Signal Processing to Improve Intelligibility in the Presence of Noise for Persons with a Ski-Slope Hearing Impairment

PAUL YANICK AND HARRIS DRUCKER, MEMBER, IEEE

Abstract—A two-band compression system was designed to improve the intelligibility of speech in noise for persons suffering from sensorineural hearing impairment. Experiments were carried out at 0 and +6 dB signal-to-noise (S/N) ratio with expansion below compression threshold. Six subjects had mean discrimination scores at S/N of 6 dB of 46, 77, and 87 percent with unprocessed speech (present hearing aid), compression but linear below threshold, and compression with expansion below threshold, respectively. At the 0 dB S/N test score means were 21, 61 and 69 percent, respectively.

INTRODUCTION

AS NOISE LEVELS increase or as the speech becomes less redundant the deleterious effects of lower frequency noise on word intelligibility increases drastically. Possession of a sensorineural hearing loss increases the masking effect of noise which can be demonstrated 1) by a shift in the signal-to-noise (S/N) ratio at which the speech reception threshold (SRT) is reached [1], and 2) by a change in the intelligibility function [2]. In addition, the use of a hearing aid greatly increases the disruptive effect of competing background noise. Many subjects wearing hearing aids complain that they can function better without their hearing aids when in the presence of background noise [3].

There is a critical need for concentrated research to develop a hearing aid which can give its user more meaningful information in the presence of noise. The immediate problem lies in defining how and to what extent do the various electroacoustical properties of hearing aids relate to the pathological ear. Subjects with sensorineural hearing impairments demonstrate a loss of frequency discrimination (i.e., an increase in the difference limen for pure tones and speech like sounds [4]), and exhibit a reduced dynamic range of hearing, recruitment, and other cochlear distortions. Fig. 1 shows threshold discomfort levels [19] for both a subject suffering from sensorineural hearing loss and an average of normal hearing subjects. Included is the range of conversational speech [27]. If the speech is amplified to bring the high-frequency sounds up to the deaf person's perception levels, the discomfort level may be exceeded at low frequencies, reducing intelligibility. Even with amplification, the dynamic range of speech will be unaltered and greater than the deaf subject's perceptual dynamic range at high frequencies. Amplitude compression could be used on the speech to reduce the dynamic range.

Manuscript received December 23, 1975; revised June 16, 1976.
P. Yanick is an Audiologist at 673 Wood Avenue, Edison, NJ 08815.
H. Drucker is with the Department of Electrical Engineering, Monmouth College, West Long Branch, NJ 07764.

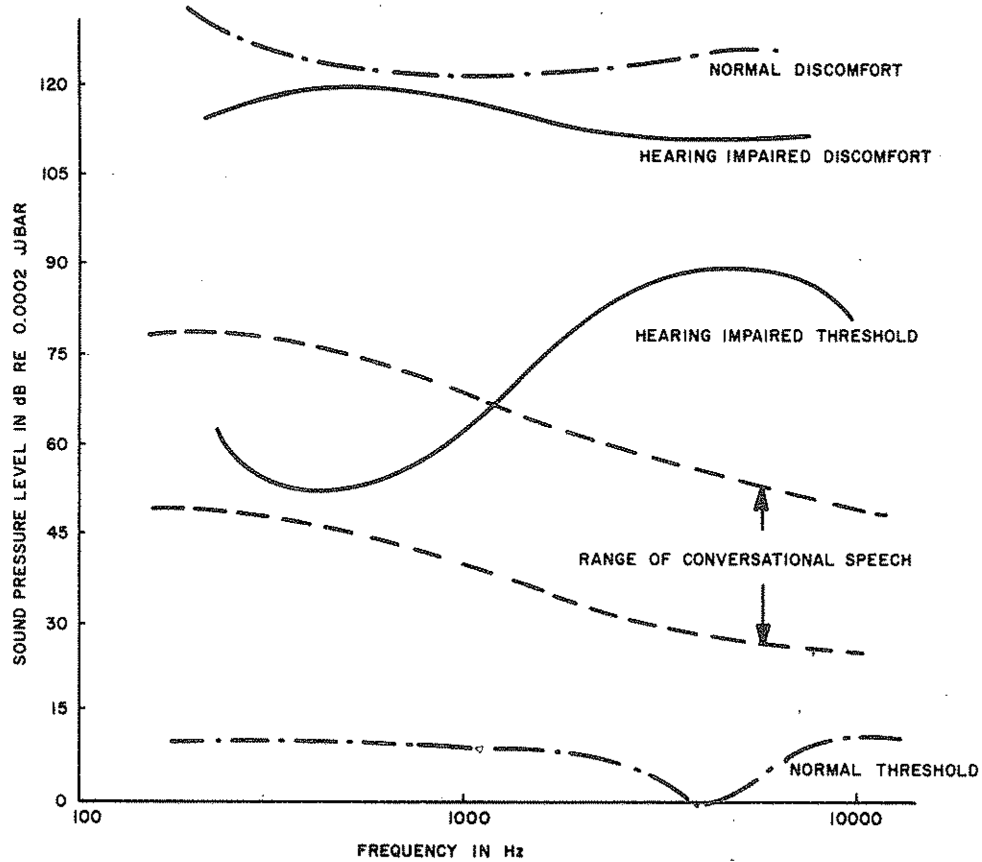

Fig. 1. Threshold-discomfort levels for hearing impaired subject and normal subjects [19]. Superimposed is the approximate range for conversational speech (from [27]). The dynamic range of conversational speech is greater than the hearing impaired subject's perceptual dynamic range at high frequencies. Amplification to raise the high-frequency sounds to within perceived levels would cause the low-frequency sounds to exceed the low-frequency discomfort level.

Signal Processing by Clipping and Related Techniques

Some researchers have concluded that the intelligibility of clipped speech is dependent on the relative weighting given to the various frequency components of speech by filter networks prior to clipping [5], [6]. Thomas and Sparks have demonstrated infinite amplitude clipping with low-frequency suppression on patients with sensorineural hearing loss and found that discrimination scores improved significantly [7]. Thomas and Neiderjohn (1970) reported that a speech signal which is highly immune to noise can be produced by passing normal speech through a high-pass filter with a cutoff frequency of 1100 Hz and an asymptotic slope of 12 dB/octave and then by infinitely clipping the speech signal [8]. In a recent experiment Thomas and Ravindran using the same technique found higher intelligibility scores even though the speech at the very input to the process is already seriously contaminated by noise [9].

Other attempts in the literature to improve discrimination scores (DS) in the presence of noise included the acoustic modification of earmolds [10], [11] and the use of a directionally sensitive microphone [12]. Results show considerable improvements of DS in the presence of noise.

Signal Processing by Amplitude Compression and Frequency Equalization

The use of amplitude compression to compensate for recruitment and to provide speech discrimination enhancement and relief from the masking effect of the stronger components of speech (i.e., vowels) and background noise has received limited attention. Experimental evidence on amplitude compression have been rather contradictory and are summarized elsewhere [13]. The findings of Carhart and his colleagues demonstrate that compression offered no benefit to the impaired ear [14]-[16]. Others [17]-[19], [22] present evidence that amplitude compression can improve speech intelligibility for listeners with sensorineural hearing losses.

Villchur demonstrated dramatic improvements in intelligibility with a two-channel amplitude compression system combined with frequency equalization for subjects with sensorineural hearing loss [19]. More recently, Yanick [13] processed speech with the same system design used by Villchur on 24 subjects with sensorineural hearing losses. Subjects were divided into two groups, one with ski-slope hearing loss and the other with a flat loss hearing loss. A person with ski-slope loss has degradation in the threshold curve of 6-25 dB/octave typically starting around 750 Hz. Each subject was

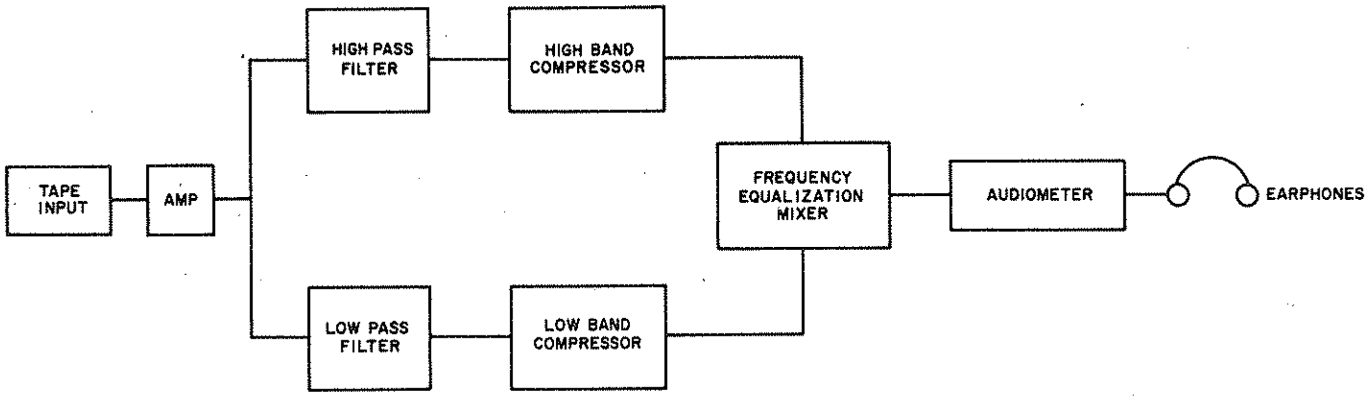

Fig. 2. Block diagram of signal processing unit.

tested with speech in the presence of cafeteria noise at a S/N ratio of +6 dB and 0 dB at preferred listening levels. Results demonstrate dramatic improvements over their present hearing aids in key words identification for processed speech which ranged from 12-46 percent, mean 28.2 percent for the ski-slope group and 6-74 percent, mean 37.4 percent for the flat loss group at a S/N ratio of 6 dB. The range of improvement was 8-86 percent, mean 38.5 percent for the ski-slope group and 8-72 percent, mean 39.4 percent for the flat loss group at 0 dB. Improvements are attributed to 1) the use of compression to reduce the dynamic range of the speech to that of the individual's dynamic range of perceived sounds, 2) the use of two-band compression to compensate for hearing loss characteristics within each band, 3) the presence of a threshold below which the compressor is linear to prevent the compressor placing itself in a high-gain state when low-level noise only is present, and 4) the choice of the proper attack and release times.

Previous research on amplitude compression that failed to demonstrate any benefit over linear amplification can be attributed to poor electronic design techniques, inappropriate attack and release times disguising transitional clues in the speech waveforms, and the use of one-band compression.

In the following sections, we will show how the use of a combination of a two-band system, compression with expansion below threshold, and equalization can improve intelligibility.

Subjects

Six subjects (aged 30-48, mean 39 years) were chosen for this study. An audiologic examination was first performed on each subject consisting of monaural pure tone air conduction and bone conduction thresholds, SRT's, impedance measurements and measurements of minimum audible pressure (MAP), most comfortable loudness levels (MCL), and equal loudness contours to determine the degree of recruitment present at each frequency.

The criteria for selection of the six subjects were 1) degree of loss: 25-60 dB HL (mild-to-moderate sensorineural hearing loss cases), obtained by averaging the loss at 500, 1000 and 2000 Hz, 2) configuration of loss: ski-slope hearing loss with 15-25 dB per octave slopes starting in the mid frequencies, 3) recruitment present as determined by measurements of MAP, MCL, and equal loudness.

Measurements for wavetrain signals were taken according to the method described by Victoreen [26]. The degree of recruitment present was determined by the difference in responses of the normal and abnormal ear to intensity levels near or above threshold.

Test Materials

Tape recordings consisted of 30 lists of Harvard sentences as revised by IEEE (1969) recorded in a reverberant environment on one track and cafeteria noise recorded on the other track. Each track has a 1000-Hz calibration. The level of the calibration tone was set to equal the average level or the peaks of speech for the sentence material and the background noise as determined with a standard UCL meter. A group of 20 different sentences were presented for each test condition. Scoring was based on five italicized words in each of the seven to nine word sentences. The five key words consisted of four monosyllables and one disyllable.

Experiment 1

A pair of dBX compressors were modified to provide compression thresholds 30 dB below maximum peak speech levels in the low-band compressor and 36 dB below maximum peak speech levels in the high-band compressor (Fig. 2). Compressor output at maximum speech input level was independent of the compression ratio. The compression release time was changed to 20 ms, while the attack time remained at 1 ms.

The compressor has signal inputs fed from high-pass and low-pass filters with a channel crossover frequency of 1500 Hz. Tape input consisted of a Pioneer RT-1050 $\frac{1}{2}$ track tape recorder, which was fed into a Kenwood KA14006 stereo amplifier to provide the proper speech levels and different S/N mixtures.

Post-frequency equalization was accomplished by bass and treble controls which were built into the subject control panel. The output of the signal processor was fed into a Beltone 200c audio meter to a telephonic Model 556 earphone with a hearing aid receiver (Knowles 1710) housed in it. The

TABLE I
COMPARISON OF KEY WORD IDENTIFICATION SCORES FOR THE UNPROCESSED AND PROCESSED TEST CONDITIONS FOR S/N RATIO OF 0 dB AND 6 dB

	S/N Ratio +6 dB			S/N Ratio 0 dB		
Subj.	Unprocessed	Experiment 1 Processed	Experiment 2 Processed	Unprocessed	Experiment 1 Processed	Experiment 2 Processed
1	52%	78%	86%	16%	58%	64%
2	36%	72%	90%	12%	68%	76%
3	58%	86%	92%	34%	78%	86%
4	62%	90%	96%	28%	52%	58%
5	24%	58%	76%	10%	46%	60%
6	44%	78%	84%	30%	68%	72%
mean	46%	77%	87.3%	21.6%	61.6%	69.3%

high-frequency cutoff of the system was set at 6 kHz. The earphones (according to the manufacturer) provide the closest approximation to free-field threshold pressure by minimizing cushion induced physiological noise. An otometric wavetrain signal generator developed by Victoreen Laboratory was used for equal loudness measurements. Measurements were obtained free field and with circumaural earphones.

Subjects were seated in a 10 × 10 ft sound booth so that the control panel was directly in front of them and they could adjust the controls individually. Each subject first adjusted the output control for maximum clarity and a comfortable listening level. Subjects were instructed to turn the knob to the right or left several times while listening to a sentence which kept repeating itself on an endless tape. Then in the following order subjects adjusted: the high-channel compression ratio (in discrete values), the low-channel compression ratio, and overall level for the highest intelligibility consistant with a comfortable listening level. Then choices were made with bass and treble controls for frequency equalization. Five different low-frequency rolloffs at 2 kHz were available for the low band and five different treble responses for the high band.

Unprocessed speech consists of frequency equalization on bass controls only, with the subjects own hearing aid used. An optimum listening level was chosen in the same manner as the processed speech, the subjects adjusting the output for comfort and highest intelligibility. According to previous experiments in a preliminary investigation of signal processing, unprocessed speech with a treble boost did not change the intelligibility scores of unprocessed speech. Subjects were tested first at a S/N ratio of 0 dB and then at a S/N of +6 dB. Three subjects listened to the processed speech first and three listened to the unprocessed speech first. Subjects task were to repeat what he heard even if it was only one word in the sentence.

EXPERIMENT 1 (RESULTS)

The results show an improvement in key word identification which ranged from 26 to 36 percent, mean 31 percent at a S/N ratio of 6 dB, and an improvement which ranged from 24 to 56 percent, mean 41 percent, at a S/N of 0 dB (Table I). The average compression ratio chosen was 1.5:1 for the low-band compressor and 2.5:1 for the high-band compressor. The average bass rolloff these subjects choose was 12 dB per octave at 1500 Hz.

DISCUSSION

The higher percentage of key work identification demonstrates the importance of processing speech with a two-channel amplitude compression system. These results and the findings of Villchur [19] and Yanick [13] suggest that by compensating for recruitment more acoustic cues are provided in the speech signal which makes it more resistant to the disruptive effect of competing background noise.

EXPERIMENT 2

Many of the subjects commented that the speech signal had more clarity, but they also complained that the relative level of background competition increased especially during periods of no speech signals. This form of degradation is inherent in compression systems. The increase in gain provided for low-level speech sounds also amplifies the low-level background noise. This degradation may be detrimental to speech intelligibility in situations where there is a steady-state background noise. More so, it may be annoying to persons wearing these hearing aids in a noncommunicative situation where there is a low-level background noise.

Methods and electronic design techniques have been devised that can reduce this degradation [21], [23], [25]. One method is to modify the compressor to provide expansion below the compression threshold [16], [26]. Expansion reduces the gain and causes low-level ambient noise to fall away rapidly thus reducing the relative level of background noise.

Since frequency components of environmental noise are predominantly in the lower frequency region (below 1500 Hz) and most subjects with sensorineural hearing losses have good low-frequency hearing but poor high-frequency hearing, it was arbitrarily decided to provide expansion below the compression threshold only in the low band of compression.

In order to present the taped Harvard sentences and cafeteria noise with an expansion ratio below the compression

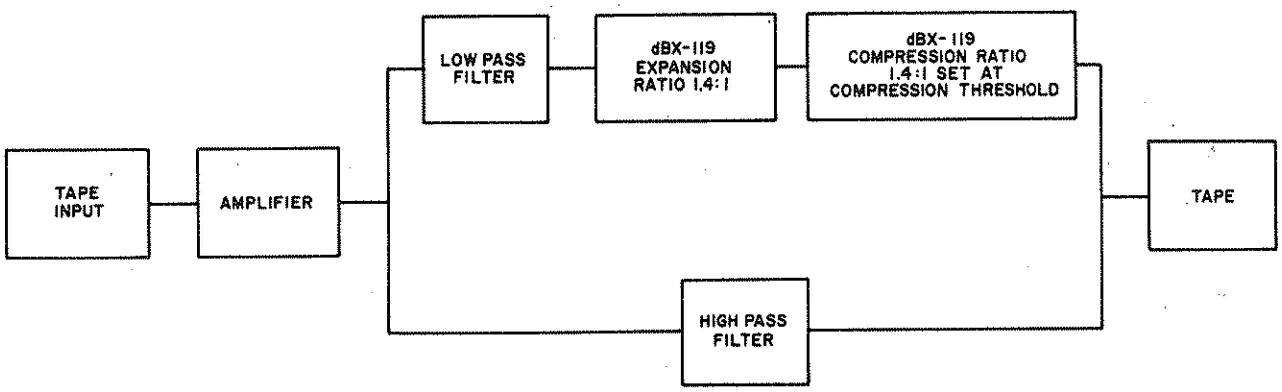

Fig. 3. Block diagram of method used to tape test materials with an expansion mode below the compression threshold in the low band.

threshold in the low band, the tapes were run first at a S/N ratio of +6 dB and then at a S/N ratio of 0 dB. A Pioneer RT-1050 $\frac{1}{2}$-track tape recorder played the tape through a Kenwood KA 14006 stereo amplifier at the two S/N ratios. The output of the amplifier was fed into the low and high bandpass filters with the crossover frequency at 1500 Hz. The low-band filter fed a dBX 119 set in an expansion mode with an expansion ratio of 1.4:1 (no threshold) and another dBX 119 set with a compression ratio of 1.4:1 above the compression threshold. The high bandpass filter and the output of the dBX 119 set in the compression mode were fed into a TEAC A-1220-U $\frac{1}{2}$-track tape recorder (Fig. 3). This produces a tape with 1.4:1 expansion below threshold and linear above threshold in the 1500 Hz low band.

The same six subjects were then retested for processed speech with the compression ratios and frequency equalization settings chosen in Experiment 1 at a S/N ratio of +6 dB and 0 dB. Fifteen different lists of Harvard sentences were used for this experiment and test presentation levels were set the same as in Experiment 1.

Experiment 2 (Results)

Key word identification scores at a S/N ratio of +6 dB showed significant improvement with scores that ranged from 76 to 96 percent, mean 87.3 percent as compared to a range of 58-90 percent, mean 77 percent obtained in Experiment 1. At a S/N ratio of 0 dB scores ranged from 58 to 86 percent, mean 69.3 percent as compared to 46-78 percent, mean 61.6 percent obtained in Experiment 1 (Table I).

Discussion

Many listeners with sensorineural hearing loss have little difficulty hearing in "quiet" situations. They purchase a hearing aid to alleviate difficulties they have in noisy situations. Unfortunately, they find they can hear better in the presence of noise with their aids off.

Speech communication often takes place in the presence of interfering noise. A speaker usually adjusts his voice on the basis of the difficulty he experiences as a listener. When the interfering noise level is high he attempts to preserve the S/N ratio by speaking louder to the listener. The problem of providing the impaired ear with more meaningful information in the presence of background noise is no easy task. Noise, like speech has varying sound pressures, frequencies, and duration which can change the S/N ratio from moment to moment.

Compensating for recruitment by signal processing and reducing the relative level of background noise levels by providing expansion below the compression threshold can alleviate some of the difficulties faced by these listeners with sensorineural hearing loss and provide them with more meaningful information as well as comfort and protection from the deleterious effects of noise.

References

[1] R. Wilson and R. Carhard, "Influence of pulsed masking on the threshold for spondees," *J. Acoust. Soc. Amer.*, vol. 46, pp. 998-1010, 1969.

[2] R. Carhart and T. W. Tillman, "Interaction of competing speech signals with hearing loss," *Arch. Otolaryng.*, vol. 91, pp. 273-279, 1970.

[3] T. W. Tillman, R. Carhart, and W. O. Olsen, "Hearing aid efficiency in a competing speech situation," *J. Speech Hear. Res.*, pp. 789-811, 1970.

[4] J. N. Pickett and E. S. Martin, "F^1 masking of F^2 transition discrimination in hearing impaired listeners," *J. Acoust. Soc. Amer.*, vol. 48, p. 93, 1970.

[5] J. C. R. Licklider and I. Pollack, "Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech," *J. Acoust. Soc. Amer.*, vol. 20, pp. 42-51, 1947.

[6] I. Pollack and J. M. Pickett, "Intelligibility of peak-clipped speech at high noise levels," *J. Acoust. Soc. Amer.*, vol. 31, pp. 14-16, 1959.

[7] I. B. Thomas and D. W. Sparks, "Discrimination of filtered/clipped speech by hearing impaired subjects," *J. Acoust. Soc. Amer.*, vol. 49, no. 6, pp. 1881-1887, 1971.

[8] I. B. Thomas and R. J. Neiderjohn, "The intelligibility of filter clipped speech in noise," *J. Audio. Eng. Soc.*, vol. 18, pp. 299-303, 1970.

[9] I. B. Thomas and A. Ravindram, "Intelligibility enhancement of already noisy signals," *J. Audio Eng. Soc.*, vol. 22, no. 4, pp. 234-236, 1974.

[10] J. Northern and K. Hattler, "Earmold influence on aided speech discrimination tasks," *J. Speech Hear. Res.*, vol. 13, 1970.

[11] M. McClellan, "Aided speech discrimination in noise with vented and unvented earmolds," *J. Audio Res.*, vol. 7, 1967.

[12] W. E. Lentz, "Speech discrimination in the presence of background noise using a hearing aid with a directional sensitive microphone," *Maico Aud. Lib. Series*, vol. 10, no. 9, 1972.

[13] P. Yanick, "The effects of signal processing on the intelligibility of speech in the presence of competition for subjects with sensorinerual hearing loss," *J. Amer. Aud. Soc.*, to be published.

[14] B. Caraway, "Effects on speech intelligibility of dynamic range reduction with a compressor amplifier," Ph.D. dissertation, Northwestern Univ., Evanston, IL, unpublished.

[15] B. J. Caraway and R. Carhart, "Influence of compressor action on speech intelligibility," *J. Acoust. Soc. Amer.*, vol. 41, pp. 1424-1438, 1967.

[16] S. W. Vargo and R. Carhart, "Amplitude compression on speech intelligibility in quiet with normal and pathological groups," *J. Acoust. Soc. Amer.*, vol. 53, p. 327(A), 1972.

[17] S. B. Burchfield, "Perception of amplitude compressed speech by persons exhibiting loudness recruitment," Ph.D. dissertation, Michigan State Univ., East Lansing, unpublished.

[18] W. Ruhrberg and G. Esser, "Die Wirkung der Dynamik-kompression bie Innenohrschwerhorighkeit und retrocochlearer Horstorung," *HNO*, vol. 21, pp. 247-255, 1973.

[19] E. Villchur, "Signal processing to improve speech intelligibility in perceptive deafness," *J. Acoust. Soc. Amer.*, vol. 53, pp. 1646-1657, 1973.

[20] R. P. Lippmann, "The effect of multichannel amplitude compression on the intelligibility of speech for persons with sensorineural hearing loss," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge (in progress).

[21] P. Yanick, "Hearing aid amplifiers employing selective gain control circuits," U.S. Patent 508998.

[22] —, "Improvement in speech discrimination with compression vs. linear amplification," *J. Aud. Res.*, vol. 13, pp. 333-338, 1973.

[23] D. E. Blackmer, "A wide dynamic range noise reduction system," *DB*, vol. 6, pp. 54-56, 1972.

[24] B. S. Blesser, "Audio dynamic range compression for minimum perceived distortion," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 22-32, Mar. 1969.

[25] B. S. Blesser and A. R. Kent, "Analysis of a feedback-controlled limiter using a logarithmic measuring scale," *IEEE Trans. Audio Electroacoust.*, vol. AU-16, pp. 481-485, Dec. 1968.

[26] J. A. Victoreen, *Basic Principals of Otometry*. Springfield, IL: Thomas, 1973.

[27] H. K. Dunn and S. D. White, "Statistical measurements on conversational speech," *J. Acoust. Soc. Amer.*, vol. 11, pp. 278-288, 1938.