INTERNAL SPEECH REPRESENTATIONS IN NORMAL AND HEARING-IMPAIRED SUBJECTS

Christopher W. Turner and Lenore A. Holte

Syracuse University

Psychoacoustical input-filter pattens were obtained from individual subjects and were used to calculate the "internal spectra" of acoustic speech and speech-like sounds. The validity of the calculated representations is demonstrated by predictions of the discrimination and recognition performance by normal and hearing-impaired subjects.

INTRODUCTION

Klatt [1] and Chistovich [2] have presented evidence that suggest that the typical listener to a steady-state or slowly varying speech sound derives the phonemic identity of the sound from the detection of local spectral peaks in the overall spectral envelope of the speech sound. The peripheral auditory system has often been modeled as a bank of parallel bandpass filters capable of resolving these spectral characteristics important for speech perception. Specification of the auditory filter characteristics across the frequency range in human subjects allows one to calculate a spectral representation of a peripherally-filtered speech stimulus that the subject's central processing system would then operate upon. This approach has been utilzed by several previous researchers [1,3,4]. Turner and Robb [5] have recently used such representations to predict the audibility of stop consonants.

It is known that persons with sensorineural hearing loss often suffer from, in addition to their elevated sensitivity thresholds, some degree of reduced frequency resolution. This is evidenced by abnormal spread of masking [6,7] and also by broadened psychophysical tuning curves [8,9,10]. More directly relevant to the bandpass-filter models of the auditory system are results showing that sensorineural hearing-loss subjects often display abnormally wide auditory filters [11,12]. Comparisons of vowel-masking patterns from normal and hearing-loss subjects suggest that the formant peaks are less clearly resolved in the "internal representations" of the hearing-loss subjects than in normals [13,14,15].

The present paper describes our measures of auditory filter characteristics in normal and impaired subjects and their relation to the subjects' detection and discrimination of spectral features in speech-like stimuli. The measured filter characteristics are also used to calculate the "internal representations" of speech stimuli, in order to provide an insight to the recognition performance of the subjects. One advantage of using subjects' filter characteristics to calculate "internal representations" of speech signals, in comparison to the vowel-masking pattern approach, is that specification of the filter characteristics in a subject subsequently allows for the calculation of "internal representations" for any arbitrary speech sound. By using both normal-hearing and hearing-loss subjects, whose auditory filtering characteristics vary across a wide range of resolving powers, the dependency of discrimination and recognition performance upon the resolution of the spectral cues can be investigated.

MEASURES OF AUDITORY FILTERING

The masking paradigm chosen for this study has been termed an "input filter pattern" [16]. In this paradigm, the masked threshold for a probe signal at a fixed frequency is measured in the presence of fixed-level maskers across a range of frequencies. The resulting pattern, of probe threshold plotted as a function of masker frequency, is taken to represent 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 constant-level input signals following auditory filtering. Thus the input-filter pattern can theoretically be used to yield a measure of the output of each auditory filter. In contrast, the more familiar psychophysical tuning curve reflects the input signal level required to produce a constant output following the filter.

Methods

A forward-masking paradigm was employed for input filter pattern measures in order to reduce potential artifacts such as beats or distortion products. The center frequency for each measured input-filter pattern is determined by the probe frequency; the six probe frequencies used were 250 Hz, 500 Hz, 1000 Hz, 1820 Hz, 3000 Hz and 4000 Hz. Eleven masker frequencies were used to define the shape of the input filter pattern for each probe frequency. The pure-tone maskers were 204-msec in duration including 10-msec rise-fall ramps. All maskers were presented at 95 dB SPL. The probe signal was an pure tone of 25-msec duration including 5-msec rise-fall ramps. Probe onset occured at masker offset.

Results

Figure 1 shows input filter patterns at several probe frequencies from a normal-hearing subject. We interpret these data as depicting a filter shape by viewing the level of the probe threshold at each masker frequency, relative to the peak of each pattern, as the auditory filter attenuation for signals at that particular masker frequency. For the normal-hearing subject's data shown in the first figure, the filters are asymmetrical bandpass in shape. The shallower slope seen on the low-frequency side of the filter is consistent with the well-known phenomenon of upward spread of masking.

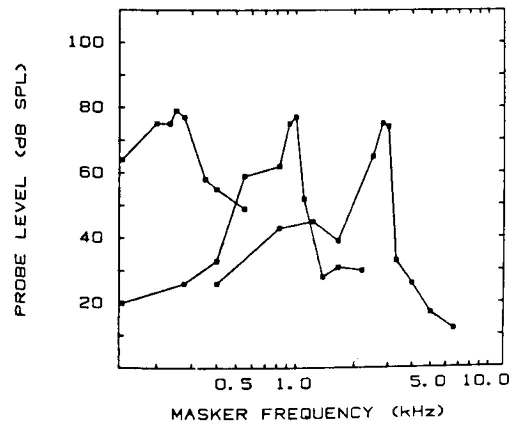

Figure 1. Input filter patterns from a normal-hearing subject

Figure 2 shows input filter patterns at three frequencies from a moderately hearing-impaired subject. The filter patterns are broader in shape than in Figure 1, showing in particular more of a low-pass characteristic, consistent with increased spread of masking in this subject.

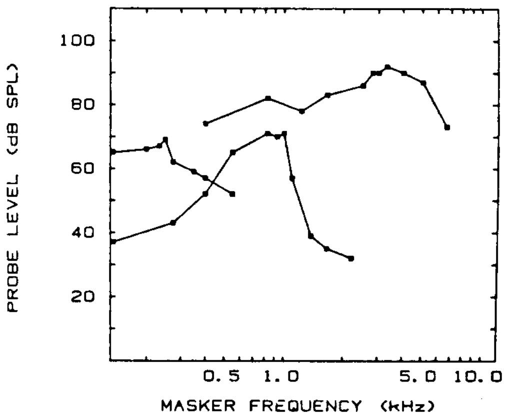

Figure 2. Input filter patterns from a hearing-impaired subject

RELATION OF FILTERING TO FORMANT PERCEPTION

The previous 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 for normal speech recognition to occur. Turner and Holte [17] demonstrated that a more prominent second-formant spectral peak was required for some hearing-impaired subjects than for the normal-hearing subjects in a discrimination task using speech-like spectral shapes. If auditory filtering characteristics in individual subjects determine the abilities of subjects to resolve the spectral characteristics of speech sounds, then the results of the input filter pattern experiments should offer predictive value for the formant-height detection task described above. Indeed, as described in Turner et al. [12], the bandwidth of subjects' input filter patterns were highly correlated with the minimum detectable formant height for the same subjects (r=-.95). That is, the narrower the measured filter shape for an auditory filter centered at the second-formant frequency, the smaller the spectral peak in the speech-like spectral shape required for equivalent levels of detectability. We therefore concluded that the psychoacoustical input-filter pattern measure is a valid indicator of the perceptual strength of spectral characteristics in individual subjects.

INTERNAL REPRESENTATIONS OF VOWELS

Although the recognition of vowels by both normal and moderately hearing-impaired subjects is generally quite accurate, we determined four vowels for which confusions are obtained. The synthetic vowels /æ, ɛ, ʌ, ɑ / were constructed using the Klatt synthesis program. Fundamental frequency was varied slightly throughout the stimulus in order to yield a more natural utterance. All stimuli were presented at 95 dB SPL, at which level the entire stimulus was above each subjects' sensitivity thresholds. The formant frequencies of the vowels were held constant during each vowel and are listed below:

	F1	F2	F3
æ	620(Hz)	1660	2430
ɛ	530	1680	2500
ʌ	620	1220	2550
ɑ	660	1200	2550

Both hearing-loss and normal subjects were tested in a closed-set recognition task using these four tokens. Recognition scores are expressed as the percentage of information transmitted for a 2x2 matrix, which contrasted correct versus incorrect identification for each vowel. Confusions were obtained primarily between the pairs /æ,ɛ/ and /ʌ,ɑ/. For those pairs, the primary distinguishing feature between the vowels was most likely the small differences in F1 frequency, although some small differences at the higher formants may have contributed to recognition. The subjects, both hearing-impaired and normal, rarely confused members of the /æ,ɛ/ pair with members of the /ʌ,ɑ/ pair.

An FFT of each vowel was obtained from the acoustic signal and was then used as the input to a model of the peripheral auditory system based upon input filter pattern measures. The input filter patterns from each subject were then used, along with interpolated filter shapes at intermediate frequencies, to calculate a 200-point "internal spectrum" of each vowel for each subject. This was accomplished via laboratory computer, in which the outputs of each filter were calculated as the sum of the stimulus power passed by each filter.

Figure 3 displays "internal representations" of two vowels /ɑ/ and /ʌ/ for a normal-hearing subject (same subject as in Figure 1). The three formants are quite well represented , due to the narrow auditory filters measured in the subject. The difference in F1 frequency is visible near 600 Hz. This example shows a vowel pair which was sometimes confused by that subject, although his recognition was still quite accurate (96% and 97% respectively). As was the case for all normal-hearing subjects, vowel recognition was near 100% for all tokens. This case demonstrates that fine distinctions in formant frequency location are preserved and available to a subject with normal auditory filtering.

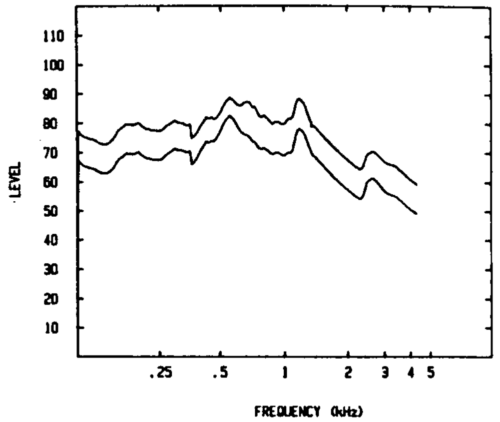

Figure 3. Internal representations for the vowels /ɑ/ (upper) and /ʌ/ (lower) from a normal-hearing subject. The lower curve has been displaced by 10 dB for ease of viewing.

In Figure 4, the calculated "internal representations" of the previous vowel pair, /ɑ,ʌ/ are displayed for the hearing-impaired subject of Figure 2. These two vowels were often confused by this subject; recognition scores were 69% and 73% respectively. The abnormal frequency resolution of this subject resulted in poorer resolution of the formants in general and in particular, a minimal difference in the internal representations of F1 frequency.

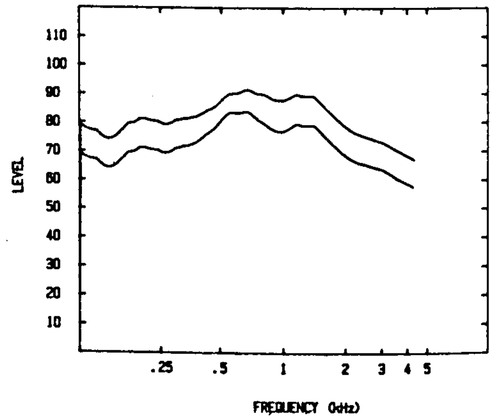

Figure 4. Internal representations for the vowels /ɑ/ (upper) and /ʌ/ (lower) from a hearing-impaired subject.

Figure 5 displays the calculated "internal representations" of the vowel pair /ɑ,æ/ for the same hearing-impaired subject as above. These vowels were never confused by the subject, due to the large difference in F2 frequency, which was preserved even after the abnormal peripheral filtering of this hearing-impaired subject.

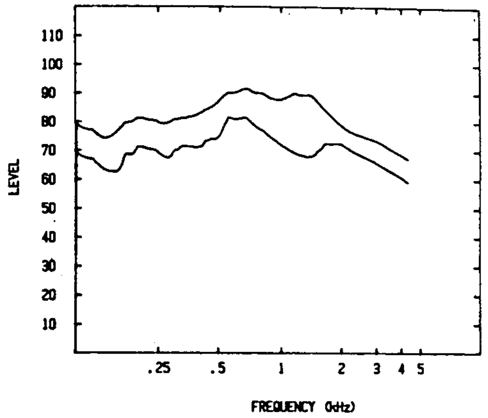

Figure 5. Internal representations for the vowels /ɑ/ (upper) and /æ/ (lower) from a hearing-impaired subject.

Conclusions

The input filter patterns provide a promising approach to expressing the representation of the spectral cues of speech in individual subjects. We have shown that measures of auditory filtering can predict the strength of percept for formant peaks in individual subjects, suggesting that auditory filter measures could be used to dictate the types and degrees of speech-cue enhancement that hearing-impaired subjects may require. The calculated "internal representations" of speech sounds provide a theoretically attractive mechanism for investigating the dependence of speech recognition upon the resolution of spectral detail.

ACKNOWLEDGEMENTS

This research was supported in part by a grant from the Deafness Research Foundation. The authors thank Carol Henn and Swati Lotlikar for assistance in data collection.

REFERENCES

1. Klatt, D.H. Prediction of perceived phonetic distance from crtical-band spectra: A first step. in Proceedings of the International Congress on Acoustics, Speech, Signal Processing. Paris, IEEE pp. 1278-1281 (1982).

2. Chistovich, L., Central auditory processing of peripheral vowel spectra., J. Acoust. Soc. Am., 77, 789-805 (1985).

3. Searle, C.L., Jacobsen, J.F. and Rayment, S.G. (1979) Stop-consonant discrimination based upon human audition., J. Acoust. Soc. Am., 65, 799-809 (1979).

4. Kewley-Port, D. Time-varying features as correlates of place and articulation in stop consonants., J. Acoust. Soc. Am.,73, 322-335 (1983).

5. Turner, C.W. and Robb, M.P. Audibility and recognition of stop consonants in normal and hearing-impaired subjects., J. Acoust. Soc. Am., in press (1987).

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

7. Trees, D.A. and Turner, C.W. Spread of masking in normal and high-frequency hearing-loss subjects., Audiology, 25, 70-83 (1986).

8. Wightman, F.L., McGee, T. and Kramer, M. Factors influencing frequency selectivity in normal and hearing-impaired listeners., in Psychophysics and Physiology of Hearing, Evans, E.F. and Wilson, J.P. (eds), Academic Press (1977).

9. Zwicker, E. and Schorn, K. Psychoacoustical tuning curves in audiology., Audiology, 17, 120-140 (1978).

10. Nelson, D.A. and Turner, C.W. Decay of masking and frequency resolution in sensorineural and hearing-impaired listeners., in Psychoacoustical, Physiological and Behavioral studies in Hearing., Van den Brink and Bilsen (eds), Delft University Press (1981).

11. Glasberg, B.R. and Moore, B.C.J. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments., J. Acoust. Soc. Am., 79, 1020-1033 (1986).

12. Turner, C.W., Holte, L.A. and Relkin, E.R. Auditory filtering and the discrimination of spectral shapes by normal and hearing-impaired subjects., J. Rehab. Res. Devel., 24 (1987).

13. Bacon, S.P. and Brandt, J.F. Auditory processing of vowels by normal-hearing and hearing-impaired listeners., J. Speech Hear. Res.,25, 339-347 (1982).

14. Sidwell, A. and Summerfield, Q. The effect of enhanced spectral contrast on the internal representation of vowel-shaped noise., J. Acoust. Soc. Am., 78, 495-506 (1985).

15. Van Tasell, D.J., Fabry, A.A. and Thibodeau Vowel identification and vowel masking patterns of hearing-impaired subjects., J. Acoust. Soc. Am., in press (1987).

16. Vershuure, J. Pulsation patterns and nonlinearity of auditory tuning I. Psychophysical results., Acustica, 49, 288-306 (1981).

17. Turner, C.W. and Holte, L.A. Discrimination of spectral-peak amplitude by normal and hearing-impaired subjects., J. Acoust. Soc. Am., 81, 445-451 (1987).