Sound Quality Measures for Speech in Noise through a Commercial Hearing Aid Implementing “Digital Noise Reduction”

Todd A. Ricketts*
Benjamin W.Y. Hornsby*

Abstract
This brief report discusses the affect of digital noise reduction (DNR) processing on aided speech recognition and sound quality measures in 14 adults fitted with a commercial hearing aid. Measures of speech recognition and sound quality were obtained in two different speech-in-noise conditions (71 dBA speech, +6 dB SNR and 75 dBA speech, +1 dB SNR). The results revealed that the presence or absence of DNR processing did not impact speech recognition in noise (either positively or negatively). Paired comparisons of sound quality for the same speech in noise signals, however, revealed a strong preference for DNR processing. These data suggest that at least one implementation of DNR processing is capable of providing improved sound quality, for speech in noise, in the absence of improved speech recognition.

Key Words: Digital noise reduction, paired comparison, sound quality

Abbreviations: BTE = behind the ear; CST = Connected Speech Test; DNR = digital noise reduction; MDNR = modulation-based digital noise reduction; MSSR = modulated-to-steady-state ratio; NAL-NL1 = National Acoustics Laboratory—Nonlinear 1; SNR = signal-to-noise ratio

Sumario
Este breve reporte discute el efecto del procesamiento por medio de la reducción digital del ruido (DNR) sobre el reconocimiento del lenguaje con amplificación y sobre las medidas de calidad de sonido, en 14 adultos usando auxiliares auditivos comerciales. Las medidas de reconocimiento del lenguaje y de la calidad de sonido se obtuvieron en dos diferentes condiciones de lenguaje en ruido (lenguaje a 71 dBA + 6 dB de SNR, y lenguaje a 75 dBA + 1 dB de SNR). Los resultados revelaron que la presencia o ausencia del procesamiento con DNR no influyó en el reconocimiento del lenguaje en ruido (ni positiva ni negativamente). Las comparaciones en pares sobre la calidad de sonido para las mismas señales de lenguaje en ruido, sin embargo, revelaron una fuerte preferencia para el procesamiento con DNR. Estos datos sugieren que al menos una implementación del procesamiento con DNR es capaz de aportar una calidad de sonido mejorada, para el lenguaje en ruido, en ausencia de un reconocimiento mejorado del lenguaje.

*Dan Maddox Hearing Aid Research Laboratory, Department of Hearing and Speech Sciences, Vanderbilt Bill Wilkerson Center

This study was made possible by a grant from Siemens Hearing Instruments and through the support of the Dan Maddox Foundation.

Todd A. Ricketts, Ph.D., CCC-A, Director, Dan Maddox Hearing Aid Research Laboratory, Vanderbilt Bill Wilkerson Center, 1114 19th Avenue South, Nashville, TN 37212; Phone: 615-936-5258; Fax: 615-936-5013; E-mail: todd.a.ricketts@vanderbilt.edu
The term “digital noise reduction” (DNR) can be used to describe hearing aid processing that has the general goal of providing less amplification, over a specified frequency range, for noise than for speech. One of the most popular forms of DNR implemented in current hearing aids is based on identification of modulation in multiple channels allowing for an estimation of the modulated-to-steady-state ratio (MSSR). These systems assume that signals that are primarily steady-state are “noise,” while signals with greater modulation are more “speechlike.” Gain is then reduced in channels for which the MSSR indicates the incoming signal is steady-state (van Dijkhuizen et al, 1991). In most of these systems, increasingly more gain (up to some predetermined target) is associated with greater modulation depth and higher modulation frequency (e.g., Powers et al, 1999). The label “modulation-based digital noise reduction” (MDNR) will be used herein to describe this processing. While MDNR is implemented by several manufacturers, the specific characteristics including time constants/analysis time, magnitude of gain reduction, and rules for estimating MSSR and implementing gain reduction vary significantly (Bentler, 2004).

The results of studies investigating changes in speech understanding due to implementation of DNR processing in modern hearing aids have been mixed (e.g., Boymans et al, 1999; Boymans and Dreschler, 2000; Walden et al, 2000; Alcantara et al, 2003). Limited data suggest that one specific implementation of MDNR processing may slightly improve speech recognition performance in the presence of steady-state noise (e.g., Bray and Nilsson, 2001; Bray et al, 2002). More commonly, however, one investigation has concluded that implementation of MDNR processing may lead to improved general sound quality (Boymans et al, 1999) in the absence of significantly improved speech recognition. In this study, however, speech recognition and sound quality judgments were not based on the same materials. More commonly, no degradation in speech recognition or sound quality have been reported for MDNR processing implemented in commercial hearing aids (Ricketts and Dahr, 1999; Boymans and Dreschler, 2000; Walden et al, 2000; Alcantara et al, 2003). Although the actual reasons for the discrepant findings across studies are unclear, it is assumed that differences in the speed and magnitude of gain reduction for steady-state signals as well as differences in experimental methodology (e.g., type of competing signal) play a role. The details concerning MDNR processing parameters are important since different goals for this processing may be in conflict. For example, it might be speculated that MDNR processing with aggressive gain reduction (large reductions based on overly broad definitions of “noise”) may optimize sound quality but could lead to reduced speech recognition if the audibility of speech is negatively affected. It is therefore of special interest to evaluate MDNR systems that are designed to greatly reduce (>10 dB) gain for “noise.”

A second type of DNR referred to as “adaptive Wiener filtering” has more recently been introduced in hearing aids. The Wiener filter was first described by Norbert Wiener in...
the 1940s (Benvenuto and Cherubini, 2002). Simply stated, a Wiener filter is a theoretically derived filter that has the goal of estimating the original signal from a degraded version of the signal. In hearing aid applications, this translates into a goal of providing the greatest attenuation for frequencies not containing speech. That is, the goals of MDNR and adaptive Weiner filtering are somewhat similar in that both are intended to provide more gain, at any point in time, for frequency ranges containing speech information than those containing “noise.” The Wiener filter is optimally derived when the speech and noise spectra are known; the success of adaptive Wiener filters in hearing aid applications, however, is highly dependent on accurate estimation of the speech and noise power in degraded samples with no a priori knowledge. Consequently, it is not surprising that multiple iterations of the adaptive Wiener filter have been developed with varying degrees of success.

It was the purpose of this study to evaluate the sound quality and speech recognition provided by DNR processing based on a combination of MDNR and adaptive Weiner filtering implemented in a commercial hearing aid. The impact of this combined DNR processing was evaluated separately in omnidirectional and directional microphone modes and in two different noise level/signal-to-noise ratio (SNR) conditions. This methodology allowed the stability of any DNR effects to be examined across a variety of hearing aid and environmental conditions. With two notable exceptions (Boymans and Dreschler, 2000; Alcantara et al, 2003), this study differs from previous investigations in that both speech recognition and sound quality measures were obtained using the same speech-in-noise materials. Directly comparing speech recognition and sound quality using the same signal was of interest because the test hearing aid model was designed to provide as much as 18 dB of gain reduction for “noise” signals. This magnitude of gain reduction is among the largest across commercial hearing aids (Bentler, 2004) and, consequently, has the greatest potential for reducing speech audibility along with reductions in noise. The paired comparisons of sound quality reported herein were collected as part of a larger experiment. Further details of the experiment related to speech recognition evaluation are described by Ricketts and colleagues (2005).

**METHODS**

Speech recognition and sound quality were evaluated across two competing noise conditions. The conditions were designed to simulate multisource noisy environments where the overall level of the noise was dominated by a single higher level noise source. In each condition, the competing noises consisted of three “lower level” noises and one “higher level” noise that were played simultaneously. In one condition (low noise/good SNR), the speech and competing noise were presented at 71 and 65 dBA, respectively (+6 dB SNR). In the second condition (high noise/poor SNR), the speech and competing noise were presented at 75 and 74 dBA, respectively (+1 dB SNR). These conditions represent a range of difficult listening situations and were chosen in part because data reveal that SNR generally decreases in real environments as noise level increases (e.g., Pearsons et al, 1977). Active and inactive DNR modes were compared in both noise conditions and for both adaptive directional and omnidirectional microphone modes using speech recognition and paired comparisons of sound quality. Specifically, there were a total of eight conditions (two competing noise conditions x two microphone modes x two DNR modes).

Fourteen (eight male, six female) bilaterally fit, adult listeners with symmetrical sloping, moderate, sensorineural hearing loss were evaluated. Participants were between 42 and 83 years old (mean age 62) and had between 1 and 35 years of hearing aid experience. Symmetry between ears was defined as exhibiting no more than a 15 dB difference in pure-tone thresholds at any octave frequency from 250 Hz through 4000 Hz. Pure-tone thresholds and acoustic immittance measures were obtained prior to the test session. All subjects exhibited no significant air-bone gap at any frequency (≤10 dB).

The Siemens Triano S behind-the-ear (BTE) hearing aid was selected as the test instrument. The Triano S can operate as either a two microphone directional or an omnidirectional mode. All listeners were fit using a modified National Acoustics Laboratory—Nonlinear 1 (NAL-NL1)
prescriptive method (Dillon, 1999), and fittings were verified in the ear using a Fonix 6500 real ear analyzer and interrupted composite noise signals. The modified NAL-NLI prescriptive gain targets were generated for the Siemens Triano by the Siemens Connex fitting Software (Version 4.1a). Gain for the omnidirectional and directional modes was individually matched as closely as possible to the same targets. The hearing aids were coupled to the ear using a custom full shell earmold with standard #13 tubing and an appropriate select-a-vent (SAV) size. The SAV sizes used ranged from one to two millimeters. Individual participant’s thresholds of discomfort (TD) were measured at 500 and 3000 Hz using a narrow band noise. These data were then used to set nominal, frequency-specific, compression limiting thresholds that were fixed across hearing aid conditions for each individual participant. The tested digital hearing aids also incorporated 16 overlapping channels of low-threshold compression. Average nominal compression ratios across all patients ranged from a low of 1:1 to a high of 3:2:1. Average, frequency-specific compression thresholds (ANSI 3.22-1996) ranged from 34 dB SPL to 57 dB SPL.

The test instruments’ digital noise reduction scheme, referred to as the “Speech Comfort System™” (SCS), had four available settings (off, minimum, medium, and maximum). Settings of “off” and “maximum” were selected for evaluation. The test hearing aid used both a 16-channel MDNR system and an adaptive Wiener filter in its DNR processing scheme. In a standard fitting, the adaptive Wiener filter is unaffected by the speech comfort system setting (that is, it is always active). For the purposes of this experiment, Version 4.1a of the commercial CONNEX software was modified by the manufacturer so that a setting of “off” disabled both types of DNR processing (this is equivalent to turning “adaptive parameters” and the speech comfort system “off” in CONNEX 4.2 and newer versions of the software).

The test MDNR processing reached maximum gain reduction approximately five to seven seconds after the introduction of a steady-state signal (combined analysis and gain reduction time). Gain returned to target levels after the introduction of speech within approximately 0.5 seconds. The magnitude of gain reduction of the MDNR processing implemented can be as great as 14–18 dB (depending on signal intensity) when using the maximum setting. In practice, however, gain reduction for a 70 dB SPL input using the maximum setting varies from 0 to 14 dB depending specifically on the modulation rate and modulation depth of the input (see Powers et al, 1999, for further details).

Speech Recognition Testing

Speech recognition was evaluated using the Connected Speech Test (CST; Cox et al, 1987, 1988). This test consists of 24 pairs of speech passages produced conversationally by a female speaker. For each subject, all test conditions were evaluated using one pair of CST passages that were administered in accordance with the authors’ recommendations. All passages were randomly selected without replacement. For all testing, the single competing noise source that accompanied the CST materials was replaced with four uncorrelated noise sources. These four competing sources were developed by randomly varying the start time of the original competing source material. The speech and competing noise signals were streamed off the hard drive of a Pentium 4 class computer using commercial, multitrack sound editing and presentation package (Syntrillium Cool Edit Pro™) and delivered to the presentation loudspeakers through the use of a multichannel digital signal processing (DSP) sound card (Echo Digital Audio Darla24™). The noise signal was amplified using a 12-channel distribution amplifier (Russound DPA-6.12), and the level of the speech signal was controlled by an audiometer (Grason Stadler Model 61). The speech stimuli were presented from a Tannoy loudspeaker (Model 600), and competing noises were presented from four Definitive Technologies loudspeakers (Model BP-2X). All testing was completed in a room with low reverberation (average reverberation time [RT60] across octave frequencies from 250 Hz through 8000 Hz was approximately 400 msec). The single speech loudspeaker and the four uncorrelated noise loudspeakers were placed at approximately ear level using 32” speaker stands at a distance of 1.25 meters from the listener. The speech source loudspeaker was always placed at 0˚ azimuth.
Three “lower-level” competing loudspeakers were equally spaced around the listener at 300˚, 180˚, and 60˚, and a single “higher-level” loudspeaker was placed at 160˚. The levels of each of the three “lower-level” loudspeakers were always equal to each other. In addition, the level of the “higher-level” loudspeaker was always 10 dB greater than the level of each of the individual “lower-level” loudspeakers. These competing noise conditions were chosen specifically for evaluation of adaptive directional processing (see Ricketts et al, 2005, for further details).

**Sound Quality**

Sound quality was assessed using paired comparisons in both test environments using both omnidirectional and adaptive directional microphone modes. Paired comparisons of sound quality were made in the same listening conditions used for assessment of speech recognition; however, sentences from practice lists of the CST were used as test stimuli during the sound quality assessment. Paired comparisons of sound quality for the two MDNR conditions were repeated three times in both test conditions (low noise/good SNR or high noise/poor SNR) in random order. The DNR conditions (“maximum” or “off”) were randomly assigned to the numbers one and two for each trial. The listener was told to tell the investigator to switch to “one” or “two” as often as they liked and then report which setting they preferred. A “strength of preference” rating was also recorded. Paired comparisons across noise and DNR conditions were evaluated separately for the two microphone settings (adaptive directional or omnidirectional) in a counterbalanced design. A second investigator (the “test investigator”) provided all instructions and was responsible for all interaction with the patients. The test investigator and the participants were blinded to the DNR condition. The following subject instructions were provided.

In this test session you will be listening to a continuous talker in the presence of background noise. We will be changing the settings on the hearing aids, and you will be asked to report which setting you prefer. This will require you to let us know when you would like to switch between the settings in order for you to make your decision. You may ask us to switch between settings as many times as you like by telling us “setting 1” or “setting 2.” The talker will be repeating the same three sentences. You must choose one setting or the other. Once a setting is preferred, let us know which setting it is, and then tell us, on a scale of 1 to 10, how strongly you preferred that particular setting over the other. This process will be repeated over multiple sessions. Any questions? The subject was provided a preference scale as a visual marker to make their decision on how strongly they preferred one setting over another. A preference of 1 was indicated as “no or very little preference,” and 10 indicated a “very strong preference.”

**RESULTS AND DISCUSSION**

The average speech recognition performance in omnidirectional and adaptive directional modes measured across the two competing noise and two DNR conditions is shown in Figure 1. These speech recognition data represent a subset of those previously described by Ricketts and colleagues (2005). Speech recognition performance across these four conditions was reanalyzed using a three-factor repeated-measures analysis of variance (ANOVA). The within-subjects factors (independent variables) were noise condition (65 dBA and 74 dBA), DNR condition (off and maximum), and microphone mode (omnidirectional and directional). Analysis revealed a significant main effect of microphone (F<sub>1, 13</sub> = 103.8, P < 0.0001) and noise condition (F<sub>1, 13</sub> = 46.6, P < 0.0001). No other significant main effects or interactions were present. This statistical treatment confirmed the conclusions previously described by Ricketts and colleagues (2005). Specifically, significant directional benefit and better performance at the more positive SNR were found; however, DNR setting did not significantly affect speech recognition performance.

The average number of times the two DNR conditions were selected during paired comparison testing (without considering “strength of preference”) across the omnidirectional and adaptive directional microphone modes and two competing noise conditions is shown in Figure 2. The maximum number of times a DNR condition could be selected was three. These data were
analyzed using four separate Wilcoxon Matched Pairs Tests (one for each of the four microphone and competing noise level combinations). These analyses indicated that the DNR “on” condition (SCS set to maximum) was selected significantly more often than the DNR “off” condition (SCS set to off) regardless of noise level or microphone mode. Specifically, this preference was present when listening in omnidirectional mode in low noise levels ($T = 18$, $Z = 2.4$, $P < 0.017$), in omnidirectional mode in high noise levels ($T = 4$, $Z = 2.6$, $P < 0.001$), in directional mode in low noise levels ($T = 19.5$, $Z = 2.3$,

![Graph 1](image1.png)

**Figure 1.** The average speech recognition performance measured across the four listening conditions (omnidirectional and adaptive directional modes; low [71 dBA speech, +6 dB SNR] and high [75 dBA speech, +1 dB SNR] competing noise levels) as a function of DNR processing (SCS Off versus SCS Maximum).

![Graph 2](image2.png)

**Figure 2.** The average number of times the DNR active (SCS off) and DNR inactive (SCS Max) conditions were selected across the four listening conditions (omnidirectional and adaptive directional modes; low [71 dBA speech, +6 dB SNR] and high [75 dBA speech, +1 dB SNR] competing noise levels).
Audiology of the hearing aid processed signals is currently being completed in an attempt to identify potential causes.

The preference for DNR processing was highlighted by at least two factors. First, it is clear that on the average a similar strong preference was measured across all four microphone/noise conditions. More importantly, the vast majority of listeners, 10 of the 14, appeared to prefer the DNR on condition regardless of the microphone/noise condition. Persons with hearing loss, however, are not a homogenous group. Specifically, one listener preferred the “DNR off” condition across all four listening conditions, another listener preferred the “DNR off” condition in two of the four listening conditions, and finally, two listeners preferred the “DNR off” condition in only one of the four listening conditions.

Acknowledgments. This study was made possible by a grant from Siemens Hearing Instruments and through the support of the Dan Maddox Foundation. The authors would also like to thank Earl Johnson for his assistance in data collection.

REFERENCES


