patients are poorer than would be indicated by the audiogram alone, then speech signals that are filtered to simulate a particular hearing loss and presented to normal-hearing subjects should serve as an enlightening control condition with which hearing-impaired subjects might be compared. Fabry and Van Tasell (9), pursuing this line of reasoning, performed comprehensive experiments comparing the performance of actual hearing-impaired ears with normal ears that had hearing losses simulated by filtering. The research of Fabry and Van Tasell demonstrated that approximately one-half of the hearing-impaired subjects did not perform as predicted by the simulations. From which specific psychoacoustic deficits this lack of predictability resulted is not evident from the present state of research.

An alternate approach to determining the contribution of audibility to speech recognition performance in hearing-impaired subjects has been to apply the Articulation Index (AI) calculations to speech recognition data. Kamm, Dirks and Bell (16) found that the AI calculations for mild to moderate hearing-impaired subjects were as valid as equivalent AI calculations for normal-hearing subjects. This implies that audibility was a sufficient factor to deter-

* This research was supported in part by a research grant from the Deafness Research Foundation to the first author.

mine speech recognition under these conditions. A more severely impaired patient was also tested by Kamm et al.; in that case the AI significantly overpredicted performance. The subject's results showed that, contrary to the AI assumptions, increases in audible acoustic information did not improve speech intelligibility.

Dugal, Braida and Durlach (8) also used the AI to predict speech recognition performance of hearing-impaired subjects and found that the method provided an orderly and relative prediction of results. However, they needed to employ a "proficiency factor" to adjust (reduce) a hearing-impaired subject's predicted performance for factors they termed "deficits in suprathreshold recognition," supporting the existence of deficits beyond audibility. Similar conclusions were reached by Fletcher (10), Wilbur (42), Milner (24) and Pavlovic (29). Pavlovic and Studebaker (30) have formulated an improved AI scheme that incorporates a "desensitivity" factor at each frequency, related to the degree of hearing loss, and also a correction for abnormal spread of masking. This model provided improved prediction accuracy over the standard AI scheme.

In summary, the evidence from Articulation Index studies tend to support the existence of suprathreshold psychoacoustic deficits in addition to the obviously important audibility factors; however, this line of research has not singled out specific auditory deficits beyond that of abnormal spread of masking.

Elevated masked thresholds (relative to normal-hearing ears) for probe tones above and below the masker frequency have often been observed in impaired ears: e.g. Jerger et al. (14), Rittmanic (32), Keith and Anderson (17), deBoer and Bouwmeester (6), Leshowitz and Lindstrom (20), Gagne (12), Scharf and Florentine (33), and Trees and Turner (34).

The issue of abnormal masking observed in hearing-impaired subjects has also been investigated by means of other psychophysical paradigms, these have been termed measures of frequency resolution or frequency selectivity. Abnormal psychophysical tuning curves have been noted in high-frequency hearing-loss patients: e.g., Wightman, McGee, and Kramer (41), Zwicker and Shorn (43), Nelson and Turner (27). Abnormal critical ratios have been reported: e.g., Palva, Goodman, and Hirsh (28), Margolis and Goldberg (22), Tyler, Fernandes, and Wood (37). Discrepant results in hearing-impaired subjects for masking produced by noise with a rippled spectrum are reported: e.g., Pick, Evans, and Wilson (31), Wightman (40). These experiments have been interpreted as indicating an abnormal shape of an "auditory filter" in hearing-loss patients, and a subsequent deficit in the ability of the patient to perform a frequency analysis of speech sounds is predicted. These models, which present sensorineural hearing loss as a decrease in auditory frequency resolution, or abnormally broadened auditory filters, also contradict the purely attenuative model of sensorineural hearing loss.

A number of researchers have attempted to demonstrate a relation between measures of frequency resolution and measures of speech recognition ability: e.g., Hoekstra and Ritsma (13), Bonding (4), Dreschler and Plomp (7), Tyler, Fernandez, and Wood (36). Although speech recognition and frequency resolution appear to be generally correlated, the degree of correlation varies across the studies, with most researchers reporting a correlation coefficient near 0.5. These mild correlations, and the resulting inability to assign a more direct role to frequency resolution in speech recognition by the hearing-impaired subjects, most likely results from the large number of uncontrolled variables in such studies. In particular, the choice of a specific frequency at which resolution is measured and the wideband frequency content of the speech materials used may have served to obscure any direct relation.

As summarized in the literature review, the most consistently identified deficits in hearing-impaired subjects, other than the lack of audibility of the less intense speech sounds, has been poorer frequency resolution. These deficits suggest that under certain conditions, some hearing-impaired subjects' perception of the individual spectral shapes of speech segments may be inaccurate. Pavlovic and Studebaker's (30) finding that a correction for abnormal upward spread of masking improved AI predictions, supports the view that impaired frequency resolution is an important factor. Other deficits, such as poorer intensity discrimination or temporal resolution, although not ruled out by the existing literature, can, in many cases, be attributed to sensitivity loss, e.g., Florentine and Buus (11), Bacon and Viemeister (2); or, they may be caused by the growth of masking in the impaired ear, Jesteadt et al., (15). For the purposes of this investigation, which was concerned with steady-state signals, it was decided to focus on those problems that hearing-impaired subjects may have in recognizing the acoustic features of speech due to poor frequency resolution.

Current speech perception theory suggests that the phonemes of speech can be recognized on the basis of given acoustic characteristics. The primary spectral cue for many speech sounds is the presence of spectral peaks corresponding to the frequency location of the vocal tract formants. Blumstein and Stevens (3) and Kewley-Port (18) have provided evidence that spectral peaks in short-term acoustic representations provide information as to the phonemic identity of stop consonants. Klatt (19), Miller (23) and Chistovich (5) have presented findings that suggest that the typical listener to a speech sound derives the phonemic identity of the sound from the detection, as well as from frequency location, of the formant peaks in the overall spectral envelope of the stimulus. One factor that may be involved in hearing-impaired patients' poor speech recognition is the poorer-than-normal frequency resolution that has been documented to exist in some sensorineural hearing-impaired subjects. This poorer frequency resolution, or abnormal masking, would in theory tend to obscure the detectability of neighboring spectral peaks. Such effects would presumably reduce the accuracy with which such listeners could identify the presence of, and frequency locations of, local spectral peaks.

This line of reasoning suggests that, for some hearing-impaired patients, the spectral peak characteristics of some normally-presented speech sounds are not distinct enough to allow normal speech recognition to occur. Turner and Holte (35) demonstrated that a more prominent second-formant spectral peak was required for a subgroup of their hearing-impaired subjects than for the normal-hearing subjects, in order to achieve equal levels of performance in a discrimination task using speech-like spectral shapes. This deficit in performance was observed in those subjects even when the stimuli were presented at levels for which the critical spectral regions were at levels above the subjects' quiet thresholds. The results of the Turner and Holte study suggest that some hearing-impaired subjects may benefit from a speech processing scheme that would enhance the spectral peaks in certain speech sounds, in order for them to perceive the relevant acoustic characteristics of speech in a satisfactory fashion. The long-range goal of the present research program is to attempt to identify which hearing-impaired subjects might benefit from such speech processing, under what circumstances the processing of the spectral shape of the speech would be appropriate, and also to be able to predict the necessary degree of spectral-peak enhancement that a given subject might require for any presented speech sound.

The present study utilizes results from a psychoacoustical masking study in an attempt to predict the data on second-formant spectral peak discrimination from the Turner and Holte (35) study. In light of the previously mentioned low correlations found by previous investigators between frequency resolution as measured by masking studies and general speech recognition performance, this study was designed to match the frequencies at which frequency resolution is measured with the frequencies important to the spectral-peak discrimination task. Thus the present effort addresses the following question: Can a standardized measure of frequency resolution be used to predict the need for spectral-peak enhancement to improve speech perception in a given subject?

METHODS

Data were obtained on subjects who participated in both a spectral-peak discrimination experiment and a frequency-resolution experiment. The masking paradigm chosen for this study has been termed an "input filter pattern," e.g., Verschuure (38). In this paradigm, the masked threshold for a probe signal at a fixed frequency is measured in the presence of fixed-level maskers at several frequencies. The resulting pattern (probe threshold plotted as a function of masker frequency) represents a subject's auditory filtering characteristic at the probe frequency location. The level of the (input) masker is held constant, thus the probe threshold reflects the relative output of the constant-input masker following auditory filtering. Such a pattern describes the relative effectiveness of a subject's auditory filtering at a given frequency in rejecting acoustic interference from surrounding frequency regions. In contrast, the more familiar psychophysical tuning curve (PTC) represents the input signal required to produce a constant output following the filter. Because the spectral-peak discrimination experiment is described in more

Journal of Rehabilitation Research and Development Vol. 24 No. 4 Fall 1987

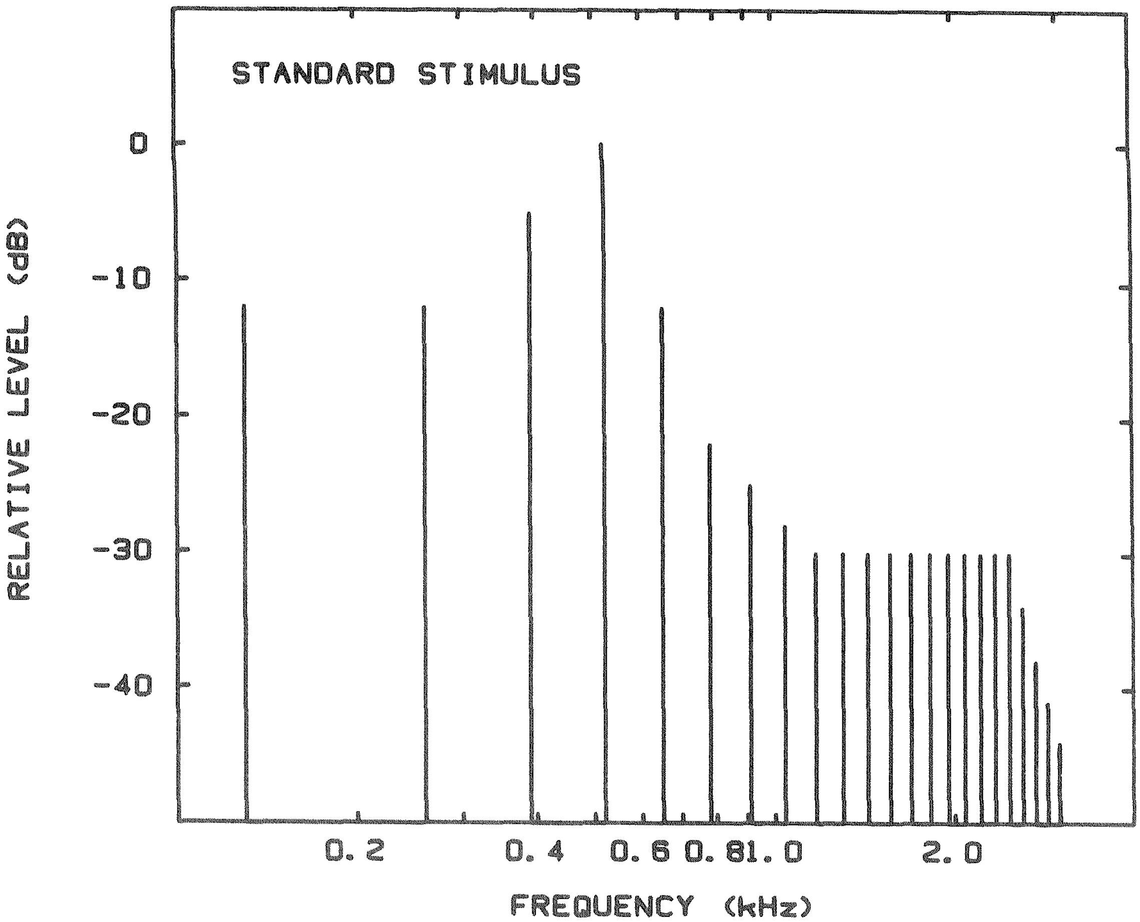

Figure 1.
Spectral representation of the standard (0-dB increment second-formant) stimulus.

detail in another publication (35) that experiment's methods and results will be presented only briefly here.

Spectral Peak Discrimination

Stimuli were synthesized versions of the steady-state vowel /ε/. Pure-tone harmonic components of a fundamental frequency of 130 Hz were added with appropriate amplitude values to yield the standard stimulus (**Figure 1**) and various comparison stimuli (for example, **Figure 2** shows one such comparison stimulus). A peak in the spectral envelope corresponding to a first formant was centered at 520 Hz. The standard stimulus had a flat spectral shape in the second formant region (referred to as the 0-dB spectral peak stimulus), while a series of comparison stimuli were constructed with second formant peaks (centered at 1820 Hz) with incrementally larger amplitudes. The amplitudes of 6 stimulus components were set to yield a spectral envelope peak at 1820 Hz. The formant amplitude was expressed as the ratio (in dB) of the incremented spectral amplitude of the 1820 Hz component to the standard stimulus value at 1820 Hz. All signals were 204 ms in duration, including 25 ms linear rise-fall ramps.

Subjects discriminated between a standard (0-dB amplitude spectral peak) stimulus and a comparison stimulus with an incremented-amplitude spectral peak in a 4-alternative forced-choice procedure. Visual feedback displaying the correct interval was presented to the subject following each trial. A laboratory computer controlled an adaptive procedure which presented discrimination trials between standard and comparison stimuli with various second-formant amplitudes in a search for the difference limen for the spectral peak discrimination. A two-down, one-up tracking rule was employed to yield