CRITICAL BANDS
IN THE PERCEPTION OF SPEECH SIGNALS
BY NORMAL AND SENSORINEURAL HEARING LOSS LISTENERS

ROBERT D. CELMER

Acoustics Program and Laboratory
Department of Mechanical Engineering
University of Hartford
West Hartford, CT, USA 06117

GORDON R. BIENVENUE

Communications Department
State University of New York
College at New Paltz
New Paltz, NY, USA 12561

ABSTRACT

The results of a novel speech intelligibility test have been correlated to a tonal complex test, estimating critical bandwidth for normal and sensorineural hearing impaired listeners.

INTRODUCTION

Existing auditory theory suggests a major role for critical bands. Scharf [1] has defined the critical band empirically as "...that bandwidth at which subjective responses rather abruptly change...". In general, two stimuli separated in frequency by less than a critical bandwidth will interact in one of a number of ways, while two stimuli separated by more than a critical bandwidth will not. The critical band phenomenon has been observed in such perceptual phenomena as masking [1,2], loudness [3], and musical consonance [4].
Speech is the most pervasive and significant acoustic stimulus for the human listener, and evidence suggests that the critical band may serve in the analysis of speech [1]. In several studies, bandwidths contributing equally to the perception of speech were approximately equal to critical bands found in pure tone psychoacoustic studies [5,6,7,8,9].
The works of Fletcher [2], Zwicker [3], and Greenwood [10] imply that the critical band serves to band-limit background noise. The narrower the passband of the ear as a filter, the more noise the ear can reject. Thus, a listener may be able to correctly perceive a spoken communication despite background noise simply because much of the energy associated with the noise lies outside the critical bands surrounding the formant frequencies of the speech.
Discrimination of the formant and harmonic content of both speech and non-speech signals requires that these components be separated by at least one critical band [4,11,12]. Synthetic vowels presented to listeners by Remez [13] showed an abrupt changeover from speech-like to non-speech-like sounds as the formant bandwidth increased to greater than a critical bandwidth.

Several researchers have reported evidence of distorted or widened critical bands in subjects with sensorineural hearing loss [14,15,16,17,18,21]. In addition, these data demonstrate that the width of the widened critical band is independent of the magnitude of threshold hearing loss amongst those sensorineurals with critical bandwidth distortion.
The purpose of the present study was to directly test the hypothesis that the critical band is an essential element in auditory speech discrimination.

METHODS AND MATERIALS

A. Subjects
Forty-eight normal hearing listeners, aged 19 to 31 years, and sixty-eight sensorineural hearing impaired listeners, aged 18 to 67 years, participated in the study. Each subject was classified using conventional audiometric techniques. The right ear was the test ear for all subjects.
In addition, subjects' critical bandwidths were independently measured using a loudness of complexes test [1]. The center frequencies of each tonal complex were located at 700 Hz, 1000 Hz, 1600 Hz and 2150 Hz. For each trial a sub-critical tonal complex was presented to the listener and the bandwidth of the tonal complex was increased in time by small amounts. Ten repetitions of each center frequency were performed at a presentation level of 50 dB HL. Subjects were asked to listen to each test signal episode, and indicate the moment they perceived a change in the stimulus. The signal bandwidth required to elicit a perceptual change was recorded for each of the trials for each subject.

B. Taped Stimulus Materials
The stimulus materials used in this study were a set of limited resolution bandwidth speech signals. Prerecorded stimuli (in analog form) were digitized by a 12-bit A/D converter at 20,000 samples per second, and stored in a computer. The software program separates the signals into 13 ms segments. The average spectrum of each segment is computed by first applying a Hamming window and then performing a Fast Fourier Transformation. The resolution bandwidth of this discrete spectrum is then limited by the software program. The frequency limits used for the normal critical bandwidth condition were those recommended by Scharf [1]. The discrete frequency amplitudes that fell within each bandwidth are averaged; each of the discrete amplitudes of that band are then set equal to this r.m.s. value, limiting the resolution allowed to the preselected bandwidth. See Figure 1. Coarser and narrower filtering schemes were realized by multiplying the bandwidth limits used by a chosen factor (retaining the original center frequency), and averaging the amplitudes contained within these widened limits. Each processed spectrum is then inverse transformed into the time domain, an inverse Hamming window is applied, and the output is converted to analog form via a D/A converter and recorded on audio tape.

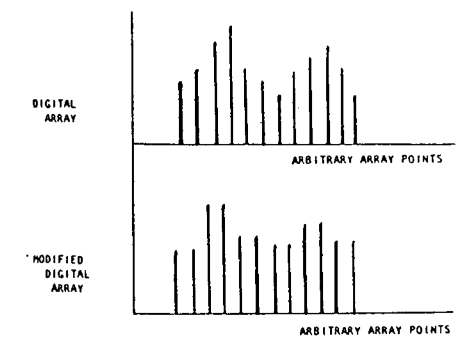

Figure 1. Effect of Digital Filtering

The NU#6 word list, a clinical audiometric word list, was used as the input audio material for the speech processing algorithm, since it includes (consonant-consonant-nucleus-consonant) sounds as opposed to only (consonant-consonant-vowel-consonant) sounds [19].
The tapes generated have seven frequency resolutions: an unprocessed list (UP); a bandwidth equal to one-half the resolution of the normal critical bandwidth (HX) [1]; a bandwidth equal to the resolution of the normal critical bandwidth (1X); two times normal (2X); three times normal (3X); five times normal (5X); and seven times normal (7X).

C. Equipment
The processed signals described above were generated using a hybrid computer system comprised of an EAI (Electronic Associates Incorporated) Model 680 analog computer interfaced with a DEC (Digital Equipment Corporation) digital computer, Model PDP-10. The audio output was recorded via a Crown Model BP824 tape deck.
The discrimination tasks were performed using an Ampex AG-440B tape recorder, a Maico Model MA-18 audiometer calibrated to ANSI 1969 standards, and TDH-39 earphones fitted with MX-41/AR cushions. The tests were performed in a Suttle Corporation Model B1 quiet room.

D. Test Procedure
After the audiometric threshold test and the independent critical bandwidth test, the subjects were presented with the seven fifty-word lists in the following resolution order: 7X, 5X, 3X, 2X, 1X, HX, UP. This sequence was chosen in order to minimize learning effects. The signal reached the earphone at a level of 50 dB HL and a signal-to-noise ratio of +10 dB. Pink noise was utilized as the masking source. Masking was used to minimize ceiling effects. Subjects were provided with answer sheets and were asked to write down the word they felt was said, guessing when necessary. Word lists were scored on a percentage basis.

RESULTS

A. Normal Hearing Subjects
The normal hearing group's mean pure-tone thresholds ranged from 1.5 to 5.6 dB for the frequencies from 500 Hz, to 4000 Hz. On the loudness-of-complexes test for independent critical bandwidth measurement, this group showed a mean critical bandwidth of 0.94 times the normal critical band as reported by Scharf; that is, 0.94X. The range of values was from 0.7X to 1.2X over the four test frequencies measured, confirming that this group did indeed have normal critical bands.
The plotted means for processed speech intelligibility scores are presented in Figure 2. A regression line was computed for the 2X through 7X portion of the data. The bandwidth condition vs the group's speech intelligibility scores exhibited a significant correlation of $r = -.75$ with a probability of error less than .01.

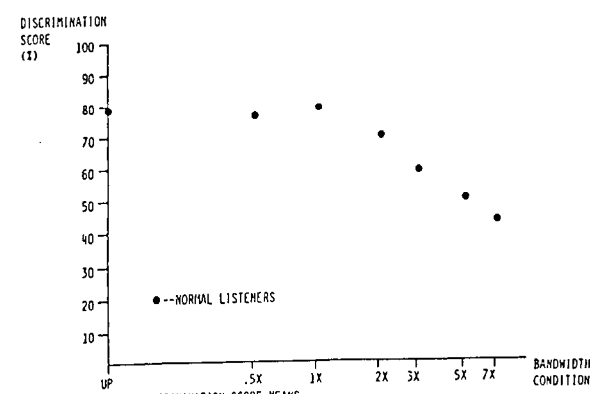

FIGURE 2. PLOTTED DISCRIMINATION SCORE MEANS.

As may be seen from Figure 2, the discrimination score was a direct function of the log of bandwidth resolution, for those bandwidths wider than one critical band, but was independent of bandwidth for those bandwidth resolution conditions equal to or narrower than the critical band. The average discrimination score for the UP, HX and 1X condition was 77.56%. Thus, the masker did eliminate a ceiling effect.

A single factor analysis of variance indicated a significant main effect of the speech processing. A Newman-Keuls follow-up test [20] demonstrated a significant decreasing trend for discrimination scores as the bandwidths were varied from 2X through 7X. However, no significant decreases in intelligibility scores were observed for the UP (unprocessed) through 1X conditions. In addition, the group scores at the 1X condition were significantly higher than those at the 2X condition.

B. Sensorineural Hearing Impaired Subjects

The sensorineural hearing impaired group's mean pure-tone threshold audiogram ranged from 22.4 dB at 250 Hz to 71.2 dB at 8000 Hz (See Table 1). The mean speech reception threshold was 37.6 dB and the mean speech intelligibility in quiet at 50 dB HL was 74.7 %. On the loudness of complexes test for independent critical bandwidth measurement, this group showed a range of widened critical bands, from 1.1X to 4.23X, with a mean bandwidth of 2.43X. The subjects fell into four critical bandwidth groups: a 1X group (range of bandwidths 1.1X to 1.5X), a 2X group (1.5X to 2.5X range), a 3X group (2.5X to 3.5X range), and a 4X group (3.5X to 4.23X range).

Table 1.

MEAN AUDIOMETRIC DATA FOR
HEARING IMPAIRED LISTENERS

f (Hz)	250	500	1000	2000	4000	8000
TL (dB)	22.4	29.3	38.3	60.2	65.6	71.2

The relationship between the critical bandwidth rating determined from the loudness-of-complexes test and the "knee" of the processed speech intelligibility data was of particular interest in the results of the study. This "knee" continued to represent the processing bandwidth at which the intelligibility scores dropped significantly from the unprocessed score. See Figure 3. These observations were made by performing a similar statistical analysis as used on the normal data.

It is interesting to note that the mean speech intelligibility functions for the subjects in the present study show a "knee" that coincides with their critical bandwidth test. This is especially easy to observe when the data are grouped according to the result of the loudness-of complexes test. The correlation coefficient between the loudness-of-complexes bandwidth measurement and the "knee" of the speech intelligibility function was 0.8749. This is a high correlation and indicates a very strong relationship between tonal estimates of the critical band and speech intelligibility of sensorineural, hearing impaired listeners.

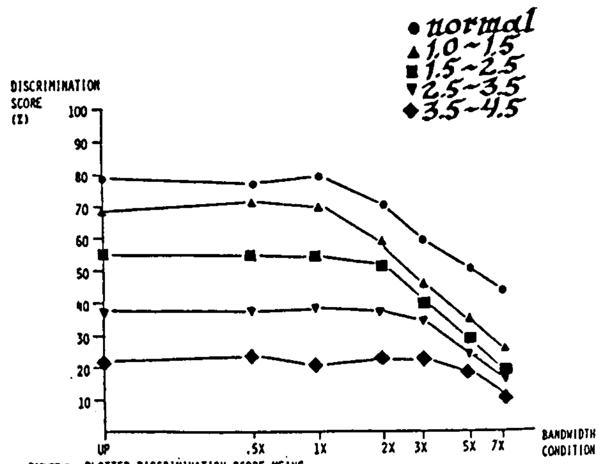

FIGURE 3. PLOTTED DISCRIMINATION SCORE MEANS.

DISCUSSION AND CONCLUSIONS

As noted above, the peripheral auditory system has been described to perform a preliminary frequency analysis of incoming acoustic signals. The limit to which frequency information may be gated is called the critical band and has been observed in a variety of psychoacoustic contexts. In particular, the critical band mechanism performs noise-band limiting and harmonic discrimination, both of which are crucial for the correct perception of such complex acoustic stimuli as speech. Thus, it was hypothesized that the critical band is a contributing factor to normal auditory speech intelligibility.

The purpose of the present study was to test the hypothesis that the critical band is an essential element in the process of speech listening. The performance of normal and sensorineural hearing impaired listeners with the processed speech test indicated the same three distinct trends:

1) A plateau effect was noted for the processing conditions UP through the condition equal to their independent critical band rating. For example, the group whose critical band was measured independently as 2X had a plateau effect occur from UP to 2X.

2) The score of the highest processing condition in the plateau was significantly higher than the next higher processing condition score.

3) A monotonic, decreasing trend was observed as the allowed bandwidth resolution was widened beyond the "knee" in the curve.

The scores above the observed "knee" in the curve demonstrated a close approximation to a logarithmic curve with a negative slope. Also significant is the correlation between the independent critical bandwidth rating and the location of the "knee" in the curve. In other words, the independent critical bandwidth rating corresponded to the observed point of inflection demonstrated by the speech intelligibility scores. This high correlation indicates a very strong relationship between tonal estimates of the critical band and speech intelligibility of normal and sensorineural hearing impaired listeners. It is reasonable to conclude that the integrity of the critical band is an important factor in the understanding of speech signals.

In summary, a correlation has been demonstrated between the performance of listeners on an auditory speech intelligibility test and a test with tonal complexes as the stimuli. The tonal complex test results yielded a bandwidth resolution value that was correlated to the probable point of inflection of the auditory speech intelligibility test results. Existing literature has implied, but has not clearly demonstrated the presence of such a correlation. The authors suggest that any consideration of auditory, speech intelligibility among normal or sensorineural, hearing impaired listeners must include an examination of the integrity of the critical band phenomenon in the subject population.

REFERENCES

[16] Bennett, T., and others. "Procedures for Characterizing Certain Effects of Prolonged Noise Exposure." Journal of the Acoustical Society of America (JASA), 63, Supplement 1, 1978.

[15] Bienvenue, G., and P. Michael. "The Temporary Effects of Short Term Noise Exposure on Masking Phenomenon." Unpublished EAL Research Project, 1977.

Bienvenue, G., and others. "Digital Processing of Speech Materials in the Study of Sensorineural Hearing Impairment." JASA, 67, Supplement 1, 1980.

[22] Bonding, P. "Critical Bandwidth in Presbycusis." Scandinavian Audiology, 8, 1979.

[8] Castle, W. "Effects of Selective Narrow Band Filtering on Perception of PB Word Lists." JASA, 36, 1964.

[9] Chari, N. "Perception of 1/3 Octave-filtered Speech." JASA, 61, 1977.

[21] deBoer, E., and Bouwmeester, J., "Critical Bands and Sensorineural Hearing Loss," Audiology, 13, 236-259, 1974.

[2] Fletcher, H. "Auditory Patterns." Review of Medical Physiology, 12, 1940.

[5] French, N., and J. Steinberg. "Factors Governing the Intelligibility of Speech Sounds." JASA, 19, 1947.

[10] Greenwood, D. "Auditory Masking and the Critical Band." JASA, 33, 1961.

[12] Haggard, M. "Feasibility of Rapid Critical Bandwidth Measurements." JASA, 55, 1974.

[7] Kryter, K. "Speech Bandwidth Compression through Spectrum Selection." JASA, 32, 1960.

[14] Michael, P., and G. Bienvênue. "A Procedure for the Early Detection of Noise Susceptible Individuals." American Industrial Hygiene Association Journal, 1976.

[11] Morton, J. and A. Carpenter. "Experiments Relating to the Perception of Formants." JASA, 35, 1963.

[4] Plomp, R. "The Ear as a Frequency Analyzer." JASA, 36, 1964.

Rabiner, L. and R. Schafer., Digital Processing of Speech Signals, ed. A. Oppenheim, (Englewood Cliffs: Prentice-Hall, 1978).

[13] Remez, R. "Adaption of the Category Boundary Between Speech and Nonspeech: A Case Against Feature Detectors." Haskins Laboratories: Status Report on Speech Research, SR-50, 1977.

[6] Richards, D. and R. Archbald. "A Development of the Collard Principle of Articulation Calculation." Proceedings of the IEE, 103B, 1956.

[1] Scharf, B. "Critical Bands." Foundations of Modern Auditory Theory, ed. J. Tobias, (New York: Academic Press, 1970).

[19] Tillman, T., and R. Carhart. "An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words (Northwestern University Auditory Test No. 6)." Technical Report, SAM-TR-66-55, United States Air Force School of Aerospace Medicine, Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas, 1966.

[3] Zwicker, E. "Uber Psychologische und Methodische Grundlagen der Lautheit." Acustica, 8, 1958.