Spectrotemporal Modulation Sensitivity as a Predictor of Speech-Reception Performance in Noise With Hearing Aids

Joshua G. W. Bernstein[1], Henrik Danielsson[2], Mathias Hällgren[2,3], Stefan Stenfelt[2,3], Jerker Rönnberg[2], and Thomas Lunner[2,4]

Abstract

The audiogram predicts <30% of the variance in speech-reception thresholds (SRTs) for hearing-impaired (HI) listeners fitted with individualized frequency-dependent gain. The remaining variance could reflect suprathreshold distortion in the auditory pathways or nonauditory factors such as cognitive processing. The relationship between a measure of suprathreshold auditory function—spectrotemporal modulation (STM) sensitivity—and SRTs in noise was examined for 154 HI listeners fitted with individualized frequency-specific gain. SRTs were measured for 65-dB SPL sentences presented in speech-weighted noise or four-talker babble to an individually programmed master hearing aid, with the output of an ear-simulating coupler played through insert earphones. Modulation-depth detection thresholds were measured over headphones for STM (2cycles/octave density, 4-Hz rate) applied to an 85-dB SPL, 2-kHz lowpass-filtered pink-noise carrier. SRTs were correlated with both the high-frequency (2–6 kHz) pure-tone average (HFA; $R^2 = .31$) and STM sensitivity ($R^2 = .28$). Combined with the HFA, STM sensitivity significantly improved the SRT prediction ($\Delta R^2 = .13$; total $R^2 = .44$). The remaining unaccounted variance might be attributable to variability in cognitive function and other dimensions of suprathreshold distortion. STM sensitivity was most critical in predicting SRTs for listeners < 65 years old or with HFA <53 dB HL. Results are discussed in the context of previous work suggesting that STM sensitivity for low rates and low-frequency carriers is impaired by a reduced ability to use temporal fine-structure information to detect dynamic spectra. STM detection is a fast test of suprathreshold auditory function for frequencies <2 kHz that complements the HFA to predict variability in hearing-aid outcomes for speech perception in noise.

Keywords

hearing aids, amplitude modulation, cognitive processing, temporal fine structure, noise

Introduction

Difficulty in understanding speech in noisy environments is one of the chief complaints among individuals with sensorineural hearing loss who use hearing aids (Kochkin, 2000). The most well-understood—and treatable—source of this deficit is a lack of signal audibility due to elevated audiometric thresholds. The reduced audibility of sounds due to hearing loss is captured by the standard audiological test, the audiogram. For speech sounds presented without amplification, audiometric sensitivity can explain as much as 50% to 75% of the variance across listeners in speech understanding in noise (Amos & Humes, 2007; Humes, 2007; Smoorenburg, 1992). However, with well-fit

amplification that increases the audibility of the speech spectrum substantially above threshold, the audiogram becomes a poor predictor of speech understanding in

[1]National Military Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, MD, USA
[2]Linnaeus Centre HEAD Graduate School, The Swedish Institute for Disability Research, Linköping University, Sweden
[3]Department of Clinical and Experimental Medicine, Linköping University, Sweden
[4]Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark

Corresponding author:
Joshua G. W. Bernstein, National Military Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, 4954 N Palmer Rd., Bethesda, MD 20889, USA.
Email: joshua.g.bernstein.civ@mail.mil

noise, in some cases no longer accounting for a significant proportion of the variance (Amos & Humes, 2007; Humes, 2007). This suggests that once audibility is restored by amplification, factors other than audibility are likely to blame for the poor speech understanding in noise experienced by some hearing-impaired (HI) listeners. Plomp (1986) categorized these other sources of the speech-perception deficit as the *distortion component*, to distinguish them from the audibility component of hearing loss. However, this general definition does not identify the precise mechanisms underlying the nonaudibility deficit. The general consensus in the literature is that this distortion could include deficits in cognitive processing ability, the distorted encoding of suprathreshold signals in the auditory periphery or midbrain, or a combination of the two factors.

Studies investigating the role of cognitive processing have found that measurements of working-memory capacity can account for 15% to 35% of the variance in speech-reception performance in noise for HI listeners supplied with well-fit amplification (e.g., Akeroyd, 2008; Foo, Rudner, Rönnberg, & Lunner, 2007; Lunner, 2003). Rönnberg et al. (Rönnberg, 2003; Rönnberg, Rudner, Foo, & Lunner, 2008) proposed that the perception of running speech involves the simultaneous processing and storage of auditory information. The idea is that some individuals are better able to compensate for an impoverished auditory signal, thereby freeing up working-memory capacity for the storage task. A more detailed account of the roles played by working memory can be found in the recent update of the Ease of Language Understanding model (Rönnberg et al., 2013).

Suprathreshold distortions have also been implicated in limiting the perception of speech in noise for HI listeners. The idea is that in addition to reducing the audibility of portions of the speech signal, damage to peripheral structures such as outer hair cells or auditory-nerve fibers can also result in the distortion of the neural representation of the audible components of the speech signal. These distortions can include reductions in temporal (e.g., Nelson, Schroder, & Wojtczak, 2001) or spectral resolution (e.g., Glasberg & Moore, 1986) that smear the features of the speech signal in these dimensions. HI listeners might also have a reduced ability to make use of cycle-by-cycle variation of fine-timing information (on the order of 100s to 1000s of Hz) in the stimulus waveform to extract speech information (e.g., Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006).

A standard clinical audiological approach that combines a pure-tone detection threshold with a measure of speech-reception performance in noise can provide information to distinguish the relative contributions of reduced audibility and the distortion component to impaired speech perception in noise—any reduction in

speech-reception performance relative to that predicted by audibility factors is attributed to the distortion. However, this approach cannot distinguish between cognitive and peripheral suprathreshold impairments as the causes of this distortion. To accomplish this, separate tests of cognitive function and suprathreshold processing relevant to the speech-perception task are required. Distinguishing between these two possible causes of speech-reception deficits is important because the specific nature of the distortion might require very different treatment once the audibility component of the hearing loss has been addressed with hearing aids. Individuals with a cognitive-processing deficit might benefit from auditory training to improve their ability to process speech cues (Song, Skoe, Banai, & Kraus, 2012), their cognitive function might benefit from long-term use of hearing aids (Amieva et al., 2015), or they might benefit from individualized signal processing depending on their level of cognitive function (Lunner, Rudner, & Rönnberg, 2009; Rudner, Rönnberg, & Lunner, 2011; Souza, Arehart, Shen, Anderson, & Kates, 2015). There is also the possibility that training working memory could improve speech-in-noise understanding. Individuals with suprathreshold processing deficits might benefit from future signal-processing algorithms designed to offset the particular distortions introduced by the hearing loss (see Bernstein, Summers, Grassi, & Grant, 2013). Furthermore, these individuals could be counseled as to the nature of their hearing loss, allowing the clinician to set realistic expectations about the likely efficacy of their treatment.

The current study was mainly concerned with the suprathreshold-distortion aspects of impaired speech-reception performance, leaving aside the question of cognitive processing for the time being. Previous studies in the literature report mixed results regarding how well individual psychophysical tests can predict speech-reception performance in noise. It is well established that frequency selectivity—the ability to hear out a signal at one frequency in the presence of a masker at a nearby frequency—is negatively affected by hearing loss (e.g., Glasberg & Moore, 1986). However, the literature is inconclusive regarding the role that frequency selectivity might play in limiting speech-reception performance in noise, with some studies showing a relationship between the two types of measure (e.g., Davies-Venn, Nelson, & Souza, 2015; Dreschler & Plomp, 1985) and others finding weak or nonsignificant correlations once the audiogram is factored out (e.g., Hopkins & Moore, 2011; Summers, Makashay, Theodoroff, & Leek, 2013; ter Keurs, Festen, & Plomp, 1993). Although temporal resolution has been shown to be negatively affected by hearing loss due to a loss of compressive gain (Nelson et al., 2001), there is little evidence that reduced speech-reception performance in noise is related to reduced temporal

resolution, except in situations involving relatively high-rate (16–32 Hz) modulated-noise maskers (Dubno, Horwitz, & Ahlstrom, 2003; George, Festen, & Houtgast, 2006). Several studies have pointed to a relationship between speech-reception performance and sensitivity to temporal fine structure (TFS) in the stimulus waveform—fluctuations on the order of 100s to 1000s of Hz that are encoded via phase locking in the auditory nerve fiber response (Buss, Hall, & Grose, 2004; Gnansia, Péan, Meyer, & Lorenzi, 2009; Hopkins & Moore, 2007; Neher, Lunner, Hopkins, & Moore, 2012; Strelcyk & Dau, 2009). But others have suggested that this relationship might reflect individual differences in age (Grose & Mamo, 2010), rather than hearing loss (Moore, Glasberg, Stoev, Füllgrabe, & Hopkins, 2012; Sheft, Shafiro, Lorenzi, McMullen, & Farrell, 2012).

One possible reason that previous studies of suprathreshold measures of auditory function have not shown convincing power to predict speech-reception performance in noise by HI listeners is that the stimuli are often tonal or narrowband in nature and lack the spectrotemporal characteristics of speech. Bernstein et al. (Bernstein, Mehraei, et al., 2013; Mehraei, Gallun, Leek, & Bernstein, 2014) examined the relationship between speech-reception performance in noise and a measure of sensitivity to combined spectrotemporal modulation (STM). Chi, Gao, Guyton, Ru, and Shamma (1999) and Elhilali, Chi, and Shamma (2003) showed that any speech spectrogram can be broken down into constituent STM components covering a range of spectral-modulation densities (cycles per octave, c/o) and temporal-modulation rates (Hz). Bernstein, Mehraei, et al. (2013) found that a measure of sensitivity to STM applied to a broadband carrier accounted for a substantial portion (∼40%) of the variance in speech-reception performance in noise for HI listeners, in addition to the 40% of the variance account for by the standard audiogram. Mehraei et al. (2014) measured sensitivity to STM applied to octave-band carriers and found that most of this predictive power could be attributed to STM sensitivity for a low-frequency carrier (1,000-Hz center frequency). Critically, Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) found that STM sensitivity was reduced for HI relative to NH listeners only for certain combinations of spectral-modulation density, temporal-modulation rate, and carrier frequency. This suggests that impaired STM processing does not reflect a general cognitive deficit, but instead a specific deficit related to the stimulus parameters, and is therefore likely of peripheral rather than central origin. The primary goal of the current study was to test an STM-sensitivity metric as a psychophysical measure of suprathreshold distortion, examining whether it can account for intersubject variability in speech-reception performance for HI listeners fitted with hearing aids. If

so, the long-term goal would be to develop the STM sensitivity metric as a clinical tool, allowing the audiologist to identify the extent to which suprathreshold distortion limits speech-reception performance in noise once audibility limitations have been overcome with a well-fit hearing aid.

Although the results of Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) show some promise that the STM sensitivity metric could eventually serve this purpose, there were several shortcomings of these previous studies that would need to be overcome before this measure could be considered as a possible clinical tool. First, they tested a very large number of STM conditions, yielding a psychophysical test that was prohibitively long for consideration for clinical use. However, one particular STM condition—2 c/o and 4 Hz—showed the largest difference in sensitivity between NH and HI listener groups, and STM sensitivity for this rate–density combination was at least as highly correlated to speech-reception scores as any other combination. Furthermore, Grant, Bernstein, and Summers (2013) found that STM detection required very little training—just one or two measurement blocks—before performance asymptoted to maximum levels. The current study focused on this single rate–density combination and used a short training period, thereby reducing testing time to a reasonable duration (∼ 15 min). Second, Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) did not test speech-reception performance with well-fit frequency-dependent amplification to overcome audibility limitations. Instead, they simply presented stimuli at a high overall level (92 dB SPL) with generic frequency shaping. This left a considerable portion of the speech dynamic range inaudible for frequencies above 2 kHz. The current study measured speech-reception performance for HI listeners fitted with individualized hearing-aid gain to optimize the audibility of the speech signal within the constraints of the device. Third, the ability of the test to account for variability in speech-reception performance would require validation in a larger group than the 12 HI listeners tested by Bernstein, Mehraei, et al. and Mehraei et al. (2014) before it could be considered for clinical use. The current study tested a much larger group of HI listeners who were part of a larger study of 200 HI participants evaluated on a range of speech-perception, psychoacoustic, and cognitive measures (the "n200" study; Rönnberg et al., 2016). The data reported here focus on the relationship between speech-reception performance in noise and STM sensitivity for the subgroup of 154 HI listeners who were tested on both of these measures.

A secondary goal of the study was to determine to what extent the audiogram and the measure of STM sensitivity could account for individual differences in the relative benefit obtained from hearing-aid

compression and noise-reduction algorithms. Previous work has shown that the benefit that listeners received from listening to a hearing aid with fast when compared with slow compression was significantly correlated to a measure of cognitive function, especially in modulated noise (Lunner & Sundewall-Thorén, 2007; Rudner et al., 2011). Others have shown that a measure of cognitive capacity can predict how well a listener can adapt to a new hearing-aid algorithm (fast or slow compression) that is different from their previous experience (Foo et al., 2007; Rudner, Foo, Rönnberg, & Lunner, 2009). Relatively, little is known about the extent to which measures of suprathreshold distortion correlate with the amount of benefit a listener might receive from a given signal-processing algorithm, although such a relationship has been theorized. For example, Moore (2008) argued that listeners who show poor sensitivity to TFS might experience difficulty in benefitting from fast compression hearing aids, because these listeners will rely more heavily on temporal-envelope information to understand speech, and fast compression tends to disrupt temporal-envelope cues (Stone & Moore, 2008). The current study compared the audiogram and STM sensitivity with aided speech-reception performance in noise for three different hearing-aid signal-processing algorithms: linear gain, nonlinear gain with fast compression, and linear gain with noise reduction.

Methods

The general approach was to measure speech-reception performance in noise for a group of listeners with sensorineural hearing loss fitted with hearing aids with individualized gain and to relate speech-reception performance to a measure of STM sensitivity and to the audiometric thresholds.

Listeners

This study examined the subset of 154 of the 200 participants (65 female) from the $n = 200$ study (Rönnberg et al., 2016) who completed the full battery of speech-in-noise and STM sensitivity testing. These individuals were experienced hearing-aid users with bilateral, symmetrical mild to moderate sensorineural hearing loss (HL), recruited from the patient population at the University Hospital of Linköping. All participants were bilaterally fitted with hearing aids and had used the aids for more than 1 year at the time of testing. Audiometric testing included ear-specific air-conduction and bone-conduction thresholds. Participants with a difference of more than 10 dB between the bone-conduction and air-conduction threshold at any two consecutive frequencies were considered to have conductive lesions and were not eligible for the study.

The participants were on average 59.9 years old (range 33–74, $SD = 8.5$). They had an average four-frequency pure-tone average at 500, 1,000, 2,000, and 4,000 Hz) in the better ear of 37.3 ($SD = 10.6$) dB HL. Figure 1 shows the audiograms for each individual listener averaged across the ears (gray lines) together with the mean audiogram across the population (circles and black lines). The study was approved by the regional ethics committee (Dnr: 55-09 T122-09), and all participants provided written informed consent. All participants were native Swedish speakers.

Hearing-Aid Fitting

Individualized hearing-aid gain settings were based on the average of the left- and right-ear tone-detection thresholds at each audiometric frequency (octave frequencies between 250 and 8,000 Hz plus 1,500 and 3,000 Hz). The gain and frequency response were fitted for individual listeners through linearization of the voice-aligned compression (VAC) rationale (Le Goff, 2015), with the working point defined according to the input spectrum of speech presented at 65 dB SPL. The VAC rationale can be classified as curvilinear wide-dynamic range compression. This compression model is partly based on loudness data by Buus and Florentine (2001) and is intended to ensure improved sound quality without loss of speech intelligibility, rather than loudness compensation per se. All automatic features were disabled (e.g., directional microphones and native noise reduction).

Three separate hearing-aid processing algorithms were tested: linear, linear with noise reduction, and fast compression. For the fast-compression algorithm, the compression ratio was set to a constant value of 2.0 across the input dynamic range (35–75 dB equivalent sound level, Leq) in four separate frequency bands,
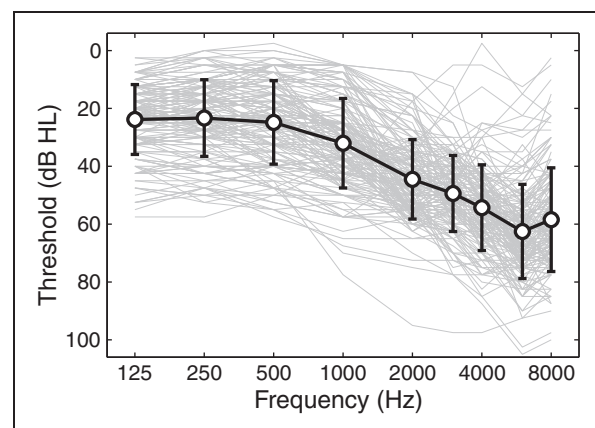

Figure 1. Audiograms for each individual listener in the study (gray lines), averaged across the left and right ears (circles and black lines), and the mean audiogram ± 1 standard deviation across the population.

with attack and release time constants of 10 and 80 ms, respectively. For the linear-processing algorithm with noise reduction, the noise reduction was based on the ideal binary mask (IBM) defined by Wang, Kjems, Pedersen, Boldt, and Lunner (2009), but adapted to realistic processing with nonideal masks (Boldt, Kyems, Pedersen, Lunner, & Wang, 2008). The IBM separates the signal from the noise by dividing the input spectrogram into a grid of time-frequency bins and selecting only those bins where the signal-to-noise ratio (SNR) exceeds a criterion value. However, this *ideal* processing requires prior knowledge of the separate signal and noise signals. The adaptation of Boldt et al. (2008) estimates the binary mask for a well-defined scenario involving a target speech signal and a single masking noise that are spatially separated from one another in a nonreverberant environment.

To control variability in the physical fit of the hearing aids and for direct control over the properties of the acoustics reaching each listener's ears, listeners did not wear the hearing aids used for speech testing. Instead, a master behind-the-ear hearing aid (Oticon Epoq XW) was placed inside an anechoic test box (Brüel & Kjær 4232) containing a loudspeaker, and the output of the hearing aid was measured by an ear-simulating coupler (Brüel & Kjær 4157). The output of the ear simulator microphone was then relayed to an amplifier (NAD 2400) that delivered the resulting stimulus diotically to the listener via ER-3A insert phones. The ER-3A response was controlled with an equalizer (Behringer ULTRACURVE PRO DEQ2496) to offset the nonflat response for frequencies below 200 Hz.

Speech Reception in Noise

Speech-reception performance was estimated using Matrix sentences (Akeroyd et al., 2015; Hagerman & Kinnefors, 1995). The Matrix sentences have a consistent structure with five key words: proper noun, verb, number, adjective, and noun (e.g., "Ann had five red boxes."). The sentences, spoken by a single female talker, were presented in randomized order, and the listener was asked to repeat as many of the words as possible, with the experimenter entering the number (0–5) of words that were correctly identified. An adaptive-tracking algorithm estimated the threshold SNRs required to achieve 50% and 80% correct levels of performance (Brand & Kolmeier, 2002) with two interleaved tracks of 15 trials each. Following each response, the SNR was adapted upward or downward for the following trial, with a variable step size based on the number of keywords correct on the previous trial. The threshold SNR was based on the average of the SNRs for the final 10 sentences in each track. The speech level was held fixed at 65 dB SPL, and the noise level was adjusted

to yield the desired SNR. For each of the three hearing-aid algorithms, four speech conditions were tested, involving combinations of two different masker types (speech-spectrum shaped noise or four-talker babble) and two different performance levels (50% or 80% correct).

STM Sensitivity

The STM stimulus consisted of a spectrally rippled spectrum with peaks that shift over time. STM sensitivity was measured in a two-alternative forced choice task. Listeners were presented with two sequential 500-ms broadband-noise stimuli separated by an interstimulus interval of 300 ms. In one interval, the noise was unmodulated. In the other interval, STM was applied. The listener's task was to identify which of the two intervals contained the modulation (Bernstein, Mehraei, et al., 2013; Chi et al., 1999; Mehraei et al., 2014). The modulation depth was varied in a three-down, one-up adaptive procedure (Levitt, 1971) to estimate the depth required for the listener to correctly identify the interval containing the presence of STM 79.4% of the time. The STM stimulus was generated by applying sinusoidal modulation to each closely spaced random-phase tone (1,000/octave) that makes up the noise spectrum, but applying a shift in the relative phase of the modulation for each successive carrier tone (for details, see Bernstein, Mehraei, et al., 2013; Mehraei et al., 2014). Modulation depth was defined in decibels (dB) as $20 \log_{10} m$, where m is the depth defined in linear terms. For example, 0 dB represented full modulation ($m = 1$). STM sensitivity was measured for a for a single rate–density combination (4 Hz and 2 c/o), with spectral ripples always moving in the upward direction. This combination of STM rate and density was chosen because it yielded the greatest difference in performance between NH and HI listener groups (Bernstein, Mehraei, et al., 2013). The STM was applied to a noise carrier that was low-pass filtered at 2 kHz. This was done, rather than employing a broadband carrier, for two reasons. First, Mehraei et al. (2014) investigated the carrier-frequency dependence of the relationship between STM sensitivity and speech-reception performance in noise and found the relationship to be strongest for a low-frequency STM carrier. Second, the goal of the study was to identify a psychophysical test that could account for variance in speech performance not accounted for by the audiogram, and the high-frequency components of the audiogram (i.e., 2 kHz or greater) are typically most highly correlated with speech perception. An example spectrogram of the 4-Hz 2-c/o STM stimulus employed in this study is shown in Figure 2, with the black and white regions representing peaks and valleys in the stimulus energy.

STM sensitivity was measured independently for each ear—first for the left ear and then for the right ear. For
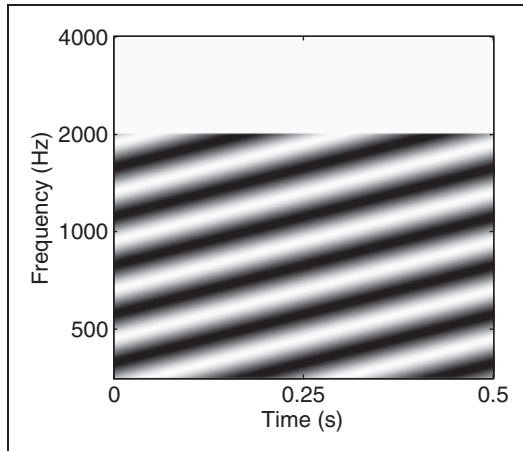
Figure 2. Example spectrogram for an STM stimulus (2 c/o, 4 Hz) with full modulation depth.
Note. STM = spectrotemporal modulation.

each ear, listeners completed one training block and three test blocks. For the training block, the temporal rate was set at 4 Hz, but the spectral density was set to 1 c/o to yield a more salient percept than the 2-c/o test stimulus. This first block was limited to a total of 13 trials. For the three test blocks, STM sensitivity was measured for a 4-Hz, 2-c/o stimulus. In each block, modulation depth was initially set to 0 dB (i.e., full modulation), changed by 6 dB until the first lower reversal point (i.e., when the direction of the adaptive track changed from decreasing to increasing modulation depth following an incorrect response), changed by 4 dB for the next two reversals, and then by 2 dB for the final six reversals. The STM detection threshold was estimated to be the mean modulation depth for the last six reversal points.

The stimulus was physically limited to a maximum modulation depth of 0 dB (full modulation). If the tracking procedure required a modulation depth greater than this value on any given trial, the modulation depth was kept at 0 dB for that trial. If there were more than three incorrect responses at full modulation depth during any single block, it was assumed that the listener could not achieve the target of 79.4% correct performance for a fully modulated stimulus. In these cases, the test reverted to a method of fixed stimuli, whereby the listener completed an additional 40 trials with the modulation depth held constant at 0 dB.

For each listener, STM sensitivity was characterized by averaging the STM detection thresholds measured across the three test runs for each of the two ears (six runs in total). For runs where a threshold could not be estimated due to more than three incorrect responses at full modulation depth, percentage-correct scores were transformed to equivalent detection thresholds using a method similar to that described by Hopkins and Moore (2010a). This transformation required an analysis of the

slope of the psychometric function describing the relationship between STM depth and d′ across the population of listeners tested. The slope of the psychometric function was estimated for each individual listener in the experiment (excluding runs where an adaptive threshold was not measured) by fitting a line to the d′ values (converted from the percentage-correct scores by assuming unbiased responses, Hacker & Ratcliff, 1979) as a function of modulation depth. This analysis confirmed that the slope of the psychometric function was constant for modulation depths between −10 and 0 dB. The slopes were then averaged across the listeners in the study, yielding a mean slope of 0.238 d′ points per dB. A particular percentage-correct score at full modulation depth (0 dB) was transformed into an equivalent threshold modulation depth by taking the difference between the measured d′ value and the d′ value of 1.16 associated with the tracked percentage score of 79.4% in a two-alternative forced choice task. This difference was divided by the mean slope (0.238 d′ points per dB), with the resulting dB value taken as the equivalent modulation depth threshold. Thus, percentage-correct scores below the tracked value of 79.4% were transformed into equivalent modulation depth values greater than 0 dB, while scores greater than 79.4% correct were transformed into values less than 0 dB.

Analysis

The main goal of the study was to determine the extent to which STM sensitivity could account for variance in speech-reception performance in noise with hearing aids that could not be accounted for by the audiogram. A series of correlation analyses were carried out to examine how audiometric thresholds and STM sensitivity could jointly account for intersubject variability in speech-reception performance. First, separate correlation analyses determined the extent to which standard audiometric data (i.e., pure-tone detection thresholds at each frequency), listener age, and the measure of STM sensitivity could each independently account for variance in speech-reception performance. Second, a multiple regression analysis was carried out to determine how much additional variance was accounted for by the STM sensitivity metric that was not accounted for by the audiogram and age variables.

A global measure of speech-reception performance was established by averaging the threshold SNRs measured across all 12 combinations of (3) hearing-aid processing algorithms, (2) noise types, and (2) performance levels. This global speech-reception threshold (SRT) measure was designed to reduce some of the measurement noise assumed to exist for any one combination of these test parameters. The audiogram and STM metrics were compared with speech-reception performance for

each individual combination of test parameters as well as to the global SRT.

The first step of the analysis was to determine the proportion of the variance in speech-reception performance with hearing aids that could be accounted for by the audiogram. To accomplish this, the audiometric data were examined in several ways to determine the maximum amount of the variance that could be accounted for. Typically, an approach of testing various combinations of the predictor variables (i.e., the audiogram) with the goal of finding the maximum correlation with the predicted variable (i.e., the SRT) is to be avoided due to the possibility of an artificially inflated correlation due to chance. However, in this case, the goal was to identify the extent to which the STM measure improved the prediction. To be conservative about the proportion of the variance in SRT that the STM measure could account for beyond the audiogram, we wanted to find the greatest amount of variance in SRT that could be accounted for by a reasonable treatment of the audiogram data. The first set of analyses employed the Speech Intelligibility Index (SII) (American National Standards Institute, 1997), which is often considered the *gold standard* for predicting speech-reception performance based on the audiogram. The second set of analyses directly compared audiometric thresholds with the SRT.

Several different versions of the SII were tested. First, the standard SII was used to predict the SRT in noise. Because no SII weighting function is available for the Swedish Matrix sentences, the predictions employed the standard default weighting function. To characterize the audibility of the speech, the audiogram was combined with hearing-aid gain data for each individual listener, with equal efficiency assumed for all listeners. Then the SNR was manipulated by adjusting the level of the speech-shaped noise in the SII model until the SII generated predicted scores of 50% and 80% correct. These two resulting SRT estimates were then averaged together to produce the predicted global SRT.

A second SII analysis examined the *desensitized* SII in noise (Ching, Dillon, Lockhart, van Wanrooy, & Flax, 2011; Johnson, 2013) that also simulates some degree of suprathreshold distortion. While the standard SII only models the audibility aspects of hearing loss, the audiogram might contain additional power to account for variability in speech-reception performance beyond audibility, because pure-tone thresholds are likely to correlate to some extent with suprathreshold distortion. Because the SII tends to overpredict speech-reception performance for HI listeners, the inclusion of a *desensitization* factor has been proposed as an addition to the SII to account for reduced speech-reception performance that is related to audiometric thresholds but is not predicted by estimates of audibility (Ching et al., 2011; Johnson, 2013).

Because the calculation of the SII in noise is dominated by the noise level, the audiometric differences across listeners played a relatively minor role in the SII calculation. To focus more on the contribution of individual differences in audiometric thresholds to the SII prediction, a third analysis calculated the standard SII in quiet. Individual audiometric thresholds and hearing-aid gain data were used to calculate the audibility of the speech, generating an SII value between 0 (*no audibility*) and 1 (*full audibility*). A fourth analysis calculated the desensitized SII in quiet, whereby a desensitization factor based on the audiogram was applied to the audibility calculation.

In addition to the SII analyses, speech-reception performance was also compared directly to the audiometric thresholds. With nine test frequencies, a very large number of frequency combinations could theoretically be tested in a series of multiple regression models. To limit the number of combinations, the approach we took was to first examine the relationship between pure-tone thresholds (averaged across the two ears) at each frequency (125, 250, 500, 1,000, 2,000, 3,000, 4,000, 6,000, and 8,000 Hz) and speech-reception performance in noise. This analysis showed that frequencies of 2,000 Hz and above tended to correlate to speech-reception scores, while lower frequencies did not. This was expected based on previous literature (e.g., Humes, 2007; Smoorenburg, 1992), and is consistent with the fact that audiometric thresholds were more elevated for frequencies of 2 kHz and above than for lower frequencies (Figure 1). On the basis of this analysis, a low-frequency average (LFA; 125–1,000 Hz) and a high-frequency average (HFA; 2,000–6,000 Hz) audiogram were computed for each listener for comparison with speech-reception performance. (The highest frequency of 8,000 Hz was excluded from the HFA calculation because an audiometric threshold was not available at this frequency for all listeners in the study.)

Results

The Relationship Between Audiogram and Speech-Reception Performance

The main goal of this study was to determine the extent to which the STM metric could predict variance in speech-reception scores not predicted by the audiogram. The first step in this process was to determine how much of the variance in speech-reception scores could be attributed to the audiogram. We tested four formulations involving the audiogram—the raw and desensitized SII in quiet and in noise—as well as the LFA and HFA audiogram—to identify the particular formulation that accounted for the largest share of the variance. The global SRT was compared with various formulations

involving the audiogram. The global SRT was moderately correlated with the SRT predicted by the standard SII in noise, $R^2 = .118$, $F(1, 152) = 20.3$, $p < .0005$, 95% CI [0.024, 0.212], or by the SII in noise modified to include hearing loss desensitization, $R^2 = .153$, $F(1, 152) = 27.5$, $p < .0005$, 95% CI [0.050, 0.256]. In both cases, the SII failed to capture the wide variation in speech-reception performance across the group of 154 listeners. Figure 3(a) shows the results for the desensitized version of the SII in noise. Whereas the measured SRTs varied over an 8-dB range (-6 to $+2$ dB), the vast majority of the predicted SRTs fell in a 3-dB range (between -6 and -3 dB). Similar results were observed for the standard SII (not shown). Thus, neither the standard SII nor the desensitized SII could account for much of the wide variability in speech-reception performance in noise across listeners.

The SII predictions in noise are dominated by the influence of the noise on the predicted audibility of the speech. To focus more on the contribution of differences in audiometric thresholds to the SII prediction, the global SRT was also compared with the SII calculated in quiet. The SII in quiet was moderately correlated with the speech scores and accounted for a somewhat larger proportion of the variance than the SII in noise for both the standard, $R^2 = .213$, $F(1, 152) = 41.0$, $p < .0005$, 95% CI [0.100, 0.326], data not shown, and desensitized, $R^2 = .231$, $F(1, 152) = 45.5$, $p < .0005$, 95% CI [0.116, 0.346], Figure 3(b), versions of the SII.

The correlations shown in Figure 3 could have been influenced by the inclusion of outlier subjects with much lower SII scores than the rest of the population.
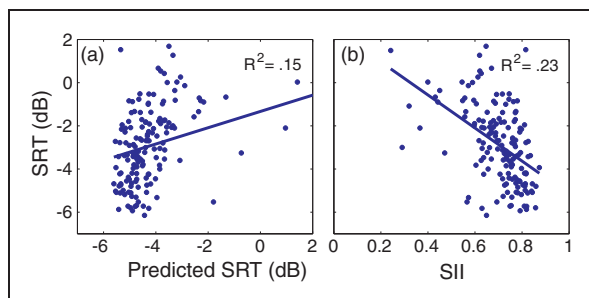
Removing these outliers (whereby the SII was more than 1.5 times of the interquartile range below the first quartile) had a small effect on the results, decreasing the correlation between the SII and the global SRT for the standard SII in noise ($R^2 = .07$), the standard SII in quiet ($R^2 = .19$) and the desensitized SII in quiet ($R^2 = .18$), but increasing the correlation for the desensitized SII in noise ($R^2 = .26$). But in all four cases, the SII metrics were still inferior to the HFA in accounting for variance in SRTs.

Figure 4 plots the variance in global SRT accounted for by the pure-tone threshold for each of the nine audiometric frequencies tested (black squares). Audiometric thresholds for frequencies 2 kHz and higher each accounted for a significant proportion of the variance in global SRT ($p < .05$, asterisks), with the highest correlation observed at 3 kHz, $R^2 = .289$, $F(1, 152) = 61.7$, $p < .0005$, 95% CI [0.170, 0.408]. Low-frequency thresholds (1 kHz and below) did not correlate significantly with the global SRT ($p > .05$). As a result, the LFA was only weakly correlated with the SRT, $R^2 = .036$, $F(1, 152) = 5.62$, $p < .05$, 95% CI [-0.021, 0.092], not shown. In contrast, the HFA was strongly correlated with the SRT, accounting for a larger proportion of the variance in SRT than any of the SII models, $R^2 = .309$, $F(1, 152) = 68.0$, $p < .0005$, 95% CI [0.189, 0.429], Figure 5(a). Thus, the HFA was used as the audiometric variable in the multiple-regression analysis in the following sections.
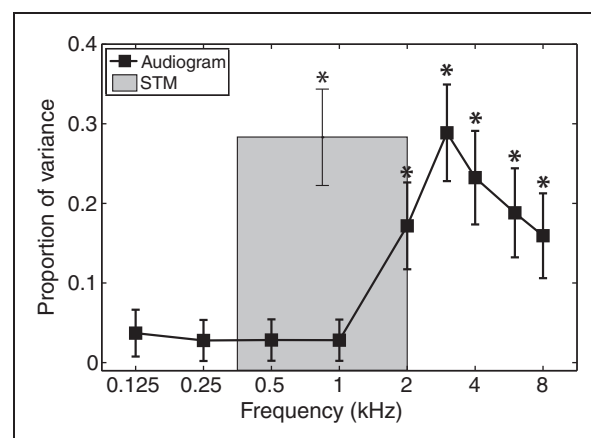

Figure 3. Scatterplots showing the relationship between the global SRT (averaged across test conditions) and the SII derived from individual audiometric thresholds and hearing-aid gain measurements. (a) The horizontal axis represents the SRT predicted from the SII calculated in noise. (b) The horizontal axis represents the SII calculated for speech presented in quiet. The diagonal lines represent a linear fit to the data. In both panels, a desensitization factor (Ching et al., 2011) was included based on each listener's audiogram; however, the inclusion of this factor had little effect on the correlations.
Note. SII = Speech Intelligibility Index; SRT = speech-reception threshold.

Figure 4. The relationship between speech-reception performance and both the audiogram and the STM sensitivity metric. Black squares denote the proportion of variance in global SRT accounted for by the audiometric threshold is plotted as a function of audiometric frequency (black squares). The gray-shaded region denotes the bandwidth of the STM stimulus and proportion of the variance in global SRT accounted for by the STM sensitivity metric. Asterisks (*) represent conditions showing significant correlations with the global SRT ($p < .05$). Error bars indicate ± 1 standard error of the R^2 estimate.
Note. SII = Speech Intelligibility Index; SRT = speech-reception threshold; STM = spectrotemporal modulation.

Age

There was a moderate correlation between the global SRT and age, $R^2 = .173$, $F(1, 152) = 31.9$, $p < .0005$, 95% CI [0.066, 0.280], Figure 5(b). When combined with the HFA in a multiple regression analysis, listener age accounted for a small but significant amount of additional variance, $\Delta R^2 = .047$, $p < .005$; total $R^2 = .356$, $F(2, 151) = 41.7$, $p < .0005$, 95% CI [0.237, 0.475], not shown.

STM Sensitivity and Its Relationship to Speech-Reception Performance

An analysis of variance conducted on the STM-sensitivity data with two within-subject factors (ear of presentation and block number) revealed a significant main effect of ear of presentation, $F(1, 148) = 15.6$, $p < .0005$, reflecting the overall better performance for the right ear (mean STM detection threshold $= -2.74$ dB) than for the left ear (-1.54 dB). There was no significant main effect of block number ($p = .92$) or interaction between block number and ear of presentation ($p = .37$), suggesting an absence of a training effect in the STM detection task. While the effect of ear of presentation is consistent with the right-ear advantage often observed for dichotic speech listening tasks (e.g., Grimshaw, Kwasny, Covell, & Johnson, 2003), STM detection does not fall into this category. It might instead reflect a learning effect, as the left ear was tested before the right ear. Therefore, the STM data were combined across the ears, with STM sensitivity defined for each listener as the average of the thresholds across the six test blocks (2 Ears × 3 Test blocks).

Table 1 shows R^2 values for pairwise correlations conducted between the major variables examined in the study: the global SRT, the LFA and HFA audiograms,

age, and STM sensitivity. STM sensitivity was strongly correlated with speech-reception performance, $R^2 = .282$, $F(1, 152) = 59.8$, $p < .0005$, 95% CI [0.164, 0.499], Figure 5(c), comparable to the strength of the correlation between the global SRT and the HFA (Figure 5(a)). Adding STM to the HFA in a multiple regression analysis significantly increased the proportion of the variance accounted for, $\Delta R^2 = .131$, $p < .0005$; total $R^2 = .440$, $F(2, 151) = 68.0$, $p < .0005$, 95% CI [0.325, 0.555], Figure 5(d). Adding STM as a third predictor after HFA and age also significantly increased the proportion of the variance accounted for, $\Delta R^2 = .090$, $p < .0005$; total $R^2 = .446$, $F(3, 150) = 72.6$, $p < .0005$, 95% CI [0.332, 0.560], not shown. In contrast, adding age as the third predictor after HFA and STM did not significantly improve the prediction ($p = .21$). Thus, all of the additional variance (beyond the HFA) accounted for by age was shared with STM, but only a portion of the additional variance accounted for by STM was shared with age. Figure 4 shows that that the frequency range of

Table 1. R^2 Values for the Pairwise Correlations Between the Main Variables Examined in the Study.

	LFA	HFA	STM	Age
Global SRT	.04	.31*	.28*	.17*
LFA		.02	.05	.00
HFA			.12*	.15*
STM				.24

Note. LFA = low-frequency average; HFA = high-frequency average; STM = spectrotemporal modulation; SRT = speech-reception threshold. Asterisks indicate significant correlations ($p < .05$) after Bonferroni correction for multiple (10) comparisons, although in all of these cases p-values were less than .005.
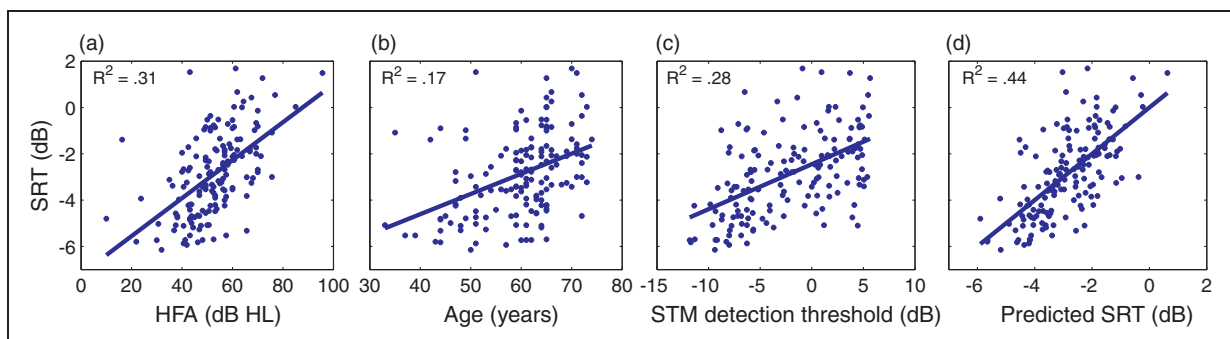

Figure 5. Scatterplots showing the relationship between the global SRT and age, HFA, and STM sensitivity. (a) The horizontal axis represents the HFA audiogram. (b) The horizontal axis represents listener age. (c) The horizontal axis represents the STM detection threshold. In cases where the STM detection threshold exceeded 0 dB, the horizontal axis represents percentage-correct scores measured for full-modulation stimuli that were converted to equivalent modulation-depth thresholds. (d) The horizontal axis represents the predicted SRT based on a multiple-regression model incorporating the STM and HFA metrics as independent variables. Diagonal lines represent linear fits to the data. *Note.* HFA = high-frequency average; SRT = speech-reception threshold; STM = spectrotemporal modulation.

the STM stimulus (353–2,000 Hz) had very little overlap with the frequency region where audiometric thresholds were correlated with speech-reception performance (2,000–8,000 Hz).

Population Subgroups

An additional analysis was conducted to determine whether the STM measure was particularly useful in accounting for variance in speech-reception performance for certain subsets of the tested population. The idea was to determine whether there were particular segments of the population for whom it did not add any predictive power to measure and include their individual STM score in the analysis. If this was found to be the case, this would suggest that time could be saved in the clinic by only measuring STM sensitivity for the critical subgroup. The population of 154 listeners was rank ordered based on age (Figure 6(a)) or HFA (Figure 6(b)), and an individualized STM sensitivity measurement was added to the multiple regression analysis one at a time for each successive listener in the population. For example, in Figure 6(a), following the solid line from left to right shows the effect of adding individual STM scores to the analysis starting with the youngest listener. For the participants below the cutoff age represented on the horizontal axis, both the HFA and the individual STM score were included as inputs to the multiple regression analysis. For the remainder of the listeners, the mean STM

score for the excluded subgroup was substituted for the individual STM score. This analysis was then repeated in the opposite direction, adding individual STM scores starting with the oldest listener followed by subsequently younger listeners (Figure 6(a), dashed curve). A similar analysis was carried out with the listeners sorted based on their HFA (Figure 6(b)), with individual STM scores added to the analysis for listeners with successively poorer (solid curve) or successively better HFA (dashed curve).

Figure 6 shows that the overall proportion of the variance generally increased as the individualized STM scores were included for more and more of the listeners in the tested population, but this was not universally the case. The horizontal lines in Figure 6 depict the minimum R^2 value (.327) representing a statistically significant increase ($p < .05$) in the proportion of the variance accounted for beyond that obtained with the HFA alone ($R^2 = .309$). Following the dashed curves from right to left shows the effect of adding individual STM scores to the analysis for successively younger or better HFA participants. This process did not yield a significant increase in the proportion of variance until individual STM scores were included for listeners below the 70th percentile for age (65 years old) or the 52nd percentile for HFA (53 dB HL; Figure 6, vertical lines). In contrast, following the solid curves from left to right shows adding individual STM scores into the analysis for successively older or poorer HFA participants significantly increased the
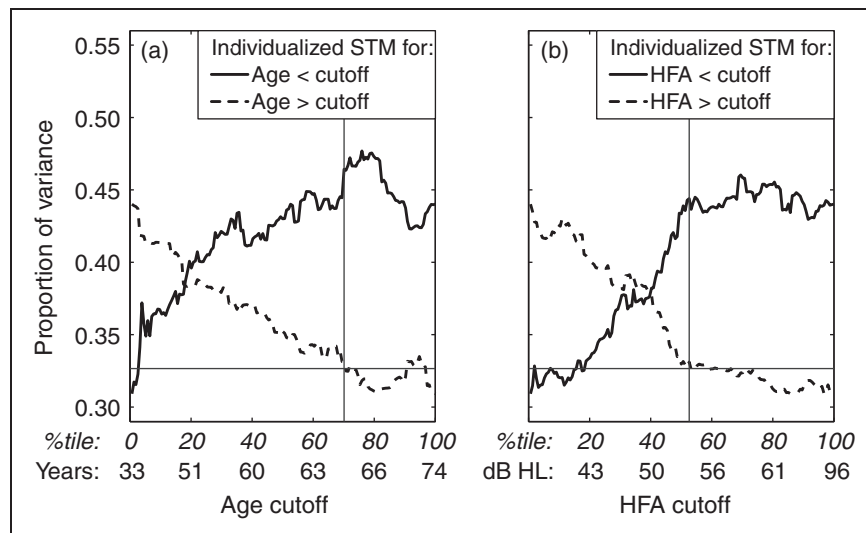

Figure 6. The proportion of the variance in global SRT accounted for by the combination of the HFA and STM metrics as individualized STM scores are included for successive participants in the group of 154 listeners. For the remainder of the listeners, the STM score for each individual was replaced by the average STM score for this group of listeners. Horizontal lines indicate the threshold R^2 value for which the STM metric significantly increased the proportion of variance accounted for beyond the HFA alone. Vertical lines indicate the point at which including additional individualized STM scores for successively younger listeners or listeners with lower HFA (i.e., following the dashed curves from right to left) yielded an R^2 value that exceeded this threshold. (a) Listeners sorted according to age. (b) Listeners sorted according to HFA.
Note. HFA = high-frequency average; SRT = speech-reception threshold; STM = spectrotemporal modulation.

proportion of the variance accounted for almost immediately (for age) and at the 15th percentile (for HFA). In summary, this analysis suggests that the STM sensitivity metric was mainly of value in accounting for individual differences in speech-reception performance with hearing aids for individuals younger than 65 years old or with HFA below 53 dB HL.

Duration of the STM Test

Each listener completed eight blocks of the STM detection threshold measurement, four with each ear (one training block and three test blocks). The mean test time (± 1 standard deviation) for each block, including rest time between blocks, was 1.14 ± 0.19 min for the training blocks and 2.19 ± 0.59 min for the test blocks. Thus on average, each listener completed the two training blocks and six test blocks in 15.4 min.

The analyses presented earlier averaged the thresholds across the six test blocks. This test would be of greater clinical utility if its duration could be shortened by eliminating some of the test blocks. The correlation between STM sensitivity and speech-reception performance was reevaluated while considering only the first test block in each ear (two blocks total), the first two test blocks (four blocks total), or all three test blocks (six blocks total). The resulting R^2 values were .264 with one test block, .275 of the variance with two test blocks, and .283 with three test blocks. Combining the STM metric with HFA in a multiple regression analysis yielded R^2 values of .438 with one test block, .440 with two test blocks, and .440 with three test blocks in each ear. Thus, presenting only a single test block in each ear should not diminish the predictive power of the STM metric. The total expected mean test time for one training and one test block in each ear would be 6.7 min.

Signal-Processing Algorithms and Noise Types

The analyses presented earlier focused only on the global SRT, averaged across all of the signal-processing conditions and noise types tested. Correlations between speech-reception performance, the HFA audiogram, and STM sensitivity were examined separately for the three signal-processing algorithms (with data averaged across masker type) and for the two masker types (with data averaged across signal-processing algorithm). A test for the difference between dependent correlations (Steiger, 1980) examined whether the HFA or STM metrics accounted for significantly different proportions of the variance in speech scores across the signal-processing algorithms or noise types, with Bonferroni corrections applied for eight multiple comparisons (three for the signal-processing algorithms plus one for the noise types, for both the STM and HFA metrics). Although

the R^2 values ranged from .171 to .271 for the STM metric, and from .232 to .293 for the HFA, none of the differences in R^2 were significant ($p > .38$ in all cases). The same trend held in each case: treated independently, the HFA and STM metrics each accounted for a roughly similar amount of variance, whereas in combination, the STM accounted for an additional approximately 15% of the variance that was not accounted for by the HFA audiogram.

An additional analysis was carried out to determine whether the HFA or STM metrics could account for individual differences in the amount of benefit listeners obtained from the various signal-processing algorithms (relative to the linear algorithm) or in stationary noise (relative to four-talker babble). None of the correlations between HFA or STM and the benefit scores were found to be significant ($p > .05$).

Discussion

The results of this study confirm the main findings of Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) that a measure of STM sensitivity can account for a significant proportion of the variance in speech-reception performance in noise for HI listeners beyond that accounted for by the audiogram. The current results extend these previous findings in a small group of listeners ($N = 12$) presented with speech stimuli over headphones with generic gain, to a much larger group of $N = 154$ HI listeners fitted with real hearing aids and individualized frequency-dependent gain.

The ability to understand speech in noise depends on at least three factors: the audibility of the speech information, the extent to which the encoding of the speech information is distorted by damaged peripheral processes, and the cognitive ability to make sense of the speech information received. While the audiogram obviously captures the audibility factor, the audiometric thresholds have also been shown to correlate to suprathreshold distortion. For example, a greater degree of hearing loss tends to yield poorer frequency selectivity (Glasberg & Moore, 1986), increased forward masking (Nelson et al., 2001), and a reduced ability to use temporal-fine-structure information to discriminate interaural time differences (Strelcyk & Dau, 2009) or detect inharmonicity (Hopkins & Moore, 2007). While audiometric frequencies of 2 kHz and above accounted for about 30% of the variance in speech-reception performance (Figure 4), the SII was a poor predictor of speech-reception performance in noise (Figure 3(a)). Our interpretation of this finding is that the SII measures the contribution of audibility to individual differences in speech-reception performance. With the use of hearing aids to overcome some of the audibility limitation, and the fact that in many cases the presence of noise limits

the audibility of the speech spectrum rather than absolute threshold, relatively little of the intersubject variability was attributable to audibility differences, consistent with previous findings (e.g., Humes, 2007). Therefore, the fact that the HFA audiogram correlates with speech-reception in noise (Figure 5(a)) likely reflects the role of nonaudibility factors (correlated with the audiogram) in limiting speech-reception performance in noise. Including a desensitization factor (Ching et al., 2011; Johnson, 2013) that attempts to incorporate this audiogram-correlated suprathreshold distortion improved the prediction somewhat, but still did not yield a correlation with speech-reception performance that was as strong as that achieved by the HFA audiogram. This suggests that the desensitization model is not as good as the audiogram itself at capturing the nonaudibility factors that limit speech-reception performance in noise. For that reason, we used the strongest audiogram-based predictor—the HFA—as the baseline audiogram metric to determine how much additional variance could be accounted for by the STM metric.

While the audiogram and the STM each accounted for a similar proportion (approximately 30%) of the variance in speech scores, when added together in a multiple regression, there was some shared variance between the two metrics. Of the 44% of the variance accounted for, roughly a third was attributable to the audiogram alone, a third was attributable to STM sensitivity alone, and a third was attributable to shared variance between the two measures. The fact that audibility (i.e., the SII) predicted a very small amount of the individual variability in speech-reception performance suggests that a good deal of the correlation between the audiogram and speech scores reflected suprathreshold distortion. While the current study did not provide any information regarding the mechanisms governing the distortion at high frequencies, one possibility is that some of the suprathreshold distortion captured by the HFA could reflect intersubject variability in frequency selectivity (Mehraei et al., 2014; Summers et al., 2013).

At the same time, it is possible that there was some effect of audibility on performance in the STM task, given that the STM stimulus contained energy in the 2-kHz frequency region where audiometric thresholds begin to rise fairly dramatically for many of the listeners in the study (Figure 1). Therefore, of the 44% of the variance accounted for by the audiogram and STM measures, it is difficult to pinpoint exactly how much could be attributed to audibility and how much could be attributed to suprathreshold distortion factors. The amount of the variance accounted for by suprathreshold distortion factors could have ranged anywhere from the 13% accounted for by the STM metric in addition to the variance accounted for by the audiogram, to the 28% accounted for by the STM metric in isolation. One

possibility is that a strategy of compensating for audibility differences in the STM stimuli might increase the sensitivity of the STM test.

Audiometric thresholds were significantly correlated with aided speech-reception performance only for audiometric frequencies of 2 kHz and above (Figure 4). This result is consistent with the audiograms shown in Figure 1, whereby absolute thresholds were substantially poorer in this region than for the low frequencies (1 kHz and below). This range of frequencies is nearly orthogonal to the stimulus frequency range encompassed by the STM stimulus (353–2,000 Hz), except for some overlap at 2,000 Hz. This suggests that for the purposes of predicting speech-reception performance based on a combination of audibility and suprathreshold factors, a measure of STM sensitivity could replace the pure-tone audiogram for frequencies below 2,000 Hz. This is not to say that the low-frequency audiogram could be eliminated altogether; identifying audibility loss below 2,000 Hz is likely to be important for the purposes of selecting hearing-aid gain settings at these frequencies. Rather, this result suggests that to more thoroughly characterize the role of low-frequency processing in limiting speech perception in noise with hearing aids, a measure of STM sensitivity should be added to the standard measurement of the high-frequency audiogram. The analyses shown in Figure 6 suggest that a clinical test of STM sensitivity could be particularly useful for this purpose for individuals below 65 years old or with HFA better than 53 dB HL. For individuals who were both older than 65 and with HFA higher than 53 dB HL, the STM metric did not add any predictive power to the HFA metric, suggesting that this additional clinical test would not be worth conducting for this subgroup of the population. It is not clear why the STM metric did not provide any predictive power for individuals in this subgroup of 25 listeners. One possibility is that most of these listeners had considerable difficulty with the STM task, such that in many cases, performance was measured in percentage-correct terms at full modulation depth rather than adaptively. Fourteen of these 25 listeners (56.0%) performed poorly enough that the STM procedure abandoned the adaptive track for more than three of the six adaptive runs (this was the case for only 39 of the remaining 129 listeners, 30.1%). It could be that the accuracy of the performance metric was reduced in these cases such that any measured differences between individual listeners were not meaningful.

It remains an open question as to the physiological mechanism that underlies STM sensitivity deficit for some HI individuals. Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) argued that reduced STM sensitivity might reflect an inability to use TFS information to detect the presence of changes in spectral-peak frequencies. This argument was based on evidence of reduced STM sensitivity for HI listeners for only a

subset of modulation conditions: those involving relatively slow modulation rates (i.e., 4 Hz) and relatively high spectral densities (2–4 c/o), especially for low carrier frequencies (1,000 Hz and below). This pattern matched the pattern of low FM rate and low carrier frequency for which Moore and Sek (1996) and Moore and Skrodzka (2002) argued that the detection of frequency modulation is mediated by the TFS of the signal. In contrast, NH and HI listeners showed very little difference in performance for conditions involving faster modulation rates (32 Hz), where Moore and Sek and Moore and Skrodzka argued that listeners rely on amplitude modulation (AM) rather than TFS cues for the detection of frequency modulation. Further supporting the idea that TFS cues for the detection of frequency modulation are disrupted by hearing loss, HI listeners have been shown to be impaired in detecting low-rate frequency modulation even when AM cues are masked by the introduction of random AM to the stimuli (Johannesen, Pérez-González, Kalluri, Blanco, & Lopez-Poveda, 2016; Kortlang, Mauermann, & Ewert, 2016). In this view, the correlation between STM sensitivity and speech-reception performance in noise is consistent with other studies in the literature that have identified a relationship between speech-reception performance in noise and the ability to use TFS information (Buss et al., 2004; Gnansia et al., 2009; Hopkins & Moore, 2007, 2010b; Johannesen et al., 2016; Neher et al., 2012; Strelcyk & Dau, 2009).

On the other hand, we cannot rule out the possibility that the reduced STM sensitivity observed for HI listeners might reflect reduced frequency selectivity. In fact, Bernstein, Mehraei, et al. (2013) and Mehraei et al. (2014) identified a correlation between STM sensitivity and frequency selectivity, at least in the high frequencies (4,000 Hz). Furthermore, the fact these studies found STM sensitivity to be impaired for higher (2–4 c/o) but not for lower (0.5–1 c/o) is consistent with an explanation based on reduced frequency selectivity. Classical measurements of frequency selectivity involve the characterization of peripheral tuning in terms of the bandwidth of a putative *auditory filter* at a particular location along the cochlea (e.g., Glasberg & Moore, 1986, 1990). Studies attempting to relate frequency selectivity to speech-reception performance in noise have had mixed results, with some studies revealing a significant correlation (e.g., Dreschler & Plomp, 1985) and others failing to do so (e.g., Hopkins & Moore, 2011). Recently, a number of studies that have used broadband spectral-ripple discrimination tests to assess frequency selectivity have identified a correlation with speech-reception performance in noise (e.g., Davies-Venn et al., 2015; Henry, Turner, & Behrens, 2005; Sheft et al., 2012). These spectral-ripple stimuli are similar to the STM stimuli employed in the current study, except that they lacked

temporal modulation. Won et al. (2015) identified a correlation between STM sensitivity and speech-reception performance for cochlear-implant listeners. As cochlear implants do not relay TFS information, it is likely that this correlation reflects the influence of spectral resolution. For both cochlear-implant and HI listeners, it could be that the more speech-like nature of the spectral-ripple stimulus is critical to the identification of a correlation between frequency resolution and speech scores. Further work comparing STM and static spectral modulation is needed to determine the extent to which the moving aspect of the STM stimulus is important for predicting speech-reception performance.

One of the goals of this study was to determine whether the STM metric could account for individual differences in the effect of the fast compression or noise-reduction hearing-aid algorithms on speech-reception performance. Moore (2008) proposed that individuals with poor TFS processing ability might have difficulty benefitting from fast-acting compression, because of the increased reliance on temporal envelope. This prediction was not borne out in the speech-perception results measured here. Perez, McCormack, and Edmonds (2014) found that a binaural measure of TFS processing ability (interaural time-difference sensitivity, Hopkins & Moore, 2010a) was predictive of subjective reports of both unaided hearing difficulties and the benefit provided by amplification. Further work would be needed to determine whether the STM metric could also predict the subjectively reported benefits provided by different hearing-aid algorithms.

In the current study, the combination of audiometric information and STM sensitivity was able to capture 44% of the variance in speech-in-noise scores for HI listeners fit with hearing aids. This still leaves the other 56% of the variance that still has not been accounted for. It is possible that some of the remaining variance reflects other suprathreshold-distortion processes, such as, for example, frequency selectivity at frequencies above the 2 kHz cutoff frequency of the STM stimulus employed here (Mehraei et al., 2014). Some of the variance might also be attributable to measurement noise in the psychoacoustic or speech-perception tests. Finally, it is also possible that individual differences in cognitive function might play a role in causing the variability in speech scores in noise (Akeroyd, 2008; Foo et al., 2007; Lunner, 2003; Lunner & Sundewall-Thorén, 2007; Neher et al., 2012). One hypothesis is that cognitive effects might be particularly important in predicting speech scores for older listeners or listeners with relatively poor hearing, for whom the STM did not provide much predictive power. The data reported here form part of a larger study that also evaluated cognitive function in addtion to the STM and speech-reception measures reported here; the possible influence of cognitive

effects on speech-reception performance with hearing aids are addressed separately (Rönnberg et al., 2016). In any case, the current results demonstrate that by including the STM sensitivity metric in addition to the audiogram, we have begun to narrow down the proportion of the variance that might be attributable to nonpsychoacoustic factors. While the proportion of the variance explained might grow even larger with the inclusion of certain other psychoacoustic metrics, we are nevertheless now in a better position to begin to differentiate between the contributions of psychoacoustic factors (e.g., modulation detection) and nonpsychoacoustic factors (e.g., cognitive function) in limiting speech-reception performance for an individual listener.

Ultimately, an understanding of the extent to which audibility, suprathreshold distortion, and cognitive factors limit performance for an individual hearing-aid user could inform patient care by guiding the audiologist in counseling the patient or a devising a treatment plan. Standard-of-care audiometric evaluations do an excellent job at diagnosing audibility limitations and remedying these limitations with amplification. Current clinical practice incorporating speech testing combined with audibility (SII) predictions of speech performance can already predict the likely limitations of the hearing aid. However, the current results suggest that the incorporation of an STM test in the audiological test battery has the potential to identify whether the limitation is a suprathreshold encoding issue. If suprathreshold distortion is found to be a major limiting factor, then at least this could be explained to the patient to understand the possible limitations of the hearing aid. If on the other hand, cognitive processing is identified as a likely contributor, then cognitive training might be recommended to improve communication skills.

Knowledge of the particular limitation (suprathreshold distortion vs. cognitive function) that affects speech perception in noise also has the potential to lead to different approaches in the choice of signal-processing strategies for a given patient. One possible interpretation of the correlation observed between STM sensitivity and speech-reception performance is that because speech consists of STM (Chi et al., 1999; Elhilali et al., 2003), internal distortion of the STM information due to hearing loss disrupts the salience of the available speech information. If so, then it is possible that signal-processing solutions that enhance STM in the relevant range of spectral densities and modulation rates might improve speech perception. In a related example, Apoux, Tribut, Debruille, and Lorenzi (2004) showed that the enhancement of certain temporal modulations through expansion could enhance the perception of consonant voicing cues in the speech stimulus. On the other hand, an understanding of the cognitive processing abilities for a given patient could lead to different recommended compression strategies. In contrast to the measure of STM sensitivity that did not correlate to the benefit obtained for different signal-processing strategies in the current study, previous work has suggested that a measure of cognitive function can correlate to the benefit obtained from a fast compression algorithms when listening to speech in modulated backgrounds (Gatehouse, Naylor, & Elberling, 2006; Lunner & Sundewall-Thorén, 2007).

Conclusions

Previous results showed in a small group of HI listeners that a psychoacoustic measure of STM detection performance can account for a significantly greater proportion of the variance in speech-reception performance in noise than the audiogram alone. The current study extends this result to a much larger group of 154 HI listeners fitted with individualized frequency-dependent amplification. Together with the audiogram, the STM sensitivity metric accounted for nearly half of the variance in speech-reception scores with hearing aids. The STM metric was most critical in predicting speech-reception scores for listeners who were less than 65 years old or with high-frequency pure-tone average audiograms better than 53 dB HL. The STM sensitivity metric provides information about psychoacoustic abilities below 2 kHz. Thus, this test could be employed clinically to complement the standard audiogram that correlates to speech-perception for pure-tone frequencies of 2 kHz and above. While the version of this test employed here with two training blocks and two test blocks took an average of 15 min to complete, an analysis suggested that similar predictive power could be obtained by presenting only two training blocks and two test blocks for an average test time of 7 min. STM sensitivity is a fast, easy test that could be employed clinically to identify the extent to which audibility and suprathreshold processing deficits account for poor speech-reception performance in noise for a given hearing-aid user. Any additional impairment not accounted for by the audiogram and STM measures is likely to be attributable to other factors, such as cognitive abilities or suprathreshold processing deficits in other dimensions not captured by the STM metric.

Author Notes

This study was facilitated by a Data Transfer Agreement between Walter Reed National Military Medical Center and

Linköping University. Portions of this work were previously presented at the International Hearing Aid Research Conference, Tahoe City, CA, August 15, 2014, and the 171st Meeting of the Acoustical Society of America, Salt Lake City, UT, May 25, 2016. The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or U.S. Government. Additionally, the identification of specific organization logos, products, or scientific instrumentation does not constitute endorsement, implied endorsement, or preferential treatment on the part of the authors, DoD, or any component agency where inclusion is an integral part of the scientific record.

Declaration of Conflicting Interests

Funding

References

Akeroyd, M. A. (2008). 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. *International Journal of Audiology*, *47*(Suppl 2): S53–S71. doi:10.1080/14992020802301142

Akeroyd, M. A., Arlinger, S., Bentler, R. A., Boothroyd, A., Dillier, N., Dreschler, W. A.,...Kollmeier, B. (2015). International Collegium of Rehabilitative Audiology (ICRA) recommendations for the construction of multilingual speech tests. *International Journal of Audiology*, *54*(Suppl 2): 17–22. doi:10.3109/14992027.2015.1030513

American National Standards Institute. (1997). *Methods for calculation of the speech intelligibility index, S3.5*. New York, NY: Author.

Amieva, H., Ouvrard, C., Giulioli, C., Meillon, C., Rullier, L., Dartigues, J. F. (2015). Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: A 25-year study. *Journal of the American Geriatrics Society*, *63*(10), 2099–2104. doi:10.1111/jgs.13649

Amos, N. E., & Humes, L. E. (2007). Contribution of high frequencies to speech recognition in quiet and noise in listeners with varying degrees of high-frequency sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *50*(4), 819–835. doi:10.1044/1092-4388(2007/057)

Apoux, F., Tribut, N., Debruille, X., & Lorenzi, C. (2004). Identification of envelope-expanded sentences in normal-hearing and hearing-impaired listeners. *Hearing Research*, *189*(1-2), 13–24. doi:10.1016/S0378-5955(03)00397-6

Bernstein, J. G. W., Mehraei, G., Shamma, S., Gallun, F. J., Theodoroff, S. M., Leek, M. R. (2013). Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners. *Journal of the American Academy of Audiology*, *24*(4), 293–306. doi:10.3766/jaaa.24.4.5

Bernstein, J. G. W., Summers, V., Grassi, E., & Grant, K. W. (2013). Auditory models of suprathreshold distortion and speech intelligibility in persons with impaired hearing. *Journal of the American Academy of Audiology*, *24*(4), 307–328. doi:10.3766/jaaa.24.4.6.

Boldt, J., Kyems, U., Pedersen, M. S., Lunner, T., & Wang, D. L. (2008). Estimation of the ideal binary mask using directional systems. In *Proceedings of the 11th International Workshop on Acoustic Echo and Noise Control*. Seattle, WA.

Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *Journal of the Acoustical Society of America*, *111*(6), 2801–2810. doi:10.1121/1.1479152.

Buss, E., Hall, J. W., & Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear and Hearing*, *25*(3), 242–250. doi:10.1097/01.AUD.0000130796.73809.09

Buus, S., & Florentine, M. (2001). Modification to the power function for loudness. In E. Sommerfield, R. Kompass, & T. Lachmann (Eds.), *Fechner day 2001*. Berlin, Germany: Pabst.

Chi, T., Gao, Y., Guyton, M. C., Ru, P., & Shamma, S. (1999). Spectro-temporal modulation transfer functions and speech intelligibility. *Journal of the Acoustical Society of America*, *106*(5), 2719–2732. doi:10.1121/1.428100

Ching, T. Y. C., Dillon, H., Lockhart, F., van Wanrooy, E., & Flax, M. (2011). Audibility and speech intelligibility revisited: Implications for amplification. In T. Dau, M. L. Jepsen, T. Poulsen, & J. C. Dalsgaard (Eds.), *Speech perception and auditory disorders* (pp. 11–19). Ballerup, Denmark: Danavox Jubilee Foundation.

Davies-Venn, E., Nelson, P., & Souza, P. (2015). Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing. *The Journal of the Acoustical Society of America*, *138*(1), 492–503. doi:10.1121/1.4922700

Dreschler, W. A., & Plomp, R. (1985). Relation between psychophysical data and speech perception for hearing-impaired subjects. II. *The Journal of the Acoustical Society of America*, *78*(4), 1261–1270. doi:10.1121/1.392895

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2003). Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing. *The Journal of the Acoustical Society of America*, *113*(4 Pt 1), 2084–2094. doi:10.1121/1.1555611

Elhilali, M., Chi, T., & Shamma, S. A. (2003). A spectro-temporal modulation index (STMI) for assessment of speech intelligibility. *Speech Communication*, *41*(2–3), 331–348. doi:10.1016/S0167-6393(02)00134-6

Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology*, *18*(7), 618–631. doi:10.3766/jaaa.18.7.8

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings – 2. Patterns of candidature. *International Journal of Audiology*, *45*(3), 153–171. doi:10.1080/14992020500429484

George, E. L. J., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *120*(4), 2295–2311. doi:10.1121/1.2266530

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, *79*(4), 1020–1033. doi:10.1121/1.393374

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1–2), 103–138. doi:10.1016/0378-5955(90)90170-T

Gnansia, D., Péan, V., Meyer, B., & Lorenzi, C. (2009). Effects of spectral smearing and temporal fine structure degradation on speech masking release. *The Journal of the Acoustical Society of America*, *125*(6), 4023–4033. doi:10.1121/1.3126344

Grant, K. W., Bernstein, J. G. W., & Summers, V. (2013). Predicting speech intelligibility by individual hearing-impaired listeners: The path forward. *Journal of the American Academy of Audiology*, *24*(4), 329–336. doi:10.3766/jaaa.24.4.7

Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: Effects of lexical and prosodic factors. *Neuropsychologia*, *41*(8), 1008–1019. doi:10.1016/S0028-3932(02)00315-9

Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and Hearing*, *31*(6), 755–760. doi:10.1097/AUD.0b013e3181e627e7

Hacker, M. J., & Ratcliff, R. (1979). A revised table of d' for M-alternative forced choice. *Perception and Psychophysics*, *26*(2), 168–170.

Hagerman, B., & Kinnefors, C. (1995). Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. *Scandinavian Audiology*, *24*(1), 71–77. doi:10.3109/14992029509042213

Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America*, *118*(2), 1111–1121. doi:10.1121/1.1944567

Hopkins, K., & Moore, B. C. J. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *The Journal of the Acoustical Society of America*, *122*(2), 1055–1068. doi:10.1121/1.2749457

Hopkins, K., & Moore, B. C. J. (2010a). Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International Journal of Audiology*, *49*(12), 940–946. doi:10.3109/14992027.2010.512613

Hopkins, K., & Moore, B. C. J. (2010b). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *127*(3), 1595–1608. doi:10.1121/1.3293003

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, *130*(1), 334–349. doi:10.1121/1.3585848

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology*, *18*(7), 590–603. doi:10.3766/jaaa.18.7.6

Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & Lopez-Poveda, E. A. (2016). The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners. *Trends in Hearing*, *20*, pii: 2331216516641055. doi:10.1177/2331216516641055

Johnson, E. E. (2013). Modern prescription theory and application: Realistic expectations for speech recognition with hearing aids. *Trends in Hearing*, *17*(3), 143–170. doi:10.1177/1084713813506301

Kochkin, S. (2000). MarkeTrak V: ''Why my hearing aids are in the drawer'': The consumers' perspective. *The Hearing Journal*, *53*(2), 34–41.

Kortlang, S., Mauermann, M., & Ewert, S. D. (2016). Suprathreshold auditory processing deficits in noise: Effects of hearing loss and age. *Hearing Research*, *331*, 27–40. doi:10.1016/j.heares.2015.10.004

Le Goff, N. (2015). *Amplifying soft sounds—A personal matter* [White paper]. Retrieved from Oticon A/S http://www.oticon.global/professionals/evidence

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2 Part 2), 467–477. doi:10.1121/1.1912375

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–18869. doi:10.1073/pnas.0607364103

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology*, *42*(s1), 49–58. doi:10.3109/14992020309074624

Lunner, T., Rudner, M., & Rönnberg, J. (2009). Cognition and hearing aids. *Scandinavian Journal of Psychology*, *50*(5), 395–403. doi:10.1111/j.1467-9450.2009.00742.x

Lunner, T., & Sundewall-Thorén, E. (2007). Interactions between cognition, compression, and listening conditions: Effects on speech-in-noise performance in a two-channel hearing aid. *Journal of the American Academy of Audiology*, *18*(7), 604–617. doi:10.3766/jaaa.18.7.7

Mehraei, G., Gallun, F. J., Leek, M. R., & Bernstein, J. G. W. (2014). Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility. *The Journal of the Acoustical Society of America*, *136*(1), 301–316. doi:10.1121/1.4881918

Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical considerations and the role of individual differences. *Trends in Amplification*, *12*(2), 103–112. doi:10.1177/1084713808317819

Moore, B. C. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., & Hopkins, K. (2012). The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies. *The Journal of the Acoustical Society of America*, *131*(2), 1003–1006. doi:10.1121/1.3672808

Moore, B. C. J., & Sek, A. (1996). Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *The Journal of the Acoustical Society of America*, *100*(4), 2320–2331. doi:10.1121/1.417941

Moore, B. C. J., & Skrodzka, E. (2002). Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation. *The Journal of the Acoustical Society of America*, *111*(1 Pt 1), 327–335. doi:10.1121/1.1424871

Neher, T., Lunner, T., Hopkins, K., & Moore, B. C. J. (2012). Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *131*(4), 2561–2564. doi:10.1121/1.3689850

Nelson, D. A., Schroder, A. C., & Wojtczak, M. (2001). A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *110*(4), 2045–2064. doi:10.1121/1.1404439

Perez, E., McCormack, A., & Edmonds, B. A. (2014). Sensitivity to temporal fine structure and hearing-aid outcomes in older adults. *Frontiers in Neuroscience*, *8*, 7. doi:10.3389/fnins.2014.00007

Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *Journal of Speech and Hearing Research*, *29*(2), 146–154.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, *42*(Suppl 1), S68–S76. doi:10.3109/14992020309074626

Rönnberg, J., Lunner, T., Ng, E. H. N., Lidestam, B., Zekveld, A. A., Sörqvist, P., Stenfelt, S. (2016). Hearing impairment, cognition and speech understanding: Exploratory factor analyses of a comprehensive test battery for a group of hearing aid users, the n200 study. *International Journal of Audiology*, in press. doi:10.1080/14992027.2016.1219775

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*(July), 31doi:10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, *47*(Suppl 2): S99–S105. doi:10.1080/14992020802301167

Rudner, M., Foo, C., Rönnberg, J., & Lunner, T. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology*, *50*(5), 405–418. doi:10.1111/j.1467-9450.2009.00745.x

Rudner, M., Rönnberg, J., & Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology*, *22*(3), 156–167. doi:10.3766/jaaa.22.3.4

Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., & Farrell, C. (2012). Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. *Ear and Hearing*, *33*(6), 709–720. doi:10.1097/AUD.0b013e31825aab15

Smoorenburg, G. F. (1992). Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *The Journal of the Acoustical Society of America*, *91*(1), 421–437. doi:10.1121/1.402729

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex*, *22*(5), 1180–1190. doi:10.1093/cercor/bhr196

Souza, P. E., Arehart, K. H., Shen, J., Anderson, M., & Kates, J. M. (2015). Working memory and intelligibility of hearing-aid processed speech. *Frontiers in Psychology*, *6*(May), 526doi:10.3389/fpsyg.2015.00526

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251. doi:10.1037/0033-2909.87.2.245

Stone, M. A., & Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *The Journal of the Acoustical Society of America*, *123*(2), 1063–1076. doi:10.1121/1.2821969

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *The Journal of the Acoustical Society of America*, *125*(5), 3328–3345. doi:10.1121/1.3097469

Summers, V., Makashay, M. J., Theodoroff, S. M., & Leek, M. R. (2013). Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners. *Journal of the American Academy of Audiology*, *24*(4), 274–292. doi:10.3766/jaaa.24.4.4

ter Keurs, M., Festen, J. M., & Plomp, R. (1993). Limited resolution of spectral contrast and hearing loss for speech in noise. *The Journal of the Acoustical Society of America*, *94*(3 Pt 1), 1307–1314. doi:10.1121/1.408158

Wang, D., Kjems, U., Pedersen, M. S., Boldt, J. B., & Lunner, T. (2009). Speech intelligibility in background noise with ideal binary time-frequency masking. *The Journal of the Acoustical Society of America*, *125*(4), 2336–2347. doi:10.1121/1.3083233

Won, J. H., Moon, I. J., Jin, S., Park, H., Woo, J., Cho, Y.-S., Chung, W.-H., ... Hong, S. H. (2015). Spectrotemporal modulation detection and speech perception by cochlear implant users. *PLoS One*, *10*(10), 1–24. doi:10.1371/journal.pone.0140920