Acoustic Evaluation of Noise Reduction in a Hearing Aid*

Yingyong Qi and Chul-Hee Choi. Acoustic Evaluation of Noise Reduction in a Hearing Aid. Korean Journal of Communication Disorders, 2000, Vol. 5, No.1, 119-132. The purpose of this study is to acoustically evaluate the noise reduction function of one specific type of hearing aid (Phonak, Piconet2 AZ). Recording were made using a microphone that was mounted on the Zwislocki coupler inside KEMAR. The signal-to-noise ratio (SNR) of the recorded signals were obtained for a variety of recording environments. These include: (1) high (0 dB SNR) and low (15 dB SNR) noise level, (2) three different types of noise (white, cafeteria, and multitalker), and (3) three different hearing aid programs (Peak Clipping, Wide Dynamic Range Compression, and Super Compression with adaptive recovery Time). The SNRs were calculated using the signal-to-noise ratio estimation algorithm proposed by Qi et al. (in press). In this algorithm, the signal was computed as the energy of correlated components of speech. Noise was computed as the energy of the remaining uncorrelated or random components of speech. The uncorrelated component was obtained by systematically removing any short- and long-term correlation exhibited by the original signal. The differences in SNR when the AudioZoom function was turned on or off was computed to evaluate the noise reduction function of the hearing aid. Results indicated that the SNR was about 1-2 dB higher when AudioZoom was turned on than when it was turned off at the high noise level. This magnitude of the SNR increase depended on the type of noise presented in the environment. The SNR increase was significantly higher for white noise than for cafeteria noise or multitalker noise. No increase of SNR was observed when the AudioZoom function was turned on at the low noise level. The effect of hearing aid program is also not significant. It appears that the AudioZoom function is most effective when high-level white noise is present. It is not as useful when low-level, speech-like noise is present, no matter what hearing aid program is used.

*This work was supported, in part, by a Graduate Research Grant from the University of Arizona. The authors thank Phonak for providing the necessary equipment for this work. They also thank James Dean for providing helpful assistance.
1. INTRODUCTION

Understanding speech in noise is one of the most difficult problems for people with sensorineural hearing loss (Bettie & Edgerton, 1976; Bronkhorst & Plomp, 1989; Sammeth & Ochs, 1991). All major hearing aid manufacturers have some type of hearing aid with a noise-reduction feature. Not all of them, however, have a true noise reduction function. Only two products, the PiCS AudioZoom system from Phonak and the Sharp Plus circuit from Miracle-Ear, have received clearance from the Food and Drug Administration (FDA) to make various hearing-in-background-noise performance claims (HJREPORT, 1997). Most other noise-reduction hearing aids are designed to enhance speech information so that it may be less susceptible to noise interference. For example, adaptive noise-reduction hearing aids mainly function by adjusting the amplification frequency response adaptively. One potential benefit of this adjustment is that background noise alone is unlikely to be amplified. By not amplifying background noise between speech signals, a perception of reduced noise interference would be created.

A true noise reduction scheme would require the hearing aid to reduce the level of noise specifically (Preves, 1990a, 1990b). Because it is extremely difficult to separate speech signals from noise, true noise reduction can be accomplished only under some special circumstances (Sammeth & Ochs, 1991). For example assuming the desired signal originates mostly in front of the listener whereas unwanted noise comes from behind, the use of a directional microphone array could minimize noise interference by picking up input from the front only.

A number of studies have been undertaken to document the benefit of using a directional microphone or microphone array (Bächler & Vonlanthen, 1995). It has been reported that there is about a 7 dB gain in SNR when a directional microphone is used under the most favorable conditions, that is, the reverberation time of the environment is relatively short and the noise originates from behind (Bächler & Vonlanthen, 1995). The gain in SNR for a listener is typically measure by determining the SNR needed to obtain a 50 % word recognition score (Valente, Fabry & Potts., 1995). Because of differences in procedures, human listeners and other experimental factors, it is often difficult to compare
the reported results. It would be helpful to have an objective, repeatable method of measuring the SNRs from the output of a hearing aid. In this work, a method for estimating the SNR of continuous speech is used to evaluate the noise reduction hearing aid from Phonak (Phonak Piconet-P2-AZ). Specifically, experiments were undertaken to determine:

1. Does the use of directional microphone improve the acoustic signal-to-noise ratio of speech?
2. If so, how is such improvement influenced by the noise level, the type of noise, and the type of hearing aid program?

II. METHOD

1. Equipment

A Knowles Electronics Manikin for Acoustic Research (KEMAR) with Zwislocki coupler was used for all acoustic recordings. KEMAR simulates the body, head, pinna, and ear canal of an average human adult approximating the diffraction effects and resonance properties of the pinna and ear canal to those of the average adult (Burkhard & Sachs, 1975). The Zwislocki coupler represents the acoustic immittance and standing wave soundfield of a median adult human ear (Skinner, 1988). KEMAR was situated in an audiometric sound booth (Industrial Acoustics Company, IAC) at the University of Arizona Hearing Clinic at 0 and 180 degree azimuth in reference to two loudspeakers, no less than one meter away from either speaker.

A behind-the-ear (BTE) hearing aid (Phonak Piconet-P2-AZ) was used. The hearing aid has three signal processing programs: Peak Clipping (PC), Wide Dynamic Range Compression (WDRC), and Super Compression with Adaptive Recovery Time (SC+aRT). Each hearing aid program can work with or without the AudioZoom. AudioZoom refers to the directional function of the microphone array in the hearing aid. The hearing aid programs can be controlled by a remote control.
The hearing aid was attached to 13 medium tubing and directly connected to the Zwislocki coupler to minimize interference due to earmold style and feedback. No vent option for fitting the hearing aid was selected. The hearing aid has three electroacoustic settings. In this project, the basic frequency response was used for the PC program to match the Revised National Acoustic Laboratories (NAL-R) prescriptive formula. The comfort frequency response was used for the WDRC and SC+aRT programs. The volume control of the hearing aid was adjusted to match the NAL-R prescribed target gain values.

2. Hearing Loss Configuration

One hearing loss configuration was used to evaluate each hearing aid program. The configuration represented sloping mild to severe high frequency hearing loss. The configuration is shown in Figure 1. A high frequency sensorineural hearing loss configuration was selected because it is the most common hearing loss configuration found in the adult population (Skinner, 1988). A software program (NOAH 2.0) was used to program the hearing loss configuration.

![Figure 1. The hearing loss configuration](image)

3. Signal Presentation and Recording

Speech signals, prerecorded on a digital audio tape (SONY, TCD-D3), were presented through the front loudspeaker at 0 degree azimuth. Noise was presented through the back
loudspeaker at 180 degree azimuth. Intensity levels were adjusted and monitored through a two-channel audiometer (Grason-Stadler, GSI-61) before both channels on the individual audio tracks were routed into the two GSI speakers.

Two microphones were used to record signals. An one-inch microphone (B&K, 4003), located above the pinna on a vertical line directly above the tagus, was used to record signals before passing through the amplification. This signal is referred to as unprocessed. A half-inch microphone (B&K, 4134), inserted into the Zwislocki coupler, was used to record signals passing through the amplification system. This signal is referred to as processed. Both the unprocessed and the processed signals were then routed through microphone amplifiers (Mackie, 1202 VLZ) and recorded on the two different channels of the digital audio tape (DAT) recorder. The setup is illustrated in Figure 2.

A sound level meter (B&K, 2204) along with its octave filter set (B&K, 1963), was used for calibrating the IAC booth, recording and presentation equipment. The IAC booth was calibrated with KEMAR in position. Correction factors were determined for GSI audiometer to correspond to sound pressure level in dB SPL. The noise floor of the IAC booth was also checked to ensure that ANSI S3.1-1991 standards were met.
One sentence, *Don’t ask me to carry an oily rag like that*, was used as the experimental speech signal. The sentence was chosen from TIMIT database. This sentence is phonemically balanced. In addition, all phonetic events have been pre-identified and labeled. These labels simplify the computation of SNR for different types of speech events, i.e., vowels and voiced consonants.

Recordings of one man (TIMIT, /test/ dr1 mdab0) and one woman (TIMIT, /test/ dr1 felc0) were used. The speech signals was presented at 65 dB SPL via a CD player. This intensity was chosen because it approximates the comfortable speech level for listeners with normal hearing. Three types of noise (cafeteria, multitalker, and white) were presented at two different levels (50 dB SPL and 65 dB SPL). Thus, the SNRs at which the speech and noise were presented were 0 dB and +15 dB, respectively.

A total of 36 runs were recorded. Each run was distinctive in either the noise level, the noise type, or the type of hearing aid program. The combination of experimental environments is shown in Figure 3.

![Figure 3. An illustration of recording conditions](image)

4. **Acoustic Analysis**

Samples recorded on the DAT were redigitized into a computer workstation (SUN, Sparc 10/30) at a sampling frequency of 16 kHz and a quantization level of 16 bits. Prior to digitizing, samples were low pass filtered with the cut off frequency set at 7.5 kHz. Approximate beginnings and endings of sentence were determined visually from
the acoustic waveform using a waveform editor on the computer.

The SNRs of the sentences were calculated using the signal-to-noise ratio estimation algorithm proposed by Qi, Hillman & Milstein (in press). In this algorithm, the signal was computed as the energy of correlated components of speech. Noise was computed as the energy of the remaining uncorrelated or random components of speech. The uncorrelated component was obtained by systematically removing any short- and long-term correlation exhibited by the original signal.

The decomposition of speech signals into short- and long-term correlations plus Gaussian noise has been successfully applied in telecommunication systems (Schroeder & Atal, 1985). In the CELP-based speech coders of some cellular phone system, for example, only parameters related to the short- and long-term correlations are transmitted. Speech signals are reconstructed at the receiver by adding (filtering) random Gaussian noise using the transmitted short- and long-term correlation coefficients. Although it is necessary to synthesize a random noise in the receiver that has similar variance and temporal distribution as that in the transmitter, the unpredictable, noise component of speech is not transmitted in the cellular system. The adequacy of decomposing speech signals into short- and long-term correlated components plus a noise component is demonstrated by the fact that cellular phones provide adequate speech quality for normal communication.

Linear prediction (LP) was used to determine both short- and long-term correlations (Markel & Gray, 1976; Ramachandran & Kabal, 1989). For short-term correlation, LP analysis was made on a window-by-window (no overlap) basis. The LP filter was obtained using a Hamming window with window length of 20 ms. The order of the LP filter was 14 (Markel & Gray, 1976). To remove short-term correlation, the original signal was inversely filtered by the LP filter using overlap save to ensure continuity during filter update. The residual signal of this LP inverse filtering was the short-term, decorrelated signal which was then further processed for long-term decorrelation.

During long-term decorrelation, linear prediction was made based on samples that were not immediately preceding the current sample of the short-term decorrelated, residue signal. The window length for minimizing prediction error was 2.5 ms, which is long enough to include the pulse-like peaks of short-term LP residual signals that often occur during voiced segments of speech. Because the exact location of the next residual
peak varies somewhat from cycle to cycle, the closet sample used for making prediction were slid between 1.25 ms to 17.5 ms prior to the first sample to be predicted. Thus, this included the fundamental frequency range from 60 to 800 Hz in the predictive analysis. The LP filter that produced the minimal prediction error over this sliding range was chosen as the final long-term LP filter. The order of the filter was 3 (Ramachandran & Kabal, 1989). To remove the long-term correlation, the short-term décorrelated residual signal was inversely filtered by the long-term LP filter. Overlap save was used again to ensure continuity during filter update. The output of this second stage of inverse filtering was considered to be the final (short- and long-term decorrelated) noise component of the speech signal. A flow chart of the short- and long-term decorrelation processes is shown in Figure 4. Example signals are shown in Figure 5.

The final SNR was computed as the ratio of average rms amplitude between the original signal and its corresponding short- and long-term decorrelated signal. Strictly speaking, this is a ratio of signal plus noise to noise, rather than signal to noise. The added noise term, however, should not significantly influence the use of the SNR to define the relative strength of the correlated component to the uncorrelative component of the signal because its contribution to the SNR should be the same no matter what the level of the true signal-to-noise ratio is.

![Flow chart of the short- and long-term decorrelation process](image-url)

**Fig. 4.** Flow chart of the short- and long-term decorrelation process.
Fig. 5. Example of original signal (top), short-term decorrelated signal (middle), and short- and long-term decorrelated signal (bottom). The original signal was recorded from a male talker saying the word "choice".

## III. RESULTS

The means and standard deviations of the measured SNR for each of the recording conditions are shown in Table 1. The means and standard deviations of the difference in the measured SNR (DSNR) when AudioZoom was turned on and off are shown in Table 2.

These results indicated that the mean SNR was greater when AudioZoom was turned on than that when AudioZoom was turned off at the recording environment of 0 dB SNR. The magnitude of the difference, however, depended on the type of noise used.
Table 1. Means and standard deviations of the measured SNR

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Noise Type</th>
<th>AudioZoom</th>
<th>Mean (dB)</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR 0 dB</td>
<td>White</td>
<td>off</td>
<td>-4.57909</td>
<td>0.1428376</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>-2.409239</td>
<td>0.655728</td>
</tr>
<tr>
<td></td>
<td>Cafeteria</td>
<td>off</td>
<td>-1.205905</td>
<td>0.1691547</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>0.7284309</td>
<td>0.4093448</td>
</tr>
<tr>
<td></td>
<td>Multitalker</td>
<td>off</td>
<td>1.4014411</td>
<td>0.2695117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>1.8110808</td>
<td>0.4368454</td>
</tr>
<tr>
<td>SNR 15 dB</td>
<td>White</td>
<td>off</td>
<td>10.08167</td>
<td>0.52654</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>9.706401</td>
<td>0.593045</td>
</tr>
<tr>
<td></td>
<td>Cafeteria</td>
<td>off</td>
<td>11.255421</td>
<td>0.512264</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>10.454224</td>
<td>0.522187</td>
</tr>
<tr>
<td></td>
<td>Multitalker</td>
<td>off</td>
<td>11.589355</td>
<td>0.478623</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>10.460239</td>
<td>0.409846</td>
</tr>
</tbody>
</table>

The mean SNR difference was the largest for white noise. It was larger for cafeteria, noise than that for multitalker noise. At the recording environment of +15 dB SNR, the mean SNR was smaller when AudioZoom was turned on than when it was turned off. The mean SNR difference was the largest for multitalker noise. It was larger for cafeteria noise than for white noise.

An analysis of variance (ANOVA) was undertaken to determine if the differences in SNR between signals recorded using omni-directional and directional microphones were significant. ANOVA was also used to determine if these SNR differences were significantly influenced as a function of noise level, noise type, and hearing aid program type.

Table 2. Means and standard deviations of DSNR

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Noise Type</th>
<th>Mean Diff. (dB)</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR 0 dB</td>
<td>White</td>
<td>2.1698561</td>
<td>0.6799521</td>
</tr>
<tr>
<td></td>
<td>Cafeteria</td>
<td>1.9343353</td>
<td>0.5163017</td>
</tr>
<tr>
<td></td>
<td>Multitalker</td>
<td>0.4096397</td>
<td>0.381723</td>
</tr>
<tr>
<td>SNR 15 dB</td>
<td>White</td>
<td>-0.3752687</td>
<td>0.5124821</td>
</tr>
<tr>
<td></td>
<td>Cafeteria</td>
<td>-0.8011972</td>
<td>0.5934346</td>
</tr>
<tr>
<td></td>
<td>Multitalker</td>
<td>-1.129116</td>
<td>0.6790365</td>
</tr>
</tbody>
</table>
Results of statistical analysis indicated that the measured SNR difference was significantly influenced by the use of AudioZoom \((F = 85.42, p < .001)\). Results of statistical analysis also indicated that the measured SNR difference was significantly influenced by noise level \((F = 365.33, p < .001)\) and noise type \((F = 40.03, p < .001)\). However, the measured SNR difference was not significantly influenced by hearing aid program type \((F = 1.38, p > .05)\). Post-hoc test (Turkeys Studentized Range Test) revealed that the SNR increase when using AudioZoom was significantly greater \((\alpha = .01)\) for white noise than for either cafeteria noise or multitalker noise at the recording environment of 0 dB SNR.

### IV. DISCUSSION AND CONCLUSION

Results of the word seem to indicate that the SNR was about 1-2 dB higher when AudioZoom was turned on than when it was turned off at the high noise level. This magnitude of the SNR increase depended on the type of noise presented in the environment. The SNR increase was significantly higher for white noise than for cafeteria noise or multitalker noise. No increase of SNR was observed when the AudioZoom function was turned on at the low noise level. The effect of hearing aid program is also not significant. It appears that the AudioZoom function is most effective when high-level, white noise is present. It is not as useful when low-level, speech-like noise is presented, no matter what hearing aid program is used.

The measured SNR increase when using the AudioZoom function (about 1-2 dB) is smaller than those obtained from human listener based experiment (about 7 dB). This difference could be attributed, in part, to the difference between acoustic and perceptual methods of SNR measurement. In listener based experiment, the SNR was obtained by estimating the SNR of the input speech signal to a hearing aid when a listener can recognize the speech stimuli 50.

Acoustic measurement is fundamentally different form the perceptual-based measurement. For example, a hearing aid typically amplifies more on the high-
frequency components of a speech signal. This amplification makes it easier for people with sensorineural hearing impairment to differentiate between consonants and vowels and, thus, improve their speech intelligibility score. High-frequency emphasis, however, will reduce the overall SNR of the signal acoustically.

Obviously, the listener based measurement appears to be more relevant to listeners hearing than the proposed acoustic measurement. However, it is also subject to larger intra- and inter-listener variability due to listeners’ biases. As stated earlier, it is often difficult to compare the results of human-listener based experiments on the noise reduction function of hearing aid due to procedural and human-listener differences. It would be more informative to first determine how a hearing aid has modified the stimuli before determining how listeners will respond to those stimuli. In traditional psychoacoustical experiment, for example, the physical properties of the acoustic stimuli often are well defined. This work represents an initial attempt to obtain priori knowledge of the characteristics of the hearing aid processed signals reaching the listener’s auditory system before soliciting their responses. Comprehensive investigations on the relationship between the measured acoustic properties and perception are needed to ultimately validate the use of such measurements for hearing aid evaluations.

REFERENCES


Acoustic Evaluation of Noise Reduction in a Hearing Aid

Seminars in Hearing, 11, 39-65.


Yingyoung Qi* 1

(Phonak, Phonet2 AudioZoom)

(1) 0°–15 dB SNR
(2) white, cafeteria and multiterk
(3) Peak Clipping, Wide Dynamic Range Compression, and Super Compression with Adaptive Recovery Time)

* E-mail: chulee@u.arizona.ac
Acoustic Evaluation of Noise Reduction in a Hearing Aid