Published in final edited form as:

Int J Audiol. 2017 August ; 56(8): 568–579. doi:10.1080/14992027.2017.1305128.

Output signal-to-noise ratio and speech perception in noise: Effects of algorithm

Christi W. Miller¹, Ruth A. Bentler², Wu Yu-Hsiang², James Lewis³, and Kelly Tremblay¹

¹Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA

²Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, USA

³Department of Audiology and Speech Pathology, University of Tennessee Health Science Center, Knoxville, TN, USA

Abstract

Objective—The aim of this study was to: 1) quantify the amount of change in signal-to-noise ratio (SNR) as a result of compression and noise reduction processing in devices from three hearing aid (HA) manufacturers, and 2) use the SNR changes to predict changes in speech perception. We hypothesized that the SNR change would differ across processing type and manufacturer, and that improvements in SNR would relate to improvements in performance.

Design—SNR at the output of the HAs was quantified using a phase-inversion technique. A linear mixed model was used to determine whether changes in SNR across HA conditions were predictive of changes in aided speech perception in noise.

Study Sample—Two groups participated: 25 participants had normal hearing and 25 participants had mild to moderately-severe sensorineural hearing loss.

Results—The HAs programmed for both groups changed the SNR by a small, but statistically significant amount. Significant interactions in SNR changes were observed between HA devices and processing types. However, the change in SNR was not predictive of changes in speech perception.

Conclusion—Although observed significant changes in SNR resulting from compression and noise reduction did not convert to changes in speech perception, these algorithms may serve other purposes.

Keywords

hearing aids; speech perception; behavioral measures; noise; assistive technology

Address correspondence to: Christi W. Miller, Department of Speech and Hearing Sciences, University of Washington, 1417 NE 42nd Street, Seattle, WA 98105-6246, tel 206-685-7456, fax 206-543-1093, christim@u.washington.edu.

Portions of this article were presented at the American Auditory Society meeting, Scottsdale, Arizona, March 2012 and at the International Hearing Aid Research Conference, Lake Tahoe, California, August 2012.

INTRODUCTION

Many factors may contribute to aided speech understanding in noisy environments, including device-related variables (e.g., directional microphones) and patient-related variables (e.g., age, attention). Although some predictors of aided speech understanding are well documented (e.g., audibility), a large portion of the variance in outcomes remains unknown. As suggested in the literature, the signal-to-noise ratio (SNR) being delivered to the HA user (i.e., the SNR post HA-processing) may be an important variable in aided speech-in-noise perception (e.g., Naylor & Johannesson, 2009; Tremblay & Miller, 2014). Metrics of SNR, measured at the output of the HA, have been shown to negatively interfere with the neural representation of sound (Billings, et al., 2007). Hearing aid algorithms and their interactions may yield different output SNRs depending on their implementation in the design of the HA. For example, wide dynamic range compression (WDRC) tends to decrease the SNR at positive input SNRs due to the speech signal dominating the compressor behavior; whereas at negative input SNRs, the noise is dominating compressor behavior and WDRC tends to increase output SNR. Noise reduction (NR) algorithms tend to decrease the gain under certain conditions; however, if WDRC is applied after NR the WDRC algorithm may reverse the gain decisions made by the NR algorithm, turning gain back up in noisy channels. Therefore, manufacturer-dependent decisions, such as the NR time constants associated with the estimation of speech and noise, and whether the signals are sequentially or simultaneously processed through the compression and noise reduction algorithms, can lead to variability in SNR changes between manufacturers. Therefore, we examined how output SNR changed across algorithms from three manufacturers, and whether improvements in SNR were associated with improvements in aided speech perception.

The positive correlation between speech perception and long-term SNR (i.e., levels averaged over 30 seconds) has been well documented when the SNR is modified at the input (e.g., Cox et al., 1987, 1988; for a review, see MacPherson and Akeroyd, 2014). However, measuring the SNR at the output of a HA is difficult to achieve due to mixing of the speech and noise signals. Historically, calculation of the output SNR has been achieved by comparing the speech and noise signals after individually presenting each signal to the HA (e.g., Henning and Bentler, 2005). But a limitation of this technique is that advanced HA processing algorithms, such as those utilized for compression and NR, react to the combined speech and noise characteristics in a way that is not predicted by the response to each signal individually. Therefore, a technique to analyze the individual contributions of the speech and noise signals at the summed output of the HA would more accurately quantify output SNR.

One technique described by Hagerman and Olofsson (2004) to separate the mixed speech and noise signal at the output of the HA uses phase-inversion. In this approach, two recordings are made, both of which involve simultaneously presenting speech and noise signals to the HA. For the first measurement, the original-phase speech and noise signals are presented; for the second measurement, either the noise or speech signal is phase-inverted. The two measures have identical modulation and magnitude spectra yet differ in phase. When the noise is phase-inverted, an estimation of the processed speech signal is obtained by adding the HA outputs across the two measurements. An estimate of the noise is achieved

by subtracting the two HA outputs. The root mean square levels of the extracted speech and noise signals can then be calculated and the output SNR derived.

Several studies have used the phase-inversion approach proposed by Hagerman and Olofsson to quantify acoustic changes to speech and noise post-HA processing, including estimation of the SNR of the HA's output (Hagerman & Olofsson, 2004; Olsen et al., 2005; Souza et al., 2006; Naylor & Johannesson, 2009; Chong & Jenstad, 2010; Gustafson et al., 2014; Alexander & Masterson, 2015). Using the phase-inversion technique, researchers have demonstrated that WDRC can result in either increases or decreases of the long-term output SNR, depending on both the input SNR and the HA compression settings (Olsen et al., 2005; Souza et al., 2006; Naylor & Johannesson, 2009). Naylor & Johannesson, (2009) used the phase-inversion approach to measure the output SNR of a single-channel HA while systematically varying the compression parameters and input SNR. In general, for positive SNR inputs, compression reduced (worsened) the output SNR. However, for negative SNR inputs, compression increased (improved) the output SNR. The magnitude of improvement or decrement in the output SNR depended on the compression ratio and time constants, such that higher compression ratios and faster time constants resulted in larger differences between the output and input SNRs. Further, the modulation properties of the noise also interacted with SNR changes in that speech interference had the smallest effect on SNR compared to other noise interference at poorer input SNRs. These results are consistent with those found in other studies, with the change in SNR (measured as the difference between the output and input SNR) dependent upon the compression settings used in the HAs, as well as the input SNR (Olsen et al., 2005; Souza et al., 2006; Alexander & Masterson, 2015). Across these studies, the change in SNR caused by WDRC ranged from -4 dB to 3 dB. Furthermore, Alexander and Masterson (2015) demonstrated that output SNR decreased as number of compression channels increased.

Hagerman's phase-inversion technique has also been used to measure long-term SNR changes made by NR processing (Hagerman & Olofsson, 2004; Wu & Stangl, 2013; Smeds et al., 2015; Brons et al., 2015). Hagerman and Olofsson (2004) tested the NR algorithms of five HAs using several types of speech and noise signals presented at various input SNRs. The NR algorithms showed a range of increases in output SNR relative to input SNR when activated, up to ~5 dB in one case, depending on the input SNR and type of noise signal. More current NR algorithms also have shown small, but positive changes in SNR (Smeds et al., 2015, Brons et al., 2015), ranging from 0 to 4 dB. The amount of improvement seemed to depend on whether NR was combined with linear or WDRC processing. For instance, Brons et al. (2015) demonstrated that the improvement in SNR with NR processing was partially negated by WDRC processing, leading to a smaller benefit of NR when combined with WDRC. Differences in the number of independent HA channels may also account for some of the variation in reported SNR changes.

Given that NR algorithms may differ among manufacturers, it's possible that the particular NR method used also contributed to SNR changes. Historically, NR algorithms have fallen into one of two categories: modulation based or spectral subtraction. Modulated based algorithms use information about the modulation rate and depth of the incoming signal to estimate whether speech is present and at what SNR (Chung, 2004; Bentler & Chiou, 2006;

Kates, 2008; Chung, 2010). Once these decisions are made, the algorithm determines how much gain reduction to implement. If the modulation rate is low (e.g., speech dominated) and the modulation depth is high (likely at favorable SNRs), then the gain is minimally reduced (Chung, 2010). If the modulation rate is high and the modulation depth is low, then gain reduction is greater. These gain changes generally occur slowly within each channel designated by the NR algorithm. Spectral subtraction methods (e.g., Wiener filtering) use the estimated spectra of the noise alone and the mixed speech-and-noise to implement decisions about gain reduction. The estimated spectrum of the noise is made during portions of the signal without speech present. The gain of the mixed speech-and-noise is then adjusted to closely match that of more “clean” speech (Chung, 2004; Bentler & Chiou, 2006; Chung, 2010; Dillon, 2012). The amount of gain reduction depends on the estimated SNR and noise spectrum, and the gain reduction is typically faster than that for modulation-based methods (Kates, 2008). One manufacturer (Oticon) also uses synchrony detection. For this method, speech is assumed to be present when energy across frequencies is synchronous and periodic. When speech is present, gain reduction does not occur (Chung, 2004). Across methods and manufacturers the amount of gain reduction and speed at which changes are applied is not straightforward, but depends on a number of factors including the modulation depth, rate, overall level, frequency weighting, number of channels, and type of noise (Chung, 2004; Bentler & Chiou, 2006; Chung, 2010; Dillon, 2012).

An unanswered question is whether individual changes in SNR caused by WDRC and NR processing affect speech perception. To understand how speech perception is related to output SNR, Alexander and Masterson (2015) compared speech perception scores at equal output SNRs, post-phase-inversion processing, across multiple compression conditions. If speech perception was dictated by output SNR, then within-participant speech perception scores should have been the same across compression conditions given that output SNR was equated across these conditions; however, this was not the case. Statistical analysis suggested that release time, number of channels, and type of masker also influenced speech perception scores. Therefore, acoustic variables other than SNR (e.g., temporal envelope) may have contributed to the variability in speech perception scores. Although Brons and colleagues (2015) found that WDRC and NR algorithms significantly modified the output SNR, in some cases up to ~2 dB, they did not find any significant within-participant differences in speech intelligibility between HA processing conditions when the input SNR was fixed at +4 dB. Across the different conditions, intelligibility ranged from 89–96% and it’s possible that many participants were performing at ceiling level such that relatively small changes in the output SNR had minimal effect on performance.

The first goal of this study was to determine the extent that WDRC and NR algorithms modify the SNR in individually programmed HAs across devices from three manufacturers (Oticon, Phonak, and Widex). The second goal of this study was to evaluate whether changes in output SNR caused by WDRC and NR were associated with changes in speech-in-noise perception at input SNRs that avoid both floor and ceiling effects. It was hypothesized that the speech perception change experienced with a change in SNR would be mediated by the participant’s underlying psychometric function, and that a stronger relationship between changes in SNR and changes in speech perception would be found if the input SNR occurs along a sensitive portion of the psychometric function. Therefore,

participants were tested at SNRs that corresponded to approximately 50% correct because this is where psychometric functions exhibit the steepest slope. Cox and colleagues (1988) indicated that for individuals with hearing loss, the SNR corresponding to 50% correct using the Connected Speech Test materials would occur at positive input SNRs, which would fall within the range where HA processing has been shown to change output SNR. Speech perception changes may differ between NH and HI participants with an equal change in SNR due to differences in the slope of their underlying psychometric function, with steeper slopes for NH participants (e.g., Wilson & McArdle, 2012, MacPherson & Akeroyd, 2014). Consequently, our goals were addressed in two participant groups: hearing-impaired (HI) and normal-hearing (NH).

MATERIALS AND METHODS

A cross-sectional, repeated-measures design was used. All participants completed the same protocol, regardless of group. The study was approved by the University of Washington Institutional Review Board.

PARTICIPANTS

Fifty participants (25 NH, 25 HI) took part in this study. Participants were recruited from an existing Communication Participant Pool (P30 DC004661). Informed consent was obtained from all participants. The median ages of the NH and HI groups were 65 years (range: 23–78 years) and 74 years (range: 29–87), respectively. The NH group had air- and bone-conduction thresholds at or below 25 dB HL from 250–4000 Hz. One participant had a threshold of 30 dB HL at 3000 Hz. The HI group consisted of individuals with mild to moderately-severe sensorineural hearing loss. A conductive component was ruled out by observation of less than a 10-dB air-bone gap and a detailed case history asking about prior ear infections or surgeries. In both groups, the better-hearing ear (based on the pure-tone average at 500, 1000, and 2000 Hz) was selected for testing. Figure 1 shows the mean and standard deviation of the thresholds for the better-hearing ear of both groups. Pure tone averages (500, 1000, and 2000 Hz) for the NH and HI groups were 12.6 (SD=5.9) and 40.7 (SD=12.5) dB HL, respectively. Of the 25 HI participants, 21 (84%) were current HA users. The average years of HA use based on the case history was 11.3 years (SD=9.8).

STIMULI

The audio-only version of the Connected Speech Test (CST; Cox et al., 1987, 1988) was used for both acoustic (i.e., HA measurements made on a KEMAR) and behavioral testing. The CST was developed to simulate real world listening situations more closely than other speech tests. The test material is composed of passages of speech about a familiar topic. Each passage includes either 9 or 10 sentences and the participant is informed of the topic prior to sentence presentation. The stimuli were transferred from the compact disc to the computer hard drive at a 44.1 kHz sampling rate using Adobe Audition. The CST consists of a female speaker on one channel and 6-talker babble (3 male talkers, 3 female talkers) on a second channel. Six-talker babble was chosen as a compromise between reflecting some properties of realistic listening conditions (i.e., multi-talker babble) and the desire for the NR algorithms to detect the stimuli as noise (i.e., steady-state noise). For behavioral testing,

which used all practice and test passages included in the CST, the long-term average spectra of the speech and babble were measured in $1/3^{\text{rd}}$ octave bands and found to be within 2.5 dB of each other from 160 to 8000 Hz, with one frequency exception (7 dB difference at 315 Hz). We did not want to violate the parameters used for validating the test (Cox et al., 1987; 1988), therefore adjustments to better equate the speech and babble levels were not made for behavioral testing. For acoustic recordings of HA processing, 1-minute samples of both stimuli were extracted. The silent periods between sentences (when participants would usually repeat the sentence they heard) were removed since including these sections would lead to underestimating the long-term SNR. Since a shorter sample of the stimuli were extracted for acoustical testing than for behavioral testing, long-term average levels were re-equated. The long-term average spectra of the babble was filtered to match the speech sample to within 1 dB from 160 to 8000 Hz, in $1/3^{\text{rd}}$ octave bands.

Speech and babble signals were routed from the computer to a Crown D-45 amplifier and presented from a single loudspeaker at 0 degrees azimuth in a double-walled audiometric test booth, located 1-meter from the center of the participant's head. The speech was presented at an average conversational level of 65 dB SPL and the noise was varied to achieve the desired SNR (described later). The signals were calibrated using the respective calibration noises from the CST. Correct or incorrect responses to 1–5 key words within each sentence were scored. Two pairs of lists were presented in each condition (different combinations of WDRC and NR setting, as described later).

HEARING AIDS

Three behind-the-ear HAs were used for all 50 participants, chosen to reflect a range of channels, time constants, and NR methods. They were: Oticon Acto Pro (8 channels, 118 dB SPL maximum output, 51 dB maximum gain), Phonak Ambra (16 channels, 129 dB SPL maximum output, 60 dB maximum gain) and Widex Mind 440 (15 channels, 124 dB SPL maximum output, 56 dB maximum gain). All aids included programming options for linear gain control, compression gain control, and NR. Manufacturers implement algorithms differently; therefore, it was possible that the implementation of a NR or WDRC algorithm in one device would yield different results than an algorithm implemented in another device. The Oticon Acto Pro changes compression characteristics, including attack and release time, based on the “identity” chosen for the individual. The time constants range from 75 ms to 1 s, depending on the “identity” (Oticon, 2009). The active “identity” was programmed to the HA for all participants, which presumably is on the faster side of the time constant range. The Phonak Ambra uses an attack time of 1 ms and a release time of 20 ms (Phonak, 2010). The Widex Mind 440 uses adaptive, channel-specific time constants, with attack times ranging from 10 ms to 2 s and release times ranging from 10 ms to 20 s (Widex, 2013). Regarding the NR algorithms, all three manufacturers have used some form of modulation detection, while Oticon also uses synchrony detection (Chung, 2004). Widex uses slower NR time constants (e.g., 5 seconds to reach maximum gain reduction), Phonak uses faster NR time constants, while Oticon's NR time constants vary depending on whether speech and/or noise is detected (e.g., 0.2 – 7 seconds; Chung, 2004). The number of channels used for NR processing across manufacturers has not been published, nor have more specific algorithm details.

The HAs were coupled to the participant's ear with an unvented Comply tip (Oakdale, MN) and programmed to NAL-NL1 (Byrne et al., 2001) targets for all participants (regardless of NH or HI group status), using individual threshold levels and real-ear measures. The Frye FP35 (Tigard, Oregon) was used for real-ear measures and coupler matching of a digital speech signal to a 65 dB SPL NAL-NL1 target with one randomly chosen HA among the three. The randomization led to approximately even distribution, with a slight bias towards the Widex aid (19) being chosen more times than the Phonak (15) or Oticon (16) aid. The match between the target and real-ear response was within 5 dB from 250–3000 Hz, but deviated by up to 10 dB at 4000 Hz and 15 dB at 6000 Hz. A 90 dB SPL pure tone sweep was compared to predicted uncomfortable levels and verified with participant judgments of loudness. Adjustments were made if the sweep was judged to be too loud. A second memory was used to create a linear program that was adjusted to match the WDRC real-ear response at 65 dB SPL. The outputs of the remaining two test aids were matched in the test box to the response of the first aid using a 65 dB SPL speech signal. On a group level, the average output responses between the 3 hearing aids were within 1.5 dB of each other from 500–6000 Hz, with the outputs at 250 Hz being within 4.2 dB of each other. On an individual level, the outputs between devices, averaged across frequency, were within 10 dB. Deviations greater than 5 dB usually only occurred at 1–2 frequencies within an individual. The greatest deviation between devices occurred at 250 Hz. The Widex aid tended to have higher outputs at 250 Hz, even when the gain was at the lowest possible setting. Limitations in matching outputs between devices likely occurred due to differences in the number of channels and crossover frequencies between aids, and depending on which aid was used to match target first. All aids included four programs: Linear (LIN), LIN+NR, WDRC, WDRC +NR. These were programmed in the same memory position for all aids, but randomly presented to each participant (see details in Procedure). Feedback cancellation was disabled during behavioral testing for all except four participants. For these participants, amplification targets could not be matched without excessive feedback occurring when feedback cancellation was deactivated for all HAs. During acoustical measures, feedback cancellation was always disabled. Frequency lowering and directional microphone algorithms were disabled for all testing.

PROCEDURE

Behavioral and acoustical procedures were performed on the same day.

Behavioral Testing—Participants were seated in the center of a double-walled, 8 foot × 8.8 foot, Industrial Acoustics Company sound-treated booth. Speech and babble were presented from a speaker at 0 degrees azimuth and elevation, 1 meter from the participant's head. A microphone was placed near the participant's mouth allowing the experimenter to judge the participant's response online. A second microphone and a digital voice recorder were placed around the participant's neck to record responses, thereby permitting offline analysis in the event that the experimenter did not understand a response at the time of testing. A foam earplug was placed in the participant's non-test ear.

The speech was presented at 65 dB SPL, as measured at the location of the center of the participant's head without the participant present. The SNR of the stimuli was chosen

individually for each participant to correspond to 50% correct performance, or the speech reception threshold in noise (i.e. SRTn). The SRTn was measured in two stages using a single HA set to the linear program. The chosen aid was counterbalanced across participants so each aid had equal chance of being used for practice. First, the listening passages of the CST were used to provide a general estimate of the level where the participant scored 50% correct (SRTn). The order of the listening passages was randomized for each participant, which resulted in approximately equivalent usage of each passage across participants. The listening passages started at an SNR of +15 dB and dropped in 5-dB decrements. The SNR was reduced until the participant began to miss more than half of the words in the passage. On average, each participant heard 4–5 listening passages. Second, the four CST practice pairs were used to fine-tune the SRTn estimate. The first pair was presented at the SNR estimated to correspond to their SRTn, as determined by the listening passages. Additional pairs were then presented at better and poorer SNRs in 1-dB step sizes to find the SNR where approximately 50% of the words were understood. In cases of ambiguity, the SNR where the participants scored above 50% was specified as the SRTn. For example, if a participant understood 28/50 words (56% correct) at 0 dB SNR and 23/50 words (46% correct) at –1 dB SNR, the more favorable level would be used (0 dB SNR in this example). Use of the SRTn metric avoids floor or ceiling effects that can otherwise occur when a fixed SNR is used for all participants.

Testing was performed with the speech level fixed at 65 dB SPL and the noise set to the level that yielded an input SNR that corresponded to the individual's estimated SRTn. The noise remained on during breaks between sentences in an effort to stabilize the behavior of the NR algorithm. The HA devices (Oticon, Phonak, Widex) were blocked and randomly presented, and the processing conditions (LIN, LIN + NR, WDRC, WDRC +NR) were randomized within each block, for each participant. This procedure led to equivalent distribution among the presentation order of HA device and processing conditions. Participants heard two pairs of CST lists in each condition. The order of CST test sentences was also randomized for each participant. The average score was calculated between the two pairs of CST lists in each condition, allowing for identification of a 1–2 dB SNR difference in speech perception when compared across conditions (Cox et al., 1987, 1988). The participant was encouraged to take a brief break from testing and walk around between each aid tested (every ½ hour).

Acoustical Testing—A modified version of Hagerman's phase-inversion technique was used to extract the individual speech and babble signals from the mixed signal at the output of the HA, which were then used to estimate the output SNR. Four versions of the CST stimuli were generated for this technique:

1. Original-phase speech + original-phase noise (+Sp +N)
2. Phase-inverted speech + original-phase noise (–Sp +N)
3. Original-phase speech + phase-inverted noise (+Sp –N)
4. Phase-inverted speech + phase-inverted noise (–Sp –N)

The original description of this phase-inversion technique given by Hagerman & Olofsson (2004) used only versions 1–2 of stimuli. Souza, Jenstad, and Boike (2006) included version

3 for ease of combining files (Jenstad; personal communication September 2, 2010). Condition 4 was included to measure a type of error produced when using the Hagerman phase-inversion technique with non-linear processing (described later). The initial 30 seconds of the 1-minute stimuli were used to allow the hearing aid algorithms time to activate; the adjacent 30 seconds were used for the analysis.

Recording of Stimuli: Because Hagerman's phase-inversion technique requires several recordings that have identical amplitudes to be made sequentially, real-ear measurements were ruled out for this procedure since maintaining patient stance would be difficult for long periods of time. After the participant completed the experiment and left the lab, each hearing aid was coupled to a Knowles Electronic Manikin for Acoustic Research (KEMAR; Knowles Electronics, 1972; Itasca, IL, USA). KEMAR has the physical dimensions of an average human adult and is equipped with an ear simulator (including a pinna). Differences between HA output responses measured on KEMAR and those measured in real-ears are assumed to be constant across input level and frequencies such that a change in SNR measured on a real-ear yields an identical SNR change when measured on KEMAR. KEMAR was situated in the same position as the participants during speech perception testing. To measure HA output, an ER-11 1/2" microphone (Etymotic Research, Elk Grove Village, IL, USA) was coupled to the terminating portion of the ear simulator. Adobe Audition (San Jose, California) was used to control presentation of the stimuli from the loudspeaker in the booth, and simultaneously record the output of the HA. Four output recordings were made for each HA condition: HA(+Sp +N), HA(-Sp +N), HA(+Sp -N), and HA(-Sp -N).

Extraction of Speech and Noise Signals: In each HA condition, the recordings were added in Adobe Audition to extract the speech or the babble as follows:

1. extracted noise = $[HA(+Sp +N) + HA(-Sp +N)] / 2$
2. extracted speech = $[HA(+Sp +N) + HA(+Sp -N)] / 2$

Subsequent analysis on the extracted speech and noise signals was done in MATLAB (The MathWorks, Inc.). Each speech and noise file was bandpass filtered using a bank of 18 1/3-octave band filters, according to the approximate frequencies used in the ANSI S3.5-1997 (R2012) Speech Intelligibility Index standard, from 160–8000 Hz. The root mean square (RMS) levels in each band and across bands were calculated over 120 ms time windows and used to derive the narrowband SNR and wideband SNR, respectively, for each test condition. Only the wideband SNR was used to describe the output SNR of the hearing aid.

Efficacy of Hagerman's phase-inversion technique: Hagerman's phase-inversion technique was designed for use with linear or quasi-linear systems. Due to the use of this technique with non-linear processing in the current study, it is important to quantify the accuracy of the measures and account for any error generated from the inversion technique itself. To quantify error, the 180° inverted recordings were added to the 0° phase recordings to obtain an error estimate:

1. error = $HA(+Sp + N) + HA(-Sp -N)$

The error should be below the softest level of the extracted speech or noise signal (Johannesson, 2006a, 2006b). One way to estimate the difference between the error and the extracted signal of interest is by measuring the amount of attenuation obtained (Hartling et al., 2012). We measured the long-term RMS level of the speech, babble, and error signals using 120 ms windows, averaged over the second half of the 1-minute recording. The difference between the level of the error and the level of the speech or noise was calculated, and referred to as “attenuation.” Previous results using these techniques have shown attenuation values of 20 dB or greater as providing sufficient separation of the extracted signals (Johannesson, 2006a, 2006b; Naylor & Johannesson, 2009; Jenstad & Zakis, 2011; Hartling et al., 2012). Therefore, a criterion attenuation value was adopted to evaluate the efficacy of the extraction of the measurements. If the difference was less than 20 dB, then the recordings for that condition were considered unreliable, and were checked for errors. If a human error was made in the extraction, it was corrected. If not, the recording was made again and the levels were re-extracted. Following this procedure, if a recording did not pass the criterion attenuation value of 20 dB, then the recording and results were discarded from further analysis.

RESULTS

Efficacy of Hagerman Separation

The efficacy of the Hagerman separation was analyzed for each condition by assessing the error levels and attenuation values. The extracted speech and noise levels were compared to the error levels across frequency for the NH and HI groups (collapsed across participants). The error in the high frequencies (1000 Hz and above) ranged from 11–18 dB SPL, whereas the error generated in the low frequencies (800 Hz and below) ranged from 0–4 dB SPL. The higher error values at 1000 Hz and above may be from greater sensitivity to non-linearity (i.e. phase misalignment).

The error level was subtracted from the extracted speech and noise levels to obtain the amount of “attenuation” achieved, which ranged from 22–31 dB. The average attenuation relative to the extracted speech was 27.3 dB (standard deviation = 2.7) and the average attenuation relative to the extracted noise was 27.9 (standard deviation = 3.1). The attenuation was measured in each condition for each participant and compared to the criterion of 20 dB. Only 19 out of 1200 analyses did not pass the criterion amount of attenuation, suggesting our technique was efficacious for separating speech and noise from a mixed output signal. For 7 of the 19 conditions, the Oticon LIN condition did not pass the criterion attenuation amount by 1 to 1.5 dB, even after re-checking the analysis, which suggested that the extracted speech and noise signals may have been inaccurate. Six of these were from the NH group, while one was from the HI group. When error files were extracted using a different combination of recordings (+Sp–N and –SP+N), the error improved by about 3 dB. Six of the seven conditions passed the criterion when these error levels were used instead. These results suggest that perhaps the Oticon aid, in the LIN condition, was not reacting to the signals the same way every time (e.g., time variant processing), which is an assumption of using Hagerman’s phase-inversion technique. It’s also possible that other algorithms were active beyond our control in the fitting software (e.g., high frequency

directionality), or since most cases were from the NH group, that we did not have enough gain to overcome the noise floor. The data were not discarded for these cases. The other 12 conditions that did not pass were all from a single participant (participant 3), who, in contrast to the other 24 HI participants, exhibited a reverse-slope severe hearing loss. Even after re-recording these conditions, the attenuation did not improve. Therefore, this participant's data were discarded from the remainder of the analysis.

To what extent do HA processing conditions change the SNR across different HA devices?

Participants were tested at different input SNRs, based on their approximate SRTn, to (1) avoid floor and ceiling effects of testing at a fixed SNR, and (2) maximize the likelihood of seeing a change in speech perception with small changes in SNR. The mean of the input SNR for the NH group was 0.2 dB (−2 to 3 dB range) compared to 3.1 dB (0–8 dB range) for the HI group. Due to the differences in input SNR between participants and in implementation of HA processing between manufacturers, we referenced the extracted speech and noise levels (and the resulting output SNR) to those for the linear condition (re: LIN) for each HA device and participant. Therefore, if a negative change was observed, that means the output SNR was reduced compared to that for the LIN condition.

For the NH and HI groups, the mean (and standard deviation) output SNR (re: LIN) for each condition averaged across participants are displayed in Table 1. Figure 2 plots output SNR (re: LIN) against input SNR for each processing condition, and for each HA device. On average, the LIN+NR processing improved output SNR compared to the LIN condition by up to 0.25 dB. However, when combined with WDRC processing, NR resulted in lower output SNRs, compared to LIN processing, in some cases by up to −1.57 dB. This effect was driven by the WDRC processing, as WDRC without NR decreased output SNR by up to −1.75 dB. All effects of processing on output SNR are greater for the HI group than for the NH group. To answer our first question of whether changes in SNR from the input to the output of the HA depend on algorithm, a repeated-measures ANOVA was performed within each group with independent variables of HA device and processing condition and a dependent variable of SNR (re: LIN). Planned paired-samples t-tests were performed to evaluate significant differences and Bonferroni adjustments were used for multiple comparisons, reducing p to 0.025.

For the NH group, Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for type of processing ($\chi^2(2)=53.97$, $p<.0001$) and for the device by type of processing interaction ($\chi^2(9)=98.36$, $p<.0001$). Therefore the degrees of freedom were corrected using the Greenhouse-Geisser estimates. The main effects of HA device [$F(1.58, 37.83)=12.45$, $p<0.0001$] and type of processing [$F(1.05, 25.21)=156.88$, $p<0.0001$] were significant. The significant interaction term [$F(1.48, 35.63)=51.70$, $p<0.0001$] suggests that the extent of changes to SNR made by HA processing depends on the HA device for aids programmed for the NH group. Post-hoc paired-samples t-tests were performed on all combinations of devices and processing, which can be observed in Table 2, along with the size and direction of change. Only two out of 18 comparisons were not significant, after Bonferroni corrections for multiple comparisons. The non-significant conditions were

between Oticon's WDRC and WDRC+NR (top panel); and within the WDRC+NR condition between Oticon and Widex (second panel).

For the HI group, Mauchly's test of sphericity indicates that the assumption of sphericity had been violated for device ($\chi^2(2)=8.13$, $p<.017$) and for the device by type of processing interaction ($\chi^2(9)=85.51$, $p<0.0001$). Therefore the degrees of freedom were corrected using the Greenhouse-Geisser estimates. The main effect of HA device [$F(1.53, 35.14)=55.92$, $p<0.0001$] and type of processing [$F(1.93, 44.45)=95.07$, $p<0.0001$] were significant. The significant interaction term [$F(1.98, 45.57)=26.88$, $p<0.0001$], suggests that the magnitude of changes in SNR made by HA processing depended on the HA device for the HI group. Post-hoc paired-samples t-tests were performed on all combinations of devices and processing, which can be observed in Table 2, along with the size and direction of change. Only five out of 18 comparisons were not significant after Bonferroni corrections for multiple comparisons. The non-significant conditions included the WDRC and WDRC +NR comparison within Oticon and Phonak (third panel). Non-significant comparisons also occurred between Widex and both Oticon and Phonak under the LIN+NR processing condition (bottom panel), and between Widex and Oticon under the WDRC+NR processing condition.

Are changes in SNR from HA processing associated with changes in aided speech perception?

Percent correct scores were converted to rationalized arcsine units (RAU) to stabilize the error variance (Studebaker, 1985). The mean (and standard deviation) speech perception scores across listeners are displayed in Table 1. To answer our question of whether change in SNR could predict speech performance, a linear mixed model approach was used due to both within-participant (repeated measures) effects (HA device, processing and change in SNR) and between-participant effects (groups of NH and HI). The model allowing for random effects of the regression intercept yielded the best fit (i.e., minimum Akaike Information Criterion). By including the random intercept for each participant, the model accounted for within-participant correlations. SNR change was treated as a continuous linear term, and all other independent variables were treated as categorical. The non-significant main effects included change in SNR [$F(1, 385) = 0.26$; $p=0.610$], HA device [$F(2,94) = 1.65$; $p=0.199$], and processing condition [$F(2,96) = 0.89$; $p=0.416$]. The main effect of group was significant [$F(1,47) = 21.35$; $p < 0.0001$]. All three-way and two-way interactions between the independent variables were examined, but only the significant interaction was kept in the final model. This significant interaction was between HA device and hearing status group [$F(2, 94) = 15.1$; $p < 0.0001$], suggesting that the HA device effect was not consistent across the two groups. The results observed in Table 1 show that speech perception was significantly lower for Oticon than for Phonak within the NH group [-3.97 , $t(94)=-3.16$; $p=0.0021$], but the reverse was true for the HI group [3.70 , $t(94)=2.62$; $p=0.0104$]. Similarly, speech perception was also significantly worse for Oticon than for Widex in the NH group [-3.24 , $t(94)=-2.59$; $p=0.0111$], but the effect was reversed for the HI group [5.97 , $t(94)=4.68$; $p<0.0001$].

DISCUSSION

This research aimed to answer two questions: 1) to what extent do HA algorithms across different devices change the SNR when programmed for an individual?; 2) do changes in SNR from HA processing contribute to differences in speech perception across those same conditions? Because HI participants potentially do not benefit from changes in SNR as much as NH participants (MacPherson & Akeroyd, 2014), these questions were addressed in both groups. Averaged across participants and HA devices, WDRC and NR processing significantly modified the SNR between -1.75 to 0.25 dB, with a mean change of -0.25 dB. However, differences in output SNR across different HA processing conditions were not correlated with differences in speech perception across the same conditions. The lack of correlation with speech perception may be due to: 1) little variability in SNR made by HA processing; 2) the CST not being sensitive enough to capture the small changes in SNR at the output of the HA; and/or 3) the HA processing modifying the long-term SNR; but not the short-term SNR when speech was present. The reasons for and implications of these results are discussed, and individual differences are addressed.

Little Variability in SNR Changes

The changes in SNR from HA processing were smaller than previously reported in the literature. The modifications to SNR in the current study depended on the HA device and type of processing, with WDRC showing the greatest effect on SNR. The differences between previous results (e.g., Hagerman & Olofsson, 2004; Souza et al., 2006; Naylor & Johannesson, 2009; Chong & Jenstad, 2010; Wu & Stangl, 2013; Alexander & Masterson, 2015; and Brons et al., 2015) and the current results are likely due to differences in test parameters, such as input SNR and type of background noise. The input SNR varied based on individual needs (i.e., SRTn), which could have concealed more systematic effects had multiple fixed SNRs been tested across the psychometric function. Most of the participants required a SRTn (i.e., input SNR) between 0 and 4 dB. Previous work showed greater changes in SNR with WDRC processing when the SNR input was more favorable (Souza et al., 2006; Naylor & Johannesson, 2009). One option would have been to use input SNRs that reflect real world listening. Pearsons and colleagues (1977) suggested that SNRs in real world environments vary from -2 dB in trains and airplanes to $+9$ dB inside urban homes (and even more favorable SNRs in suburban environments and school classrooms). However, the disadvantage of using a fixed SNR is the potential for floor or ceiling effects within a given participant. Furthermore, the 50% point on the psychometric function is the most sensitive to small changes in SNR, and generally variation in human speech perception is greater than variation in electroacoustic measures at this point. It's also possible that the choice of background noise prevented greater changes in output SNR from being observed. Naylor and Johannesson (2009) showed that when stimuli were processed with WDRC, smaller changes in SNRs were observed when using speech backgrounds than when using steadier noise backgrounds.

Previous work suggested NR algorithms would improve the SNR by 0–5 dB (Hagerman & Olofsson, 2004; Wu & Stangl, 2013; Brons et al., 2015). Hagerman and Olofsson (2004) used a simulated NR algorithm and showed SNR improved by 3–5 dB for all five algorithms

tested, but only in a steady-state background noise, not in multi-talker babble. In fact, Hagerman and Olofsson (2004) found that in multi-talker babble, only a -1 to $+1$ dB SNR change was observed. Furthermore, for speech in a background of babble in one HA tested, Hagerman and Olofsson found that the NR improved output SNR only when input SNRs were negative. Therefore, the effect of input SNR and type of noise on the WDRC results in this study may also hold true for the NR results, likely explaining why the NR algorithms only showed an average of 0.15 dB SNR improvement. It is also possible that some of the NR algorithms were not engaged to their full potential due to the 6-talker babble used in this study, which may not have been detected as noise due to some temporal modulations inherent in the stimuli. On the other hand, the babble-noise represents a noise condition in which HA users report great communication difficulty. Another possibility is that the NR algorithm measured a more favorable SNR during the time the participant was repeating back the sentence, despite the babble remaining on between sentences, which could have led to smaller NR effects than previously reported. Under these test conditions, the effect that the NR algorithms had on the SNR was significant, yet small, whether it was combined with linear or WDRC processing.

Differences in the size of SNR changes were observed between the two groups, with the aids programmed for the HI group showing larger output SNR changes than those programmed for the NH group. This was likely due to the compression ratio differences between groups, with the aids for the HI group applying higher compression ratios than for the NH group. Naylor and Johannesson (2009) demonstrated that stronger compression ratios at more favorable SNR inputs would result in larger declines in SNR.

Significant Interactions in SNR between HA Device and Processing

It is interesting to note that SNR changes depended on HA device and processing condition, suggesting that the differences in how WDRC and NR algorithms are implemented in the 3 HA devices may be clinically relevant. One characteristic that differentiates algorithms between devices is the compression time constants. For WDRC processing, both Oticon and Phonak use faster attack and release times than Widex. The effect on SNR expected from using fast time constants is that when the input level decreases, faster compressors will increase the gain more quickly, leading to higher output levels. Since most of the input SNRs in this study fell between 0–4 dB, we would expect the gain changes of the compressor to be dominated by the speech signal. Therefore, one would expect a less favorable SNR with faster time constants, which was demonstrated by Naylor and Johannesson (2009). In this study, it was expected that Widex WDRC would behave closer to linear processing than Oticon or Phonak WDRC, which was not the case. In fact, Oticon WDRC was the closest to linear processing and Phonak WDRC was the furthest from linear processing, while Widex fell in between (Figure 2). It's possible that the compression ratios and thresholds were different enough between Oticon and Phonak to cause significant differences. While all aids were programmed to NAL-NL1 targets in their respective fitting software, and gain was verified at one input level, compression ratios were not empirically measured. Furthermore, the Phonak aid used twice as many channels (16) than the Oticon aid (8), and when combined with faster processing, an interaction is possible. Alexander and Masterson (2015)

showed that with faster release times, 16 channels led to slightly poorer output SNRs compared to 8 channels, which is in agreement with our results.

Noise reduction algorithms can also differ in their speed of processing, as well as in the method they use for speech and noise detection, and most of these details remain proprietary to the manufacturer. All of the algorithms used in this study depended on the temporal modulations of the signal, and the combination of speech and 6-talker babble may have had enough fluctuations to prevent the NR algorithm from fully engaging. This is particularly true of the Oticon algorithm which also uses synchrony detection, as the NR algorithm improved the SNR by 1 dB (with LIN processing); while NR improved SNR by 0.18 – 0.25 for Phonak and 0.20 – 0.28 for Widex. Regardless of the explanation for these effects, they highlight the need for standardized objective measures of acoustic effects from digital signal processing to assist in making clinical decisions.

Sensitivity of the Connected Speech Test

Using two passage pairs from the CST, Cox et al. (1987, 1988) demonstrated a 23% change in speech performance for NH participants as the SNR was increased by 2 dB. The HI participants showed slightly less benefit in that speech perception increased by 17% for a 2 dB increase in SNR. The changes observed in SNR in the current study often did not exceed 2 dB (Figure 2), so it's possible that the CST was not sensitive enough to reflect changes in speech perception. The CST was chosen because it reflects real world situations, where the participant usually knows the topic of conversation, but subsequent sentences about that topic become progressively more detailed. This means that participants benefitted from contextual cues, which may have masked comparatively smaller changes in improvement due to audibility. If a test that was low in context was chosen and forced a high reliance on audibility (e.g., nonsense syllables or words in isolation), a stronger relationship between output SNR and speech perception might have been observed.

Regardless of the lack of SNR output effects on CST scores, significant effects of HA device were found for both groups. In the NH group, performance with the Oticon device was significantly poorer than with the Phonak or Widex devices; however, within the HI group, the reverse was true, where the Oticon device outperformed the Phonak or Widex devices. Despite these statistical differences, the magnitude of the effects were only 5% or less. Furthermore, the effect of processing on the CST was not significant. The lack of benefit from WDRC compared to LIN processing with a speech input of 65 dB SPL is consistent with previous work, which supports greater benefits with WDRC processing for lower level inputs (e.g., 50 dB SPL; Moore et al., 1992). Although the algorithms tested in this study did not change speech perception, they may serve other purposes. For example, WDRC may act as an automatic volume control to normalize loudness, and NR may reduce listening effort and increase acceptance of noise.

Acoustic Changes to Short-Term SNR and other Speech Cues

Changes in SNR resulting from HA processing over the long term (i.e., averaged over 30 seconds of speech) was measured. Speech perception changes are associated with changes in the long-term SNR when modified in the environment, or at the input (e.g., Cox et al., 1987,

1988). While the WDRC and NR processing did modify the long-term SNR, they may not have modified the short-term SNR. For example, if gain is reduced during speech segments, both noise and speech will be reduced and the SNR will not change. However, when gain is reduced during noise-only segments, only noise levels will decline and when averaged over the long term, will appear as an improvement in output SNR.

It is also possible that the algorithms used in this study made other acoustic modifications to the signal that contributed to the variability in speech perception scores (Rhebergen et al., 2009). Work by Jenstad & Souza (2005) has shown that WDRC processing modifies the temporal envelope, and that greater distortions in temporal envelope predicted declines in speech perception benefit with HA processing. Only fast release times (e.g., less than 100 ms) and high input levels (65 dB SPL or greater) created a significant amount of envelope distortion (Jenstad & Souza, 2005). Both of these criteria were within the scope of the HA parameters used in this study, particularly for the HI group. These undesirable temporal envelope changes could explain why most of the participants did not benefit from small changes in SNR. In fact, Alexander and Masterson (2015) showed that variability in speech perception was probably explained by both output SNR and temporal envelope changes, modified by length of release time. Other quantifications of HA output have been shown to better predict speech intelligibility changes, such as the Coherence Speech Intelligibility Index (Kates & Arehart, 2005; Smeds et al., 2015).

Individual Variability in Benefit from SNR Changes

Figure 2 displays some trends consistent with results from other studies (e.g., Naylor & Johannesson, 2009). It appears that while input SNR (i.e., the participant's SRTn) didn't affect the LIN+NR condition very much, the effect was much larger for WDRC and WDRC +NR processing. As a participant's SRTn increased, the more likely their HAs were to make the SNR poorer. This seems counterintuitive, given that people with poorer speech-in-noise abilities are those who need improvements in SNR from their aids the most. Yet these results are reasonable given that at higher SNRs the target speech will assume greater responsibility for driving the WDRC. In turn, WDRC processing will cause temporal envelope distortion associated with reduced speech perception, as described previously. At the same time, the higher test SNR will likely reduce the strength of the NR algorithm, further limiting the effect of this processing on SNR.

Despite small changes in SNR and the method of selecting the input SNR to correspond to the participant's 50% performance level, CST scores were highly variable between participants at a given SNR with different HA conditions (up to 60%). Alternative measures may better differentiate between participants' benefit from WDRC and NR processing, such as listening effort (Rudner et al., 2012) and/or working memory (e.g., Humes, 2002; Ohlenforst et al., 2015). It could be argued that asking participants to perform at 50% correct for an extended period of time might have been too hard for some participants (i.e., require too much effort). This idea is supported by results from Rudner et al. (2012) showing a significant, negative correlation between the environmental SNR and effort ratings. However, in the current study, post-hoc analysis showed no systematic decline in speech perception scores when arranged by presentation order for either group. Working memory

differences may have contributed to individual variability in the ability to take advantage of SNR improvements such that those with poorer working memory capacity were less able to take advantage of the audibility (i.e., improvements in SNR) provided by fast compression processing (Gatehouse et al., 2003; Lunner and Sundewall-Thoren, 2007). The influence of working memory on speech perception in noise is particularly highlighted in more complex background noises, such as the 6-talker babble used in this study (Rudner et al., 2012).

Future work might include measures of working memory and individual tolerance of noise.

CONCLUSIONS

Based on the devices and conditions used here, the SNR was only slightly (yet significantly) affected by WDRC and NR processing, but was unrelated to changes in speech perception under monaural listening conditions using sentences in 6-talker babble. The influence of changes in SNR on other hearing aid outcomes (e.g., self-report questionnaires on aided performance) is still unanswered. Future studies could include SNR measurements that reflect real world listening situations, working memory quantification, individual tolerance to noise, as well as outcomes sensitive to HA use and satisfaction in everyday life.

Acknowledgments

All authors contributed to this work. C.M. designed and performed the experiment, analyzed the data, and wrote the manuscript as part of her dissertation work at the University of Washington. K.T., R.B., and Y.W. reviewed the design, data, and analysis. J.L. helped with design and wrote the analysis Matlab code, and reviewed the manuscript. All authors discussed the results and implications, and commented on the manuscript at various stages. The authors thank Jessica Sullivan for the use of her lab space, Lorie Jenstad for training on implementation of the inversion technique, and Xuyang Zhang for statistical assistance. This work was partially supported by the National Institutes of Health (NIDCD 1R01DC012769 and P30DC004661).

Abbreviations

NR	Noise Reduction
LIN	Linear
WDRC	Wide Dynamic Range Compression
HA	Hearing Aid
SNR	Signal-to-Noise Ratio
RMS	Root-mean-square
CST	Connected Speech Test

References

- Alexander JM, Masterson K. Effects of WDRC release time and number of channels on output SNR and speech recognition. *Ear Hear*. 2015; 36(2):35–49.
- ANSI S3.5-1997. Methods for Calculation of the Speech Intelligibility Index. American National Standards Institute; New York:
- Bentler R, Chiou LK. Digital noise reduction: An overview. *Trends Amplif*. 2006; 10(2):67–82. [PubMed: 16959731]

- Billings CJ, Tremblay KL, Souza PE, Binns MA. Effects of hearing aid amplification and stimulus intensity on cortical auditory evoked potentials. *Audiol Neurotol*. 2007; 12:234–246.
- Brons I, Houben R, Dreschler WA. Acoustical and perceptual comparison of noise reduction and compression in hearing aids. *J Speech Lang Hear Res*. 2015; 58(4):1363–76. [PubMed: 26090648]
- Byrne D, Dillon H, Ching T, Katsch R, Keidser G. The NAL-NL1 procedure for fitting non-linear hearing aids: Characteristics and comparisons with other procedures. *J Am Acad Audiol*. 2001; 12:37–51. [PubMed: 11214977]
- Chong FY, Jenstad LM. The effects of digital noise reduction algorithms on speech and noise. Poster presented at the International Hearing Aid Research Conference; Lake Tahoe, CA: 2010 Aug.
- Chung K. Challenges and recent developments in hearing aids. Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. *Trends Amplif*. 2004; 8(3):83–124. [PubMed: 15678225]
- Chung K. Reducing Noise Interference: Strategies to Enhance Hearing Aid Performance. The ASHA Leader. 2010 Apr 06.
- Cox RM, Alexander GC, Gilmore C. Development of the Connected Speech Test (CST). *Ear Hear*. 1987; 8(5S):119S–126S. [PubMed: 3678650]
- Cox RM, Alexander GC, Gilmore C, Puskulich KM. Use of the Connected Speech Test with hearing-impaired listeners. *Ear Hear*. 1988; 9(4):198–208. [PubMed: 3169400]
- Dillon H. *Hearing Aids*. Turrumurra: Boomerang Press; 2012.
- Gatehouse S, Naylor G, Elberling C. Benefits from hearing aids in relation to the interaction between the user and the environment. *Intl J Audiol*. 2003; 42:S77–S85.
- Gustafson S, McCreery R, Hoover B, Kopun J, Stelmachowicz P. Listening effort and perceived clarity for normal-hearing children with the use of digital noise reduction. *Ear Hear*. 2014; 35(2):183–194. [PubMed: 24473240]
- Hagerman B, Olofsson A. A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. *Acustica*. 2004; 90:356–361.
- Hartling C, Wu YH, Bentler RA. Hearing aid algorithm stability: Hagerman's phase inversion technique. Poster presented at the American Auditory Society Conference; Scottsdale, Arizona: 2012 Mar.
- Henning RW, Bentler R. Compression-dependent differences in hearing aid gain between apeech and nonspeech input signals. *Ear Hear*. 2005; 26:409–422. [PubMed: 16079635]
- Humes LE. Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *J Acoust Soc Am*. 2002; 112(3 Pt 1):1112–1132. [PubMed: 12243159]
- Jenstad LM, Souza PE. Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J Speech Lang Hear Res*. 2005; 48(3):651–667. [PubMed: 16197279]
- Jenstad LM, Zakis J. Using an inversion technique to quantify non-linear hearing aid processing. Poster presented at the International Hearing Aid Research Conference; Lake Tahoe, CA: 2011 Aug.
- Johannesson RB. Output SNR Manual. Oticon. 2006a Report 052-08-05.
- Johannesson RB. Output SNR Measurement Method. Oticon. 2006b Report 052-08-04.
- Kates JM. *Digital Hearing Aids*. San Diego: Plural Publishing; 2008.
- Kates J, Arehart K. Coherence and the speech intelligibility index. *J Acoust Soc Am*. 2005; 117:2224–2237. [PubMed: 15898663]
- Lunner T, Sundewall-Thorén E. Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J Am Acad Audiol*. 2007; 18(7):604–17. [PubMed: 18236647]
- MacPherson A, Akeroyd MA. Variations in the slope of the psychometric functions for speech intelligibility: A systematic survey. *Trends Hear*. 2014:1–26.
- Moore B, Johnson J, Clark T, Pluvineau V. Evaluation of a dual channel full dynamic range compression system for people with sensorineural hearing loss. *Ear Hear*. 1992; 13(5):349–370. [PubMed: 1487095]
- Naylor G, Johannesson RB. Long-term signal-to-noise ratio at the input and output of amplitude-compression systems. *J Am Acad Audiol*. 2009; 20(3):161–171. [PubMed: 19927686]

- Ohlenforst B, Souza PE, MacDonald EN. Exploring the relationship between working memory, compressor speed, and background noise characteristics. *Ear Hear*. 2015; 37:137–143.
- Olsen HL, Olofsson A, Hagerman B. The effect of audibility, signal-to-noise ratio, and temporal speech cues on the benefit from fast-acting compression in modulated noise. *Int J Audiol*. 2005; 44(7):421–433. [PubMed: 16136792]
- Oticon. Oticon Clinical Update. 2009. Retrieved from <http://www.oticon.com/~asset/cache.ashx?id=10212&type=14&format=web>
- Pearsons KS, Bennett RL, Fidell S. EPA Report No. 600/1-77-025. 1977
- Phonak. Phonak Ambra Technical Data. 2010. Retrieved from http://www.phonakpro.com/content/dam/phonak/b2b/C_M_tools/Hearing_Instruments/Phonak_Ambra/Documents/02-gb/com_datasheet_Phonak_Ambra_microP_GB_V1.00.pdf
- Rhebergen K, Versfeld N, Dreschler WA. The dynamic range of speech, compression, and its effect on the speech reception threshold in stationary and interrupted noise. *J Acoust Soc Am*. 2009; 126(6): 3236–3245. [PubMed: 20000937]
- Rudner M, Lunner T, Behrens T, et al. Working memory capacity may influence perceived effort during aided speech recognition in noise. *J Am Acad Audiol*. 2012; 23(8):577–589. [PubMed: 22967733]
- Smeds K, Wolters F, Rung M. Estimation of signal-to-noise ratios in realistic sound scenarios. *J Am Acad Audiol*. 2015; 26(2):183–96. [PubMed: 25690777]
- Souza PE, Jenstad LM, Boike KT. Measuring the acoustic effects of compression amplification on speech in noise. *J Acoust Soc Am*. 2006; 119(1):41–44. [PubMed: 16454262]
- Tremblay KL, Miller CW. How Neuroscience Relates to Hearing Aid Amplification. *International Journal of Otolaryngology*. 2014; 2014:7. Article ID 641652. doi: 10.1155/2014/641652
- Widex. Frequently Asked Questions. 2013. Retrieved from http://www.widexpro.com/wps/portal!/ut/p/b1/04_SjzS2NDQwtTQyt9SP0L_KSyzLTE8syczPS8wB8aPM4s0NDLwtnAwdDfwDHd0MPD0MDIIIs_J2N_C3M9XOjHBUBz3btag!!/#FAQNUM19
- Wilson RH, McArdle R. Speech-in-noise measures: Variable versus fixed speech and noise levels. *Int J Audiol*. 2012; 51:708–712. [PubMed: 22640687]
- Wu H, Stangl E. The effect of hearing aid signal-processing schemes on acceptable noise levels: perception and prediction. *Ear Hear*. 2013; 34(3):333–41. [PubMed: 23334355]

Figure 1. Mean hearing threshold levels (dB HL) for the participants in the normal hearing (diamonds) and hearing impaired groups (circles). Error bars represent the standard deviation at each frequency.

Figure 2.

Output SNR (re: Linear) plotted as a function of input SNR (i.e., SRTn) for the normal hearing group (left column) and the impaired hearing group (right column). Each row is a different processing condition (1: LIN + NR; 2: WDRC; 3: WDRC + NR). The mean and standard deviation for each HA device is displayed. The sample size is shown as numbers in parenthesis above each input SNR.

Mean and standard deviations across participants for output SNR (re: LIN) and speech perception (CST; RAU) for the normal-hearing (NH) and hearing-impaired (HI) groups for all conditions.

Table 1

		Output SNR (re: LIN)		CST (RAU)	
		NH	HI	NH	HI
Oticon	LIN	-	-	54.49 (14.17)	48.51 (9.24)
	LIN+NR	.00 dB (.06)	.10 dB (.14)	58.28 (13.19)	53.33 (12.18)
	WDRC	-.14 dB (.12)	-.32 dB (.33)	57.00 (13.16)	48.69 (9.91)
	WDRC+NR	-.16 dB (.10)	-.32 dB (.33)	58.25 (14.37)	50.10 (10.78)
Phonak	LIN	-	-	58.19 (12.22)	50.08 (11.70)
	LIN+NR	.23 dB (.10)	.25 dB (.08)	62.76 (11.78)	49.39 (10.70)
	WDRC	-.69 dB (.38)	-1.75 dB (.93)	62.22 (12.0)	45.58 (7.03)
	WDRC+NR	-.43 dB (.33)	-1.57 dB (1.0)	60.23 (15.26)	45.12 (9.03)
Widex	LIN	-	-	58.40 (13.64)	47.13 (10.91)
	LIN+NR	.09 dB (.08)	.20 dB (.16)	61.09 (11.32)	43.97 (10.36)
	WDRC	-.34 dB (.20)	-.63 dB (.42)	59.16 (14.74)	46.38 (10.40)
	WDRC+NR	-.17 dB (.19)	-.35 dB (.45)	62.96 (14.47)	43.83 (10.62)

Table 2
Post-hoc paired comparisons of the hearing aid device-processing interaction in output SNR.

NH Group		Mean	SD	t	df	Sig. (2-tailed)
Oticon	LIN+NR	0.15	0.12	6.33	24	<0.0001
	WDRC					
	LIN+NR	0.16	0.10	7.91	24	<0.0001
Phonak	WDRC	0.02	0.06	1.40	24	0.176
	LIN+NR	0.92	0.42	10.94	24	<0.0001
	LIN+NR	0.65	0.36	9.01	24	<0.0001
Widex	WDRC	-0.26	0.07	-18.34	24	<0.0001
	LIN+NR	0.44	0.21	10.60	24	<0.0001
	LIN+NR	0.26	0.18	7.15	24	<0.0001
	WDRC	-0.18	0.08	-11.56	24	<0.0001
NH Group		Mean	SD	t	df	Sig. (2-tailed)
LIN + NR	Oticon	-0.22	0.10	-10.62	24	<0.0001
	Phonak					
	Oticon	-0.09	0.08	-5.28	24	<0.0001
	Phonak	0.13	0.09	7.07	24	<0.0001
WDRC	Oticon	0.55	0.37	7.35	24	<0.0001
	Phonak					
	Oticon	0.20	0.21	4.65	24	<0.0001
	Phonak	-0.35	0.38	-4.59	24	<0.0001
WDRC+NR	Oticon	0.27	0.33	4.03	24	<0.0001
	Phonak					
	Oticon	0.01	0.19	0.23	24	0.820
	Phonak	-0.26	0.35	-3.68	24	0.001
HI Group		Mean	SD	t	df	Sig. (2-tailed)
Oticon	LIN+NR	0.42	0.32	6.30	23	<0.0001
	WDRC					
	LIN+NR	0.42	0.32	6.30	23	<0.0001
Phonak	WDRC	0.00	0.01	0.31	23	0.760
	LIN+NR	2.00	0.93	10.52	23	<0.0001
	LIN+NR	1.83	1.03	8.68	23	<0.0001
Widex	WDRC	-0.17	1.19	-0.71	23	0.486
	LIN+NR	0.82	0.39	10.26	23	<0.0001
	WDRC					

NH Group	Mean	SD	t	df	Sig. (2-tailed)
LIN+NR	0.54	0.36	7.44	23	<0.0001
WDRC	-0.28	0.21	-6.61	23	<0.0001
HI Group	Mean	SD	t	df	Sig. (2-tailed)
LIN + NR	-0.15	0.17	-4.48	23	<0.0001
Phonak	-0.10	0.23	-2.08	23	0.048
Widex	0.05	0.17	1.55	23	0.135
WDRC	1.43	0.94	7.46	23	<0.0001
Phonak	0.31	0.40	3.78	23	0.001
Widex	-1.12	0.89	-6.16	23	<0.0001
WDRC+NR	1.26	1.00	6.14	23	<0.0001
Phonak	0.03	0.38	0.36	23	0.720
Widex	-1.23	0.95	-6.30	23	<0.0001