Interactions Between Digital Noise Reduction and Reverberation: Acoustic and Behavioral Effects

Paul Reinhart*
Pavel Zahorik†
Pamela Souza*‡

Abstract

Background: Digital noise reduction (DNR) processing is used in hearing aids to enhance perception in noise by classifying and suppressing the noise acoustics. However, the efficacy of DNR processing is not known under reverberant conditions where the speech-in-noise acoustics are further degraded by reverberation.

Purpose: The purpose of this study was to investigate acoustic and perceptual effects of DNR processing across a range of reverberant conditions for individuals with hearing impairment.

Research Design: This study used an experimental design to investigate the effects of varying reverberation on speech-in-noise processed with DNR.

Study Sample: Twenty-six listeners with mild-to-moderate sensorineural hearing impairment participated in the study.

Data Collection and Analysis: Speech stimuli were combined with unmodulated broadband noise at several signal-to-noise ratios (SNRs). A range of reverberant conditions with realistic parameters were simulated, as well as an anechoic control condition without reverberation. Reverberant speech-in-noise signals were processed using a spectral subtraction DNR simulation. Signals were acoustically analyzed using a phase inversion technique to quantify improvement in SNR as a result of DNR processing. Sentence intelligibility and subjective ratings of listening effort, speech naturalness, and background noise comfort were examined with and without DNR processing across the conditions.

Results: Improvement in SNR was greatest in the anechoic control condition and decreased as the ratio of direct to reverberant energy decreased. There was no significant effect of DNR processing on speech intelligibility in the anechoic control condition, but there was a significant decrease in speech intelligibility with DNR processing in all of the reverberant conditions. Subjectively, listeners reported greater listening effort and lower speech naturalness with DNR processing in some of the reverberant conditions. Listeners reported higher background noise comfort with DNR processing only in the anechoic control condition.

Conclusions: Results suggest that reverberation affects DNR processing using a spectral subtraction algorithm in such a way that decreases the ability of DNR to reduce noise without distorting the speech acoustics. Overall, DNR processing may be most beneficial in environments with little reverberation and that the use of DNR processing in highly reverberant environments may actually produce adverse perceptual effects. Further research is warranted using commercial hearing aids in realistic reverberant environments.

Key Words: acoustics, digital noise reduction, reverberation, speech perception

Abbreviations: ANOVA = analysis of variance; DNR = digital noise reduction; RAU = rationalized arcsine units; SNR = signal-to-noise ratio

*Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL; †Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY; ‡Knowles Hearing Center, Evanston, IL

Corresponding author: Pamela Souza, Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208; E-mail: p-souza@northwestern.edu

INTRODUCTION

Hearing aid performance in situations with background noise remains one of the most common complaints among hearing aid users. Many listeners report that their hearing aids merely amplify the background noises, potentially causing loudness discomfort (Kochkin, 2000). In response to this, digital noise reduction (DNR) algorithms have been developed to improve the amplification of speech-in-noise signals by hearing aids. Briefly, DNR operates on the principle that although hearing aids receive a combined speech-in-noise input, speech and noise are acoustically distinct. Therefore, it is possible to estimate the speech and noise signals from the combined input using a time-variant sampling and classification of the input signal. Combined speech-in-noise inputs are decomposed across a number of frequency channels, and if the estimated power of the noise is greater than that of the speech, then gain is reduced within that channel (Levitt, 2001; Bentler and Chiou, 2006). Using this process, DNR improves the long-term signal-to-noise ratio (SNR) of a speech-in-noise input (Gustafson et al, 2014).

Although these algorithms have become increasingly advanced, the reliance on estimations of the speech and noise source acoustics invariably leads to some misclassification of the signals (i.e., speech misclassified as noise). This misclassification introduces acoustic artifact which distorts the speech information (Boll, 1979) and occurs to a greater extent when the speech and noise are acoustically similar (Arehart et al, 2003). As such, behavioral research has overwhelmingly demonstrated either no change in intelligibility or even a slight decrease in intelligibility with DNR processing (Alcántara et al, 2003; Ricketts and Hornsby, 2005; Mueller et al, 2006; Bentler et al, 2008; Sarampalis et al, 2009; Desjardins and Doherty, 2014; Ng et al, 2015). Thus, the improvement in SNR observed with DNR processing is not equivalent to a direct improvement in SNR (i.e., adjusting the speech or noise levels at their sources).

Although DNR processing is no longer expected to improve speech intelligibility in noise, research has identified other perceptual benefits associated with DNR processing. Previous studies have found that listeners experience a higher tolerance for noise and report increased comfort in noisy listening situations when listening with DNR processing (Boymans and Dreschler, 2000; Mueller et al, 2006; Bentler et al, 2008). By suppressing the amplification of noise, DNR addresses one of the core complaints of hearing aid wearers that background noise is often amplified to uncomfortably loud levels. As a result, research has shown that listeners prefer listening in noise with DNR processing because of increased background noise comfort (Jamieson et al, 1995; Ricketts and Hornsby,

2005; Brons et al, 2014). Although background noise comfort can be maximized with aggressive processing parameters which suppress the noise, this may also lead to increased misclassification and distortion of the speech which would have a substantial impact on speech understanding and perceived naturalness. Previous work has identified that listeners prefer DNR when it provides optimal background noise comfort without substantially decreasing speech understanding and perceived speech naturalness (Brons et al, 2013; 2014).

DNR processing is also associated with decreased listening effort in noise. Listening effort is broadly defined as the cognitive resources required for speech recognition, with greater resources being expended during more effortful listening (Hicks and Tharpe, 2002). Listening effort is higher in listening situations with competing auditory signals (e.g., background noise) because listeners are required to use top-down processing to inhibit or suppress the nondesired, noise signals and focus on the speech (Rönnberg et al, 2008; 2013; Stenfelt and Rönnberg, 2009). When DNR processing is active, the hearing aid assists in suppressing the noise before the combined signal is relayed to the auditory system. In doing so, DNR may make available cognitive resources that would otherwise be dedicated toward effortful noise suppression. In support of this hypothesis, previous studies have demonstrated that listeners experience enhanced dual-task performance (Sarampalis et al, 2009; Desjardins and Doherty, 2014), more rapid decoding of the speech signal (Gustafson et al, 2014), and improved consolidation of information for later recall (Ng et al, 2015) when listening under highly unfavorable conditions with DNR processing. In summary, so long as DNR is able to suppress the noise while minimizing acoustic artifact and speech distortion, then DNR processing may decrease listening effort when perceiving noisy speech.

Given these perceptual benefits, DNR is a ubiquitous feature available in commercial hearing aids (Bentler and Chiou, 2006). Yet, despite the prevalence of DNR, poor hearing aid performance in noisy situations remains one of the most frequently cited reasons for discontinuing device use and hearing aid nonadoption (Kochkin, 2007; Bertoli et al, 2009; McCormack and Fortnum, 2013). Although the benefits of DNR have been relatively well researched and validated under laboratory conditions, listeners are still often unhappy with device performance in realistic listening situations.

The reason for this disparity may be that realistic listening situations frequently contain reverberation, which may interact with DNR processing. Reverberation refers to the reflection of acoustic energy off features in an environment, which causes a portion of energy to arrive at listeners substantially delayed in time relative to the direct energy. This late-arriving

energy causes the spectral and temporal contents of the signal received by the hearing aid microphone to be smeared across time (Nábělek et al, 1989; Reinhart et al, 2016). Recall that DNR must be able to accurately identify the speech and noise signals within a channel using time-varying estimations of the input signal. Any factor that disrupts the ability of DNR processing to validly distinguish the speech from noise will likely increase the amount of artifact and speech distortion caused by DNR. Because reverberation causes a smearing of acoustic energy over time, this spread of energy will potentially decrease the ability of the DNR algorithm to accurately distinguish the time-varying speech and noise signals. As a result, reverberation may cause decreased noise suppression and increased artifact rates relative to processing of anechoic signals. For these reasons, we hypothesize that reverberation reduces the benefits of DNR for noisy speech perception. Because hearing aids perform DNR processing in environments with varying reverberation, the interaction between DNR processing and reverberation is an important consideration for generalizing the effects of DNR to the real world.

The purpose of this study was to investigate acoustic and perceptual effects of DNR processing across a range of reverberant conditions for individuals with hearing impairment. To explore the potential interaction between DNR processing and reverberation, we examined the effects of DNR on signal acoustics, speech intelligibility, and subjective ratings, including perceived listening effort, speech naturalness, and background noise comfort. The specific research questions were as follows:

- What effect would DNR have on the SNR of speech-in-noise signals across a range of reverberant conditions?
- What effect would DNR have on the intelligibility of speech-in-noise signals across a range of reverberant conditions for older listeners with hearing impairment?
- What effect would DNR have on the subjective perception of speech-in-noise signals across a range of reverberant conditions for older listeners with hearing impairment?

Based on the extant literature, we predicted that without reverberation, DNR would improve the SNR. Furthermore, in the absence of reverberation, we expected either no difference or a small decrease in intelligibility and subjective speech naturalness with DNR processing, as well as subjective improvements in listening effort and background noise comfort. Last, we predicted that these acoustic and perceptual benefits of DNR would decrease or be eliminated under varying amounts of reverberation.

MATERIALS AND METHODS

Participants

Twenty-six older adults with sensorineural hearing impairment participated in the study (mean age = 73.0 years, range 60–85; 16 males, 10 females). Air-conduction thresholds were measured at 250–8000-Hz octave frequencies and interoctaves at 3000 and 6000 Hz. Bone conduction testing was performed at octave frequencies 250–4000 Hz. Participants had no more than a single air-bone gap ≥15 dB. Participants had symmetrical hearing loss defined as no more than a 10-dB difference in pure-tone average (thresholds 500, 1000, and 2000 Hz) between ears. Mean participant audiograms for both ears are depicted in Figure 1. The Northwestern University Institutional Review Board approved all study procedures. Participants completed an informed consent process before participation, and they were compensated for their time.

Speech Stimuli

Sentences

Sentence stimuli from the IEEE corpus (IEEE, 1969) were used in the present study. Overall, these sentences are low-context (e.g., "The rope will bind the seven books at once"). Sentences spoken by a single male talker from the Inland North dialect area were locally recorded to control for differences in regional dialect (McCloy et al, 2015). Sentence stimuli were selected to measure speech intelligibility because they are more indicative of real-world listening than isolated speech segments, such as nonsense syllables.

Story Passages

Story passages from the Discourse Comprehension Test (Brookshire and Nicholas, 1997) were also used
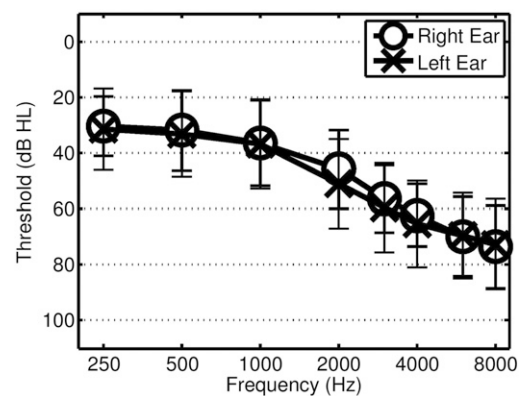
Figure 1. Mean air-conduction thresholds for participants. Error bars represent ± 1 standard deviation.

in the present study. These story passages were used because they are age-appropriate for adults, and passages are controlled for length, grammatical complexity, listening difficulty, and information content. Story passages were selected to measure subjective ratings because they are longer stimuli which have the advantage of providing listeners a longer auditory sample with which to form their subjective judgments. In addition, the length of the story passages would more fully capture the dynamic, time-varying effects of the DNR processor than shorter stimuli. This would better reflect the cumulative effects of DNR processing on listener perception in the real world.

Stimuli Processing

Sentence and story passage stimuli were processed in three stages: (a) combined with noise, (b) convolved with simulated reverberation, and (c) processed with simulated hearing aid DNR.

Noise

Both sentence and story passage stimuli were first combined with unmodulated broadband noise. The broadband noise had a flat spectrum from 0 to 20000 Hz. This noise was used because DNR may not be engaged by more speech-like noises because of the acoustic similarities between the speech and noise (Bentler et al, 2008). The noise preceded the onset of the speech stimuli by 2.0 sec to provide the DNR processor with an initial noise estimate. On the basis of pilot testing to avoid floor and ceiling effects when combined with reverberation, the background noise was added at 3- and 8-dB SNR for the sentence stimuli and at 8-dB SNR for the story passages. In addition to avoiding floor and ceiling effects, these SNRs are indicative of the generally favorable SNR conditions in which hearing aid users typically listen (Smeds et al, 2015; Humes et al, 2018). The speech level was fixed at 65-dB SPL and the noise level varied to yield the final SNR levels.

Reverberation Simulation

The second stage of signal processing was to process the speech-in-noise signals using a reverberation simulation. Reverberation was simulated using a MATLAB-based program developed and validated by Zahorik (2009), which produced binaural room response simulations. Briefly, the simulation used a three-dimensional image model (Allen and Berkley, 1979) to simulate early specular reflections within a hypothetical room. The simulated reflections were a computation of the directions, delays, and attenuations of different sound energy reflections across the frequency spectrum. The direct path and 500 early reflections were spatially ren-

dered using nonindividualized head-related transfer functions. Late response modeling was constructed for each simulated ear using independent Gaussian noise samples. Separate decay functions were applied to six octave bands ranging from 125 to 4000 Hz derived from the Sabine equation (Sabine, 1922) to estimate the time required, in seconds, for the reflected acoustic energy to decrease by 60 dB after the offset of the source signal. Overall, this method provided experimental control and flexibility while producing reverberant signals that are accurate physical and perceptual approximations of those measured in real rooms (Zahorik, 2009).

In the current simulation, source–listener distance was fixed, whereas the room size and absorptive properties of the reflective surfaces were varied to produce a range of reverberation conditions. Source–listener distance was fixed at 1.4 m, which is a typical conversational distance. Both the speech and noise signals had the same source position. In the real world, larger rooms tend to produce longer reverberation times. To reflect this, the room size was increased incrementally with increasing reverberation time conditions. In total, four room conditions were simulated, exploring a range of reverberation conditions: one anechoic control condition (without any reverberation) and three experimental conditions (containing various amounts of reverberation). See Table 1 for a summary of these conditions and the reverberation parameters.

Two metrics were used to quantify the degree of reverberant degradation: reverberation time and clarity index. Reverberation time was calculated as the time required for the impulse response to decay by 60 dB relative to its initial level averaged across octave bands from 125 to 4000 Hz. Longer reverberation times are typically indicative of greater reverberant degradation because reverberant energy persists for a longer duration. Clarity index was calculated as the logarithmic ratio of direct sound and early reflections arriving within the first 50 msec, to the late reflections arriving after 50 msec (e.g., Martellotta, 2010). A higher clarity index is expected to yield higher perceived clarity because there is less reflected energy to mask out the direct energy from subsequent speech portions. Both duration (reverberation time) and density (clarity index) quantifications may be important measures when it comes to how the reverberation will affect listener perception and interact with hearing aid processing.

DNR Simulation

The third and final stage of signal processing was to process the reverberant speech-in-noise signals through a DNR amplification simulation. DNR was simulated using a MATLAB-based hearing aid simulation developed by Kates (2008). The method uses a spectral subtraction DNR algorithm (Boll, 1979), which is a

Table 1. Simulation Condition Details for Each of the Room Conditions

Room Condition	Room Size (Length × Width × Height)	Reverberation Time (sec)	Clarity Index (C50)
Control	Free field	0.00	∞
Room 1	5.7 m × 4.3 m × 2.6 m	0.75	7.79
Room 2	8.6 m × 6.5 m × 3.9 m	1.50	4.82
Room 3	12.9 m × 9.8 m × 5.9 m	3.00	6.21

family of related algorithms widely implemented in commercial hearing aids. Both channels of the reverberant speech-in-noise signal were separately processed by the DNR simulation, similar to a bilateral hearing aid fit independently applying signal processing. Briefly, the algorithm used an adaptive procedure described by Arslan et al (1995), which operates in two stages: (a) estimate the noise spectrum across frequency bands and then (b) use the estimated SNR in frequency bands to control a time-varying gain to subtract or attenuate the estimated noise spectrum from noisy speech.

The initial noise spectrum estimate was calculated in a 100-msec time window at the beginning of the digital file. The noise spectrum estimate was then continuously updated in 50-msec time windows in which the incoming signal was windowed and the short-time Fast Fourier transform was computed from the windowed data sequence. To limit the adaptive noise estimate from fluctuating too rapidly, updated noise estimates could not exceed 1.006 times the previous estimate or be smaller than 0.978 times the previous estimate. Next, the speech envelope was estimated using a peak detector with an attack time of 3 msec and release time of 50 msec. The instantaneous SNR was calculated across frequencies using the current noise estimate and the speech estimate obtained from the peak detector. The noise in each frequency band is estimated from the noisy speech segments input to the processing. If the noisy speech segment power is greater than the current noise estimate, the noise estimate is incremented at a rate of 10 dB/sec. If the noisy speech segment power is less than the current noise estimate, the noise estimate is decremented at a rate of -25 dB/sec at that frequency region. In the current simulation, the maximum allowed speech attenuation was set to 10 dB, which is within the range of attenuation values typically used in commercial hearing aids (Kates, 2008). Last, stimuli were bandpass-filtered from 250 to 6000 Hz to represent the typical hearing aid receiver bandwidth.

Acoustic Measure

The changes to the reverberant speech-in-noise signals as a result of DNR processing were measured as the change in SNR (i.e., the difference between output SNR following DNR processing and input SNR). Change in SNR was quantified using the Inversion Method (Hagerman and Olofsson, 2004), which has previously been used to quantify the acoustic effects of DNR in hearing aids (Gustafson et al, 2014). This method used signal phase cancellation to isolate speech and noise signals following processing to compute and compare the relative root mean square values. To isolate speech and noise signals, two versions of the speech-in-noise signals were processed by the DNR simulation: (a) a phase-normal speech and noise combination; (b) a phase-normal speech and a phase-inverted noise signal. Following DNR processing, the signals were added together to isolate the processed speech component. The processed noise was then isolated by subtracting the isolated processed speech from the processed speech-in-noise. After isolating the processed speech and noise signals, the root mean square values were calculated only in the sampling window where speech was present (i.e., ignoring the effects of DNR on noise preceding and following the speech). Isolated speech and noise signals were then divided by two to correct for the doubling of the phase-normal signal. The processed SNR in dB was then calculated comparing the log ratio of the processed speech with the processed noise. Last, change in SNR was calculated as the difference between the initial SNR of the reverberant speech-in-noise signal and the output SNR of the signal following DNR processing.

Behavioral Measures

Speech Intelligibility

Speech intelligibility scores were obtained in each of the 16 test conditions (two SNRs, four reverberation conditions, and either with or without DNR processing) using 160 sentence stimuli (ten per condition). Speech intelligibility testing was conducted in a double-walled sound booth. Digital signals were converted to analog by Tucker-Davis Technologies (Alachua, FL) equipment and played through Etymotic-ER2 insert phones (Elk Grove Village, IL). Sentences were presented binaurally. The speech presentation was fixed at 65-dB SPL to represent a conversational level of speech. With the addition of noise, final presentation levels were calibrated to either 66.8-dB SPL (for the 3-dB SNR condition) or 65.6-dB SPL (for the 8-dB SNR condition). Listeners received individual NAL-NL1 shaping for each ear to mimic the individualized frequency shaping provided by wearable hearing aids (Byrne et al, 2001).

Sentences were presented and scored using custom-made MATLAB code. Each sentence contained five key words for scoring (e.g., "The *rope* will *bind* the *seven books* at *once*"). Final speech intelligibility was recorded as the percent of words correctly repeated within a given condition (50 key words per condition).

Subjective Ratings

Subjective ratings were obtained in eight test conditions (four reverberation conditions, each with/without DNR) using nine passages from the Discourse Comprehension Test (one per condition + a practice story). Because of the limited number of story passages, we only examined the effects of DNR at 8-dB SNR. Story listening was performed in a double-walled sound booth. Digital signals were converted to analog by M-Audio sound card and played through Sennheiser HD25 supra-aural headphones. Stories were presented binaurally. Similar to the sentence task, speech presentation was fixed at 65-dB SPL with the noise added for a final presentation level of 65.6-dB SPL for the 8-dB SNR condition plus NAL-NL1 frequency shaping based on individual audiograms for each ear (Byrne et al, 2001).

Following presentation, listeners were asked to rate three subjective aspects of the story listening experience: (a) how much overall listening effort it took for them to understand the story, (b) how natural the speech of the story was, (c) and how comfortable the background noise was while listening to the story. The ratings were obtained on 7-point scales (see Appendix for scales) similar to those which have been previously used in the hearing literature (e.g., Johnson et al, 2015; Reinhart and Souza, 2016). Comprehension scores were recorded using the original questions from the Discourse Comprehension Test; however, results were not analyzed given that the Yes–No answer format yields high variability when only including one story per condition. Instead, the comprehension questions were included in the test protocol to keep listeners engaged with the listening task. Stories were presented and listeners made their ratings using custom-made MATLAB code. Listeners completed 1 practice story in quiet to acclimate them to the test procedure/interface and storytelling narrative structure.

RESULTS

Acoustic Results

See Figure 2 to view the change in SNR as a result of DNR processing for the different reverberation room conditions and different input SNRs. One-sample t-tests with a Bonferroni correction indicated that DNR produced a significant change in SNR for all room conditions (all $p < 0.05$). A two-way analysis of variance (ANOVA) test was conducted with Input SNR and Room Condition as fixed, independent variables and change in SNR as the dependent variable. There was a statistically significant main effect of Input SNR [$F_{(1, 392)} = 280.013$, $p < 0.001$, partial $\eta^2 = 0.417$] in which the change in SNR as a result of DNR processing was greater at the lower input SNR. There was also a statistically significant main effect of room condition [$F_{(3, 392)} = 64.622$, $p < 0.001$, partial $\eta^2 = 0.331$]. The interaction between room condition and input SNR
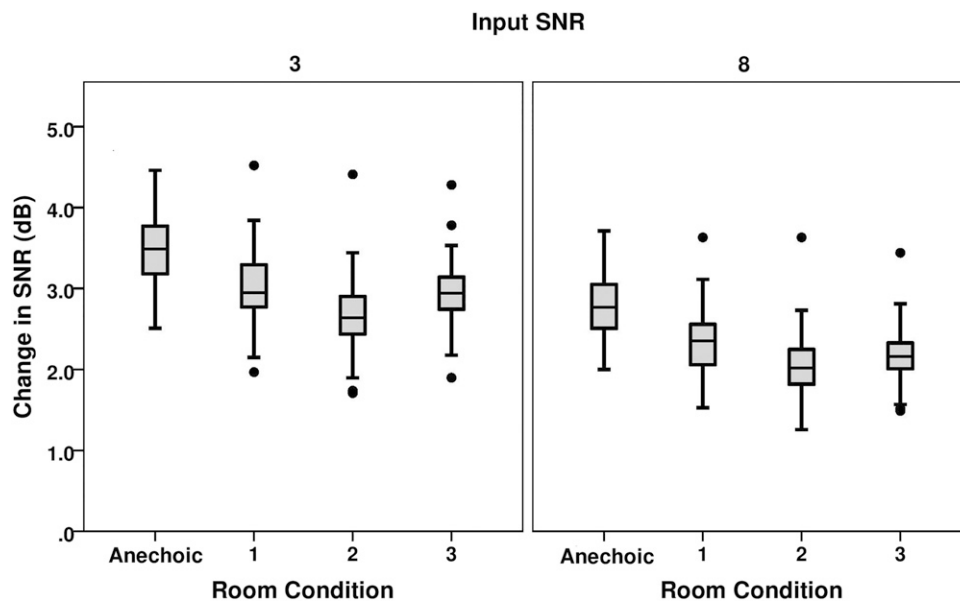

Figure 2. Change in SNR quantified using the Inversion Method as a result of DNR processing across the different room conditions. The left and right panels show the acoustic effects at different input SNRs.

was not statistically significant [$F_{(3, 392)} = 0.634$, $p = 0.593$, partial $\eta^2 = 0.005$].

To further explore the main effect of room condition, pairwise comparisons with a Bonferroni correction were conducted. Change in SNR was highest in the control room condition compared with any of the reverberant room conditions ($p < 0.001$). In addition, the change in SNR in room condition 2 was significantly lower than in either room condition 1 or room condition 3 ($p < 0.001$). There was no significant difference between room condition 1 and room condition 3 ($p = 0.228$).

Speech Intelligibility Results

Raw intelligibility scores were first transformed to rationalized arcsine units (RAU) to normalize the variance near floor and ceiling (Studebaker, 1985). See Figure 3 to view the transformed speech intelligibility scores for all the conditions. Data were analyzed using a three-way repeated-measures ANOVA with Room Condition (Anechoic, 1, 2, 3), Input SNR (3, 8), and DNR (unprocessed, DNR) as within-subject variables. All assumptions of the model were met. A summary of the results of the repeated-measures ANOVA are depicted in Table 2.

The significant Room Condition × DNR interaction is most relevant to the purpose of the study (research question 2) and was further analyzed. Results were collapsed across SNR and paired-samples t-tests were conducted for each room condition with a Bonferroni correction. Transformed data are depicted in Figure 4. The difference in speech intelligibility for processed versus DNR conditions was not significantly different in the anechoic room condition ($p > 0.05$); however, the difference was significant in all other room conditions (all $p < 0.05$).

Subjective Ratings Results

Subjective listener ratings of listening effort, speech naturalness, and background noise comfort during story listening in noise across different room conditions both with and without DNR processing can be seen in Figure 5. All of the y-axes are oriented such that larger numbers reflect better ratings (i.e., less listening effort, higher speech naturalness, and greater background noise comfort). There was a significant effect of room condition on listening effort ratings, as indicated by a significant Friedman test ($p < 0.001$). Post hoc Wilcoxon signed-rank tests with a Bonferroni correction showed a significant difference between the control condition and room condition 2 ($p = 0.010$). To analyze whether there were any benefits of DNR on subjective listening effort, Wilcoxon signed-rank tests with a Bonferroni correction were conducted between DNR and
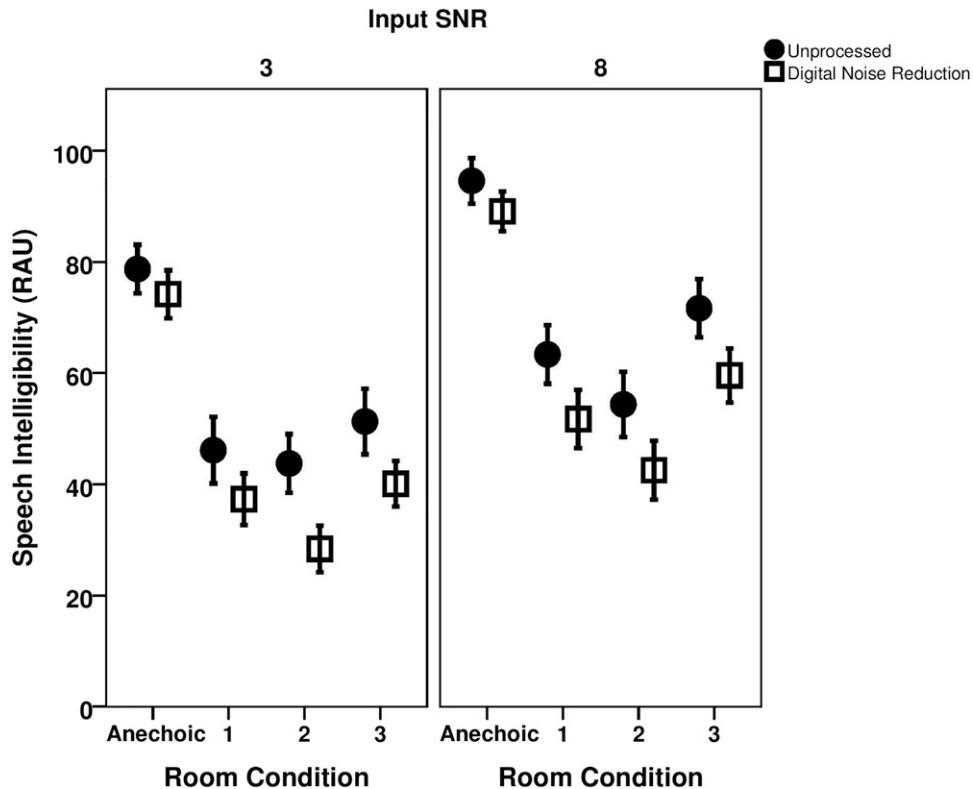

Figure 3. Speech intelligibility scores in RAUs with and without DNR processing across the different room conditions. The left and right panels show intelligibility at different input SNRs. The error bars represent ± 1 standard error.

Table 2. Results of Three-Way Repeated-Measures ANOVA on RAUs of Speech Intelligibility with Room Condition, SNR, and DNR as within-Subject Factors

		Degrees of Freedom	*F* Value	*p* Value	Partial η^2
Main effects	Room condition	**(3,72)**	**234.432**	**< 0.001**	**0.904**
	SNR	**(1,24)**	**243.769**	**< 0.001**	**0.907**
	DNR	**(1,24)**	**36.723**	**< 0.001**	**0.595**
Two-way interactions	Room condition × DNR	**(3,72)**	**2.781**	**0.046**	**0.101**
	Room condition × SNR	**(3,72)**	**2.772**	**0.047**	**0.100**
	SNR × DNR	(1,24)	0.022	0.883	0.001
Three-way interaction	Room condition × SNR × DNR	(3,72)	0.559	0.644	0.022

Note: Significant main effects and interactions are in bold.

unprocessed conditions at each level of room condition. There was a trend toward improved listening effort with DNR in the control condition, but this relationship failed to reach significance ($p = 0.161$). Listening effort was significantly poorer with DNR processing in room condition 2 ($p = 0.035$).

A Friedman test showed no effect of room condition on speech naturalness ($p = 0.316$). To analyze whether there were any effects of DNR on speech naturalness, Wilcoxon signed-rank tests with a Bonferroni correction were conducted between DNR and unprocessed conditions at each level of room condition. Speech naturalness was significantly poorer with DNR processing in room condition 1 ($p = 0.005$) and room condition 2 ($p = 0.039$). There was no effect of DNR processing on speech naturalness in the control condition or room condition 3 (both $p > 0.05$).

For background noise comfort ratings, a Friedman test showed no effect of room condition ($p = 0.078$). To analyze whether there were any benefits of DNR on background noise comfort, Wilcoxon signed-rank tests with a Bonferroni correction were conducted

Figure 4. Speech intelligibility scores in RAUs collapsed across input SNR for each of the room conditions. The error bars represent ± 1 standard error. *$p < 0.05$ after Bonferroni correction.

between DNR and unprocessed conditions at each level of room condition. Background noise comfort was significantly higher with DNR processing in the control condition ($p = 0.019$). There was no benefit of DNR processing on background noise comfort in any of the experimental room conditions (all $p > 0.05$).

DISCUSSION

The purpose of this study was to investigate acoustic and perceptual effects of DNR processing across a range of reverberant conditions for individuals with hearing impairment. Consistent with previous work (e.g., Gustafson et al, 2014), DNR processing improved the SNR of speech-in-noise signals (Figure 2). However, improvement in SNR decreased in the reverberant conditions and was smallest in room condition 2, which had the poorest ratio of direct to reverberant energy (i.e., clarity index), although the effect was small (~1 dB). This suggests that the presence of reverberant energy interferes with the ability of the DNR processor to form valid estimates of the signals and estimate the momentary SNR. If the DNR processor does not have accurate estimates of the speech and noise, then it will be more likely to misclassify parts of the speech as noise and distort those parts of the speech signal. Although the improvement in SNR is not equivalent to adjusting the speech and noise levels at their source, it is indicative of to what extent the noise is being selectively suppressed without also removing the speech. It is possible that the presence of reverberation increased the amount of artifact removal of speech by DNR, which reduced the overall output speech level. This increased acoustic artifact hypothesis would potentially explain why there was a smaller improvement in SNR in the reverberant conditions compared with the anechoic control condition.

The interaction effect of reverberation on DNR processing was also evident in the speech intelligibility data. Consistent with previous studies (e.g., Sarampalis et al, 2009), there was a slight but nonsignificant decline in speech intelligibility with DNR processing in the anechoic control condition, suggesting that DNR
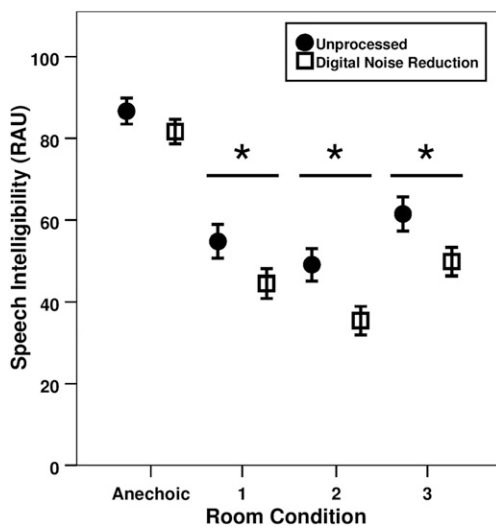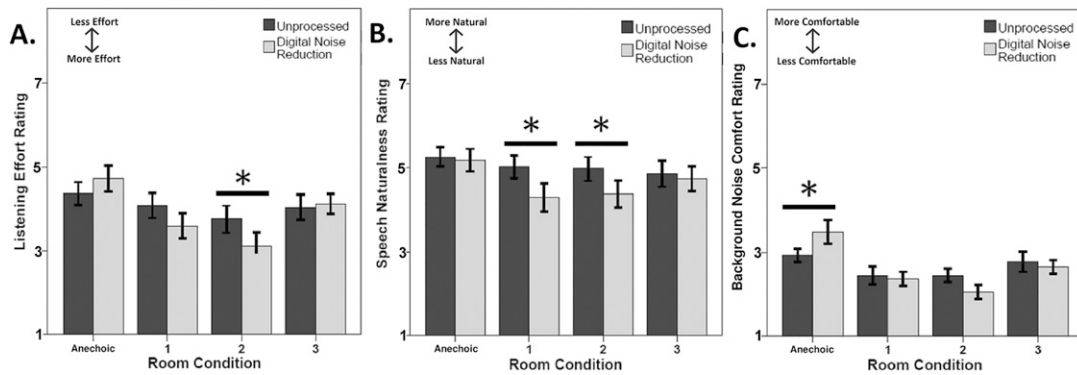
Figure 5. Subjective listener ratings after the story listening task with and without DNR processing across the different room conditions. (A) depicts subjective listening effort, (B) depicts speech naturalness, and (C) depicts background noise comfort. Y-axes of all the panels are oriented such that higher ratings reflect an improved state for the given scale. The error bars represent ± 1 standard error. *$p < 0.05$ after Bonferroni correction.

introduced some speech distortion but not enough to significantly hinder speech intelligibility. However, there was a significant decline in speech intelligibility with DNR processing (Figure 4) across all three experimental reverberant conditions which all had smaller improvements in SNR with DNR processing than the anechoic condition. Furthermore, the average decline in intelligibility was poorest in room condition 2, and there was a trend such that the declines in speech intelligibility with DNR processing increased with poorer ratios of direct to reverberant energy. Overall, the speech intelligibility results are consistent with the hypothesis that an increasing ratio of direct to reverberant energy decreases the efficacy of DNR processing leading to greater artifact and subsequent speech distortion.

Although it is generally accepted that DNR may lead to a slight decrease in speech intelligibility, it is believed that the perceptual benefits of DNR on listening effort and background noise comfort will outweigh that effect. Consistent with our prediction, listeners reported a significant increase in background noise comfort with DNR processing in the anechoic control condition. We also expected a benefit of DNR processing for listening effort; however, although we observed a trend in that direction in the anechoic control condition (Figure 5A), that trend failed to reach statistical significance. There could be several reasons for this. Previous studies have often only found significant effects of DNR on listening effort in a subset of the most difficult listening conditions (e.g., Sarampalis et al, 2009; Desjardins and Doherty, 2014). It is possible that although the 8-dB SNR is representative of real-world listening situations in which hearing aids are used, it was not challenging enough to yield a measurable change in listening effort with DNR. Another possibility is that the subjective listening effort scale used in the present study was not sensitive to changes in listening effort with DNR processing. Previous research has suggested that objective measures of listening effort (e.g., dual-

task) may be more sensitive to effects on listening effort than subjective measures that rely on listener report (Desjardins and Doherty, 2014). Thus, it is possible we would have observed a significant improvement in listening effort using an alternative, objective measure. With the growing interest in listening effort as a hearing aid outcome measure, there remains a lack of a proper method to measure listening effort (Edwards, 2007; Brons et al, 2013). Moreover, properly defining and quantifying listening effort may be particularly difficult for evaluating DNR processing in which the potential listening effort benefit may occur concurrently with a slight decrease in speech intelligibility.

The decline in speech intelligibility with DNR was greater in the reverberant conditions than in the anechoic control condition. Because the primary purpose of hearing aid processing is to facilitate speech understanding, this decrease in speech intelligibility with DNR processing may outweigh any other potential benefits of DNR processing. In the subjective ratings, listeners reported DNR-processed speech as significantly less natural in room conditions 1 and 2 (Figure 5B). This finding suggests that listeners may have been aware of the poorer overall speech intelligibility with DNR processing in reverberant environments. Listeners also indicated greater listening effort in room condition 2 presumably because DNR processing led to significantly decreased intelligibility of the story. This increase in listening effort with DNR processing is consistent with models of degraded speech perception. With the increased misclassification and degradation of speech by DNR in reverberant environments, the fragmented bottom-up representations of speech received by listeners do not match the phonological representation in their mental lexicon; thus, it is necessary for listeners to recruit top-down cognitive resources to help compensate and resolve this phonological mismatch (Rönnberg et al, 2008; 2013). This increased allocation of mental resources would lead to an increase in

listening effort. It is possible that listeners were also experiencing increased listening effort with DNR in room conditions 1 and 3 because of decreased speech intelligibility; however, the listening effort scale was not sensitive to it. Alternatively, it could be that speech intelligibility must be reduced to a certain threshold before listening effort is affected. Overall, these findings are consistent with the hypothesis that DNR introduces a greater amount of acoustic artifacts and speech distortion when processing reverberant speech-in-noise signals. Moreover, behavioral results suggest that this increased distortion may outweigh the potential benefits of DNR processing in reverberant environments.

Brons and colleagues (2013; 2014) suggested that speech naturalness and background noise comfort are determining factors for whether listeners prefer DNR processing compared with unprocessed. That is, listeners prefer DNR most when background noise comfort is maximized and speech naturalness is not degraded. Based on this, we may be able to infer whether listeners would prefer DNR processing in anechoic and reverberant environments. In the anechoic condition, there was no significant decrease in speech intelligibility or perceived speech naturalness with DNR processing; however, there was a significant increase in background noise comfort. Therefore, it is likely that listeners would prefer listening with DNR processing because DNR processing provided a significant benefit (background noise comfort) without any significant decrement (speech naturalness). However, in room condition 2, which had the poorest ratio of direct to reverberant energy and poorest speech intelligibility, listeners indicated no benefit in background noise comfort and a decrease in speech naturalness with DNR. In this situation, we may be able to infer that listeners would actually prefer listening without DNR processing because DNR processing caused a significant decrement (speech naturalness) without providing any significant benefit (background noise comfort). It is possible that if the experiment were conducted at poorer SNRs, then listeners may have preferred DNR processing across all reverberant conditions because the weighting of background noise comfort would dominate other factors. However, this scenario is less extrinsically valid as hearing aid users are not often in such unfavorable listening environments (Smeds et al, 2015; Humes et al, 2018).

We predicted that the benefits of DNR processing would be decreased or eliminated in reverberant environments. Interestingly, these results extend beyond our prediction and suggest there may even be adverse perceptual effects of using DNR processing in highly reverberant environments. Although DNR is typically applied in most hearing aid programs designed for speech listening, this would suggest the use of DNR

processing in the real world should be contingent on room acoustics. For listening in highly reverberant environments, listeners may be best served by not using a spectral subtraction DNR algorithm. However, there are several limitations, which prevent the results from being immediately generalizable to clinical contexts.

Limitations

Although the present study has important implications for the application of hearing aid DNR in realistic situations, the study is not without limitations. First, the present study used an unmodulated broadband noise, which has limited extrinsic validity. This noise was selected because it represents an ideal noise to remove and provides a theoretical framework for how DNR might operate with other noise types. For this reason, it has frequently been used in previous DNR research (e.g., Arehart et al, 2003; Gustafson et al, 2014). Nevertheless, this type of noise is not typically found in everyday listening environments. To investigate whether reverberation has a similar effect on DNR processing of more realistic noises, additional acoustic analyses were conducted post hoc using a factory noise. This noise was similar to other realistic noises in that it contained natural amplitude modulations and had a time-varying spectrum. The factory noise was combined with speech at +5-dB SNR, processed with DNR, and the change in SNR was calculated as previously described in the Materials and Methods section. The results are depicted in Figure 6 alongside comparison data using white noise calculated in the same way. As expected, DNR did not improve SNR as much for the factory noise compared with the
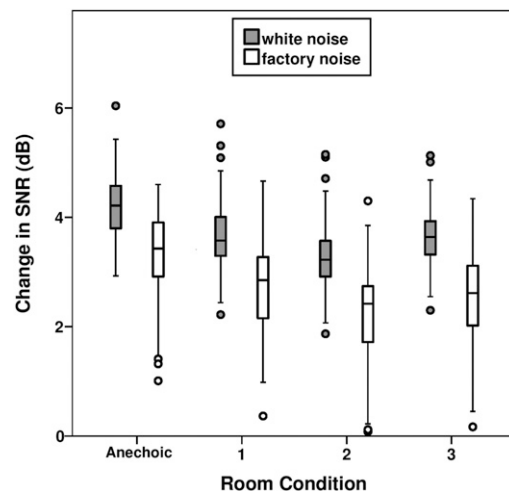

Figure 6. Change in SNR quantified using the Inversion Method as a result of DNR processing across the different room conditions. Sentences were combined with either white noise (gray boxplot) or factory noise (white boxplot) at an input SNR of +5 dB.

white noise. This is likely because the time-varying spectrum of the factory noise presents a "moving target" which is more difficult for the DNR to accurately estimate and subtract in real time. Most interestingly, the factory noise data exhibit the same effect of reverberation on change in SNR as it does using a white noise regardless of room condition. This suggests that the deleterious effect of reverberation on DNR processing is consistent for varying noise types. As such, if the behavioral experiments were to be conducted using a more realistic noise, then results would likely be similar as those in the present study obtained using unmodulated broadband noise. However, further behavioral research is warranted to examine this.

An additional limitation of the present study is that the DNR and reverberation were implemented using simulations presented over insert phones. This methodology was used to provide superior experimenter control compared with using commercial hearing aids in real reverberant environments. A spectral subtraction simulation was used because it is the basis for most commercially implemented DNR algorithms. However, the use of a nonspecific DNR algorithm and only using fixed parameters (e.g., maximum 10 dB DNR attenuation) prevents generalizing results to any specific commercial hearing aid, which may be using a different algorithm and/or parameters. In addition, the simulation and presentation of reverberation through insert phones prevents listeners from using head movements and other cues, which could be important for spatial hearing (e.g., Wightman and Kistler, 1999). The results of this study provide motivation for further study under more realistic conditions.

CONCLUSIONS

We conclude that reverberation affects the efficacy of DNR processing in such a way that likely increases the introduction of acoustic artifact and speech distortion caused by DNR processing. Although DNR processing had minimal effect on speech intelligibility without reverberation, DNR processing has a substantial effect on speech intelligibility with reverberation. Listeners also indicated greater listening effort and rated speech as less natural with DNR processing in a subset of the reverberant conditions. Last, although listeners reported listening in noise as more comfortable with DNR processing in the anechoic condition, there was no increase in comfort with DNR in any of the reverberant conditions. Overall, these results suggest that DNR processing may be most beneficial in environments with little reverberation, and that the use of DNR processing in highly reverberant environments may actually produce adverse perceptual effects. However, further research is warranted in realistic conditions using commercial hearing aids.

REFERENCES

Alcántara JI, Moore BC, Kühnel V, Launer S. (2003) Evaluation of the noise reduction system in a commercial digital hearing aid: evaluación del sistema de reducción de ruido en un auxiliar auditivo digital comercial. *Int J Audiol* 42(1):34–42.

Allen JB, Berkley DA. (1979) Image method for efficiently simulating small-room acoustics. *J Acoust Soc Am* 65(4):943–950.

Arehart KH, Hansen JH, Gallant S, Kalstein L. (2003) Evaluation of an auditory masked threshold noise suppression algorithm in normal-hearing and hearing-impaired listeners. *Speech Commun* 40(4):575–592.

Arslan L, McCree A, Viswanathan V. (1995) New methods for adaptive noise suppression. ICASSP-95. 1995 International Conference on Acoust, Speech, Signal Process 1:812–815. Detroit, MI: IEEE.

Bentler R, Chiou LK. (2006) Digital noise reduction: an overview. *Trends Amplif* 10(2):67–82.

Bentler R, Wu YH, Kettel J, Hurtig R. (2008) Digital noise reduction: outcomes from laboratory and field studies. *Int J Audiol* 47(8):447–460.

Bertoli S, Staehelin K, Zemp E, Schindler C, Bodmer D, Probst R. (2009) Survey on hearing aid use and satisfaction in Switzerland and their determinants. *Int J Audiol* 48(4):183–195.

Boll S. (1979) Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans Acoust Speech Signal Process* 27(2): 113–120.

Boymans M, Dreschler WA. (2000) Field trials using a digital hearing aid with active noise reduction and dual-microphone directionality: estudios de campo utilizando un audifono digital con reduccion activa del ruido y micrófono de direccionalidad dual. *Audiology* 39(5):260–268.

Brookshire RH, Nicholas LE. (1997) *Discourse Comprehension Test: Test Manual*. Bar Harbor, ME: BRK Publishers.

Brons I, Houben R, Dreschler WA. (2013) Perceptual effects of noise reduction with respect to personal preference, speech intelligibility, and listening effort. *Ear Hear* 34(1):29–41.

Brons I, Houben R, Dreschler WA. (2014) Effects of noise reduction on speech intelligibility, perceived listening effort, and personal preference in hearing-impaired listeners. *Trends Hear* 18: 2331216514553924.

Byrne D, Dillon H, Ching T, Katsch R, Keidser G. (2001) NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *J Am Acad Audiol* 12(1):37–51.

Desjardins JL, Doherty KA. (2014) The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear Hear* 35(6):600–610.

Edwards B. (2007) The future of hearing aid technology. *Trends Amplif* 11(1):31–45.

Gustafson S, McCreery R, Hoover B, Kopun JG, Stelmachowicz P. (2014) Listening effort and perceived clarity for normal hearing

children with the use of digital noise reduction. *Ear Hear* 35(2): 183.

Hagerman B, Olofsson Å. (2004) A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. *Acta Acustica united Acustica* 90(2):356–361.

Hicks CB, Tharpe AM. (2002) Listening effort and fatigue in school-age children with and without hearing loss. *J Speech Lang Hear Res* 45(3):573–584.

Humes LE, Rogers SE, Main AK, Kinney DL. (2018) The acoustic environments in which older adults wear their hearing aids: insights from datalogging sound environment classification. *Am J Audiol* 27(4):594–603.

IEEE. (1969) IEEE recommended practice for speech quality measurements; IEEE Report No. 297.

Jamieson DG, Brennan RL, Cornelisse LE. (1995) Evaluation of a speech enhancement strategy with normal-hearing and hearing-impaired listeners. *Ear Hear* 16(3):274–286.

Johnson J, Xu J, Cox R, Pendergraft P. (2015) A comparison of two methods for measuring listening effort as part of an audiologic test battery. *Am J Audiol* 24(3):419–431.

Kates JM. (2008) *Digital Hearing Aids*. San Diego, CA: Plural Publishing.

Kochkin S. (2000) MarkeTrak V: "why my hearing aids are in the drawer": the consumers' perspective. *Hear J* 53(2):34–36.

Kochkin S. (2007) MarkeTrak VII: obstacles to adult non-user adoption of hearing aids. *Hear J* 60(4):24–51.

Levitt H. (2001) Noise reduction in hearing aids: a review. *J Rehabil Res Dev* 38(1):111.

Martellotta F. (2010) The just noticeable difference of center time and clarity index in large reverberant spaces. *J Acoust Soc Am* 128(2):654–663.

McCloy DR, Wright RA, Souza PE. (2015) Talker versus dialect effects on speech intelligibility: a symmetrical study. *Lang Speech* 58(3):371–386.

McCormack A, Fortnum H. (2013) Why do people fitted with hearing aids not wear them? *Int J Audiol* 52(5):360–368.

Mueller HG, Weber J, Hornsby BW. (2006) The effects of digital noise reduction on the acceptance of background noise. *Trends Amplif* 10(2):83–93.

Nábělek AK, Letowski TR, Tucker FM. (1989) Reverberant overlap-and self-masking in consonant identification. *J Acoust Soc Am* 86(4):1259–1265.

Ng EH, Rudner M, Lunner T, Rönnberg J. (2015) Noise reduction improves memory for target language speech in competing native but not foreign language speech. *Ear Hear* 36(1):82–91.

Reinhart PN, Souza PE. (2016) Intelligibility and clarity of reverberant speech: effects of wide dynamic range compression release time and working memory. *J Speech Lang Hear Res* 59(6):1543–1554.

Reinhart PN, Souza PE, Srinivasan NK, Gallun FJ. (2016) Effects of reverberation and compression on consonant identification in individuals with hearing impairment. *Ear Hear* 37(2):144–152.

Ricketts TA, Hornsby BW. (2005) Sound quality measures for speech in noise through a commercial hearing aid implementing. *J Am Acad Audiol* 16(5):270–277.

Rönnberg J, Rudner M, Foo C, Lunner T. (2008) Cognition counts: a working memory system for ease of language understanding (ELU) *Int J Audiol* 47(2, Suppl):S99–S105.

Rönnberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, Dahlström O, Signoret C, Stenfelt S, Pichora-Fuller MK, Rudner M. (2013) The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci* 7:1–17.

Sabine WC. (1922) *Collected Papers on Acoustics*. Cambridge, England: Harvard University Press.

Sarampalis A, Kalluri S, Edwards B, Hafter E. (2009) Objective measures of listening effort: effects of background noise and noise reduction. *J Speech Lang Hear Res* 52(5):1230–1240.

Smeds K, Wolters F, Rung M. (2015) Estimation of signal-to-noise ratios in realistic sound scenarios. *J Am Acad Audiol* 26(2):183–196.

Stenfelt S, Rönnberg J. (2009) The signal-cognition interface: interactions between degraded auditory signals and cognitive processes. *Scand J Psychol* 50(5):385–393.

Studebaker GA. (1985) A rationalized arcsine transform. *J Speech Lang Hear Res* 28(3):455–462.

Wightman FL, Kistler DJ. (1999) Resolution of front–back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am* 105(5):2841–2853.

Zahorik P. (2009) Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J Acoust Soc Am* 126(2): 776–791.

APPENDIX: Subjective Rating Scales

Listening Effort Scale:

    7. No Effort.

    6. Very Little Effort.

    5. Little Effort.

    4. Moderate Effort.

    3. Considerable Effort.

    2. Much Effort.

    1. Extreme Effort.

Speech Naturalness Scale:

    7. Completely Natural.

    6. Natural.

    5. Somewhat Natural.

    4. Neither Natural nor Unnatural.

    3. Somewhat Unnatural.

    2. Unnatural.

    1. Completely Unnatural.

Background Noise Comfort Scale:

    7. Completely Comfortable.

    6. Comfortable.

    5. Somewhat Comfortable.

    4. Neither Comfortable nor Uncomfortable.

    3. Somewhat Uncomfortable.

    2. Uncomfortable.

    1. Completely Uncomfortable.