

Influence of Instantaneous Compression on Recognition of Speech in Noise with Temporal Dips

DOI: 10.3766/jaaa.16165

Daniel M. Rasetshwane*

David A. Raybaine*

Judy G. Kopun*

Michael P. Gorga*

Stephen T. Neely*

Abstract

Background: In listening environments with background noise that fluctuates in level, listeners with normal hearing can “glimpse” speech during dips in the noise, resulting in better speech recognition in fluctuating noise than in steady noise at the same overall level (referred to as masking release). Listeners with sensorineural hearing loss show less masking release. Amplification can improve masking release but not to the same extent that it does for listeners with normal hearing.

Purpose: The purpose of this study was to compare masking release for listeners with sensorineural hearing loss obtained with an experimental hearing-aid signal-processing algorithm with instantaneous compression (referred to as a suppression hearing aid, SHA) to masking release obtained with fast compression. The suppression hearing aid mimics effects of normal cochlear suppression, i.e., the reduction in the response to one sound by the simultaneous presentation of another sound.

Research Design: A within-participant design with repeated measures across test conditions was used.

Study Sample: Participants included 29 adults with mild-to-moderate sensorineural hearing loss and 21 adults with normal hearing.

Intervention: Participants with sensorineural hearing loss were fitted with simulators for SHA and a generic hearing aid (GHA) with fast (but not instantaneous) compression (5 ms attack and 50 ms release times) and no suppression. Gain was prescribed using either an experimental method based on categorical loudness scaling (CLS) or the Desired Sensation Level (DSL) algorithm version 5a, resulting in a total of four processing conditions: CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA.

Data Collection: All participants listened to consonant-vowel-consonant nonwords in the presence of temporally-modulated and steady noise. An adaptive-tracking procedure was used to determine the signal-to-noise ratio required to obtain 29% and 71% correct. Measurements were made with amplification for participants with sensorineural hearing loss and without amplification for participants with normal hearing.

Analysis: Repeated-measures analysis of variance was used to determine the influence of within-participant factors of noise type and, for participants with sensorineural hearing loss, processing condition on masking release. Pearson correlational analysis was used to assess the effect of age on masking release for participants with sensorineural hearing loss.

Results: Statistically significant masking release was observed for listeners with sensorineural hearing loss for 29% correct, but not for 71% correct. However, the amount of masking release was less than masking release for participants with normal hearing. There were no significant differences among the amplification conditions for participants with sensorineural hearing loss.

Conclusions: The results suggest that amplification with either instantaneous or fast compression resulted in similar masking release for listeners with sensorineural hearing loss. However, the masking release was less for participants with hearing loss than it was for those with normal hearing.

*Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE

Corresponding author: Daniel Rasetshwane, Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE; Email: rasetshwane@boystown.org

Key Words: aided audibility index, hearing aids, instantaneous compression, masking release

Abbreviations: AAI = aided audibility index; ANOVA = analysis of variance; CLS = categorical loudness scaling; DPOAE = distortion-product otoacoustic emission; DSL = Desired Sensation Level; ERB = equivalent rectangular bandwidth; GHA = generic hearing aid; HASPI = Hearing Aid Speech Perception Index; HASQI = Hearing Aid Speech Quality Index; HL = hearing level; HSD = honestly significant difference; NH = normal hearing; SD = standard deviation; SE = standard error; SHA = suppression hearing aid; SNHL = sensorineural hearing loss; SNR = signal-to-noise ratio; SPL = sound pressure level; STC = suppression tuning curve; WDRC = wide dynamic range compression

INTRODUCTION

In listening environments with background noise that fluctuates in level, listeners with normal hearing (NH) can “glimpse” speech during dips in the noise and achieve better recognition than for speech in steady noise at the same level (Peters et al, 1998; Hall et al, 2012). Listeners with sensorineural hearing loss (SNHL) are less able to improve their speech recognition in fluctuating noise (Festen and Plomp, 1990; Eisenberg et al, 1995; Peters et al, 1998; Moore et al, 1999; George et al, 2006; Hall et al, 2012; Gregan et al, 2013). Current hearing aids provide only marginal improvement in speech perception in fluctuating noise (Peters et al, 1998; Moore et al, 1999; George et al, 2006; Jin and Nelson, 2006; Brennan et al, 2016). The purpose of this study was to compare masking release for listeners with SNHL obtained with an experimental hearing-aid signal-processing algorithm having instantaneous compression (referred to as a suppression hearing aid, SHA; Rasetshwane, Gorga, et al, 2014) to masking release obtained with a simulation of a generic hearing aid (GHA) having fast compression. In this study, we focused on the instantaneous compression of the SHA.

By glimpsing speech during the dips, listeners with NH achieve better speech perception in fluctuating noise than in steady noise at identical signal-to-noise ratios (SNRs). The better performance in fluctuating noise than in steady noise is referred to as *masking release*. Listeners with NH are able to achieve masking release for maskers with dips in the time domain, frequency domain, and both time and frequency domains (Peters et al, 1998; Hall et al, 2012). Masking release was higher for maskers that fluctuated in level in both the time and frequency domains compared to maskers that fluctuated only in one domain (Peters et al, 1998; Moore et al, 1999; Hall et al, 2012). When speech or multitalker babble was used as a masker with the fluctuating level, the amount of masking release depended on the number of background talkers and the material (e.g., sentences, isolated words) used as the target speech (Freyman et al, 2004; Simpson and Cooke, 2005; Rosen et al, 2013). In studies evaluating the influence of the number of competing talkers, Freyman et al (2004) obtained the largest masking release with two talkers when using nonsense sentences, whereas Simpson and Cooke (2005) obtained the largest masking release with

eight talkers in a consonant identification task. The differences may reflect the relative contributions of factors such as informational masking, energetic masking, and modulation masking, as well as contributions of cognitive and linguistic skills to speech perception. Informational masking occurs when a listener is unable to disentangle the elements of the target signal from a similar-sounding masker signal (Brungart, 2001). Energetic masking occurs when portions of the target signal are inaudible because the masker signal has energy in the same critical bands and at the same time as the target signal (Brungart, 2001). Modulation masking occurs when a listener is unable to detect or discriminate modulations of the target signal because there is overlap in the modulation rates of the target and masker signals (Fogerty et al, 2016). Experimental control of the temporal characteristics of the masker is achieved through modulation of broadband noise signals, whereas control of spectral characteristics is achieved using filtering (Moore et al, 1999; Füllgrabe et al, 2006; Stone et al, 2011). Under certain conditions, for example, when the target speech is degraded to reduce the redundancy of speech information (Kwon and Turner, 2001) and when modulation masking is eliminated (Stone and Moore, 2014), fluctuating maskers can degrade speech perception.

Listeners with SNHL have greater difficulty understanding speech in noise than listeners with NH (Festen and Plomp, 1990; Eisenberg et al, 1995; Peters et al, 1998; Moore et al, 1999; George et al, 2006; Hall et al, 2012; Gregan et al, 2013). The difference in speech perception between listeners with SNHL and NH is larger when the competing noise is fluctuating (Duquesnoy, 1983; Peters et al, 1998) than when the noise is steady and does not have distinct spectral or temporal fluctuations (Plomp, 1994; Peters et al, 1998). Although the reasons why listeners with SNHL show less masking release are not well understood, listeners with SNHL have reduced audibility, reduced frequency selectivity, reduced temporal resolution, and abnormal response growth (e.g., loudness recruitment), all of which contribute to the difficulties experienced by these listeners. Studies have shown that elevated audiometric thresholds lead to reduced masking release (George et al, 2006; Gregan et al, 2013). However, reduced audibility did not fully account for reduced masking release when audibility was disassociated from suprathreshold deficits (deteriorations in speech intelligibility not accounted for by audibility) through simulation of hearing loss in listeners

with NH using masking noise (Bacon et al, 1998; George et al, 2006). Reduced temporal resolution adversely affects masking release (Jin and Nelson, 2006; George et al, 2006; Gregan et al, 2013). George et al (2006) found that temporal resolution, measured using a procedure in which listeners reported the number of tone sweeps they were able to detect within a temporal window, was correlated with the amount of masking release for their participants with SNHL. Similarly, Gregan et al (2013) observed a small but significant correlation between masking release and temporal resolution estimated from the slope of off-frequency temporal masking curves (Nelson et al, 2001) for participants with SNHL. A relationship between masking release and temporal resolution has been observed by others (Festen, 1993; Dubno et al, 2003). However, Jin and Nelson (2006) observed a relationship between masking release and temporal resolution estimated from the recovery of forward masking for consonant–vowel stimuli but not for sentences. Loss of cochlear compression did not explain reduced masking release in listeners with SNHL (Gregan et al, 2013). There are conflicting data regarding the effect of spectral resolution on masking release. Some studies reported less masking release with reduced spectral resolution (Baer and Moore, 1993; 1994; Nelson and Jin, 2004; Xu et al, 2005), whereas other studies did not observe this relationship (ter Keurs et al, 1993a,b; George et al, 2006). On the other hand, age is a predictor of masking release for listeners with SNHL after audiometric thresholds are taken into account (Gustafsson and Arlinger, 1994; Snell et al, 2002; Dubno et al, 2003; George et al, 2006), but not for listeners with NH (Füllgrabe et al, 2015).

Recently, Bernstein and colleagues suggested that listeners with SNHL did not show less masking release when compared with listeners with NH at the same SNR (Bernstein and Grant, 2009; Bernstein and Brungart, 2011). They showed that masking release was influenced by the SNR associated with the steady noise baseline condition and that differences in masking release between listeners with SNHL and NH are smaller when they are compared at a low SNR (<0 dB) for steady noise.

To some extent, hearing aid amplification can improve audibility and other processes that are diminished when SNHL is present, such as compression. Other processes, such as frequency selectivity, cannot be improved through amplification. Studies that have evaluated the effects of amplification on masking release have demonstrated limited improvements (Peters et al, 1998; Moore et al, 1999; George et al, 2006; Jin and Nelson, 2006; Brennan et al, 2016). Similar to results for listeners with NH, masking release was greater for noise with both spectral and temporal dips than for noise with only spectral or temporal dips (Moore et al, 1999; Hall et al, 2012). However, masking release for participants with SNHL even with amplification was less than unaided masking release for listeners with NH. When wide dynamic range

compression (WDRC) was compared to linear amplification, WDRC resulted in greater masking release (Moore et al, 1999). Unlike linear amplification that provides gain that is independent of the input sound level, WDRC hearing aids provide higher gain for low-level sounds, increasing the probability that such sounds are audible, and lower gain for high-level sounds, avoiding loudness discomfort. Compression speed, that is, the rate at which a hearing aid adjusts gain in response to changes in input level should affect the ability to listen in the dips. In theory, faster acting compression should improve the ability to listen in the dips of competing fluctuating noise by increasing the gain during the dips and providing access to speech. In fact, Brennan et al (2016) demonstrated slightly greater masking release with fast compression (5 msec attack time, 50 msec release time) than with slow compression (150 msec attack time, 1,500 msec release time) for some listeners with hearing loss.

The present study compared masking release for listeners with SNHL obtained with SHA, an experimental hearing aid signal-processing algorithm with instantaneous compression (Rasetshwane, Gorga, et al, 2014), with masking release obtained with fast compression. In addition to having instantaneous compression, the fitting rationale for SHA is to restore normal growth of loudness. SHA also simulates the effect of normal cochlear suppression, which it does by reducing the gain for a particular frequency component when other competing frequency components are present in a signal. Suppression is a nonlinear property of healthy cochleae in which basilar-membrane displacement, neural firing rate, or otoacoustic emission magnitude in response to one tone is reduced by the simultaneous presentation of other tones at different frequencies. Suppression plays an important role in the processing of complex stimuli, such as speech (Houtgast, 1974; Sachs and Young, 1980). Suppression is reduced when SNHL occurs. WDRC that simulates the effect of normal suppression, in addition to compression, may improve speech recognition over WDRC that includes compression, but does not include suppression. To restore normal growth of loudness, the gain prescription for SHA uses individual measurements of categorical loudness scaling (CLS; Brand and Hohmann, 2002; Rasetshwane et al, 2015) with the goal of providing gain such that loudness judgments provided by a listener with SNHL match those provided by listeners with NH, for a given sound pressure level (SPL). Henceforth, we refer to this gain prescription strategy as CLS. The current study focuses on evaluation of the instantaneous compression of SHA. Our rationale for including instantaneous compression in SHA was that it would allow SHA to react instantaneously to changes in noise level, amplifying speech during dips and thus improving the ability to listen in the dips. The effectiveness of SHA at restoring the ability to listen in the dips was compared with that for a GHA algorithm representative

of technology that is currently implemented in hearing aids. The GHA used WDRC with fast compression (5 and 50 msec attack and release times) but without suppression. Gain for the GHA was prescribed using Desired Sensation Level (DSL Version 5a; Scollie et al, 2005), an algorithm that is currently used in clinical settings. To separately evaluate effects of hearing aid signal-processing algorithm and gain-prescription strategy on masking release, we also fitted SHA using the DSL, and GHA using CLS, resulting in four processing conditions: CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA. Measurements were made without amplification for listeners with NH to provide a reference for the amount of masking release. Because it has been suggested that instantaneous compression can result in non-linear distortion that, in turn, diminishes speech quality (Herzke and Hohmann, 2005), we evaluated distortions for the hearing aid simulators.

METHODS

Participants

Fifty native English-speaking adults were tested. Participants were recruited from a database of potential research participants that is maintained at Boys Town National Research Hospital. There were 29 (13 female) participants with SNHL (mean age = 64, standard deviation (SD) = 12, range = 25–79 yr) and 21 (15 female) participants with NH (mean age = 33, SD = 12, range = 19–57 yr). Twenty one of the participants with SNHL wore hearing aids. For all participants, pure-tone air-conduction audiometric thresholds were measured at octave frequencies from 0.25 to 8 kHz and at interoctave frequencies of 3 and 6 kHz. Additional air-conduction thresholds were measured at interoctave frequencies of 0.75 and 1.5 kHz for participants with SNHL when thresholds at consecutive frequencies in this range differed by more than 20 dB. Bone-conduction thresholds were measured at octave frequencies from 0.25 to 4 kHz when air-conduction thresholds exceeded 15 dB hearing level (HL). Thresholds were measured in 5-dB steps, following standard clinical procedures (ASLHA, 2005). Participants with air-conduction thresholds ≤ 15 dB HL at all frequencies were considered to have NH. Participants with air-conduction thresholds > 15 dB HL at one or more test frequencies were considered to have SNHL. Participants with SNHL had bilateral hearing loss, with average thresholds at 2, 3, and 4 kHz that differed by < 15 dB between ears. Participants were excluded from the study if air-bone gaps were > 10 dB at any frequency. This study was conducted under an approved Institutional Review Board protocol and informed consent was obtained from all participants. Figure 1 shows the audiometric thresholds for the participants with SNHL in the form of box and whisker plots.

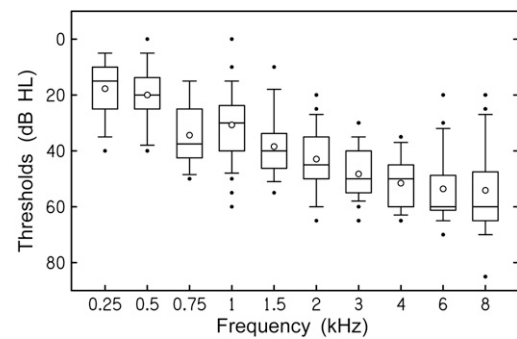


Figure 1. Audiometric thresholds for the test ear for participants with SNHL. Boxes represent the interquartile range and whiskers represent the 10th and 90th percentiles. Outliers, defined as data points that are outside the 10th to 90th percentile range, are plotted using filled circles. For each box, lines represent the median and open circles represent the mean. This convention is used in the remaining box-and-whisker plots.

Stimuli for loudness and speech-recognition measurements were presented monaurally. If both ears met the inclusion criteria for participants with NH, the ear with better hearing was selected for testing. For participants with SNHL, the test ear was randomly selected if both ears met the inclusion criteria. Data were collected from 16 right and 13 left ears with SNHL and 13 right and eight left ears with NH.

Equipment

A 24-bit sound card (Babyface, RME, Germany) was used to generate stimuli. Sampling rates of 48000 and 22050 Hz were used to generate the stimuli for CLS and speech-recognition measurements, respectively. The stimuli were presented to the participant's test ear with a headphone (HD-25-1 II, Sennheiser, Ireland). MATLAB software (MathWorks, Natick, MA) was used to control stimulus delivery and to record the responses of the participants for both CLS and speech-recognition measurements.

Hearing Aid Signal Processing and Fitting Strategies

The hearing aid simulators were implemented using MATLAB software. Brief descriptions of GHA and SHA are provided below, as are descriptions of the DSL gain-prescription strategy and the CLS measurement procedure and gain-prescription strategy.

Generic Hearing Aid

An eight-channel hearing aid simulator with attack and release times of 5 and 50 msec, respectively, served as the GHA. This hearing aid simulator included a filter bank, compression circuit, and synthesis stage. The filter bank had overlapping channels with center frequencies

and, in parenthesis, cutoff frequencies (−6 dB) of 0.25 (0, 0.36), 0.4 (0.28, 0.56), 0.63 (0.42, 0.81), 1 (0.78, 1.24), 1.6 (1.24, 1.99), 2.5 (1.99, 3.16), 4 (3.17, 5.03), and 6.3 (5.03, 11.03) kHz (ANSI, 2004). The upper frequency of amplification was limited to the Nyquist frequency (11.025 kHz) by an anti-aliasing filter. The bandwidth of 11.025 kHz is wider than the typical bandwidth (about 4–5 kHz) of hearing aids fitted in the clinic. The output of each channel was fed to a compressor that featured two knee points.

Nonlinear compressive amplification, using compression ratios specified by the gain prescription, was used for input levels between the two knee points. Linear amplification was used for input levels below the lower knee point. Input levels above the upper knee point were compressed using a compression ratio of 10:1. The upper knee point was set to 105 dB SPL. The lower knee point, compression ratio, and the gain were set using DSL or the CLS-based gain prescription method, and depended on a listener's audiometric thresholds. The processing delay of the simulator varied by less than 1 msec across frequency. This simulator has been used in several studies, often paired with the DSL gain prescription (McCreery et al, 2013; Alexander and Masterson, 2015; Brennan et al, 2016). For an extended description of this simulator, see McCreery et al (2013).

Suppression Hearing Aid

The SHA signal-processing algorithm includes three stages—analysis, suppression, and synthesis. The analysis stage uses a complex Gammatone filter bank to decompose the input into 31 channels that span the frequency range up to 12 kHz (Hohmann, 2002). The outputs of the filter bank are complex signals, which allow calculation of the instantaneous time domain level. Filters below 1 kHz were designed to have a constant equivalent rectangular bandwidth (ERB) of 0.1 kHz and linearly spaced center frequencies (f_c) with 0.1 kHz steps. The filter at 1 kHz had an ERB of 0.11 kHz. Filters above 1 kHz had constant tuning ($Q_{\text{ERB}} = 8.65$) and center frequencies that were logarithmically spaced with 1/6-octave steps. Q_{ERB} is defined as $f_c/\text{ERB}(f_c)$ and $\text{ERB}(f_c)$ is the ERB of the filter with center frequency f_c . This tuning resulted in a filter bank delay that varied by approximately 4 msec across frequency, which is within limits that are acceptable to hearing aid users (Stone et al, 2008).

SHA provides instantaneous compression with gain that implements two-tone suppression. The suppressive influence of one frequency on another frequency was based on measurements of distortion-product otoacoustic emission (DPOAE) suppression tuning curves (STCs) in humans with NH (Gorga et al, 2011b). In these experiments, DPOAEs were elicited by a pair of primary tones (f_1 and f_2), whose levels were held constant while a third tone (f_3) was presented as a suppressor (Gorga et al, 2011a,

b). The effect of the suppressor tone was defined as the amount by which its presence reduced the DPOAE level in response to the primary tones. This suppressive effect was characterized for a range of levels and frequencies of both the suppressor and primary tones. The gain applied to a particular frequency band is time varying and is based on the instantaneous level of every filter bank output in a manner based on measurements of DPOAE STCs. However, unlike in DPOAE suppression measurements where the suppressive effect of a suppressor frequency (f_3) on the DPOAE level in response to two primary tones (f_1 and f_2) was represented, the SHA signal processing represents the influence of a suppressor frequency (f_s) which is equivalent to f_3 on a probe frequency (f_p) which is equivalent to (f_2). The suppressive effect was extended to multiple suppressors by assuming that suppression is additive in the intensity domain. This assumption is a simplification and might not describe the ways in which suppressive effects add for all stimulus conditions (see Sieck et al, 2016). Suppose that the total suppressive influence on a tone at f of multiple suppressor tones at f_j can be described by

$$L_s(f) = 10 \cdot \log_{10} \left(\sum_{j=1}^N 10^{S_j(f)/10} \right), \quad (1)$$

where

$$S_j(f) = c_1(f, f_j) + c_2(f, f_j) \cdot L_j, \quad (2)$$

represents the individual suppressive level on a tone at f of a single suppressor tone at f_j , and L_j is the filter output level at f_j . The “total suppressive influence” combines the suppressive effect of all frequency components into a single, equivalent level L_s that would cause the same reduction in gain (due to compression) if it was the level of a single tone. Coefficients c_1 and c_2 are derived from the DPOAE data. The output level of each channel was limited to 110 dB SPL. Lastly, the synthesis stage combined the output of the suppressor stage across channels, producing an output signal with suppressive influences. Please see Rasetschwane, Gorga, et al (2014) for more information on SHA.

The ability of the SHA to mimic suppressive effects is demonstrated in Figure 2 in which DPOAE STC for $f_2 = 1, 2, 4$, and 8 kHz from Gorga et al (2011b) are compared with the output of the SHA for f_p at the same frequencies. In the DPOAE study, the level of f_2 (i.e., L_2) ranged from 10 dB SL (lowest STC) to 60-dB SL (highest STC) in 10-dB steps. The level of f_1 (L_1) was determined empirically and individually, using a paradigm in which both L_1 and L_2 were continuously varied, resulting in a Lissajous pattern of L_d (see Neely et al, 2005 for a description of the Lissajous procedure). For each combination of f_2 and L_2 , the suppressor frequency (f_3) was varied from about 1 or 2 octaves below f_2 to about 1/4–1/2 octave above f_2 . STCs represent the level of

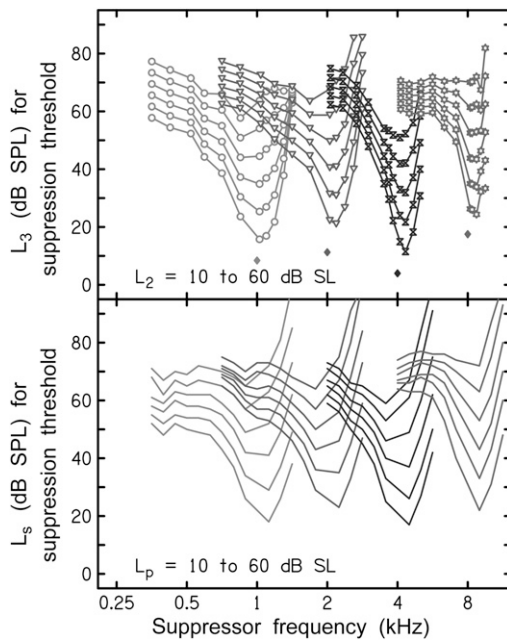


Figure 2. Comparison of measurements of DPOAE STCs of Gorga et al (2011) to the SHA simulation of STCs. The top panel shows DPOAE STC measurements for $f_2 = 1$ (circles), 2 (triangles), 4 (hourglasses), and 8 kHz (stars). The unconnected symbols below each set of STCs represent the mean behavioral thresholds for the group of participants contributing data at that frequency. All of these data came from participants with NH. STCs produced by the SHA are shown in the bottom panel. Adapted from Gorga et al (2011) and Rasetshwane, Gorga, et al (2014).

the suppressor L_3 at the suppression threshold (which was defined as a 3-dB reduction in the DPOAE level caused by the suppressor tone). Parameter values for the SHA signal processing were selected to match corresponding values for DPOAE STCs. That is, probe tone frequencies were selected to match f_2 , probe level were selected to match L_2 , and the suppression threshold was defined as the level of the suppressor tone that caused a 3-dB reduction in the level on the probe. Figure 2 shows that the STCs produced by the SHA are qualitatively similar to measurements of DPOAE STCs of Gorga et al (2011b). The SHA STCs are similar to the DPOAE STCs in both their absolute level and their dependence on probe-tone level.

Rasetshwane, Gorga, et al (2014) demonstrated that the SHA results in enhancement of spectral contrasts (see their Figures 10 and 11), which has the potential to improve speech perception in the presence of background noise (Turicchia and Sarpeshkar, 2005; Oxenham et al, 2007). However, as stated in the Introduction, the aim of this study was to evaluate speech recognition with the instantaneous compression of the SHA relative to fast compression, and not to evaluate the suppression feature of the SHA. A reader interested in the suppression feature of the SHA is referred to Rasetshwane, Gorga, et al (2014) for further details.

DSL-Based Gain Prescription

DSL gain prescription was based on the adult algorithm of DSL version 5a (Scollie et al, 2005), with frequency-dependent gain, compression ratio, and target sensation level determined from audiometric thresholds. DSL parameters at 8 kHz were set to be the same as those at 6 kHz because DSL version 5a does not prescribe gain at 8 kHz.

CLS Measurements

The loudness-based gain prescription strategy used individualized measurements of CLS with pure tones (1,000 msec duration with 20 msec rise/fall time) at the same frequencies at which audiometric thresholds were measured (Rasetshwane, Gorga, et al, 2014; Rasetshwane et al, 2015). An adaptive procedure was used to determine the levels of pure tones that corresponded to different loudness categories (Brand and Hohmann, 2002; ISO 16832, 2006; Al-Salim et al, 2010; Rasetshwane et al, 2013; 2015). The response scale included 11 loudness categories, seven of which were assigned textual labels (“can’t hear,” “very soft,” “soft,” “medium,” “loud,” “very loud,” and “too loud”). The categories were displayed using colored horizontal bars with increasing length from the softest to the loudest descriptor. Participants used a computer mouse to select the category that best matched their loudness perception and were encouraged to use both labeled and unlabeled loudness categories. The CLS procedure included two stages. In the first stage, the dynamic range of the participant was determined by presenting two sequences of stimuli, one sequence ascending in level and the other descending in level. The lower end of a participant’s dynamic range was based on the last audible level of the descending sequence, whereas the upper end was based on the last level of the ascending sequence that was not judged as “too loud.” The starting presentation level was fixed at 60-dB SPL for participants with NH. For participants with SNHL, the starting level was half way between their audiometric threshold and the maximum presentation level (110 dB SPL). In the second stage, stimuli were presented in random order at 18 equally spaced levels within the participant’s dynamic range. The second stage of the CLS procedure was repeated three times, with the dynamic range adjusted for each subsequent repetition based on the responses of the participant. Thus, the procedure iteratively adapted the level range to the participant’s responses. The data from the three repetitions were analyzed to remove outliers and to create a CLS function (i.e., loudness category as a function of level in dB SPL). Measurements were made separately for each frequency, and data collection took approximately 5 min for each frequency. A practice run at 1.25 kHz served to familiarize participants with the procedure. For more

details on the CLS measurement and analysis procedures, please see Rasetshwane et al (2013, 2015).

CLS-Based Gain Prescription

The aim of the CLS gain-prescription strategy is to restore loudness growth. In this strategy, average CLS data for listeners with NH provide reference input levels for a given loudness category and frequency. The gain required for an individual with SNHL is the difference between the normal-reference input level and the input level required by that individual to elicit the same loudness percept. Average CLS data for listeners with NH were based on previous data from 30 participants collected using pure-tone stimuli at the same octave and interoctave frequencies as those used in the present study (Rasetshwane, Brennan, et al, 2014). Gain was prescribed to restore loudness for categories of “very soft,” “medium,” and “very loud.” Calculations of gain were performed separately for each frequency. Thus, the resulting gain was both frequency and level dependent. The gain function for the CLS prescription strategy had four knee points. Levels associated with loudness categories of “very soft” and “very loud” corresponded to the lower and upper knee points of the gain function. Linear gain was applied below the lower knee point and above the upper knee point, whereas compressive gain was applied between these knee points. A mid-level knee point, corresponding to the loudness category “medium” was included to provide a characterization of loudness growth that resembled the two-segment CLS function, that has been observed for both listeners with NH and SNHL (Brand and Hohmann, 2002; Al-Salim et al, 2010). The output level was limited to 110 dB SPL.

Output Level and Gain Analysis

Audibility and speech recognition depend on presentation level. To assess the influence of output level on masking release, output levels for the four hearing aid gain prescription combinations (CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA) were estimated by presenting the “carrot passage” from the Verifit 300 Hearing Instrument Fitting System (AudioScan, Dorchester, ON) to the hearing aid simulators. The output of the simulator was filtered into 1/3 octave bands and then weighted using the frequency response of the Sennheiser HD-25-1 headphones, previously measured using KEMAR, to simulate listening conditions experienced by the study participants. Output levels were measured for input levels of 50-, 60-, and 70-dB SPL using gain prescriptions for each participant with SNHL.

Nonlinear Distortion

Instantaneous compression can result in nonlinear distortion that, in turn, diminishes speech quality

(Herzke and Hohmann, 2005). To assess this issue, nonlinear distortion was evaluated using the Hearing Aid Speech Perception Index (HASPI; Kates and Arehart, 2014a), and the Hearing Aid Speech Quality Index (HASQI; Kates and Arehart, 2014b). HASPI is a predictor of speech intelligibility and HASQI is a predictor of speech quality. Both indices are based on a model of the auditory periphery that incorporates changes due to hearing loss. Calculation of the indices is based on comparison of envelope and temporal fine-structure outputs of an auditory model for a reference signal that assumes NH to the outputs of the model for the signal under test that incorporates the peripheral hearing loss. Both indices range from 0 to 1, where HASPI values of 0 and 1 indicate poor and high intelligibility, and HASQI values of 0 and 1 indicate poor and perfect signal quality. The HASQI is the product of two independent nonlinear and linear components. The nonlinear component measures the degree to which the processing and additive noise alter the dynamics of the short-term spectrum of the test signal over time. The linear component captures effects of linear filtering and spectral changes. Just as measured speech intelligibility can be high when sound quality is low, HASPI values can be close to one when HASQI values are not. Ten sentences from the IEEE database (IEEE, 1969) and the audiograms of the 29 participants with SNHL were used in the calculation. Calculations were made using aided speech in quiet and with speech-weighted steady noise at SNRs of -3, 0, 3, and 6 dB. The noise was spectrally weighted using the international long-term average speech spectrum for combined male and female talkers (Byrne et al, 1994). HASPI and HASQI have been used in several studies to evaluate signal distortion and/or to predict speech intelligibility and quality (Arehart et al, 2013; Kressner et al, 2013; Neher, 2014). An alternative procedure for evaluating hearing aid distortions is described in Tan and Moore (2004).

Speech Recognition

Speech stimuli were 657 consonant-vowel-consonant nonwords from a set recorded by McCreery and Stelmachowicz (2011). The nonwords were phonotactically correct for the English language but did not constitute real words. The stimuli included combinations of consonants /b/, /tʃ/, /d/, /f/, /g/, /dʒ/, /k/, /m/, /n/, /p/, /s/, /ʃ/, /t/, /θ/, /ð/, /v/, and /z/, and vowels /a/, /æ/, /e/, /i/, /I/, /ɛ/, /o/, /u/, /U/, and /Λ/. The nonwords were spoken by a 22-yr-old female from the Midwest. The nonwords had a mean duration of 704 msec (SD = 80 msec; range = 486–936 msec). Stimuli were randomly drawn from the 657-nonword list without replacement for all trials for practice, screening, and testing for each condition. Within a given amplification condition, the participants did not hear the same stimulus twice, but across conditions, they may have heard the same stimulus.

The nonwords were presented in the presence of two noise maskers—steady and modulated noise. The steady noise was spectrally weighted using the international long-term average speech spectrum for combined male and female talkers (Byrne et al, 1994). To produce the modulated noise, the steady noise was 100% sinusoidally amplitude modulated with a modulation rate of 8 Hz. This rate is within the 1–10 Hz range of dominant rates for speech (Steeneken and Houtgast, 1980; Fogerty et al, 2016). In addition, there is a peak in masking release at this rate (Füllgrabe et al, 2006). Twenty noise samples for each type of noise were created and were randomly drawn for presentation.

The combined speech and noise were presented to participants with NH without amplification at 65-dB SPL. For participants with SNHL, the combined speech and noise were first processed by the hearing-aid simulator using a 65 dB SPL input level. The noise started 500 msec before and extended 500 msec after the speech signal. The input level of the combined speech and noise signal (before amplification) was fixed at 65-dB SPL. Participants with NH were tested in two conditions (unmodulated and modulated noise), whereas participants with SNHL were tested in eight conditions (unmodulated and modulated noise, each paired with CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA amplification). Each condition was presented three times for a total of six and 24 runs for participants with NH and SNHL, respectively. The conditions were blocked by noise type and amplification type was randomized within each block. The presentation order of the blocks (noise type) was randomized as well.

An interleaved, two-track, adaptive procedure was used to vary the noise level to determine the threshold SNR (Levitt, 1971). The two tracks were used to determine the SNR that corresponded to the 29% and 71% points on the performance-intensity function. A one-down, two-up procedure was used to determine the 29% point, whereas a two-down, one-up procedure was used to determine the 71% point. The starting SNR was 40 dB for the participants with SNHL and 30 dB for the participants with NH, for both the 29% and 71% performance points. The step sizes for the initial three reversals were 18, 9, and 6 dB, and a step size of 3 dB was used for the remaining reversals. Six reversals were obtained for each track. In the event that one track was completed before the other track, data collection was discontinued for the completed track. Data collection continued for the remaining track until the maximum number of presentations (50 per track) or six reversals was reached. The speech reception threshold was calculated as the mean SNR at the last four reversals. The mean across the three runs served as the final estimate of the threshold SNR.

For the participants with SNHL, a screening run in quiet was provided before data collection. The type of amplification used for the screening was randomized. Twenty nonwords were presented in quiet with an input level of

65-dB SPL to the hearing aid and the participant was required to achieve 60% (12 words) correct to be included in the study. If the participant failed the screening the first time, the screening was repeated. If they failed the screening again, they were excluded from the study. Five participants with SNHL were excluded from the study because they failed to pass the screening. Three additional participants did not finish because the task was too difficult for them (even though all three scored 13/20 on the screening). The exclusion of these eight participants resulted in a total of 29 participants from the initial 37 who enrolled in the study. The screening run was followed by two practice runs, one with each noise type (amplification type randomized), which used the interleaved, two-track, adaptive procedure described above, but with the maximum number of presentations fixed at 14 per track.

Participants were seated in a single-walled sound booth for presentation of the nonword stimuli. They were instructed to repeat back each nonsense word to the best of their ability. Each participant's response was scored and entered into the computer as correct or incorrect by one of three examiners who sat in the booth with the participant.

Before the initiation of data collection, the three examiners scored 1,773 nonword responses from audiovisual recordings of three adults with NH (two females, one male) participating in the nonword task. Interrater agreement assessed using Fleiss's Kappa was 0.923, representing an excellent level of agreement among raters (Fleiss, 1981; Geertzen, 2012).

Some participants with SNHL required an SNR greater than 40 dB to achieve the target percent correct recognition, which is equivalent to listening in quiet because the noise was not audible. This occurred for one participant at 29% correct recognition and nine participants at 71% correct recognition. When this occurred for either the steady or modulated noise masker for a given processing condition, data for that participant were excluded from further analysis. As a result, our data set was unbalanced, with 28 participants with SNHL contributing data for 29% correct recognition and 20 participants with SNHL contributing data for 71% correct recognition.

Assessing the Role of Audibility

To examine the relationship between masking release and audibility, we calculated Pearson correlations between the masking release for the different processing conditions and (1) the aided audibility index (AAI; Stelmachowicz et al, 1994; Brennan and Souza, 2009) and (2) pure-tone average thresholds at 0.5, 1, 2, and 4 kHz. The AAI is similar to the speech intelligibility index (ANSI, 1997), but accounts for the reduced dynamic range of aided speech in its calculations. The dynamic range of speech was characterized by calculating the distribution of short-term (100 msec) levels in each 1/3-octave band. The speech peak level within each band was selected as

the 99th percentile of the distribution of levels, and the speech valley was selected as the 30th percentile of the distribution. Ten sentences from the IEEE database (the same sentences that were used for the HASPI and HASQI calculation) were used for the calculation of the AAI. Calculations of the AAI were made for speech in quiet at an input speech level (before amplification) of 65-dB SPL.

Statistical Analysis

Repeated measures analysis of variance (ANOVA) was used to determine the influence of within-subject factors of noise type (steady, modulated) and, for participants with SNHL, processing condition (CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA) on masking release. Separate ANOVAs were performed on data for participants with NH and those with SNHL, as described later. Because of the unbalanced nature of the data set, separate ANOVAs were performed for performance-intensity points of 29% and 71% for the data for participants with SNHL. An additional ANOVA was performed on the AAI data to determine the influence of within-subject factors of processing condition and SNR on audibility. Mauchly's test of sphericity was used to test for sphericity, and correction for the violation of sphericity was performed using the Greenhouse-Geisser correction, where necessary. In addition to the ANOVA, pairwise comparisons were performed using paired-sample *t* tests with correction for multiple comparisons using Tukey's test for honestly significant difference (HSD). A significance level of $\alpha = 0.05$ was used during analyses. Pearson correlational analysis was used to assess the effect of age on masking release for participants with SNHL. All statistical analyses were performed using SPSS Statistics version 22 software (International Business Machines, Armonk, NY).

RESULTS

Output Level and Gain Analysis

Average output levels as a function of frequency for the four hearing aid gain prescription combinations

(CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA) are shown in Figure 3 for input levels of 50-, 60-, and 70-dB SPL. Compared with the DSL conditions, the output levels for the CLS conditions were higher for frequencies below 1 kHz and lower for frequencies above 2 kHz. As a result, the overall output levels (averaged across frequency) for a given input level were similar as shown in Table 1. DSL resulted in lower output than CLS at low frequencies because DSL prescribed negative gain for participants with audiometric thresholds in the normal range at low frequencies, whereas CLS did not.

Nonlinear Distortion

Results of the analysis of nonlinear distortion as a function of SNR are shown in Figure 4. The top and bottom panels show values for HASPI (prediction of speech perception) and HASQI (prediction of speech quality), respectively. Error bars indicate ± 1 SD. Both HASPI and HASQI increased with increases in SNR, as expected. All four processing conditions resulted in HASPI values of one for speech in quiet. For speech in noise, the four processing conditions resulted in similar HASPI values at each SNR. HASQI values were similar among the four processing conditions. They were less than 0.2 for speech in noise and about 0.7 for speech in quiet. These results indicate there were no differences in predictions of speech intelligibility and sound quality among the processing conditions.

Speech Recognition

Figure 5 shows the SNR required for 29% and 71% correct nonword recognition for modulated and steady maskers in the form of box-and-whiskers plots. Data for participants with SNHL are plotted using shaded boxes for GHA amplification and hatched boxes for SHA amplification. Participants with NH required a lower SNR than participants with SNHL to achieve both 29% and 71% correct for both modulated and steady noise. The variability in SNR for participants with NH was lower than the variability in SNR for participants with SNHL. As expected, a higher mean SNR was

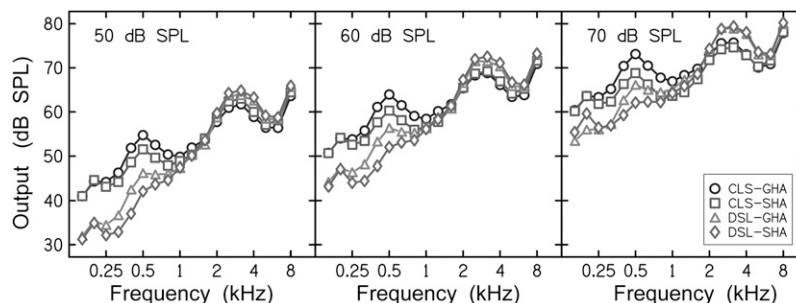


Figure 3. 1/3-octave output levels as a function of frequency for the four processing conditions when the input was speech-shaped noise. Stimulus was the “carrot” passage. Input levels were 50-, 60-, and 70-dB SPL as indicated within each panel. Results for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA are shown using circle, square, triangle, and diamond symbols, respectively.

Table 1. Average Output Level in dB SPL for the Four Processing Conditions for Input Levels of 50-, 60-, and 70-dB SPL

Input	50	60	70
CLS-GHA	66.6	73.8	81.0
CLS-SHA	66.5	73.1	79.2
DSL-GHA	64.8	72.3	79.4
DSL-SHA	65.1	72.5	79.4

Note: Stimulus was the “carrot” passage.

required to achieve 71% correct than to achieve 29% correct for all processing conditions. For a given processing condition and performance-intensity point, the average SNR required for modulated noise was lower than the average SNR required for steady noise. For each processing condition, masking release was calculated as the SNR required for the steady noise masker minus the SNR required for the modulated noise masker. Figure 6 shows masking release, plotted using the same convention as for Figure 5. Participants with NH had greater average masking release than participants with SNHL at both performance-intensity points.

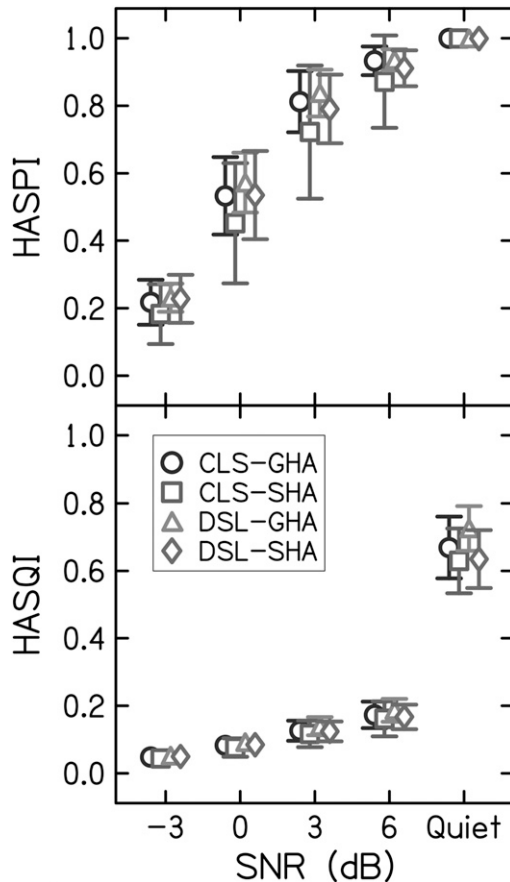


Figure 4. Analysis of nonlinear distortion using HASPI (top) and HASQI (bottom) as function of SNR. Results for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-GHA are shown using circle, square, triangle, and diamond symbols, respectively. Error bars indicate ± 1 SD.

Results of the ANOVA for participants with NH with performance-intensity point (29%, 71%) and noise type (steady, modulated) as within-subject factors are shown in Table 2. As expected, the effect of performance-intensity point was significant, indicating that participants required a lower SNR (mean = -7.1 , standard error [SE] = 0.3 dB) for 29% correct recognition than for 71% correct recognition (mean = 6.7 , SE = 1.1 dB). The effect of noise type was significant. Participants demonstrated masking release, requiring a lower SNR (mean = -3.4 , SE = 0.6 dB) for modulated noise than for steady noise (mean = 3.0 , SE = 0.8 dB). The interaction between performance-intensity point and noise type was not statistically significant. This indicates that masking release did not depend on performance-intensity point, with masking release observed for both 29% (mean = 7.7 , SE = 0 dB) and 71% (mean = 5.1 , SE = 1.2 dB) performance-intensity points.

Results of the ANOVA for participants with SNHL at the performance-intensity point of 29% are shown in Table 3. Within-subject factors were processing condition (CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA) and noise type (steady, modulated). Mauchly's test of sphericity indicated that the assumption of sphericity was violated for processing condition [$\chi^2(5) = 24.8$, $p < 0.001$]. Correction for the violation was performed using Greenhouse-Geisser correction with $\epsilon = 0.6$. The effect of processing condition was not significant. The effect of noise type was significant, indicating that participants with SNHL demonstrated masking release at the performance-intensity point of 29%, requiring a lower SNR (mean = -1.5 , SE = 1.1 dB) for modulated noise than for steady noise (mean = 1.19 , SE = 0.79 dB). The interaction of noise type \times processing condition was not significant. Mean (\pm SE) masking release was 2.6 (1.1), 3.7 (0.8), 3.5 (0.8), and 1.0 (0.8) dB for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively.

Results of the ANOVA for participants with SNHL at the performance intensity point of 71% are shown in Table 4. The factors and their interactions were not significant. Mean (\pm SE) masking release was 0.0 (1.6), 1.2 (1.6), 1.7 (1.1), and 2.4 (1.4) dB for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively.

For the analysis of the relationship between age and masking release for participants with SNHL, Pearson correlation coefficients at the 29% performance-intensity point were 0.29 , 0.10 , 0.05 , and 0.06 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. None of the correlations were statistically significant. At the 71% performance-intensity point, correlation coefficients were 0.26 , 0.18 , 0.17 , and 0.07 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. None of the correlations were statistically significant. Thus, age was not a predictor of masking release for participants with SNHL.

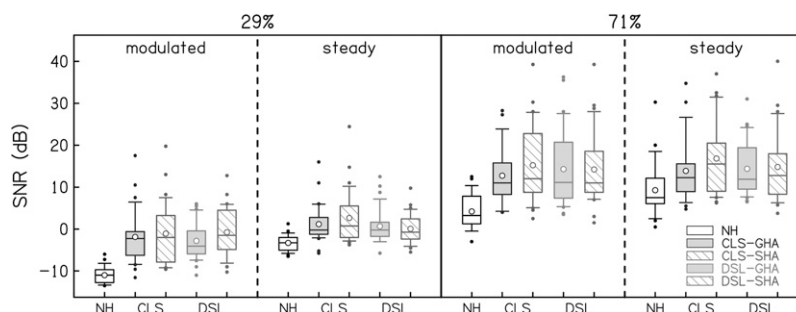


Figure 5. SNR required for 29% (two left-most panels) and 71% (two right-most panels) correct recognition for modulated and steady maskers. Unfilled boxes represent data for participants with NH, which were obtained without amplification. For participants with SNHL, shaded and hatched boxes represent GHA and SHA amplification, respectively. Results are shown for participants with NH (unfilled boxes) and for the four processing conditions for participants with SNHL.

Assessing the Role of Audibility

The calculations of the AAI were made for quiet speech at an input level of 65-dB SPL. Figure 7 shows the AAI for each processing condition for participants with SNHL, plotted using the convention used for Figures 5 and 6. As expected, amplification improved audibility. The mean (SD) AAI without amplification was 0.33 (0.12). Following amplification, the mean (SE) AAI values were 0.65 (0.15), 0.64 (0.18), 0.67 (0.07), and 0.65 (0.07) for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively.

A repeated measures ANOVA was used to assess the influence of processing condition on audibility. Mauchly's test of sphericity indicated that the assumption of sphericity was violated ($\chi^2(9) = 157.4, p < 0.001$). Correction for the violation was performed using Greenhouse-Geisser correction with $\epsilon = 0.33$. The effect of processing condition was significant ($F_{(1.3,37.2)} = 87.72, p < 0.001, \eta_p^2 = 0.76$). Pairwise comparisons indicated that differences between the unaided condition and any of the aided conditions were statistically significant (all $p < 0.001$). Differences between any pair of processing conditions were not statistically significant ($p = 1.0$).

Pearson correlation coefficients between the AAI and masking release for 29% correct recognition were $r = -0.03, -0.29, -0.22$, and -0.33 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. Coefficients between AAI and masking release for 71% correct recognition were $r = -0.16, -0.23, -0.22$, and -0.21 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. None of these correlations were statistically significant after correction for multiple comparisons using Tukey's HSD test. Pearson correlation coefficients between pure-tone average and masking release for 29% correct recognition were $r = -0.03, -0.29, -0.220$, and -0.33 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. Coefficients between pure-tone average and masking release for 71% correct recognition were $r = -0.15, -0.23, -0.22$, and -0.21 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. Again, none of these correlations were

statistically significant after correction for multiple comparisons using Tukey's HSD test. These results indicate that there was no significant relationship between audibility and masking release, irrespective of whether it was assessed using the AAI or pure-tone average.

DISCUSSION

This study evaluated whether an experimental hearing aid with instantaneous compression (SHA) improved the ability to listen in temporal dips for listeners with SNHL, relative to a hearing aid with fast compression (GHA). Masking release was observed for participants with SNHL for both SHA and GHA at the performance-intensity point of 29%, but not at 71%. At the performance-intensity point of 29%, there were no statistically significant differences in masking release between SHA and GHA regardless of whether they were fit using CLS or DSL.

Some participants with SNHL required an SNR greater than 40 dB to achieve the target percent correct nonword recognition, which was equivalent to listening in quiet because the noise was not audible. Data for these participants for the condition with SNR greater than

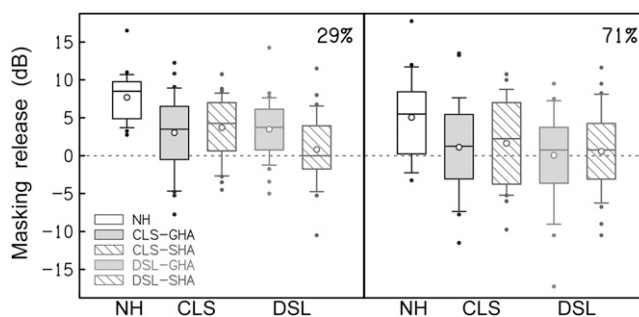


Figure 6. Masking release, defined as the SNR required for criterion speech recognition in steady noise minus the SNR required for criterion speech recognition in modulated noise. The left and right panels show data for 29% and 71% correct recognition, respectively. Results are shown for participants with NH (unfilled boxes) and for the four processing conditions for participants with SNHL, with shaded boxes indicating GHA amplification and hatched boxes indicating SHA amplification.

Table 2. Analysis of Variance for the Participants with NH to Determine the Influence of Noise Type and Performance-Intensity Point (PI) on the Speech Reception Threshold

Main Effects and Interactions	df	<i>F</i>	<i>p</i>	η_p^2
Noise type	1, 20	70.964	< 0.001	0.780
PI	1, 20	160.420	< 0.001	0.8889
Noise type × PI	1, 20	3.864	0.063	0.162

Note: Bold values indicate $p < 0.05$.

40 dB were excluded from further analysis. Our exclusion criterion is justifiable because speech recognition was no longer influenced by the type of noise when the noise was not audible. Previous studies using a similar adaptive procedure to measure speech intelligibility have used data exclusion criteria based on the SD across repeated tracks and on audibility (Brennan et al, 2016) as means of controlling data quality. In the present study, it is unclear why some participants had difficulty performing the task after passing the screening trial. Perhaps the screening condition, with a pass criterion of 12/20, was not stringent enough.

SHA and GHA resulted in similar HASPI (predictor of speech intelligibility) and HASQI (predictor of speech quality) values, irrespective of whether they were fitted using CLS or DSL (see Figure 4). These results suggest that if the instantaneous compression as implemented in SHA had a detrimental effect on sound quality, it was not greater than the effect resulting from the fast compression of GHA. The lack of a negative impact of instantaneous compression is further supported by the speech-recognition results. The threshold SNRs required for either 29% or 71% correct recognition when using SHA amplification were similar to the SNRs required when using GHA amplification, for all combinations of noise type and performance intensity point (see Figure 5). The same is true for masking release (see Figure 6).

HASQI values were less than 0.2 for speech in noise (see Figure 4). Recall that HASQI is a product of a linear component capturing filtering and spectral changes, and a nonlinear component capturing time domain changes. The HASQI values were driven by the nonlinear component, which was lower than the linear component. For example, when the SNR was 6 dB, the mean linear HASQI components were 0.79, 0.79, 0.82, and 0.75 for

CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively, and the mean nonlinear HASQI components were 0.22, 0.20, 0.23, and 0.22 for CLS-GHA, CLS-SHA, DSL-GHA, and DSL-SHA, respectively. Importantly, there were no differences in HASQI (and HASPI) across processing conditions.

The similarity of HASQI values across processing conditions is supported by a previous study in which CLS-SHA was compared with DSL-GHA in terms of preference (Rasetshwane, Brennan, et al, 2014). Participants preferred CLS-SHA to DSL-GHA when listening to music in quiet, but there were no significant differences in preference when listening to sentences in quiet. Although CLS-SHA and DSL-GHA differed in fitting procedures, the differences in preference cannot be attributed to differences in audibility or overall output level resulting from the fitting procedures because the current study demonstrated that such differences do not exist (see Figures 3 and 7). Although differences in frequency response may have contributed to the preference rating, the similarity in preference ratings for two simulators with different frequency response in the Rasetshwane, Brennan, et al (2014) study suggests that the instantaneous compression of SHA did not have a deleterious effect on speech and music quality when compared with fast compression.

In the current study, participants with SNHL were older than participants with NH (average of 64 yr compared to 33 yr). Thus, it is possible that the lower masking release observed in the former group was partly due to differences in age, as age has been shown to be correlated with masking release (George et al, 2006). However, correlational analyses showed that within the participants with SNHL, age was not related to masking release. This is in contrast to the results of George et al (2006) who observed a relationship for their participants with SNHL, but in agreement with Füllgrabe et al (2015) who did not observe an effect of age on masking release when audiograms were matched for young and older listeners with NH. The disagreement between the results of our study and the results reported by George et al (2006) may be due to the fact that amplification was provided to participants with SNHL in the current study whereas it was not provided to the participants in George et al, leading to differences in audibility. Other methodological differences between the two studies, such as differences in speech material and performance intensity point, may have also contributed to the discrepancy.

We used the AAI to assess whether audibility for participants with SNHL was related to masking release. AAI was not related to masking release for any processing condition. In addition, pure-tone average was not related to masking release.

The test ear for participants with SNHL was selected randomly if both ears met the inclusion criteria. The pure-tone average for test ear (mean = 36, SD = 8, range = 19–56 dB HL) was better than the pure-tone average

Table 3. Analysis of Variance for the Participants with SNHL to Determine the Influence of Processing Condition and Noise Type on the Speech Reception Threshold at the Performance-Intensity Point of 29%

Main Effects and Interactions	df	<i>F</i>	<i>p</i>	η_p^2
Processing	1.9, 50.7	1.825	0.174	0.063
Noise type	1, 27	21.908	< 0.001	0.448
Processing × noise type	3, 81	2.455	0.069	0.083

Note: Bold value indicates $p < 0.05$.

Table 4. Analysis of Variance for the Participants with SNHL to Determine the Influence of Processing Condition on the Speech Reception Threshold at the Performance-Intensity Point of 71%

Main Effects and Interactions	df	F	p	η_p^2
Processing	3, 19	1.965	0.129	0.094
Noise	1, 19	1.774	0.199	0.085
Processing × noise	3, 57	0.680	0.568	0.035

for the nontest ear (mean = 39, SD = 9, range = 16–56 dB HL), and a paired *t*-test revealed that the difference was statistically significant [$t_{(28)} = -2.32, p = 0.03$]. Despite these differences, it is unlikely that less masking release would have been observed had the test ear been selected as the ear with higher thresholds because amplification was provided to the listeners with SNHL. The test-ear pure-tone average selection criterion would have resulted in higher hearing aid gain. Thus, we do not believe that the pure-tone average criterion used to select the test ear influenced the outcome of the study.

Twenty one of the 29 participants with SNHL wore hearing aids in their daily lives. Information about the type of gain prescription and compression speed was not obtained, but it is possible that some of these participants used DSL gain prescription with fast compression and were accustomed to this type of processing. If that was the case, it is possible that these participants obtained better speech recognition for processing conditions that used DSL (DSL-GHA and DSL-SHA).

A limitation of the current study is that a condition with linear amplification was not included. However, the aim of this study was to compare masking release for listeners with SNHL obtained with an experimental hearing aid signal-processing algorithm with instantaneous compression to masking release obtained with a simulation of current technology using fast compression. The aim was not to evaluate the effect of WDRC on masking release relative to linear amplification, as this has been evaluated in several previous studies (Moore et al, 1999). Another lim-

itation of the current study is that there was no condition without amplification for participants with SNHL. We did not include this condition because it would have resulted in large differences in audibility making it difficult to interpret the results of masking release.

In the current study, modulated maskers that changed only in the time domain were used because the goal was to compare instantaneous compression to fast compression, and compression speed primarily affects the temporal characteristics of a signal. In future studies with spectrally modulated noise, we expect SHA to improve masking release for individuals with SNHL because the suppression in SHA signal processing leads to spectral enhancements (see Figure 11 of Rasetshwane, Gorga, et al, 2014), which in turn may make it easier to discriminate such cues.

Further refinements of SHA signal processing and the CLS-based gain-prescription strategy may lead to improvements in the ability to listen in the dips. Refinements to the signal processing could include adjustments of the cross-channel influences due to suppression. The implementation of suppression in SHA was based on measurements of DPOAEs with a single suppressor tone and a model that assumed that suppressive effects of multiple suppressor components were additive in the intensity domain. Recent DPOAE data using two suppressor tones have indicated that the suppressive effects of multiple tones are not well described by the simple additive model that was used to mimic the effects of suppression in the design of SHA, but are better described by a hybrid model that involves both additive intensity and additive attenuation models (Sieck et al, 2016). Although future refinements using this model are planned, it is unlikely that these processing changes will affect the ability to listen in noise with temporal dips because suppression mainly affects the spectral characteristics of a signal.

It is possible, however, that the ability to listen in the dips can never be restored to normal for listeners with SNHL through amplification if the ability is affected by suprathreshold hearing deficits such as reduced temporal resolution and frequency selectivity, as discussed in the Introduction.

CONCLUSION

The results of this study suggest that hearing aid signal processing that includes instantaneous compression can perform as well as fast compression in terms of the ability to benefit from temporal dips in background noise for listeners with hearing loss. However, the amount of benefit obtained with either instantaneous or fast compression was less than the masking release observed for listeners with NH without amplification.

Acknowledgments. This research was supported by the NIH-NIDCD grants R03 DC013982 (DMR), R01 DC8318

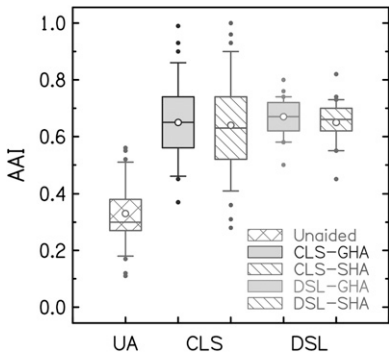


Figure 7. Mean AAI for participants with SNHL for the four processing conditions and for unaided speech. The input speech level was 65-dB SPL. Shaded and hatched boxes indicate GHA and SHA amplification, respectively, following the convention used in previous figures.

(STN), and P30 DC4662 (MPG). We would like to thank Emily C. Bosen and Lindsay N. Reuter for their help with data collection.

REFERENCES

- Al-Salim SC, Kopun JG, Neely ST, Jesteadt W, Stiegemann B, Gorga MP. (2010) Reliability of categorical loudness scaling and its relation to threshold. *Ear Hear* 31(4):567–578.
- Alexander JM, Masterson K. (2015) Effects of WDRC release time and number of channels on output SNR and speech recognition. *Ear Hear* 36(2):e35–e49.
- American Speech-Language-Hearing Association. (2005) Guidelines for manual pure-tone threshold audiometry [Guidelines]. www.asha.org/policy.
- American National Standards Institute. (1997) *American National Standard Methods for the Calculation of the Speech Intelligibility Index. ANSI S3.5-1997*. New York, NY: ANSI.
- American National Standards Institute. (2004) *Specification for Octave-Band and Fractional-Octave-Band Analog and Digital Filters. ANSI S1.11-2004*. New York, NY: ANSI.
- Arehart KH, Souza P, Baca R, Kates JM. (2013) Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear* 34(3):251–260.
- Bacon SP, Opie JM, Montoya DY. (1998) The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J Speech Lang Hear Res* 41(3):549–563.
- Baer T, Moore BCJ. (1993) Effects of spectral smearing on the intelligibility of sentences in noise. *J Acoust Soc Am* 94(3):1229–1241.
- Baer T, Moore BCJ. (1994) Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J Acoust Soc Am* 95(4):2277–2280.
- Bernstein JG, Grant KW. (2009) Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 125(5):3358–3372.
- Bernstein JG, Brungart DS. (2011) Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *J Acoust Soc Am* 130(1):473–488.
- Brand T, Hohmann V. (2002) An adaptive procedure for categorical loudness scaling. *J Acoust Soc Am* 112(4):1597–1604.
- Brennan M, Souza P. (2009) Effects of expansion on consonant recognition and consonant audibility. *J Am Acad Audiol* 20(2):119–127.
- Brennan M, McCreery R, Kopun J, Lewis D, Alexander J, Stelmachowicz P. (2016) Masking release in children and adults with hearing loss when using amplification. *J Speech Lang Hear Res* 59(1):110–121.
- Brungart DS, Simpson BD, Ericson MA, Scott KR. (2001) Informational and energetic masking effects in the perception of simultaneous talkers. *J Acoust Soc Am* 110(5):2527–2538.
- Byrne D, et al. (1994) An international comparison of long-term average speech spectra. *J Acoust Soc Am* 96(4):2108–2120.
- Dubno JR, Horwitz AR, Ahlstrom JB. (2003) Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J Acoust Soc Am* 113(4 Pt 1):2084–2094.
- Duquesnoy AJ. (1983) Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J Acoust Soc Am* 74(3):739–743.
- Eisenberg LS, Dirks DD, Bell TS. (1995) Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *J Speech Hear Res* 38(1):222–233.
- Festen JM. (1993) Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice. *J Acoust Soc Am* 94(3 Pt 1):1295–1300.
- Festen JM, Plomp R. (1990) Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am* 88(4):1725–1736.
- Fleiss JL. (1981) *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley, 598–626.
- Fogerty D, Ahlstrom JB, Bologna WJ, Dubno JR. (2016) Glimpsing speech in the presence of nonsimultaneous amplitude modulations from a competing talker: effect of modulation rate, age, and hearing loss. *J Speech Lang Hear Res* 59(5):1198–1207.
- Freyman RL, Balakrishnan U, Helfer KS. (2004) Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am* 115(5 Pt 1):2246–2256.
- Füllgrabe C, Berthommier F, Lorenzi C. (2006) Masking release for consonant features in temporally fluctuating background noise. *Hear Res* 211(1–2):74–84.
- Füllgrabe C, Moore BCJ, Stone MA. (2015) Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front Aging Neurosci* 6:347.
- Geertzen J. (2012) *Inter-rater Agreement with Multiple Raters and Variables*. Retrieved October 17, 2016. <https://nlp-ml.io/jg/software/ira/>
- George ELJ, Festen JM, Houtgast T. (2006) Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 120(4):2295–2311.
- Gorga MP, Neely ST, Kopun J, Tan H. (2011a) Growth of suppression in humans based on distortion-product otoacoustic emission measurements. *J Acoust Soc Am* 129(2):801–816.
- Gorga MP, Neely ST, Kopun J, Tan H. (2011b) Distortion-product otoacoustic emission suppression tuning curves in humans. *J Acoust Soc Am* 129(2):817–827.
- Gregan MJ, Nelson PB, Oxenham AJ. (2013) Behavioral measures of cochlear compression and temporal resolution as predictors of speech masking release in hearing-impaired listeners. *J Acoust Soc Am* 134(4):2895–2912.
- Gustafsson HÅ, Arlinger SD. (1994) Masking of speech by amplitude-modulated noise. *J Acoust Soc Am* 95(1):518–529.
- Hall JW, Buss E, Grose JH, Roush PA. (2012) Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear* 33(3):340–348.
- Herzke T, Hohmann V. (2005) Effects of instantaneous multiband dynamic compression on speech intelligibility. *EURASIP J Adv Signal Process* 2005(18):3034–3043.
- Hohmann V. (2002) Frequency analysis and synthesis using a Gammatone filterbank. *Acta Acust United Acust* 88:433–443.
- Houtgast R. (1974) Auditory analysis of vowel-like sounds. *Acta Acust United Acust* 31:320–324.

- IEEE. (1969) IEEE recommended practice for speech quality measurements. *IEEE Trans Audio Electroacoust* 17(3):225–246.
- International Organization for Standardization. (2006) *Acoustics - loudness scaling by means of categories. ISO 16832:2006*. Geneva, Switzerland: ISO.
- Jin SH, Nelson PB. (2006) Speech perception in gated noise: the effects of temporal resolution. *J Acoust Soc Am* 119(5 Pt 1):3097–3108.
- Kates JM, Arehart KH. (2014a) The hearing-aid speech perception index (HASPI). *Speech Commun* 65:75–93.
- Kates JM, Arehart KH. (2014b) The hearing-aid speech quality index (HASQI) version 2. *J Audio Eng Soc* 62(3):99–117.
- Kressner AA, Anderson DV, Rozell CJ. (2013) Evaluating the generalization of the hearing aid speech quality index (HASQI). *IEEE Trans Audio Speech Lang Process* 21(2):407–415.
- Kwon BJ, Turner CW. (2001) Consonant identification under maskers with sinusoidal modulation: masking release or modulation interference? *J Acoust Soc Am* 110(2):1130–1140.
- Levitt H. (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49(2, Suppl):467–477.
- McCreery RW, Stelmachowicz PG. (2011) Audibility-based predictions of speech recognition for children and adults with normal hearing. *J Acoust Soc Am* 130(6):4070–4081.
- McCreery RW, Brennan MA, Hoover B, Kopun J, Stelmachowicz PG. (2013) Maximizing audibility and speech recognition with nonlinear frequency compression by estimating audible bandwidth. *Ear Hear* 34(2):e24–e27.
- Moore BCJ, Peters RW, Stone MA. (1999) Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *J Acoust Soc Am* 105(1):400–411.
- Neely ST, Johnson TA, Gorga MP. (2005) Distortion-product otoacoustic emission measured with continuously varying stimulus level. *J Acoust Soc Am* 117(3 Pt 1):1248–1259.
- Neher T. (2014) Relating hearing loss and executive functions to hearing aid users' preference for, and speech recognition with, different combinations of binaural noise reduction and microphone directionality. *Front Neurosci* 8:391.
- Nelson DA, Schroder AC, Wojtczak M. (2001) A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 110(4):2045–2064.
- Nelson PB, Jin SH. (2004) Factors affecting speech understanding in gated interference: cochlear implant users and normal-hearing listeners. *J Acoust Soc Am* 115(5 Pt 1):2286–2294.
- Oxenham AJ, Simonson AM, Turicchia L, Sarpeshkar R. (2007) Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing. *J Acoust Soc Am* 121(3):1709–1716.
- Peters RW, Moore BCJ, Baer T. (1998) Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J Acoust Soc Am* 103(1):577–587.
- Plomp R. (1994) Noise, amplification, and compression: considerations of three main issues in hearing aid design. *Ear Hear* 15(1):2–12.
- Rasetshwane DM, Neely ST, Kopun JG, Gorga MP. (2013) Relation of distortion-product otoacoustic emission input-output functions to loudness. *J Acoust Soc Am* 134(1):369–383.
- Rasetshwane DM, Gorga MP, Neely ST. (2014) Signal-processing strategy for restoration of cross-channel suppression in hearing-impaired listeners. *IEEE Trans Biomed Eng* 61(1):64–75.
- Rasetshwane DM, Brennan MA, Kopun JG, Neely ST, Gorga MP. (2014) Evaluation of a hearing-aid signal-processing strategy for restoration of cross-channel suppression. Poster presented at International Hearing Aid Research Conference, Tahoe, CA.
- Rasetshwane DM, Trevino AC, Gombert JN, Liebig-Trehearn L, Kopun JG, Jesteadt W, Neely ST, Gorga MP. (2015) Categorical loudness scaling and equal-loudness contours in listeners with normal hearing and hearing loss. *J Acoust Soc Am* 137(4):1899–1913.
- Rosen S, Souza P, Ekelund C, Majeed AA. (2013) Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J Acoust Soc Am* 133(4):2431–2443.
- Sachs MB, Young ED. (1980) Effects of nonlinearities on speech encoding in the auditory nerve. *J Acoust Soc Am* 68(3):858–875.
- Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Laurnagaray D, Beaulac S, Pumford J. (2005) The desired sensation level multistage input/output algorithm. *Trends Amplif* 9(4):159–197.
- Sieck NE, Rasetshwane DM, Kopun JG, Jesteadt W, Gorga MP, Neely ST. (2016) Multi-tone suppression of distortion-product otoacoustic emissions in humans. *J Acoust Soc Am* 139(5):2299–2309.
- Simpson SA, Cooke M. (2005) Consonant identification in N-talker babble is a nonmonotonic function of N. *J Acoust Soc Am* 118(5):2775–2778.
- Snell KB, Mapes FM, Hickman ED, Frisina DR. (2002) Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. *J Acoust Soc Am* 112(2):720–727.
- Steeneken HJM, Houtgast T. (1980) A physical method for measuring speech-transmission quality. *J Acoust Soc Am* 67(1):318–326.
- Stelmachowicz PG, Lewis DE, Kalberer L, Creutz T. (1994) *Situational Hearing-Aid Response Profile Users Manual (SHARP, v. 6.0)*. Omaha, NE: Boys Town National Research Hospital.
- Stone MA, Moore BCJ, Meisenbacher K, Derleth RP. (2008) Tolerable hearing aid delays. V. Estimation of limits for open canal fittings. *Ear Hear* 29(4):601–617.
- Stone MA, Füllgrabe C, Mackinnon RC, Moore BCJ. (2011) The importance for speech intelligibility of random fluctuations in “steady” background noise. *J Acoust Soc Am* 130(5):2874–2881.
- Stone MA, Moore BCJ. (2014) On the near non-existence of “pure” energetic masking release for speech. *J Acoust Soc Am* 135(4):1967–1977.
- Tan CT, Moore BCJ. (2004) Comparison of two forms of fast-acting compression using physical and subjective measures. In: Proceedings of the 18th International Congress on Acoustics, Kyoto, Japan. II:1393–1396.
- ter Keurs M, Festen JM, Plomp R. (1993a) Effect of spectral envelope smearing on speech reception. II. *J Acoust Soc Am* 93(3):1547–1552.
- ter Keurs M, Festen JM, Plomp R. (1993b) Limited resolution of spectral contrast and hearing loss for speech in noise. *J Acoust Soc Am* 94(3 Pt 1):1307–1314.
- Turicchia L, Sarpeshkar R. (2005) A bio-inspired companding strategy for spectral enhancement. *IEEE Trans Speech Audio Process* 13(2):243–253.
- Xu L, Thompson CS, Pfingst BE. (2005) Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am* 117(5):3255–3267.