Working memory, age and hearing loss: susceptibility to hearing aid distortion

Kathryn H. Arehart, Ph.D.,
Speech, Language, and Hearing Sciences, University of Colorado Boulder

Pamela Souza, Ph.D.,
The Roxelyn and Richard Pepper Department of Communication Sciences and Knowles Hearing Center, Northwestern University

Rosalinda Baca, Ph.D., and
Speech, Language, and Hearing Sciences, University of Colorado Boulder

James M. Kates, E.E.
GN ReSound and Speech, Language, and Hearing Sciences, University of Colorado Boulder

Abstract

Objectives—Hearing aids use complex processing intended to improve speech recognition. While many listeners benefit from such processing, it can also introduce distortion that offsets or cancels intended benefits for some individuals. The purpose of the present study was to determine the effects of cognitive ability (working memory) on individual listeners' responses to distortion caused by frequency compression applied to noisy speech.

Design—The present study analyzed a large dataset of intelligibility scores for frequency-compressed speech presented in quiet and at a range of signal-to-babble ratios. The intelligibility dataset was based on scores from 26 adults with hearing loss with ages ranging from 62 to 92 years. The listeners were grouped based on working memory ability. The amount of signal modification (distortion) due to frequency compression and noise was measured using a sound quality metric. Analysis of variance and hierarchical linear modeling were used to identify meaningful differences between subject groups as a function of signal distortion caused by frequency compression and noise.

Results—Working memory was a significant factor in listeners' intelligibility of sentences presented in babble noise and processed with frequency compression based on sinusoidal modeling. At maximum signal modification (caused by both frequency compression and babble noise), the factor of working memory (when controlling for age and hearing loss) accounted for 29.3% of the variance in intelligibility scores. Combining working memory, age, and hearing loss accounted for a total of 47.5% of the variability in intelligibility scores. Furthermore, as the total

Corresponding Author: Kathryn H. Arehart; University of Colorado SLHS Department; 409 UCB; Boulder, CO 80309; Kathryn.arehart@colorado.edu; 303-492-3036 (office phone); 303-492-3274 (fax).

amount of signal distortion increased, listeners with higher working memory performed better on the intelligibility task than listeners with lower working memory.

Conclusions—Working memory is a significant factor in listeners' responses to total signal distortion caused by cumulative effects of babble noise and frequency compression implemented with sinusoidal modeling. These results, together with other studies focused on wide-dynamic range compression (WDRC), suggest that older listeners with hearing loss and poor working memory are more susceptible to distortions caused by at least some types of hearing aid signal processing algorithms and by noise, and that this increased susceptibility should be considered in the hearing-aid fitting process.

Keywords

frequency compression; distortion; hearing loss; working memory; aging

INTRODUCTION

Several studies have recently shown a link between cognitive abilities and response to hearing aid signal processing. For example, Gatehouse and colleagues (Gatehouse et al. 2003, 2006a,b) demonstrated that listeners with "poor" cognition performed better with slow-acting wide-dynamic range compression (WDRC) while listeners with "good" cognition performed better with fast-acting WDRC. Similarly, Lunner and Sundewall-Thorén (2007) and Foo et al. (2007) have shown a relationship between one specific cognitive component - working memory – and the ability of listeners to make use of fast-acting compression.

We hypothesize that the relationship between working memory and listener response to hearing aids is not unique to WDRC but rather is related to the amount of signal modification caused by the hearing aid processing. In this context, signal modification refers to anything that substantially alters the available acoustic cues of the target signal. Under realistic listening conditions, the sources of signal modification include hearing aid processing, as well as fluctuating background noise that masks some portions of the target signal. This paper considers the relationship between working memory and response to signal modifications (frequency compression) applied to noisy speech.

Working memory (Daneman & Carpenter 1980) is a limited-capacity system that involves both storage and processing; that is, working memory allows a person to actively store task-related information while concurrently carrying out other relevant processing. The role of working memory in listeners' responses to hearing aid signal processing can be considered in the context of basic information processing models for speech (e.g., Pichora-Fuller et al. 1995; Pichora-Fuller 2003; Pichora-Fuller & Souza 2003; Wingfield et al. 2005; Rönnberg et al. 2008; Rossi-Katz & Arehart 2009; Lunner et al. 2009). These models suggest that processing resources are both finite and shared. In the case of degraded speech information, listeners may have to allocate a greater share of processing resources to the recovery of degraded information at the auditory periphery, leaving fewer resources available for successfully processing and identifying the linguistic content in the message. That is, listeners may have to rely more on working memory to process the degraded speech signal, and when the working memory is reduced, this processing may be more difficult. In the context of these models, signal degradation includes any type of signal modification that substantially alters the available acoustic cues of the target signal. In support of these models, experimental studies have shown a significant relationship between working memory and the intelligibility of speech that has been degraded by peripheral hearing loss (e.g., Cervera et al. 2009), by noise (e.g., Pichora-Fuller et al. 1995; Lunner 2003; Akeroyd

2008), by spectral reduction (Schvartz et al. 2008), and as noted above, by fast-acting WDRC (Foo et al. 2007; Lunner & Sundewall-Thorén 2007; Rudner et al. 2011; Piquado et al. 2012). These sources of degradation may occur singly or in combination. For example, younger normal-hearing listeners may encounter a single source of degradation such as spectral reduction (e.g., Schvartz et al. 2008) whereas older persons with hearing loss listening to noisy or time-compressed speech may encounter multiple sources of signal degradation (e.g., Pichora-Fuller et al. 1995; Jenstad & Souza 2007).

To explore the relationship between signal modification and working memory, the present study carried out a focused analysis of data reported in Souza et al. (2011) who studied the effects of frequency compression on the intelligibility of sentences presented in quiet and in babble for older listeners with hearing loss. The rationale for frequency compression is to improve intelligibility of high-frequency speech sounds by shifting them to lower-frequency regions where listeners with high-frequency hearing loss have better hearing thresholds.[1] However, frequency compression also introduces distortion by reducing spacing between harmonics, altering spectral peak levels, and modifying spectral shape (McDermott 2011). In addition, the presence of background noise distorts the speech envelope and temporal fine structure. The purpose of the present study was to determine the effects of cognitive ability (working memory) on individual listeners' responses to distortion caused by frequency compression applied to noisy speech.

METHODS

The following section provides a) an overview of the dataset from Souza et al. (2011) b) a description of factors and groupings used in the statistical analysis and c) a description of the methods used in the statistical analysis.

Overview of Souza et al. (2011)

Listeners—The present analysis focused on the participants in Souza et al. (2011) who had sensorineural hearing loss. Specifically, the participant group included 26 individuals ranging in age from 62 to 92 years. All listeners passed the Mini-Mental State Exam (MMSE) (Folstein et al. 1975), with a score of 26 or better. The data were collected both at Northwestern University and at University of Colorado, using identical equipment and procedures. The test procedures were reviewed and approved by the local Institutional Review Boards.

Stimuli—Stimuli consisted of low-context (IEEE) (Rosenthal 1969) sentences spoken by a female talker. In Souza et al. (2011), the sentences were presented in quiet, and with 8-talker babble noise taken from a recording of the Connected Speech Test (Cox et al. 1987) at a wide range of signal-to-noise ratios (SNRs) (10, 5, 0, −5, −10 dB SNR).

Signal Processing—Frequency compression was implemented using sinusoidal modeling (McAulay & Quatieri 1986). In the present implementation, the signal was divided into low-frequency and high-frequency bands using a complementary pair of recursive five-pole Butterworth filters. The low-frequency signal was used without further modification and sinusoidal modeling was applied to the high-frequency signal. The high-frequency signal was modeled using ten sinusoids, with the sinusoid frequency, amplitude, and phase computed for overlapping 6-msec signal blocks. The ten highest signal peaks were selected and the amplitude and phase of each peak were preserved while the frequencies were

---

[1]Frequency compression processing modifies speech information above a cutoff frequency (CF) using a particular compression ratio (CR).

reassigned to lower values. Output sinusoids were then synthesized at the shifted frequencies (McAulay & Quatieri 1986) and combined with the original low-frequency signal to produce the processed output. This strategy for frequency compression has been used successfully in previous research (Aguilera Muñoz et al. 1999). It is related to the strategy reported in Simpson et al. (2005), although that approach uses all of the FFT bins rather than just the peaks and might be expected to produce different amounts of nonlinear distortion in comparison to the approach used in this paper.

The frequency compression parameters included three frequency compression ratios (1.5:1, 2:1, and 3:1) and three frequency compression cutoffs (1, 1.5, and 2 kHz). The focus in this study was on listener response to distortion rather than an attempt to validate any specific implementation of frequency compression. A control condition (i.e., no frequency compression) was also included, for a total of 10 frequency compression conditions (3 compression ratios × 3 cutoff frequencies + control condition). The control condition included the low-pass and high-pass filter signal band separation, after which the two bands were recombined without the high-frequency sinusoidal modeling. As pointed out by Humes (2007), cognitive factors may become evident in speech recognition of older adults when speech signals are audible. Therefore, following frequency compression, the speech signals for all 26 listeners with hearing loss were amplified using the National Acoustics Laboratories-Revised (NAL-R) linear prescriptive formula (Byrne & Dillon 1986) based on individual thresholds, with the goal of compensating for reduced audibility caused by poorer-than-normal thresholds.

Stimulus Presentation and Response Task—Each listener was seated in a double-walled sound booth. The digitally stored stimuli were routed through a digital-to-analog converter (TDT RX6 or RX8), an attenuator (TDT PA5), and a headphone buffer (TDT HB7) and were presented monaurally to a listener's ear through a Sennheiser HD 25-1 earphone. The stimulus level prior to NAL-R amplification was 65 dB SPL.

Using a graphical user interface displayed on a computer screen, listeners controlled the timing of the stimulus trials using a computer mouse. On each trial, listeners heard a sentence, and then repeated back the sentence that was heard. Participants were first given sixty practice sentences, which were followed by 600 test sentences (10 sentences × 10 processing conditions × 6 noise levels [−10, −5, 0, 5, 10 dB SNR and quiet]. No feedback was provided. The presentation of the sentences was ordered randomly, and differed for each listener. Scoring of the sentences was based on key words correct (5 per sentence for 50 words per condition, per listener). For statistical analysis, the percent correct scores were transformed to rationalized arcsine units (RAU) (Studebaker 1985) to normalize variance across the range of scores.

Factors and Groupings used in Statistical Analysis

Signal Modification Metric—The Hearing Aid Sound Quality Index (HASQI) ((Kates and Arehart 2010; Arehart et al. 2010) was used to quantify the total amount of signal alteration caused by the frequency compression and by the additive babble noise. HASQI measures signal envelope and spectral fidelity in comparison with an undistorted reference signal. It returns a value between 0 and 1, with 1 representing perfect fidelity and 0 indicating very low fidelity. Additive noise, for example, will fill in the valleys of the speech signal. The noise changes the envelope peak-to-valley ratio and thus reduces the envelope correlation between the noisy signal and a clean reference. Frequency compression alters the spectral regions in which formants, formant transitions, and consonant onsets occur. These frequency shifts will cause changes in the signal envelope modulation within each auditory filter band such that an event that would have been concentrated in one band is moved to

another. The shifts will also change the relationship of the modulation across bands. Both of these effects will reduce the HASQI score for the frequency-compressed speech in comparison with an unprocessed reference. As shown in Table 1, HASQI values for total distortion for signals presented in quiet ranged from 0.253 to 1.000 depending on the frequency compression parameters. In the presence of noise, HASQI values ranged 0.006 to 0.418.

Working Memory—Working memory in the 26 listeners with hearing loss was measured using the Reading Span Test (RST) (Daneman & Carpenter 1980; Rönnberg et al. 1989). The RST was designed to capture individual variability in working memory capacity in terms of coordinating simultaneous storage and processing requirements. The participants were asked to recall, in correct serial order, the first or last words of a sequence of sentences shown on a computer screen. The participants were not told whether the first or last word would be prompted prior to seeing the sentences. The scores were based on the total proportion of first or last words correctly recalled, whether or not in correct serial order. Figure 1 shows the RST scores for the 26 participants. The scores ranged from 0.17 to 0.57. There was no significant correlation between age and RST score (Pearson correlation = −0.258; p=0.203). (See discussion for further consideration of this relationship).

Group Definition—The 26 listeners were grouped into high and low working memory groups, using the sample median of the RST scores as the cutoff criterion (median=0.37). The "High-RST" group had 14 listeners with RST scores of 0.37 and above and the "Low-RST" group" had 12 listeners with RST scores below 0.37. (The unequal numbers in the two groups is due to two listeners having RST values of 0.37.) Audiograms for the two groups are shown in Figure 2. The degree of hearing loss between the High-RST and Low-RST groups was comparable, as indicated by the similar audiograms between the two groups and a lack of significant difference between the two groups for both the four-frequency pure tone average (500 Hz, 1000 Hz, 2000 Hz, 4000 Hz) (t24=1.42, p=0.168) as well as for thresholds at 4 kHz (t24=1.47, p=0.154) The mean age of the listeners in the High-RST group was also not significantly different from the mean age of the listeners in the Low-RST group (t24=1.07, p=0.298).

Statistical Analysis

A two-fold approach to the statistical analysis was used. First, a conventional approach (repeated-measures analysis of variance [ANOVA]) was used to address the experimental question of interest: namely, whether response to frequency compression signal processing and babble noise differed for adults in the High-RST and Low-RST groups. Second, we used hierarchical linear regression models to examine intra-individual and inter-individual variability across the entire continuum of HASQI scores. On a general level, the hierarchical approach is a conceptual orientation to modeling individual variability (e.g, as a function of signal degradation and cognitive function) which considers individual listener characteristics as possible predictors. The hierarchical framework explicitly represents each listener's intelligibility scores as a function of person-specific parameters plus random error and describes the variation of these parameters across individuals (Raudenbush & Bryk 2002). An important distinction between the ANOVA and the hierarchical linear modeling is that the ANOVA treats the signal modifications due to frequency compression signal processing and to noise as separate categorical variables, whereas the hierarchical modeling approach allows us to consider total signal degradation in terms of a continuous metric.

Phase 1 Approach—A mixed repeated-measures ANOVA was used to determine the effects of signal degradation caused by both processing and noise on sentence intelligibility for the Low- RST and High-RST groups. For this analysis, we had one between-subject

factor (RST group) and two within-subject factors (signal processing and SNR). The processing factor had 10 levels, corresponding to each of the 10 processing conditions (cf Table 1). The SNR factor had six levels (quiet and 10, 5, 0, −5, −10 dB SNR.). As described above, listeners were divided into High-RST and Low-RST groups, with no significant differences between the RST groups in terms of age and hearing loss. However, because both age and hearing loss have been associated with degraded intelligibility (e.g., Dubno et al. 1984), the subjects' ages and thresholds at 4 kHz were also included in the ANOVA as covariates. Greenhouse-Geisser corrections were used when Mauchley's test of sphericity was significant. The specific research questions addressed by the Phase 1 analysis were as follows:

1. Do mean intelligibility scores differ between the High-RST and Low-RST groups, when controlling for age and hearing loss (4 kHz thresholds)?

2. Does the pattern of intelligibility scores across signal processing condition or SNR differ between the High-RST and the Low-RST groups, when controlling for age and hearing loss?

Phase 2 Approach—The second phase of the analysis used a multi-level model to describe how individual listeners differ in their response to signal distortion and to identify the variability in this response between subjects that could be explained by differences in working memory. In order to investigate this question, the total amount of signal alteration caused by frequency compression and by babble noise had to be quantified for all 60 conditions onto a continuous metric. The metric we used for this purpose was HASQI, which was described in detail above. A description of the multi-level model is outlined below using the terminology provided by Raudenbush and Bryk (2002).

The relationship between HASQI and RAU Intelligibility was determined to be monotonic and nonlinear. Following the guidelines from Keene (1995), we transformed HASQI using the natural log (ln) prior to the multi-level analysis (Figure 3). Intelligibility scores and the amount of distortion quantified using a logarithmic transform of HASQI were significantly correlated (Pearson correlation = 0.928, p=0.001).

A multi-level model can be specified in a hierarchical fashion where there are two sources of variation: variation within an individual (intra-subject variance) and variation between individuals (inter-subject variance). In the first stage of the analysis, known as level-1, the within-subject variability in intelligibility as a function of HASQI is estimated and the relationship between intelligibility and HASQI for each listener is characterized. The level-1 model was written according to Eq. (1):

$$Y = P_0 + P_1{}^* \ln(HASQI) + e_1 \quad (1)$$

where Y represents the intelligibility score in RAUs, P_0 represents an intercept, P_1 represents a slope, and e_1 corresponds to the within-subject residual variance. Therefore, the intercept, P_0, represents the expected RAU intelligibility when ln($HASQI$) equals zero. Since the ln($HASQI$) is equal to zero when HASQI is at a value of 1, the intercept in this model can be interpreted as the subject's baseline intelligibility.

In the second stage of the analysis, known as level-2, we evaluated the presence of inter-subject variability in the intercept and slope estimates of level-1. Specifically, we assessed whether the relationship between ln($HASQI$) and intelligibility varied across RST groups by using the level-1 coefficients as outcome variables. The level-2 submodels were written as

$$P_0 = B_{00} + B_{01}{}^*(RST\ Group) + B_{02}{}^*(\text{age}) + B_{03}{}^*(4\ kHz\ threshold) + r_0 \quad (2)$$

$$P_1 = B_{10} + B_{11}{}^*(RST\ Group) + B_{12}{}^*(\text{age}) + B_{13}{}^*(4\ kHz\ threshold) + r_1 \quad (3)$$

The two-level model can be written in terms of a composite model by combining the level-1 and level-2 submodels as follows:

$$
\begin{aligned}
Y = {} & B_{00} + B_{01}{}^*(RST\ Group) \\
& + B_{02}{}^*(\text{Age}) \\
& + B_{03}{}^*(4\ kHz\ threshold) \\
& + B_{10}{}^*[\ln(HASQI)] + B_{11}{}^*(RST\ Group)^*[\ln(HASQI)] \quad (4) \\
& + B_{12}{}^*(\text{age})^*[\ln(HASQI)] \\
& + B_{13}{}^*(4\ kHz\ threshold)^*[\ln(HASQI)] \\
& + r_0 + r_1{}^*[\ln(HASQI)] + e_1
\end{aligned}
$$

The amount of variability explained by the level-2 variables was determined by comparing the error terms in an unconditional base model (no predictor variables at level-2) with the error terms in a conditional model (one or more predictor variables at level-2) as suggested by Kreft and de Leeuw (1998) and Singer (1998). For the current analysis we explored three conditional models beginning with a model that included RST group alone followed by the addition of each covariate.

Multilevel models provide a statistical framework specifically designed to address research questions related to hierarchical data structures (Raudenbush & Bryk 2002). This framework is particularly suited for the current analysis because it allows us to formulate and test hypotheses about how variables at one level (e.g. RST at level-2) affect the relations occurring at another (e.g. the relation of HASQI and intelligibility). Thus, similar to the ANOVA analysis, the multi-level analysis addressed whether there were differences between the High-RST and Low-RST groups and whether there was an interaction between RST groups and amount of signal distortion. In addition, multilevel models allow us to partition the variance components that occur within and between subjects. As such, this type of analysis addressed the following research questions:

1. How much variation in the intercepts and the slopes is explained by using RST group as a predictor, with age and hearing loss (4 kHz threshold) as covariates?

2. Does baseline intelligibility (P_0) differ between the High-RST and Low-RST groups, when controlling for age and hearing loss (4 kHz threshold)?

3. Does the strength of association between HASQI and intelligibility (P_1) differ between the High-RST group and the Low-RST groups, when controlling for age and hearing loss (4 kHz threshold)?

RESULTS

Phase 1: Mixed Repeated Measures ANOVA

Figure 4 shows the data submitted to the ANOVA: intelligibility scores (in RAU units) for all ten levels of processing for the High-RST and Low-RST groups for each of the six SNR conditions. The frequency compression processing conditions are labeled as conditions 1 through 10 based on a rank ordering of the HASQI values for the processing conditions in

quiet (see Table 1). The results of the mixed repeated-measures ANOVA are shown in Table 2.

The factor of RST group provides insight into whether listeners in the High-RST group have different mean intelligibility scores than listeners in the Low-RST group. Notably, the effect of RST was significant *after* controlling for age and hearing loss as covariates, such that listeners in the High-RST group have higher intelligibility scores than listeners in the Low-RST group.

To assess whether the influence of RST group on intelligibility (controlling for age and hearing loss) depended upon signal processing condition or SNR, we examined interactions between RST, signal processing, and SNR. While the intelligibility scores of listeners in both RST groups were adversely affected by frequency compression processing and by noise, the lack of significant interactions indicates that the pattern of differences between mean intelligibility for the RST groups did not change across processing conditions or across SNR.

The ANOVA shows that RST-group is a significant factor in the intelligibility scores. However, the analysis is limited in that it treats the frequency compression processing and noise as categorical variables and does not allow us to quantify the magnitude of distortion caused by the cumulative effects of the frequency compression processing and of the babble noise. The Phase 2 analysis explored the relationship between the intelligibility scores in the two RST groups and the total amount of signal degradation quantified using a continuous scale.

Phase 2: Two-Level Linear Model

The amount of variation in the intercepts and slopes explained by using RST group as a predictor, with age and 4 kHz threshold as covariates is summarized in Table 3. The model building process for level-2 revealed that the independent variability explained in baseline intelligibility (intercept) by RST group was 29.3%. When age was added to the model, it explained an additional 11.5%, and lastly, when 4 kHz threshold was added to the model, it explained an additional 6.7% of the variability in baseline intelligibility. Thus, for the intercept, a total of 47.5% of the variance was explained when considering RST group as a predictor and age and 4 kHz threshold as covariates.

We did not observe the same pattern in terms of slope. The independent variability explained in slope by RST group was 21.3%. However, the addition of age and 4 kHz threshold decreased the amount of variability explained. For this reason, age and 4 kHz threshold were not included in the final level-2 model for slope. The exclusion of these variables at this step did not change the effect of RST group. The adjusted Eq. (3) was rewritten as Eq. (5):

$$P_1 = B_{10} + B_{11}{}^*(RST\ Group) + r_1 \quad (5)$$

The adjusted two-level model can be rewritten as follows:

$$\begin{aligned} Y = &\ B_{00} + B_{01}{}^*(RST\ Group) \\ &+ B_{02}{}^*(Age) \\ &+ B_{03}{}^*(4\ kHz\ threshold) + B_{10}{}^*[\ln(HASQI)] \quad (6) \\ &+ B_{11}{}^*(RST\ Group){}^*[\ln(HASQI)] \\ &+ r_0 + r_1{}^*[\ln(HASQI)] + e_1 \end{aligned}$$

The composite form of the adjusted multi-level model illustrates that intelligibility (Y) can be viewed as a function of the overall intercept (B_{00}), the main effect of RST group (B_{01}), the covariate effect of age (B_{02}), the covariate effect of 4 kHz threshold (B_{03}), and one cross-level interaction (B_{11}) involving RST group with ln ($HASQI$).

The analysis of the adjusted multi-level model yielded the following results, which are summarized in Table 4. The High-RST group had a significantly higher baseline intelligibility than did listeners in the Low-RST group, controlling for age and 4 kHz threshold (B_{01}=13.02; t=3.36). More specifically, listeners in the High-RST group had an estimated initial intelligibility score that was 13.02 RAU units higher than listeners in the Low-RST group. We also noted that the covariates age and 4 kHz threshold were negatively related to baseline intelligibility (B_{02}=−0.43, t=4.12; B_{03}=−0.20, t=3.0). This relationship was not unexpected as these covariates are known predictors of intelligibility (Dubno et al. 1984; Divenyi et al. 2005).

The significant interaction between RST group and ln (HASQI) provides insight into whether the strength of association between HASQI and intelligibility (P_I) differs between the High-RST group and the Low-RST group, when controlling for age and for the 4 kHz threshold. There was a tendency for listeners in the High-RST group to have higher intelligibility scores as HASQI approached zero (B_{11}=2.09, t=2.36). In other words, as the total amount of signal distortion increased, listeners in the High-RST group were less affected by the distortions than the listeners in the Low-RST group.

Figure 5 is a graphical representation of the final multi-level model. It specifies four different prototypical trajectories for intelligibility as a function of total distortion beginning with a trajectory for a listener with the following characteristics: High-RST, age 65 years, and a 4 kHz threshold of 55 dB HL. The subsequent trajectory represents the effect of RST from high to low. It is followed by the additional effect of hearing loss with a change in the 4 kHz threshold of 55 to 75 dB HL. The final trajectory demonstrates the combined effects of low RST, increased hearing loss, and an increase in age from 65 to 80 years.

DISCUSSION

The current analysis showed that working memory, as measured by high or low RST scores, was a significant factor in listeners' responses to sentences presented in babble noise and processed with one form of frequency compression. Specifically, the present study showed that at maximum signal modification (caused by both frequency compression and babble noise), the factor of working memory (when controlling for age and hearing loss) accounted for 29.3% of the variance in intelligibility scores. A model combining working memory, age, and hearing loss accounted for a total of 47.5% of the variability in intelligibility scores. Furthermore, as the total amount of signal distortion increased, listeners with higher working memory performed better on the intelligibility task than listeners with lower working memory.

Signal Distortion and Working Memory

Recent studies have suggested a relationship between listeners' responses to distortion caused by hearing aid signal processing and working memory (e.g., Rönnberg et al. 2008; Lunner et al. 2009). In support of this suggestion, several studies (e.g., Lunner & Sundewall-Thorén 2007; Rudner et al. 2011) have shown that relative benefit from fast-acting WDRC (which causes more distortion than slow-acting WDRC) is reduced in listeners with poor working memory. While evidence shows that lower cognition is often associated with response to WDRC settings, the nature of this association may depend on acclimatization and sentence materials (Foo et al. 2007; Cox & Xu 2010; Rudner et al. 2011).

Based on these data, it was proposed that listeners with poor working memory capacity are less able to adapt to rapid changes to the signal. The results of the current study suggest that the relationship between response to distortion and working memory is not limited to WDRC but also extends to distortions caused by the form of frequency compression implemented here. Together, the findings from WDRC and frequency compression suggest that older listeners with hearing loss and poor working memory are quite susceptible to distortions caused by hearing aid signal processing algorithms applied to noisy speech.

A unique feature of the present study is that we related working memory to the cumulative effects of distortion caused by frequency compression and noise. Total signal distortion was quantified along a continuum using the HASQI speech quality metric. There are other objective metrics that also could have been used. These approaches include the Perceptual Evaluation of Speech Quality (PESQ) (Beerends et al. 2002), which measures changes in the signal loudness, the log-likelihood ratio of the linear prediction coefficients (Hu & Loizou 2008), cepstral coefficients (Hu & Loizou 2008), the PEMO-Q measure of envelope fluctuation (Huber & Kollmeier 2006), and the Hearing Aid Speech Quality Index (HASQI) (Kates & Arehart 2010; Arehart et al. 2010). The advantage of HASQI is that it responds to the short-time variation in the log magnitude spectrum as well as to changes in the signal envelope. Thus, time-frequency modifications to the signal caused by frequency compression, such as changes in formant transitions, will be reflected in a reduced HASQI value. The accuracy of HASQI, in comparison to other metrics, in predicting the participants' responses to frequency compression was not evaluated in this study; HASQI was used as a measure of signal distortion and not as a quality predictor.

Contributions of working memory, age and hearing loss

The present results are consistent with the idea that distortion from both the babble noise and from the frequency compression implemented in this paper, as well as hearing loss, are all sources of signal degradation that contribute to an impoverished representation at the auditory periphery. In the context of the information processing models for speech (e.g., Pichora-Fuller et al. 1995; Rönnberg et al. 2008; Rossi-Katz & Arehart 2009; Lunner et al. 2009), listeners experiencing these multiple sources of signal degradation will need to allocate more processing resources to earlier processing stages. This allocation of resources may then place a drain on later operations necessary for identifying the linguistic content of the sentence materials. That is, listeners may have to rely more on working memory to process the degraded speech signal, and when the working memory is reduced, this processing may be more difficult.

The finding that age – above and beyond the effects of working memory – contributed to variance in intelligibility scores suggests that the listeners in the present study may have age-related degradations in higher-level processing that extend beyond what is captured in the RST implemented in the present study. For example, Schvartz et al. (2008) showed that older listeners' ability to process distorted speech was related to both working memory and speed of processing. Current studies in our laboratories are considering the contribution of other higher-level processing abilities to variability in hearing-aid response by older adults.

Although not the focus of the study, an examination of subject characteristics indicated that working memory did not decline with age. A number of other studies have shown an relationship between age and working memory, but those studies generally compared a very young (20s or 30s) group to a single older group (e.g., Reuter-Lorenz et al. 2000; Cabeza et al. 2004; Bopp & Verhaeghen 2005; Waters & Caplan 2005; Hale et al. 2007; Brehmer et al. 2012). In contrast, our study examined working memory only within an older cohort with hearing loss, which may contribute to why working memory does not decline in an obvious way within our group. Indeed, the variation in working memory and the demonstrated

relationship between working memory and response to signal processing suggest that decisions based solely on age may not lead to patient-appropriate choices.

Finally, the present study considered susceptibility to frequency compression distortion without extended listening experience with the signal processing algorithm. While the listeners included in this analysis were familiarized with frequency compressed sentences, the listeners did not receive exposure to frequency compression over extended periods of time (weeks or months). It is possible that acclimatization to the frequency compression over time (Wolfe et al. 2011; Simpson 2009; Rudner et al. 2011) might affect listeners' susceptibility to total signal distortion.

Implications

The results of this study showed that distortions caused by one form of frequency compression (sinusoidal modeling) can negatively impact laboratory-based measures of intelligibility of noisy speech, and that this impact is greater in older listeners with hearing loss who have poor working memory. The results support the idea that both stimulus related and subject related factors may have cumulative negative impacts on listener intelligibility. At the signal level, both noise and signal processing may degrade the signal in ways that impact intelligibility. At the subject level, working memory, age and hearing loss are also important factors in individual listeners' responses to hearing aid processing of noisy speech. Future work should consider these factors in broader clinical contexts, including hearing aids that implement other forms of frequency compression and that consider possible acclimatization of listeners to different forms of processing.

Acknowledgments

REFERENCES

Aguilera Muñoz, CM.; Nelson, PB.; Rutledge, JC., et al. Frequency lowering processing for listeners with significant hearing loss. Paper presented at the International Conference on Electronics, Circuits and Systems; Pafos, Cypress; 1999.

Akeroyd M. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. Int J Audiol. 2008; 47:S53–S71. [PubMed: 19012113]

Arehart KH, Kates JM, Anderson MC. Effects of noise, nonlinear processing, and linear filtering on perceived speech quality. Ear Hear. 2010; 31:420–436. [PubMed: 20440116]

Beerends JG, Hekstra AP, Rix AW, et al. Perceptual Evaluation of Speech Quality (PESQ) the new ITU standard for end-to-end speech quality assessment: Part II - Psychoacoustic model. J Audio Eng Soc. 2002; 50:765–778.

Bopp K, Verhaeghen P. Aging and verbal memory span: A meta-analysis. J Gerontology Psychol Sci. 2005; 60:223–233.

Brehmer Y, Westerberg H, Backman L. Working-memory training in younger and older adults: Training gains, transfer, and maintenance. Front Hum Neurosci. 2012; 6:1–7. [PubMed: 22279433]

Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear. 1986; 7:257–265. [PubMed: 3743918]

Cabeza R, Daselaar S, Dolcos F, et al. Task-independent and task-specific age effects on brain activity during working memory, visual attention and episodic retrieval. Cerebral Cortex. 2004; 14:364–375. [PubMed: 15028641]

Cervera TC, Soler MJ, Dasi C, et al. Speech recognition and working memory capacity in young-elderly listeners: effects of hearing sensitivity. Can J Exp Psychol. 2009; 63:216–226. [PubMed: 19739905]

Cox RM, Xu J. Short and long compression release times: speech understanding, real-world preferences, and association with cognitive ability. J Am Acad Audiol. 2010; 21:121–138. [PubMed: 20166313]

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19:450–466.

Davies-Venn E, Souza P, Brennan M, et al. Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing loss. Ear Hear. 2009; 30:494–504. [PubMed: 19633563]

Divenyi PL, Stark PB, Haupt KM. Decline of speech understanding and auditory thresholds in the elderly. J Acoust Soc Am. 2005; 118:1089–1100. [PubMed: 16158663]

Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am. 1984; 76:87–96. [PubMed: 6747116]

Folstein MF, Folstein SE, McHugh PR. Mini-Mental State – Practical method for grading cognitive state of patients for clinician. J Psychiatr Res. 1975; 12:189–198. [PubMed: 1202204]

Foo C, Rudner M, Rönnberg J, et al. Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. J Am Acad Audiol. 2007; 18:618–631. [PubMed: 18236648]

Gatehouse S, Naylor G, Elberling C. Benefits from hearing aids in relation to the interaction between the user and the environment. Int J Audiol. 2003; 42(Suppl 1):S77–S85. [PubMed: 12918613]

Gatehouse S, Naylor G, Elberling C. Linear and nonlinear hearing aid fittings-- 1. Patterns of benefit. Int J Audiol. 2006a; 45:130–152. [PubMed: 16579490]

Gatehouse S, Naylor G, Elberling C. Linear and nonlinear hearing aid fittings-- 2. Patterns of candidature. Int J Audiol. 2006b; 45:153–171. [PubMed: 16579491]

Hale, S.; Myerson, J.; Emery, L., et al. Variation in working memory across the lifespan. In: Conway, ARA.; Jarrold, C.; Kane, MJ.; Miyake, A.; Towse, JN., editors. Variation in Working Memory. New York: Oxford University Press; 2007. p. 194-224.

Hu Y, Loizou P. Evaluation of objective quality measures for speech enhancement. IEEE Trans Audio Speech and Lang Proc. 2008; 16:229–238.

Huber R, Kollmeier B. PEMO-Q – A new method for objective audio quality assessment using a model of auditory perception. IEEE Trans Audio, Speech, and Lang Proc. 2006; 14:1902–1911.

Humes LE. The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. J Am Acad Audiol. 2007; 18:590–603. [PubMed: 18236646]

Jenstad LM, Souza PE. Temporal envelope changes of compression and speech rate: Combined effects on recognition for older adults. J Speech Lang Hear Res. 2007; 50:1123–1138. [PubMed: 17905900]

Kates JM, Arehart KH. The Hearing-Aid Speech Quality Index (HASQI). J Audio Eng Soc. 2010; 58:363–381.

Keene ON. The log transformation is special. Statist Med. 1995; 14:811–819.

Kreft, I.; de Leeuw, J. Introducing multilevel modeling. London: Sage Publications; 1998.

Lunner T. Cognitive function in relation to hearing aid use. I Int J Audiol. 2003; 42:S49–S58.

Lunner T, Rudner M, Rönnberg J. Cognition and hearing aids. Scand J Psychol. 2009; 50:395–403. [PubMed: 19778387]

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; 18:604–617. [PubMed: 18236647]

McAulay RJ, Quatieri TF. Speech analysis synthesis based on a sinusoidal representation. IEEE Trans Audio Speech Lang Processing. 1986; 34:744–754.

McDermott HJ. A technical comparison of digital frequency-lowering algorithms available in two current hearing aids. PLoS One. 2011; 6(7):e22358. [PubMed: 21789254]

Pichora-Fuller MK. Cognitive aging and auditory information processing. Int J Audiol. 2003; 42:S26–S32.

Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. J Acoust Soc Am. 1995; 97:593–608. [PubMed: 7860836]

Pichora-Fuller MK, Souza PE. Effects of aging on auditory processing of speech. Int J Audiol. 2003; 42:S11–S16.

Piquado T, Benichov JI, Brownell H, Wingfield A. The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. Int J Audiol. 2012; 51:576–583. [PubMed: 22731919]

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2nd ed.. Newbury Park, CA: Sage Publications; 2002.

Reuter-Lorenz P, Jonides J, Smith E, et al. Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. J Cogn Neurosci. 2000; 12:174–187. [PubMed: 10769314]

Ronnberg J, Rudner M, Foo C, et al. Cognition counts: A working memory system for ease of language understanding (ELU). Int J Audiol. 2008; 47:S99–S105. [PubMed: 19012117]

Rönnberg J, Arlinger S, Lyxell B, Kinnefors C. Visual evoked potentials: relation to adult speech reading and cognitive function. J Speech Lang Hear Res. 1989; 32:725–735.

Rosenthal S. IEEE: Recommended practices for speech quality measurements. IEEE Trans Acoust. 1969; 17:227–246.

Rossi-Katz J, Arehart KH. Message and talker identification in older adults: Effects of task, distinctiveness of the talkers' voices, and meaningfulness of the competing message. J Speech Lang Hear Res. 2009; 52:435–453. [PubMed: 19064902]

Rudner M, Foo C, Rönnberg J, et al. Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. Scand J Psychol. 2009; 50:405–418. [PubMed: 19778388]

Rudner M, Rönnberg J, Lunner T. Working memory supports listening in noise for persons with hearing impairment. J Am Acad Audiol. 2011; 22:156–167. [PubMed: 21545768]

Schvartz KC, Chatterjee M, Gordon-Salant S. Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. J Acoust Soc Am. 2008; 124:3972–3988. [PubMed: 19206821]

Simpson A. Frequency-lowering devices for managing high-frequency hearing loss: A review. Trends Amplif. 2009; 13:87–106. [PubMed: 19447764]

Simpson A, Hersbach AA, McDermott HJ. Improvements in speech perception with an experimental nonlinear frequency compression hearing device. Int J Audiol. 2005; 44:281–292. [PubMed: 16028791]

Singer J. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. J Educ Behav Stat. 1998; 24:323–355.

Souza PE, Arehart KH, Kates JM, et al. Effects of frequency compression on the intelligibility and quality of speech in noise. J Acoust Soc Am. 2011; 129:2655.

Studebaker GA. A "rationalized" arcsine transform. J Speech Hear Res. 1985; 28:455–462. [PubMed: 4046587]

Waters G, Caplan D. The relationship between age, processing speed, working memory capacity, and language comprehension. Memory. 2005; 13(3–4):403–413. [PubMed: 15952262]

Wingfield A, Tun PA, McCoy SL. Hearing loss in older adulthood - What it is and how it interacts with cognitive performance. Curr Dir Psychol. 2005; 14:144–148.

Wolfe J, John A, Schafer E, et al. Long-term effects of non-linear frequency compression for children with moderate hearing loss. Int J Audiol. 2011; 50:396–404. [PubMed: 21599615]

This study considers the relationship between working memory and intelligibility of noisy speech subjected to frequency compression. Results show that intelligibility scores for listeners with poor working memory are degraded more by signal distortions caused by frequency compression and noise compared to listeners with good working memory. These results suggest that older listeners with hearing loss and poor working memory are more susceptible to distortions caused by at least some signal processing algorithms and that this increased susceptibility should be considered in the hearing-aid fitting process.
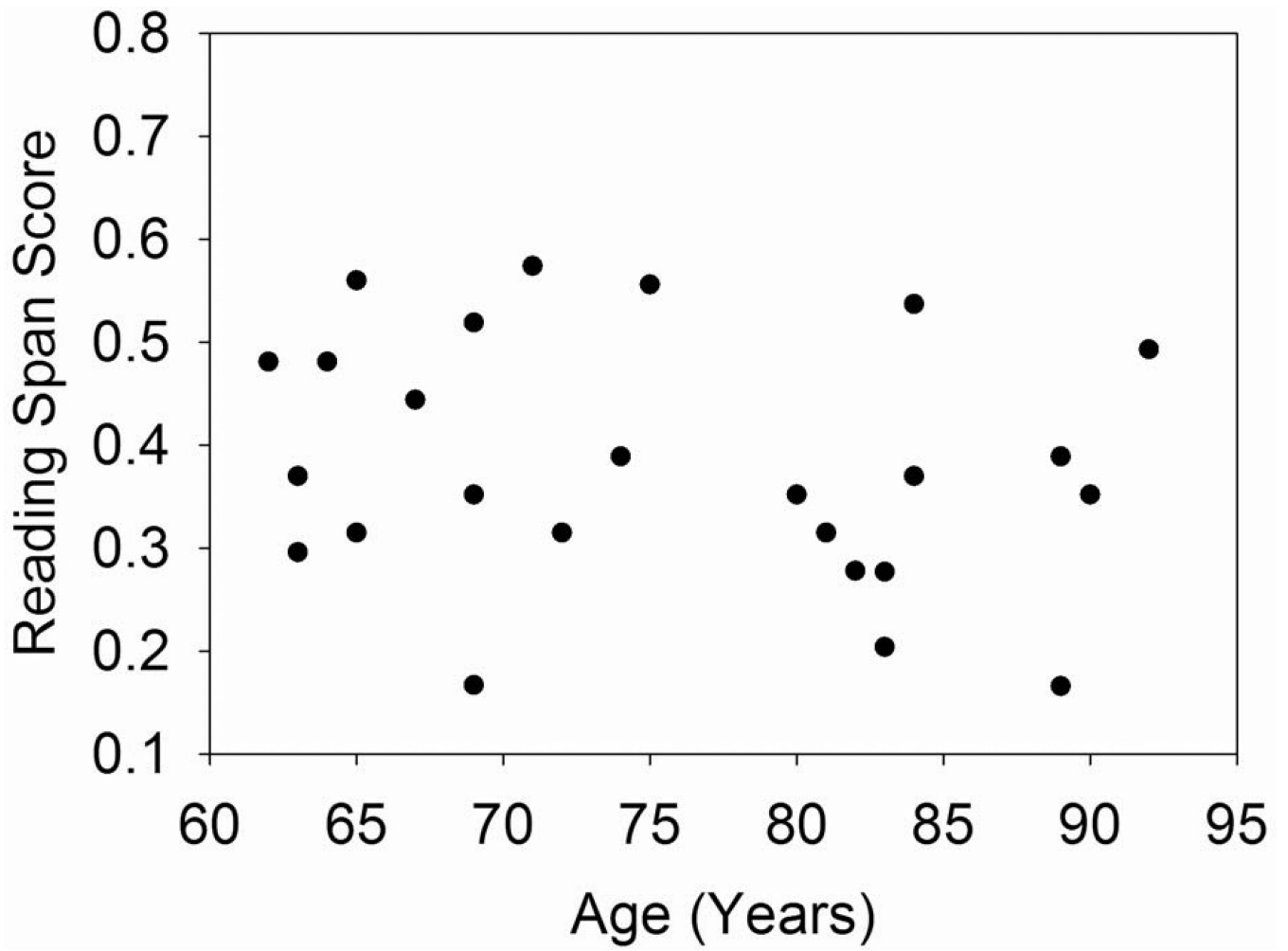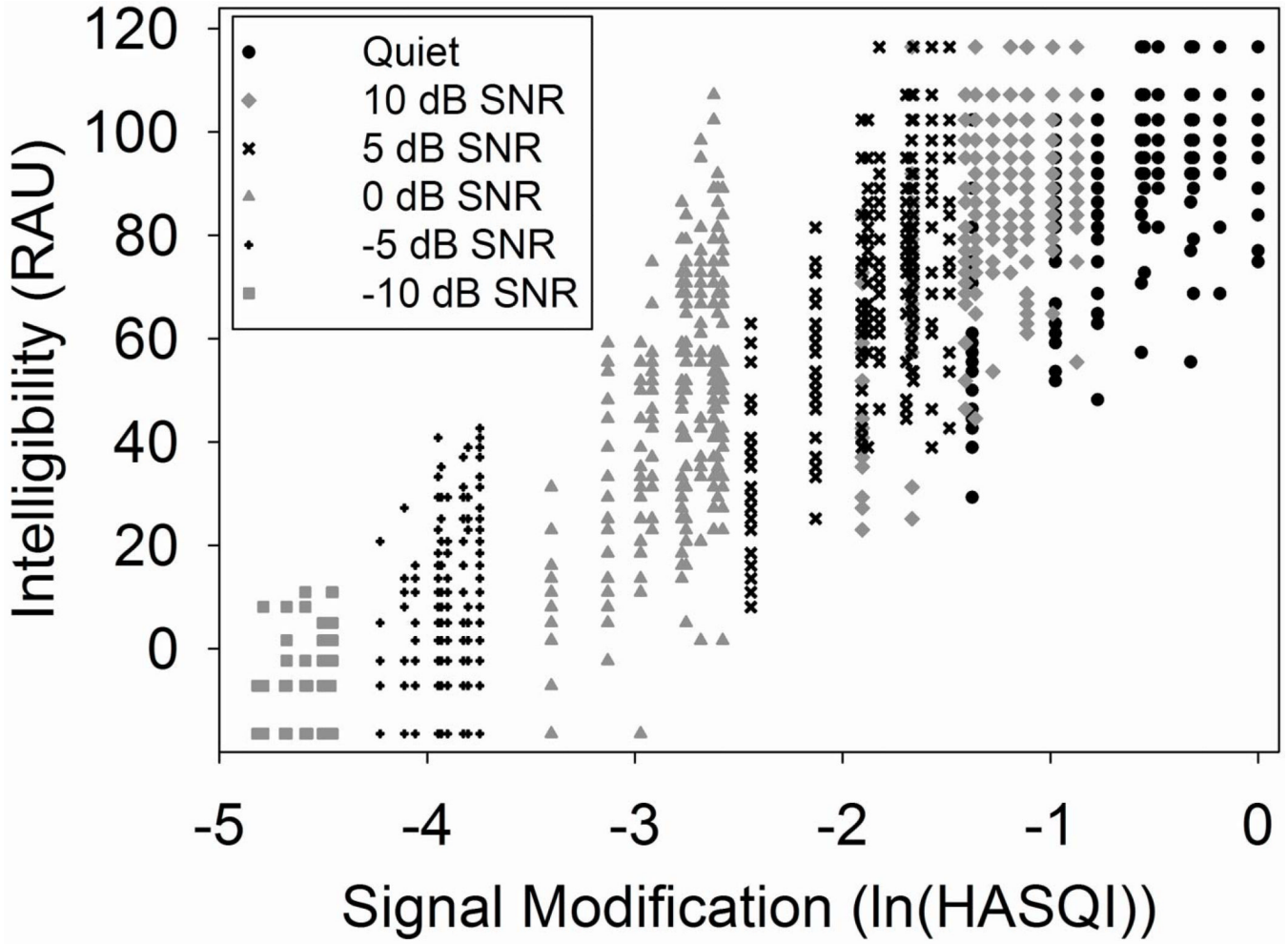
Figure 1.
Reading span test (RST) scores as a function of age are shown for the 26 older listeners with hearing loss.
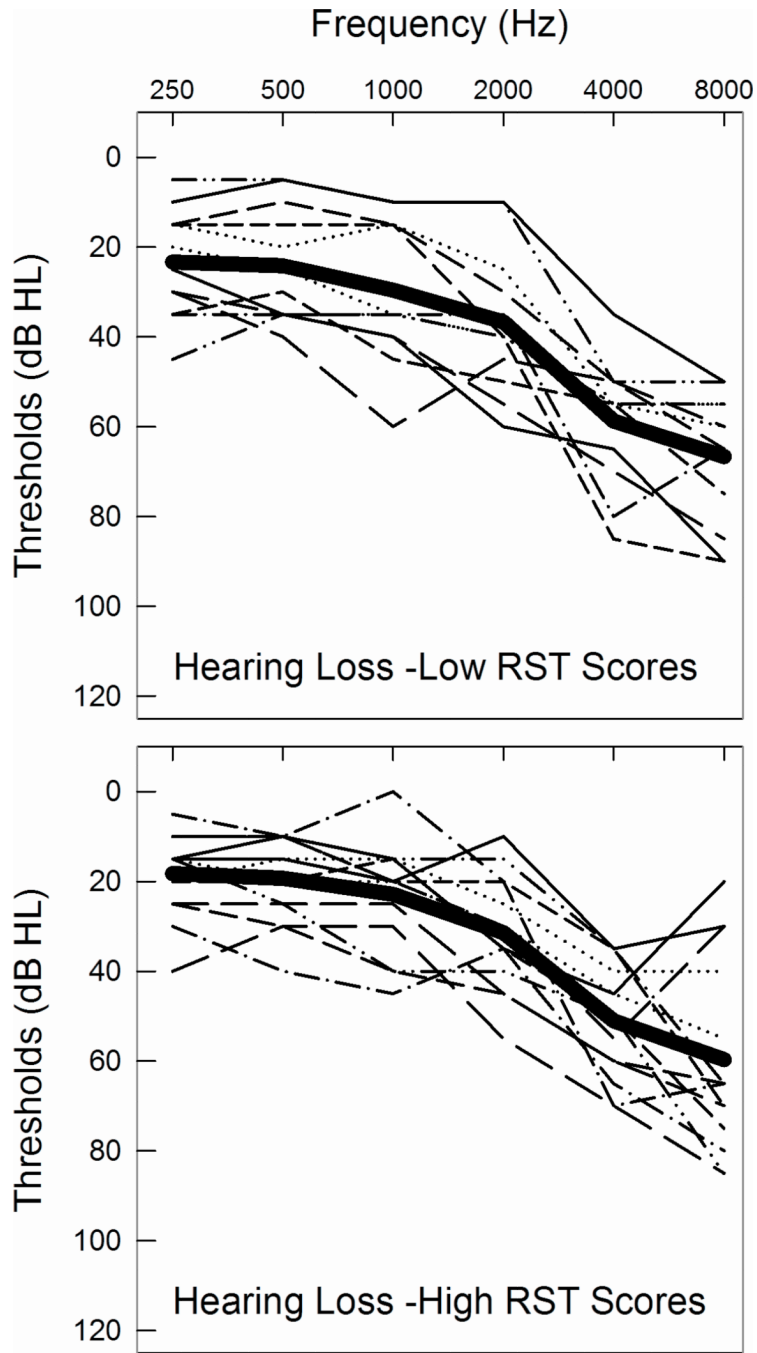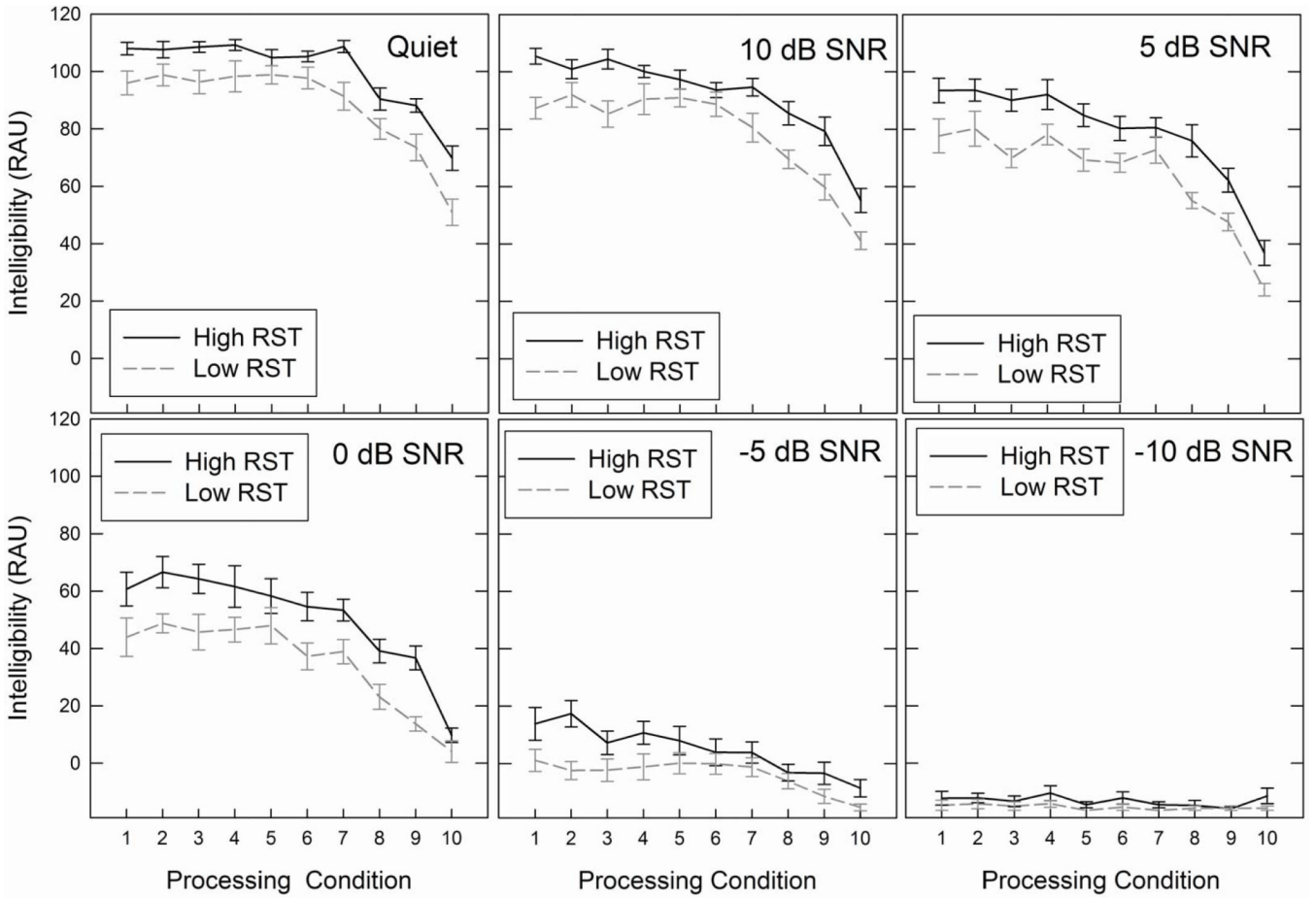
Figure 2.
Audiograms are shown for the participants in the Low-RST group (top panel) and for the High-RST scores (bottom panel). The average audiograms for the two groups are shown with heavy lines.
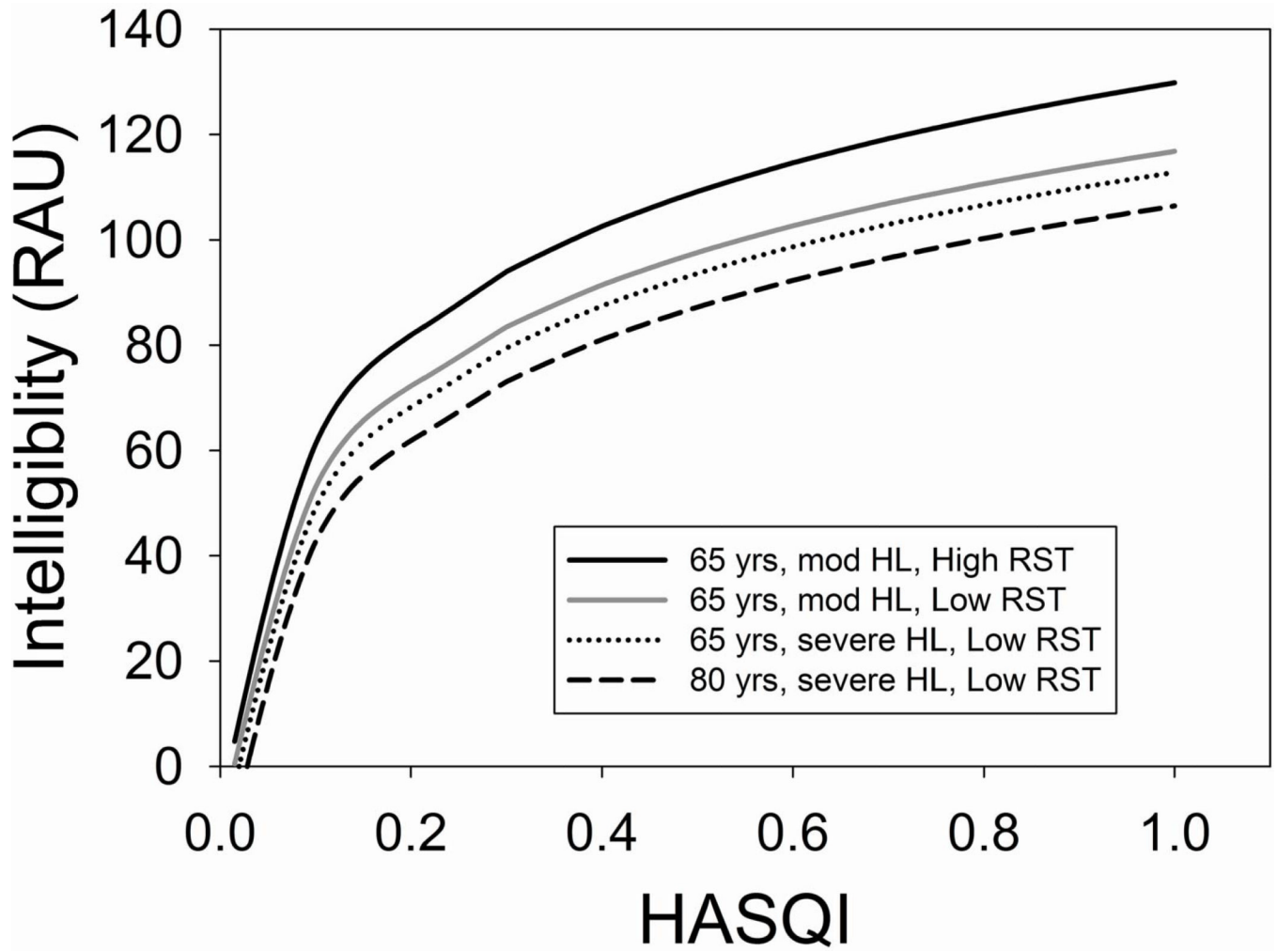
Figure 3.
Intelligibility scores (in RAU units) for all 26 listeners are shown as a function of the HASQI values on a log-transformed scale. Scores are shown for sentences presented in quiet, 10 dB SNR, 5 dB SNR, 0 dB SNR, −5 dB SNR and −10 dB SNR. Overlaps in scores preclude clear delineation of all 26 subject data points at all of the HASQI values.

Figure 4.
Intelligibility scores (in RAU units, with standard errors) are shown for all ten levels of processing for the High-RST and Low-RST groups for each of the six SNR conditions. The frequency compression processing conditions are labeled as conditions 1 through 10 based on a rank ordering of the HASQI values for the processing conditions in quiet (see Table 1).

Figure 5.
Final multi-level model graphs for different prototypical trajectories for intelligibility as a function of total distortion for listeners with the following characteristics: 1) High-RST, age 65 years, and a 4 kHz threshold of 55 dB HL; 2) Low-RST, age 65 years, and a 4 kHz threshold of 55 dB HL; 3) Low-RST, age 65 years, and a 4 kHz threshold of 75 dB HL and 4) Low-RST, age 80 years, and a 4 kHz threshold of 75 dB HL.

Table 1

Total distortion (as measured by HASQI) for the 10 stimulus conditions in quiet, including no processing and nine frequency compression conditions.

Condition #	Frequency Cutoff (Hz)	Compression Ratio	HASQI
1	No Process	No Process	1.000
2	2000	1.5	0.833
3	1500	1.5	0.733
4	2000	2	0.723
5	2000	3	0.618
6	1500	2	0.578
7	1000	1.5	0.570
8	1500	3	0.462
9	1000	2	0.377
10	1000	3	0.253

Table 2

Results of repeated measures ANOVA, including the between subject factor of RST group and the within subject factors of frequency compression processing and SNR.

Effect	df	F	p-value	Partial η^2
Between Subjects				
Reading Span Test (RST) Group	1, 22	7.1	0.014*	0.244
Within Subjects				
Signal-to-noise Ratio (SNR)	2.6, 57.9	37.4	<0.001*	0.629
SNR × RST Group	2.6, 57.9	2.6	0.070	0.105
Signal Processing (SP)	9, 198	13.9	<0.001*	0.387
SP × RST Group	9, 198	1.3	0.266	0.054
SP × RST Group × SNR	12.3, 271.3	1.1	0.333	0.049

Significant effects (p<0.05) are indicated with an asterisk (*).

Table 3

Variance explained by three conditional models that included the factor of RST Group, with and without the covariates (age and hearing loss).

	Independent Variable (*p-value*)	Covariate Variables (*p-value*)	Explained Variance
Intercept			
Model A	RST Group (*p*=0.003)		29.3%
Model B	RST Group (*p*=0.007)	Age (*p*=0.009)	40.8%
Model C	RST Group (*p*=0.011)	Age (*p*=0.056) 4 kHz Threshold (*p*=0.076)	47.5%
Slope			
Model A	RST Group (*p*=0.027)		21.3%
Model B	RST Group (*p*=0.691)	Age (*p*=0.003)	15.5%
Model C	RST Group (*p*=0.054)	Age (*p*=0.920) 4 kHz Threshold (*p*=0.371)	14.9%

Table 4

Coefficient estimations for the final two-level model.

Fixed Effects	Coefficient	se	T Ratio	p-value
Model for baseline Intelligibility				
Intercept (B_{00})	155.45	7.76	20.12	0.000 [*]
RST Group (B_{01})	13.02	4.25	3.36	0.006 [*]
Age (B_{02})	−0.43	0.10	−4.12	0.001 [*]
4kHz (B_{03})	−0.20	0.07	−3.00	0.007 [*]
Model for Intelligibility slopes				
Intercept (B_{10})	27.69	0.75	36.77	0.000 [*]
RST GROUP (B_{11})	2.09	0.88	2.36	0.027 [*]

Random Effects	Variance Component	Df	Chi-sq	p-value
Level-2 Residual Variance Intercept (r_0)	107.33	22	235.18	0.00 [*]
Level-2 Residual Variance Slope (r_1)	3.77	22	88.59	0.008
Level-1 Residual Variance (e_1)	182.10			

Significant effects (p<0.05) are indicated with an asterisk (*).