

Effects of Frequency Compression and Frequency Transposition on Fricative and Affricate Perception in Listeners With Normal Hearing and Mild to Moderate Hearing Loss

Joshua M. Alexander,¹ Judy G. Kopun,² and Patricia G. Stelmachowicz²

Objectives: The authors have demonstrated that the limited bandwidth associated with conventional hearing aid amplification prevents useful high-frequency speech information from being transmitted. The purpose of this study was to examine the efficacy of two popular frequency-lowering algorithms and one novel algorithm (spectral envelope decimation) in adults with mild to moderate sensorineural hearing loss and in normal-hearing controls.

Design: Participants listened monaurally through headphones to recordings of nine fricatives and affricates spoken by three women in a vowel-consonant context. Stimuli were mixed with speech-shaped noise at 10 dB SNR and recorded through a Widex Inteo IN-9 and a Phonak Naída UP V behind-the-ear (BTE) hearing aid. Frequency transposition (FT) is used in the Inteo and nonlinear frequency compression (NFC) used in the Naída. Both devices were programmed to lower frequencies above 4 kHz, but neither device could lower frequencies above 6 to 7 kHz. Each device was tested under four conditions: frequency lowering deactivated (FT-off and NFC-off), frequency lowering activated (FT and NFC), wide-band (WB), and a fourth condition unique to each hearing aid. The WB condition was constructed by mixing recordings from the first condition with high-pass filtered versions of the source stimuli. For the Inteo, the fourth condition consisted of recordings made with the same settings as the first, but with the noise-reduction feature activated (FT-off). For the Naída, the fourth condition was the same as the first condition except that source stimuli were preprocessed by a novel frequency compression algorithm, spectral envelope decimation (SED), designed in MATLAB, which allowed for a more complete lowering of the 4 to 10 kHz input band. A follow-up experiment with NFC used Phonak's Naída SP V BTE, which could also lower a greater range of input frequencies.

Results: For normal-hearing and hearing-impaired listeners, performance with FT was significantly worse compared with that in the other conditions. Consistent with previous findings, performance for the hearing-impaired listeners in the WB condition was significantly better than in the FT-off condition. In addition, performance in the SED and WB conditions were both significantly better than in the NFC-off condition and the NFC condition with 6 kHz input bandwidth. There were no significant differences between SED and WB, indicating that improvements in fricative identification obtained by increasing bandwidth can also be obtained using this form of frequency compression. Significant differences between most conditions could be largely attributed to an increase or decrease in confusions for the phonemes /s/ and /z/. In the follow-up experiment, performance in the NFC condition with 10 kHz input bandwidth was significantly better than NFC-off, replicating the results obtained with SED. Furthermore, listeners who performed poorly with NFC-off tended to show the most improvement with NFC.

Conclusions: Improvements in the identification of stimuli chosen to be sensitive to the effects of frequency lowering have been demonstrated using two forms of frequency compression (NFC and SED) in individuals with mild to moderate high-frequency sensorineural hearing loss. However, negative results caution against using FT for this population. Results also indicate that the advantage of an extended bandwidth as reported here and elsewhere applies to the input bandwidth for frequency compression (NFC/SED) when the start frequency is ≥ 4 kHz.

Key words: Frequency compression, Frequency lowering, Frequency transposition, Fricatives, Hearing aids.

(*Ear & Hearing* 2014;35;519–532)

INTRODUCTION

Historically, attempts to implement frequency lowering have been limited to adults and children with severe to profound sensorineural hearing loss (SNHL; e.g., Braida et al. 1978; Parent et al. 1998; McDermott et al. 1999; Simpson et al. 2005, 2006; Kuk et al. 2009a). For these individuals, arguments for the use of frequency lowering are relatively easy to make because speech perception deficits increase as high-frequency audibility decreases (Turner & Robb 1987; Dubno et al. 1989; Ching et al. 1998; Horwitz et al. 2002). However, the relationship between severity of SNHL and amount of frequency lowering seems to be give and take (cf. Alexander 2013a, pp. 97–99): while there is more speech information to be gained from frequency lowering as SNHL increases, the audible bandwidth available to recode the missing information also becomes increasingly narrower. In turn, frequency lowering becomes progressively more difficult to implement without distorting critical aspects of the low-frequency speech signal, such as formant frequencies.

However, while individuals with less severe hearing loss have fewer speech perception deficits, they also have a wider bandwidth of audibility available to implement frequency lowering. This property combined with modern techniques and advances in signal processing that allow frequency lowering to operate more selectively in frequency may make it possible to provide some benefit to these individuals with less risk of distorting important low-frequency information (for a review on modern frequency-lowering techniques and associated research findings, see Alexander 2013a; for a review on clinical issues surrounding these techniques, see Alexander 2013b, Scollie 2013, and Mueller et al. 2013). The purpose of this research study is to test the efficacy of using frequency-lowering amplification for adults with mild to moderate high-frequency SNHL in conditions with significant high-frequency information. If frequency lowering decreases speech perception in these conditions, then future research on alternative techniques or hearing losses may be suggested. If it

¹Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana, USA; and ²Boys Town National Research Hospital, Omaha, Nebraska, USA.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

increases or maintains speech perception, then future research that examines whether the findings generalize to other test conditions, subject populations, signal processing techniques, hearing losses, etc. may be suggested.

To help understand the potential benefits associated with frequency lowering for adults with mild to moderate SNHL, it is useful to examine the role of high-frequency audibility in speech perception. One consequence of reduced high-frequency audibility includes a loss of some or all of the noisy friction energy associated with the fricatives, affricates, and some stop consonant bursts. Miller and Nicely (1955) showed that the error patterns caused by progressively low-pass filtering nonsense syllables most notably include confusions between the fricatives /s/, /ʃ/, /θ/, and /f/, which can have spectral peaks above 9 kHz (Fox & Nissen 2005). Speech is linguistically and acoustically redundant and, with varying degrees of success, listeners can identify high-frequency phonemes using only the transitions from the lower-frequency formants of the coarticulated phonemes that precede and follow them (e.g., Repp & Mann 1980; Whalen 1981). Despite this, there is evidence that speech perception improves when an effort is made to preserve the high-frequency noise energy associated with friction. Studies that estimated the degree to which the two cues (formant transitions versus friction) are perceptually weighted (Heinz & Stevens 1961; Nittrouer & Studdert-Kennedy 1987; Zeng & Turner 1990; Nittrouer 1992; Nittrouer & Miller 1997; Hedrick 1997; Pittman & Stelmachowicz 2000) report that adults tend to rely more on the information provided by the high-frequency friction for identification between pairs of phonemes differing only in place of articulation. This seems to be especially true for listeners who have a decreased ability to track changes in formant frequency due to presumed spectro-temporal processing deficits associated with SNHL (Zeng & Turner 1990; Sullivan et al. 1992; Hedrick & Younger 2003).

There is further evidence that both normal-hearing (NH) and hearing-impaired (HI) adults and children perform best when the audible bandwidth of speech extends at least to 9 kHz, especially when contextual information is minimal (e.g., Stelmachowicz et al. 2001, 2007; Pittman 2008; Hornsby et al. 2011) or test conditions provide an advantage for high-frequency spatial cues (Moore et al. 2010). Stelmachowicz et al. (2001) examined the effect of increasing the cutoff frequency of a low-pass filter from 2 to 9 kHz on the identification of voiceless fricatives spoken in quiet in a consonant–vowel (CV) or vowel–consonant (VC) context. Talkers included an adult male, an adult female, and a child. Listeners were adults and children with normal hearing and moderate to moderately severe SNHL. All groups needed information up to 9 kHz to achieve their maximum performance level for the female and child talkers, but only needed high-frequency information up to 4 to 5 kHz to achieve maximum performance for the male talker. The only other exception was the NH adults who only needed information up to 6 kHz to achieve maximum performance for the female talker.

Hornsby et al. (2011) also demonstrated benefit from extended high-frequency audibility for adult HI listeners. Using successive low- and high-pass filtering techniques designed to measure importance functions for the Speech Intelligibility Index (ANSI S3.5 1997), Hornsby and colleagues examined sentence recognition in noise for HI listeners with a wide range of thresholds. Stimuli were frequency shaped and amplified to

output targets for a 65 dB SPL presentation level via insert earphones, as recommended by the DSL v4.1 prescriptive fitting algorithm (Seewald et al. 1997). They found that when the low-pass filter increased from about 3.5 to 9 kHz, speech recognition improved for about one third of the listeners. Those most likely to benefit from natural bandwidth extension had greater SNHL in the low frequencies and less SNHL (better audibility) in the high frequencies, which indicates that HI listeners can and do use information from the high frequencies when it is accessible.

The abovementioned findings, which illustrate the benefits of extended high-frequency bandwidth for NH and HI adults, together with studies that document the difficulty of providing sufficient high-frequency gain for individuals with mild to moderate high-frequency SNHL (Moore et al. 2008), suggest that some hearing aid users with less severe SNHL may be able to use frequency lowering to reduce perceptual ambiguity among fricative and affricate consonants. Considering that about half of the English consonants are fricatives or affricates, this may be a worthwhile goal. In fact, /s/ and its voiced cognate, /z/, account for about 8% of the spoken English consonants (Rudmin 1983) and together carry significant morphological importance. Rudmin (1983) highlights over 20 different linguistic uses for /s/ and /z/, including plurality, third-person present tense, past versus present tense, possession, pronouns, contractions, etc. The linguistic utility of the phoneme /s/ combined with the fact that it is one of the most frequently confused phonemes for adults and children with SNHL (Bilger & Wang 1976; Dubno & Dirks 1982; Danhauer et al. 1986) may help explain some of the communication difficulties experienced by many hearing aid users with mild to moderate SNHL.

While it would be ideal if high-frequency speech information were presented without frequency lowering, for various reasons it is not always possible to achieve sufficient high-frequency output with the current hearing aid technology, even for mild to moderate losses. One reason is that the frequency response of a typical hearing aid receiver has a sharp roll-off above 5 to 6 kHz (Dillon 2001). In addition, the response is affected by feedback control techniques and is further reduced by the filtering properties of the tubing for behind-the-ear (BTE) instruments (Killion 2003) and the increasingly popular “slim tube” fittings. To highlight the difficulty of achieving high-frequency audibility in a clinical population, Moore et al. (2008) assessed the audiograms of 31 listeners with mild to moderate SNHL through 4 kHz that progressively worsened at higher frequencies. They estimated that audibility at 10 kHz could only be partially achieved in fewer than half the ears tested. They also argued that amplification at the necessary levels could lead to loudness discomfort or unpleasant, harsh, or tinny sound quality. Therefore, it is reasonable to question whether frequency lowering might be of benefit for at least some individuals with mild to moderate SNHL and with limited high-frequency audibility despite conventional amplification.

This study was conducted to determine whether two modern frequency-lowering techniques, frequency transposition (FT) as implemented in Widex hearing aids and nonlinear frequency compression (NFC) as implemented in Phonak hearing aids, which were both originally designed for listeners with severe to profound SNHL (Simpson et al. 2005; Kuk et al. 2009a), have the potential to improve fricative and affricate perception for listeners with mild to moderate SNHL. A brief overview of the

two techniques is provided here; see Alexander (2013a) for a more extensive discussion and the methods below for illustrative examples. The part of the spectrum where information is moved *from* is the source region, and the part of the spectrum where information is moved *to* is the target region. The lowest frequency in the source region is the “start frequency”. NFC techniques can be best understood using amplitude compression as an analogy. The start frequency with NFC is similar to the compression threshold with amplitude compression—all the action occurs above it. Therefore, the start frequency is like an anchor that does not move, with spectral information above it moving down toward it. Just as dynamic range is reduced with amplitude compression, the source bandwidth is reduced with NFC such that the information after lowering (the target region) is fully contained within the source region.

With FT, information from a portion of the source region (possibly including the start frequency itself) is resynthesized at lower frequencies. With the Widex implementation of this method, the algorithm continually searches for the most intense peak in the source region and then duplicates it at a frequency that is a linear factor of the original (1/2 or 1 octave in basic mode and 1/3 in expanded mode). To avoid having to compress the lowered information, only a portion of the signal is selected. Specifically, the peak is band-pass filtered so that it only spans 1 octave after lowering. The source region actually begins a 1/2 octave *below* the nominal start frequency in the programming software and extends for 1 octave above this. Therefore, all lowering (the target region) is below the start frequency with FT, whereas with NFC all lowering is above the start frequency.

PARTICIPANTS AND METHODS

Participants

Twenty-four NH listeners, 18 to 50 years of age, (6 men and 18 women) and 24 HI listeners, 18 to 74 years of age (11 men and 13 women) participated in the study. All listeners were screened for normal middle ear function using tympanometry. Audiometric thresholds between 0.25 to 8 kHz for NH listeners were ≤ 15 dB HL (ANSI S3.6-1996) in the test ear, except for 1 listener whose 8 kHz threshold was 20 dB HL. Audiometric thresholds for the HI listeners are shown in Table 1. With a few exceptions, HI listeners were included in the study if they had normal hearing or mild SNHL (≤ 40 dB HL) for frequencies ≤ 1 kHz and moderate to moderately severe SNHL (45 to 70 dB HL) for frequencies ≥ 3 kHz. One ear of each listener was tested. The right ear was chosen unless the hearing loss in that ear did not meet the abovementioned inclusion criteria. Most of the HI listeners had acquired SNHL, and fewer than half of them regularly wore hearing aids—none of which had frequency lowering.

Stimuli

Experimental stimuli were chosen to maximize the relative information contained in the high-frequency speech spectrum. Therefore, fricative (/s/, /z/, /ʃ/, /ʒ/, /v/, /θ/, and /ð/) and affricate consonants (/tʃ/, /dʒ/) as spoken in a VC context by 3 female talkers were used (cf. Stelmachowicz et al. 2001). The front vowel /i/ was used to minimize cues provided by formant transitions. Finally, speech tokens were mixed with speech-shaped

TABLE 1. Listeners' age, test ear, whether they were regular hearing aid users, and their thresholds in dB HL at each frequency (in kHz)

Listener	Age	Ear	Aid?	Threshold, dB HL									
				0.25	0.5	0.75	1.0	1.5	2.0	3.0	4.0	6.0	8.0
HI 01	59	R	Y	20	20	5	10	15	50	50	60	60	65
HI 02	68	R	N	15	10		10	25	40	50	60		55
HI 03	47	L	N	15	25	30	35	40	45	50	50	45	40
HI 04	59	R	Y	20	25	35	35	40	40	50	45	55	65
HI 05	69	R	Y	15	15	20	30		35	50	60	50	70
HI 06	56	L	N	10	10		15		25	35	45	55	60
HI 07	50	L	N	20	20	20	25	20	20	45	45	65	65
HI 08	21	R	N	25	35	40	40	45	50	45	40	45	45
HI 09	48	R	Y	20	30	40	50		55		60		65
HI 10	34	L	N	15	20	35	45	50	45	50	50	45	45
HI 11	18	R	Y	25	30	40	50		45		55		50
HI 12	68	L	Y	15	15	25	30	45	45	45	50	55	50
HI 13	61	R	Y	15	25	35	35	40	50	60	55	55	45
HI 14	66	R	N	25	15		5	25	35	60	65	75	50
HI 15	56	R	N	35	40		45		45	45	55	55	55
HI 16	70	R	Y	10	15		25	30	45	50	60	65	60
HI 17	57	R	Y	40	40		40	40	60		65		65
HI 18	71	R	Y	35	30		40		45		55	55	70
HI 19	51	R	N	15	10		25		35		45	65	65
HI 20	59	R	N	25	30		40		45	55	65	70	80
HI 21	74	R	N	5	10		15	20	45	70	70	75	55
HI 22	64	R	N	15	30		35		35	30	35	50	55
HI 23	53	R	N	20	20		25		35		45		55
HI 24	71	R	Y	15	30		30		30	45	50		55
Mean	56.3			19.6	22.9	29.5	30.6	33.5	41.7	49.2	53.5	57.8	57.7

noise* at 10 dB SNR and then low-pass filtered at 10 kHz using a minimum-order equiripple filter with a band stop of 100 dB at 11 kHz. These cumulative efforts resulted in stimuli where the primary spectral differences were confined to frequencies above 3 kHz, as shown in Figure 1.

Stimuli were presented monaurally to listeners via Sennheiser supra-aural headphones (HD-25-1) using control software developed in MATLAB (MathWorks, Natick, MA). NH listeners received the same amplified stimuli as the HI listeners to control for differences in perception attributable to cochlear nonlinearities that accompany increased presentation levels (e.g., broadened auditory filter tuning). Several procedures were implemented to help ensure that the sounds presented were representative of the processing and output levels associated with a conventional hearing aid fitting. Stimuli consisted of recordings made through a hearing aid with FT (Widex Inteo IN-9 BTE) and a hearing aid with NFC (Phonak Naída V UP BTE) at 62 dB SPL (representing an average conversational speech level) with most advanced features deactivated (e.g., directionality, feedback management, speech enhancement, expansion, and with 1 exception, noise reduction). During recording, each VC stimulus was preceded by a 2-second speech-in-noise carrier spoken by a female talker. Furthermore, the signals played to the hearing aids were 46- to 48-second long concatenations of all the carriers and stimuli for a given VC (3 talkers times 5 renditions) with 500 msec of silence separating them. The VCs were then extracted from the carrier and saved to hard disk for later presentation to the listeners.

Recordings were made synchronously with the playback using a two-channel CardDeluxe sound card (Digital Audio Labs, Chanhassen, MN). Signals were routed to the test chamber of a Fonix 6500-CX Hearing Aid Test System (Frye Electronics, Tigard, OR) and recorded with an ER-7C probe microphone (Etymotic Research, Elk Grove, IL). The hearing aids were connected to a 2-cm³ coupler using standard #13 earmold tubing. The probe microphone was connected to the opposite coupler opening and acoustically sealed.

The hearing aids were programmed for an audiogram with 30 dB HL thresholds from 0.25 to 1 kHz and with 50 dB HL thresholds from 2 to 6 kHz—a hearing loss representative of the sample reported here. The gain of each hearing aid as determined by a 1/3 octave band analysis of the recordings was repeatedly adjusted until it approximated the prescriptive target gain for a 2-cm³ coupler reported by the DSL v5.0a GUI (National Centre for Audiology, University of Western Ontario), assuming an average adult listener, BTE hearing aid style with wide dynamic range compression, no venting, and a monaural fitting. Figure 2 displays the gain for the different conditions discussed later in the article in comparison with the DSL target gain. As can be seen, hearing aid gain closely matched the prescriptive targets for the specified audiogram through 4 to 5 kHz. It can also be seen that gain rapidly dropped above the target region for each of the frequency-lowering algorithms (see figure legend and descriptions below).

To help ensure that the amplified output was audible for individual HI listeners, sensation level at the audiometric frequencies was computed by comparing the headphone output derived

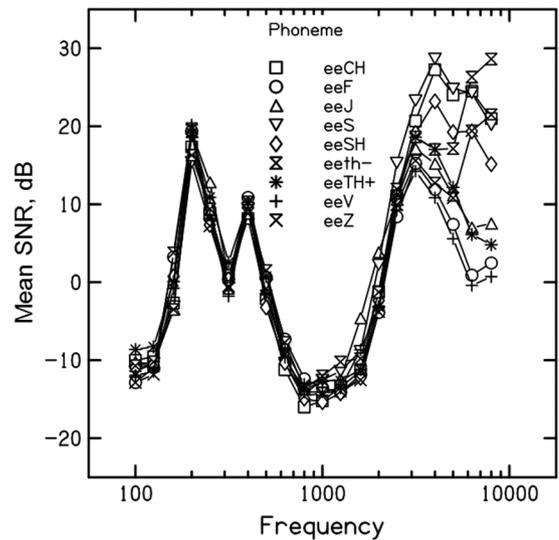


Fig. 1. Mean 1/3 octave spectra displayed in terms of SNR for the unprocessed stimuli after mixing with speech-shaped noise at 10 dB SNR. Spectra are averaged over 15 exemplars per vowel-consonant (3 female talkers times 5 renditions). SNR indicates signal-to-noise ratio.

from a 1/3 octave analysis of the digital waveform to pure-tone thresholds obtained in a separate procedure using the Sennheiser headphones, both were expressed in dB SPL relative to a flat plate coupler. As necessary on an individual-by-individual basis, an automatic routine in MATLAB increased the overall presentation level for all conditions so that the long-term average speech spectrum between 2 and 8 kHz for the wideband stimuli was equal to the average threshold in this range.† Level adjustments were ≤ 6 dB for 20 HI listeners because their audiograms were close enough to the audiogram programmed into the hearing aids. Three listeners required an 11 dB increase in presentation level and 1 listener required a 15 dB increase to make the average high-frequency speech spectrum audible. A plot of the average amplified speech spectra for both of the wideband conditions against the average thresholds obtained with the Sennheiser headphones from the HI listeners is shown in Figure 3.

Signal Processing

Frequency Transposition • The hearing aid used for FT in this study was a Widex Inteo IN-9 BTE, which has an input frequency response that rolls off around 7 kHz. Because the focus of this study was on individuals with mild to moderate high-frequency SNHL, the highest start frequency (6 kHz) was

* Speech-shaped noise was generated by filtering white noise to match the 1/3 octave spectra of the long-term average speech spectrum from the study by Byrne et al. (1994).

† In light of the previous paragraph, which emphasized the closeness of the initialized hearing aid output to DSL v5.0a targets for average listeners, two procedural details outlined here work to move real ear output away from targets, which is the reason why high-frequency audibility was double-checked using each individual's thresholds as obtained with the Sennheiser headphones. The first is the headphone response that shapes the output. When measured on a KEMAR and IEC 711 coupler, the frequency shaping provided by the headphone (re: 1.0 kHz) was estimated to be: -4.3 dB (0.25 kHz), -0.2 dB (0.5 kHz), -0.8 dB (0.75 kHz), 2.5 dB (1.5 kHz), 7.2 dB (2.0 kHz), 7.7 dB (3.0 kHz), 3.7 dB (4.0 kHz), -2.8 dB (6.0 kHz), and 7.6 dB (8.0 kHz). The second is the overall level adjustment to compensate for high-frequency audibility, which could lead to overamplification of the low-mid frequencies and possibly upward spread of masking. However, this was determined not to be of concern because the low-mid frequency parts of the signal were primarily confined to the vowel portion and the high-frequency parts to the fricative/affricate.

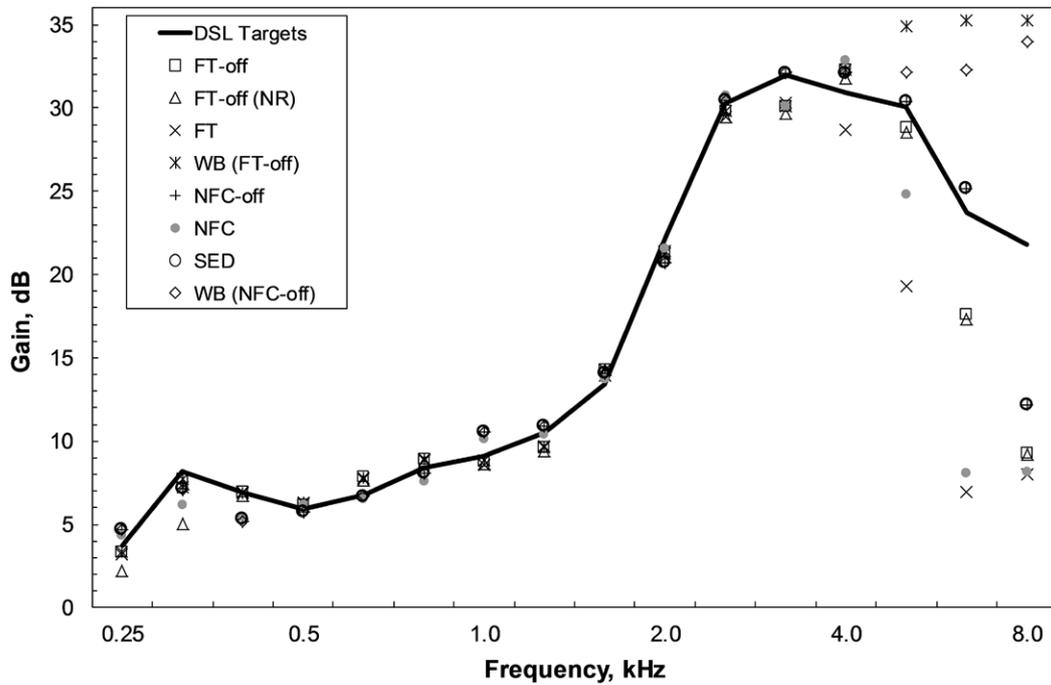


Fig. 2. Average gain applied to the stimuli for each condition (different symbols) compared with the prescribed DSL targets (solid line). FT-off (NR) corresponds to the recordings made through the Inteo with frequency transposition deactivated and noise reduction activated. Symbols for NFC-off (+) and SED (o) overlap because SED stimuli were processed off-line and recorded through the hearing aid with NFC-off. FT indicates frequency transposition; NFC, nonlinear frequency compression; NR, noise reduction activated; SED, spectral envelope decimation; WB, wideband.

chosen. This means that the source region spanned approximately 4.2–7 kHz and the target region spanned approximately 2.1–3.5 kHz. An example recording from the hearing aid with

this setting is shown in Figure 4. Displayed is a spectrogram of the carrier sentence “children like strawberries” and the stimulus “eeSH” before (top) and after (bottom) FT. The boxes in

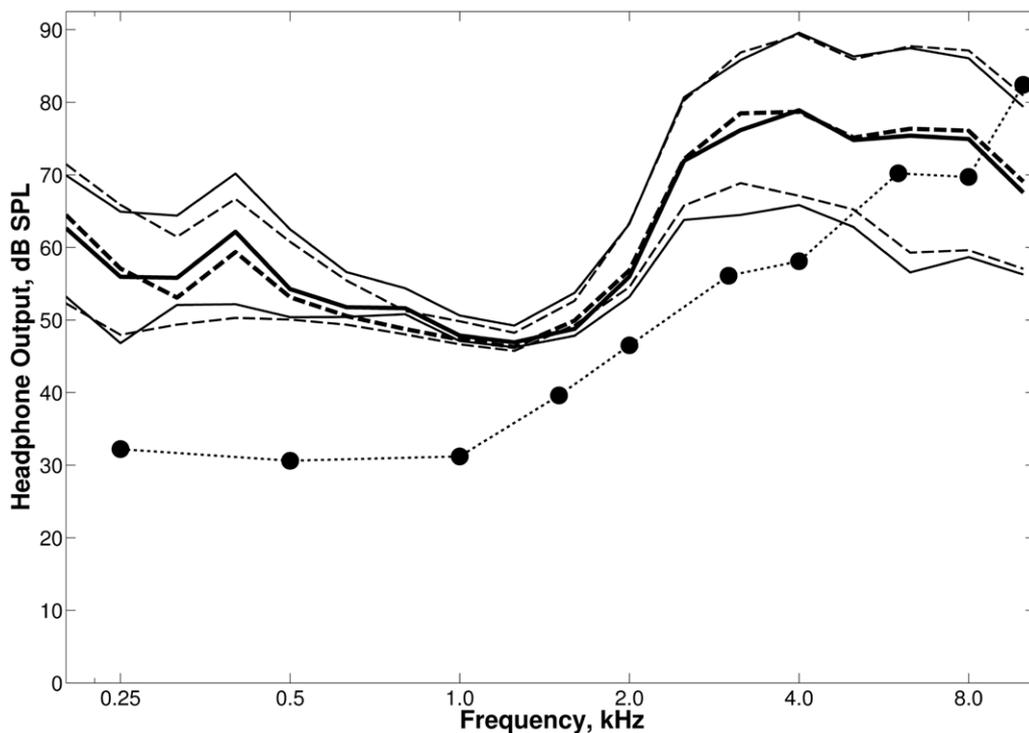


Fig. 3. Average pure-tone thresholds of the hearing-impaired listeners measured using the Sennheiser headphones (filled circles) are plotted alongside the distribution of 1/3 octave band levels for amplified speech energy in the wideband conditions. The solid and dashed lines represent conditions generated using recordings from the Inteo and Naída hearing aids, respectively. The thick line in the center of each distribution corresponds to the amplified long-term average speech spectrum of the VCs and the top and bottom thin lines correspond to the peaks (first percentile) and the valleys (70th percentile), respectively.

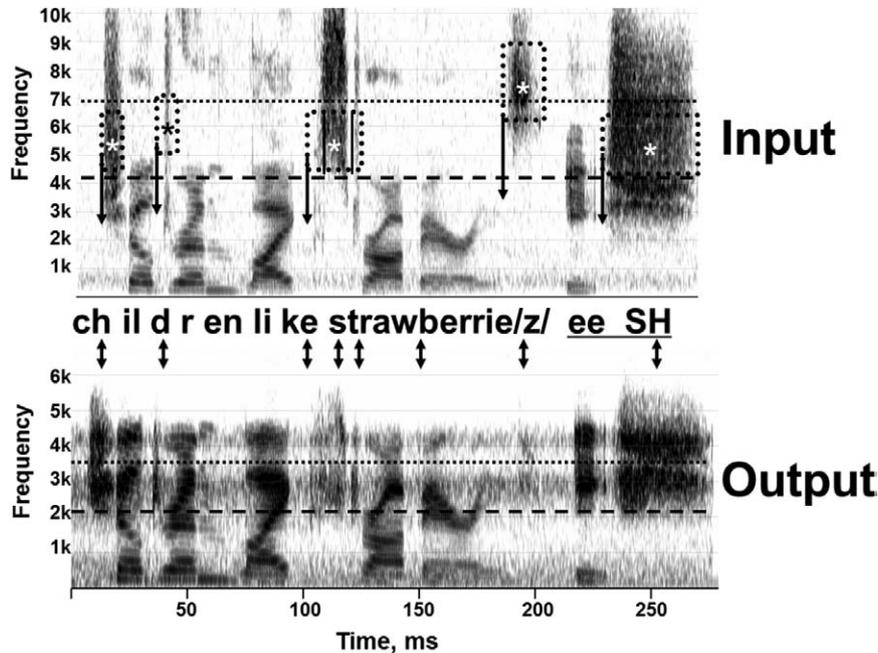


Fig. 4. A spectrogram of the carrier sentence “children like strawberries” and the stimulus “eeSH” before (top) and after (bottom) processing with frequency transposition. Arrows highlight the phonemes with significant high-frequency energy, which are most likely to be affected by the frequency-lowering settings used in this study. See text for details.

the top panel show identifiable regions where the signal will be transposed. Asterisks in each box correspond roughly to the center frequency of the transposed band and arrows indicate where it likely is moved to in the output. The dotted and dashed lines in each panel correspond to the upper and lower limits, respectively, of the source (top panel) and target (bottom panel) bands.

Nonlinear Frequency Compression • The hearing aid used for NFC in the first experiment was a Phonak Naída V UP BTE, which has an input frequency response that rolls off around 6 kHz when NFC is activated. The highest start frequency for this product, 4 kHz, and a compression ratio of 2.3:1 were chosen.‡ This means that the source region spanned approximately 4 to 6 kHz and the target region spanned approximately 4 to 4.7 kHz. An example recording from the hearing aid with this setting is shown in Figure 5. Displayed is a spectrogram of the same source stimulus as shown in Figure 4 before (top) and after (bottom) going through NFC. The dotted and dashed lines in each panel correspond to the upper and lower limits, respectively, of the source (top panel) and target (bottom panel) bands. The hatched region in the top panel shows the part of the spectrum that is not subject to frequency lowering with this hearing aid.

Spectral Envelope Decimation • Because neither hearing aid could lower frequencies above 6 or 7 kHz and because we were interested in examining whether the benefits of extending the bandwidth to 9 to 10 kHz could also be achieved using frequency lowering, a novel form of frequency compression that lowered frequencies up to 10 kHz was implemented in MATLAB. Frequencies above a 4 kHz start frequency were compressed by a factor of 2:1, such that the source region spanned approximately 4 to 10 kHz and the target region spanned approximately 4 to

7 kHz. Frequency compression was implemented by taking a 512-point Fast Fourier Transform (FFT) of overlapping segments from the source signal and decimating the amplitude spectra above 4 kHz by a factor of two while preserving phase. After converting the altered FFT back into the time domain using an inverse FFT, the output signal was reconstructed using 75% overlap and added (Allen 1977). Whereas frequency remapping above the start frequency with the Phonak hearing aid is logarithmic (nonlinear), frequency remapping with this method (spectral envelope decimation [SED]) is linear so that a signal 3 kHz above the start frequency before lowering will be 1.5 kHz above it after lowering and a signal 4 kHz above the start frequency before lowering will be 2 kHz above it after lowering, and so on. A comparison of the frequency input–output functions from the NFC and the SED algorithms are shown in Figure 6. The figure only shows input frequencies through 8.6 kHz for SED because, as discussed later in the article, the stimuli processed with this algorithm were recorded through the Naída UP V hearing aid (with NFC deactivated), where the output frequency response rolled off around 6.3 kHz (i.e., 2.3 kHz above the start frequency). Figure 7 shows an example spectrogram of speech processed with the SED algorithm.

Conditions

Stimuli were recorded through a Widex Inteo IN-9 BTE for FT and a Phonak Naída V UP BTE for NFC and SED. For each hearing aid, there were four conditions. The primary control condition for each hearing aid consisted of recordings made through the hearing aid with frequency lowering deactivated: FT-off and NFC-off. Because noise reduction with the Widex Inteo could not be deactivated when in FT mode, an additional control condition for FT was added (i.e., FT-off, noise reduction activated). The lack of a significant difference in the main experiment between the two conditions indicates that noise

‡ Because the NFC start frequency and compression ratio options made available to the dispenser depend on the audiogram entered into the iPGF v2.1 software, the setting with the lowest start frequency above 4 kHz was chosen.

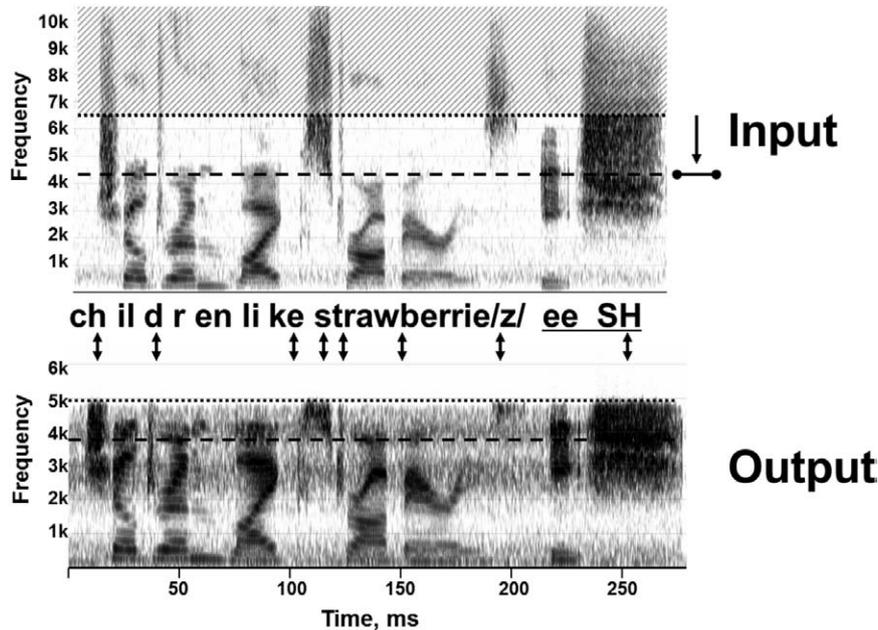


Fig. 5. A spectrogram of the carrier sentence “children like strawberries” and the stimulus “eeSH” before (top) and after (bottom) processing with nonlinear frequency compression. See text for details.

reduction did not appear to influence the outcomes associated with FT (where noise reduction in this setting was activated by default) in this experiment; therefore, this condition will not be discussed further. The frequency-lowering algorithms included FT, NFC, and SED using the settings described previously. Preprocessed SED stimuli were recorded through the Naída hearing aid with NFC deactivated to ensure equivalency for all conditions except for the preprocessing.

A final control condition for each hearing aid consisted of a wideband frequency response through 10 kHz: WB (FT-off) and

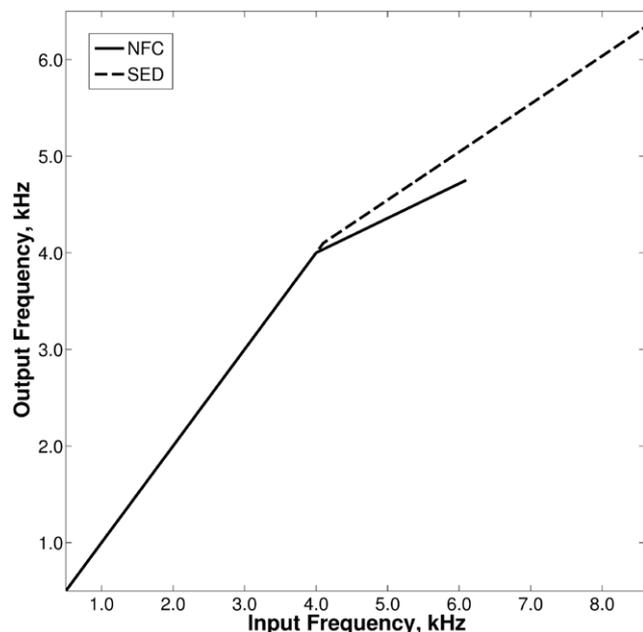


Fig. 6. A comparison of the frequency input–output functions from the nonlinear frequency compression (NFC, solid line) and spectral envelope decimation (SED, dotted line) algorithms. NFC indicates nonlinear frequency compression; SED, spectral envelope decimation.

WB (NFC-off), which were constructed using the recordings made through the Inteo and Naída, respectively. The purpose of including these conditions was to assess how much benefit potentially could be obtained by extending the hearing aid bandwidth using conventional amplification and to compare how much of this benefit could be recovered using frequency lowering. Obviously, the necessary high-frequency gain could not be obtained using the hearing aids; however, it was desirable to maintain equivalent conditions between the FT-off/NFC-off conditions and the WB conditions with the exception of the additional high-frequency information. Therefore, using MATLAB, the recordings for the FT-off and NFC-off conditions were each time-aligned and then combined with the source stimuli that had been high-pass filtered at 5 kHz using a 150-order finite impulse response filter with a band stop of 60 dB at 4.2 kHz. Figure 8 shows the spectra for the Naída UP hearing aid (solid line) and for the high-pass filtered source (dashed line) for the concatenated carriers and VCs for the stimulus /is/. Gain of the high-pass signals was adjusted separately for each set of VCs. A 1/3 octave band analysis was performed on each set of hearing aid recordings and each source signal (concatenated carriers and VCs). The gain of the high-pass source signal was adjusted so that the maximum 1/3 octave filter level was equal to the maximum filter level of the recorded stimuli (low-pass). This effectively created a flat frequency response in the long-term average speech spectrum from 3 to 10 kHz (cf. Fig. 3). As shown in Figure 3, audibility in the wideband conditions extended to around 8.6 kHz when compared with average thresholds, which conveniently corresponded to the upper input frequency made audible by SED after lowering and processing by the Naída hearing aid in NFC-off mode.

Procedures

Listeners completed the experiments in two 60- to 90-min sessions in which the task was to identify the VCs recorded from only one of the hearing aids. Half of the listeners identified

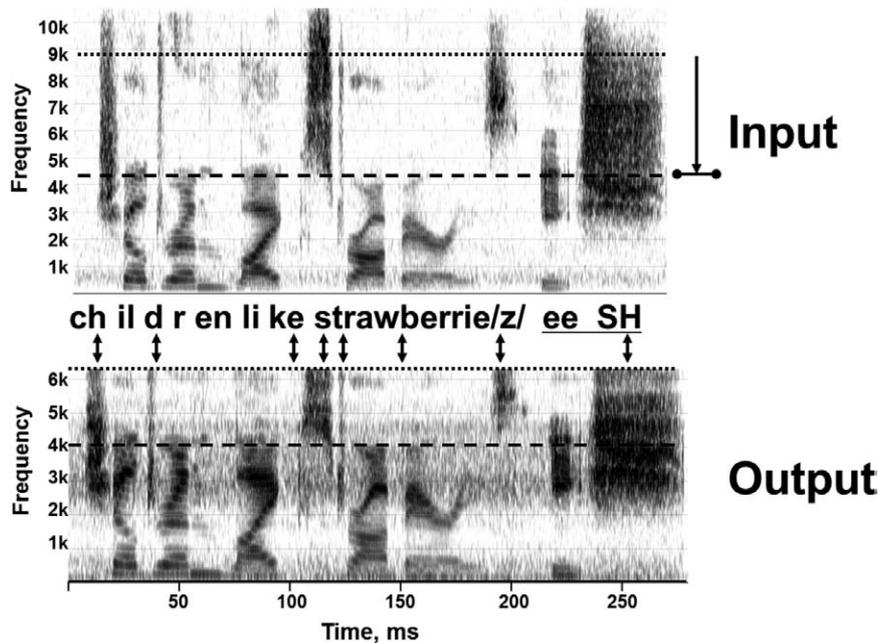


Fig. 7. Same as Figure 5, but for spectral envelope decimation.

VCs recorded through the Inteo hearing aid in the first session and identified the same VCs recorded through the Naída hearing aid in the second session. The order was reversed for the remaining listeners. The order of the four conditions for each hearing aid was counterbalanced using a Latin squares design. One of the five renditions of each VC from the talkers was used for familiarization and practice phases; the other four were used exclusively during the test phase. For the familiarization phase, each of the nine VCs spoken by one of the three talkers was presented to the listener along with the orthographic representation on the computer screen (e.g., “eeS”). The orthographic representation for /θ/ was “eeth (teeth)” and the representation for /ð/ was “eeTH (breathe).” For the practice and test phases, a grid of response options appeared on the computer screen, including “ee” in case a listener did not hear the final frication and “Other” in case a listener heard something altogether different. § Listeners responded by using the mouse or touch screen monitor to select one of the response options. The procedure for the practice phase was the same as the test phase, except that listeners received feedback about which token was the correct response. For the practice phase, each of the nine VCs spoken by one of the three talkers was presented twice. For the test phase, listeners heard each of the VCs spoken by the three talkers in blocks of 27 trials. There were four blocks of trials for a total of 108 test trials.

RESULTS

Performance for FT and its control conditions are shown in Figure 9. Results for the NH listeners are represented by the light-gray bars and results for the HI listeners are represented by the dark-gray bars. Error bars indicate the standard error of

the mean. Despite amplification, HI listeners performed significantly worse than NH listeners in every condition, which is consistent with reduced information processing associated with cochlear dysfunction (Pavlovic 1984; Ching et al. 1998). However, of primary interest is the pattern of results within each listener group. To assess differences between conditions, a Wilcoxon signed rank test was performed after controlling for possible order effects. ¶ For the HI listeners, performance with FT was significantly worse compared with the other conditions ($p < 0.001$) while performance with WB was significantly better than in the other conditions ($p < 0.001$). For the NH listeners, FT was also significantly worse compared with the other conditions ($p < 0.001$). No other differences were significant ($p > 0.05$).

Performance for NFC and its control conditions are shown in Figure 10. For the HI listeners, SED and WB were both significantly better than NFC-off and NFC ($p < 0.05$ each; except WB versus NFC, where $p < 0.01$). There was no significant difference between SED and WB, indicating that the benefit observed from an increase in the uncompressed bandwidth was also obtained when spectral envelope decimation was used to compress the high-frequency bandwidth. It is of interest to note that there was no significant difference in performance between NFC-off and NFC, suggesting that while NFC did not improve overall performance, it did not make it worse. Finally, the pattern of results for the NH listeners for NFC was similar to the results for FT; in both cases performance was significantly worse compared with the other conditions ($p < 0.001$ each; except NFC versus SED, where $p < 0.01$). No other differences were significant ($p > 0.05$).

To better understand how the different signal processing strategies influenced speech perception, confusion matrices were generated from the aggregate responses for both the HI

§ The response options “ee” and “Other” were selected only 55 times of 41,427 test trials or ~0.1% of the total trials. Thirty-two of these were made by the HI listeners and 23 by the NH listeners. These responses most often occurred when /θ/ or /ð/ were presented.

¶ Because performance on the first condition across the eight tended to be worse than the rest when results were collapsed across the different manipulations, the effect of test order (1 to 8) was partialled out of the results by using the residuals from a quadratic regression with test order as the predictor variable. The Wilcoxon signed rank test, a nonparametric alternative to Student’s t test, was used because normality of the small sample could not be assumed.

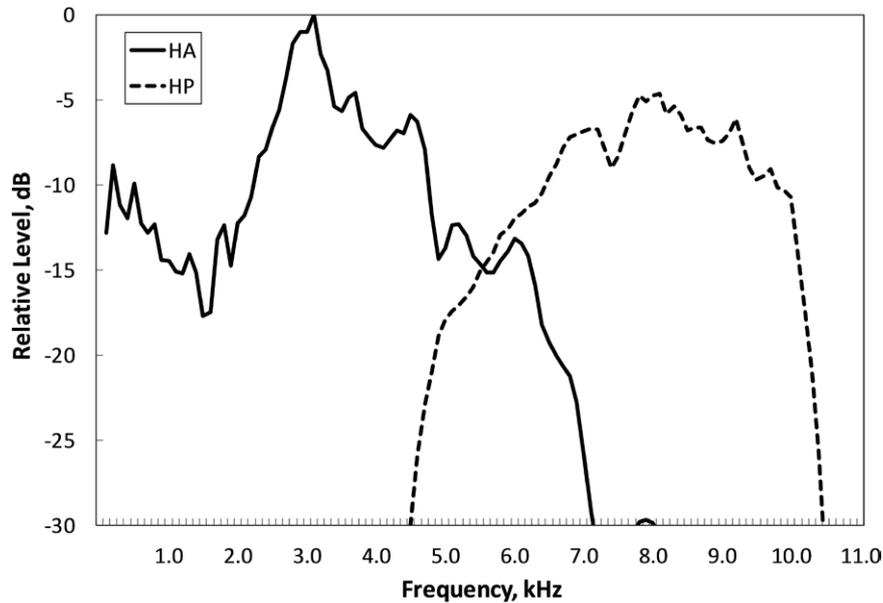


Fig. 8. An example of the HA response for the Phonak Naída UP with NFC-off when analyzed with a 256-point Fast Fourier Transform (solid line) and the response for the high-pass filtered source (dashed line) for the concatenated carriers and vowel–consonants for the stimulus /is/ that together formed the wide-band stimulus. HA indicates hearing aid; HP, high-pass filtered source.

and NH listeners. Statistical significance was evaluated using methods outlined by Thornton and Raffin (1978) and Carney and Schlauch (2007). First, proportions representing each stimulus–response combination (81 total) were converted into arcsine units to stabilize the variance [Eq. (3) in the study by Thornton & Raffin (1978)]. Then, the differences in arcsine units between the FT-off condition (representative of “typical” hearing aid processing) and its control conditions were computed separately for each listener group. The same was done for NFC-off and its control conditions. These differences were then compared with the critical value, which was computed by multiplying the inverse of the standard normal cumulative

distribution by the square root of the estimated variance of the binomial distribution in arcsine units [Eq. (4) in the study by Thornton & Raffin (1978)]. The inverse of the standard normal cumulative distribution is the z value corresponding to a given probability of a type I error (α divided by 2, for a two-tailed test of significance). A Bonferroni correction was applied to $\alpha = 0.05$ to control for family-wise error (i.e., divided by 81, the number of comparisons).

Percent correct for each phoneme in each condition is shown in Table 2, in which the bolded and underlined entries correspond to significant differences compared with the FT-off condition or the NFC-off condition. Confusion matrices are

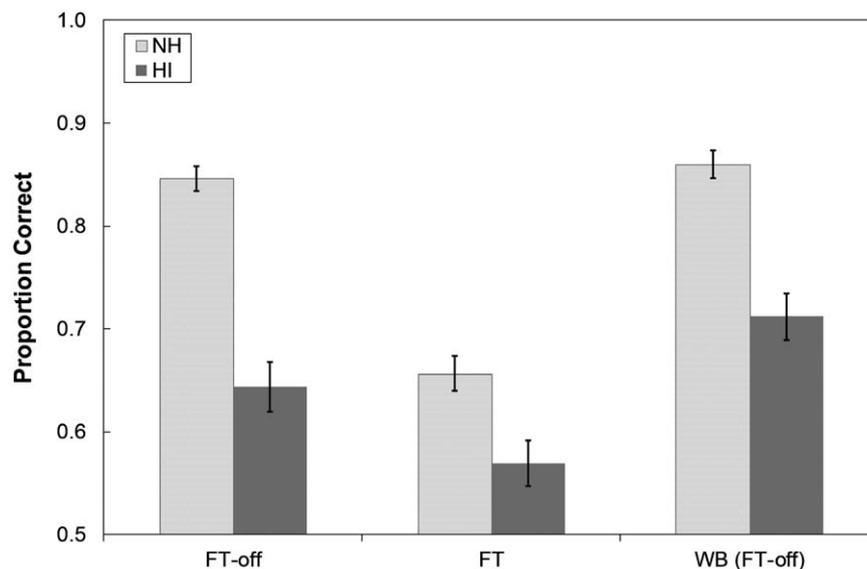


Fig. 9. Mean proportion correct for the NH listeners (light-gray bars) and HI listeners (dark-gray bars) for conditions involving the Widex Inteo hearing aid with FT activated and deactivated (FT-off). The WB control was created by mixing the recordings from the FT-off condition and the original source signal after high-pass filtering. Error bars represent the standard error of the mean. FT indicates frequency transposition; HI, hearing impaired; NH, normal hearing; WB, wideband.

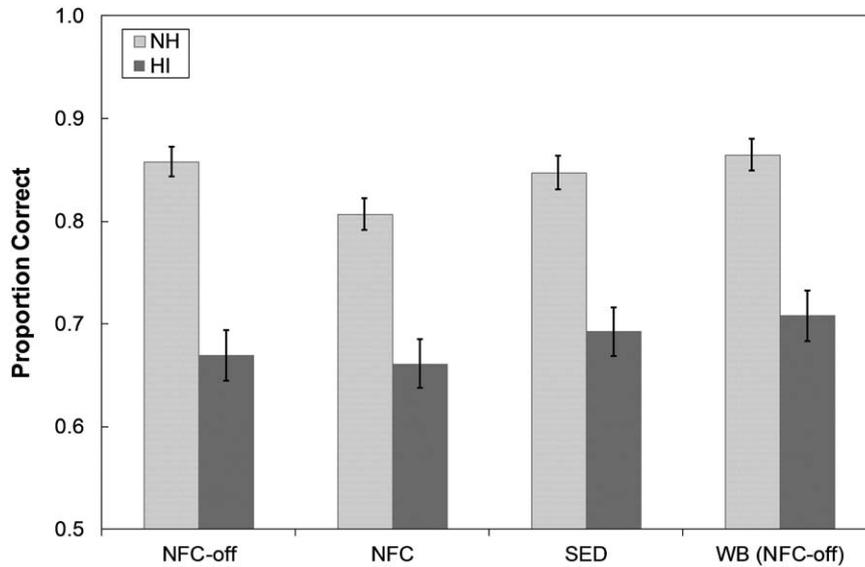


Fig. 10. Mean proportion correct for the NH listeners (light-gray bars) and HI listeners (dark-gray bars) for conditions involving the Phonak Naída V UP hearing aid with NFC activated and deactivated (NFC-off). SED was processed off-line and recorded through the hearing aid in its NFC-off setting. Whereas NFC lowered frequencies up to 6 kHz, SED lowered frequencies up to ~8.6 kHz in the band-pass range of the hearing aid. The WB control was created by mixing the recordings from the NFC-off condition and the source signal after high-pass filtering. Error bars represent the standard error of the mean. HI indicates hearing impaired; NFC, nonlinear frequency compression; NH, normal hearing; SED, spectral envelope decimation; WB, wideband.

provided in Supplementary Table 1 (see Supplemental Digital Content 1, <http://links.lww.com/EANDH/A141>). For the HI listeners, significant differences between conditions were confined to the phonemes /s/ and /z/. Identification of the labiodental and interdental fricatives across all conditions was generally poor and identification of /j/ and the affricates was generally good across all conditions. Consistent with the findings of Stelmachowicz et al. (2001), /s/ and /z/ identification was significantly better in the WB conditions compared with the FT-off and NFC-off conditions in which audible bandwidth was restricted by the maximum output power of the hearing aids. Specifically, the WB conditions significantly decreased place of articulation

errors for /s/ and /z/, which were confused for one of the labiodental or interdental fricatives and vice versa. In contrast, FT significantly decreased /s/ and /z/ identification relative to FT-off, in part, because of a significant increase in confusions of /s/ as /f/, /θ/, /ʃ/ and confusions of /z/ as /v/, /ð/, /dʒ/. Therefore, the primary effect of FT for the HI listeners was an increase in place errors with no effect on voicing errors. NFC, however, essentially had no effect on phoneme identification compared with NFC-off. Finally, the overall improvement observed for SED versus NFC-off can be attributed to a significant improvement in /s/ identification. Significant decreases in confusions of /s/ as /θ/ and confusions of /z/ as /v/ were also observed.

TABLE 2. Percent correct for each condition across the different phonemes is displayed for the hearing-impaired and normal-hearing listeners in the top and bottom tables, respectively

Listeners	/f/	/θ/	/s/	/ʃ/	/tʃ/	/v/	/ð/	/z/	/dʒ/
Hearing impaired									
FT-off *	49.3%	53.8%	61.5%	94.1%	92.4%	54.9%	31.9%	60.8%	80.9%
FT	54.5%	53.1%	21.9%	92.0%	91.7%	56.6%	26.4%	32.6%	80.9%
WB (FT-off)	53.8%	62.5%	86.1%	90.6%	91.3%	61.1%	34.7%	84.4%	80.2%
NFC-off *	57.3%	51.7%	61.8%	95.1%	93.4%	60.8%	28.1%	64.9%	89.2%
NFC	51.7%	46.5%	70.5%	92.0%	93.4%	56.9%	30.9%	63.5%	89.6%
SED	54.9%	53.5%	76.0%	92.4%	94.4%	57.3%	30.6%	71.2%	93.1%
WB (NFC-off)	53.5%	53.5%	86.8%	92.0%	92.7%	56.6%	30.2%	81.9%	89.9%
Normal hearing									
FT-off *	84.4%	66.0%	98.6%	99.7%	91.0%	85.1%	42.0%	97.6%	97.2%
FT	74.0%	61.1%	22.9%	97.6%	95.8%	71.9%	45.8%	25.3%	96.5%
WB (FT-off)	66.7%	81.9%	97.9%	97.2%	93.4%	84.7%	59.0%	97.2%	95.8%
NFC-off *	69.8%	82.6%	96.9%	99.3%	95.8%	79.2%	56.6%	93.4%	98.3%
NFC	74.7%	59.4%	81.9%	98.6%	96.5%	78.5%	52.4%	85.1%	99.0%
SED	72.9%	67.4%	98.3%	99.3%	93.4%	81.6%	59.0%	92.7%	98.3%
WB (NFC-off)	70.5%	79.2%	96.5%	98.3%	94.4%	79.2%	66.3%	95.8%	97.9%

The values in italics denote significant differences relative to the performance of the recordings made with the hearing aid with FT and NFC deactivated, FT-off and NFC-off, respectively (denoted with asterisks).

FT frequency transposition; NFC, nonlinear frequency compression; SED, spectral envelope decimation; WB, wideband.

Similar analyses are also helpful for interpreting how the signal processing strategies influenced speech perception in NH listeners. Recall that performance with FT was significantly worse compared with its control conditions. This can be primarily attributed to a substantial decrease in /s/ and /z/ identification (same as the HI listeners) and a decrease in /f/ and /v/ identification compared with FT-off. It is interesting to note that the types of confusions for /s/ and /z/ were similar for the two groups of listeners. NH listeners also performed more poorly with NFC compared with NFC-off. This occurred because of a decrease in identification of /s/, /z/, and /θ/. Significant increases in confusions for /s/ and /z/ occurred only with the voiceless and voiced labiodentals and interdental, respectively. This set of place errors is more restricted than the set for FT. Additional increases in place confusions associated with NFC occurred for the interdental, which were identified more often as labiodentals.

FOLLOW-UP EXPERIMENT

The results of the previous experiment indicate that for HI listeners, performance with SED was equivalent to WB (NFC-off) and that performance with NFC was equivalent to NFC-off. The effective input bandwidth for the former conditions was about 8.6 kHz while it was only about 6.0 to 6.3 kHz for the latter two conditions, suggesting that when the start frequency is at least 4 kHz, frequency compression can transmit the same information as an uncompressed signal with the same effective bandwidth. In addition, the results indicate that the advantage of extending bandwidth as reported here and elsewhere (e.g., Stelmachowicz et al. 2001) applies to the effective bandwidth for frequency compression. It is not known whether this finding is unique to the form of frequency compression implemented by SED (linear frequency remapping) or to other details related to the signal processing in MATLAB. Therefore, the purpose of this follow-up experiment was to examine whether the outcomes observed with SED could be obtained using the signal processing in an actual hearing aid.

PARTICIPANTS AND METHODS

Twelve of the 24 HI listeners from the first experiment participated in this experiment. The stimuli, recording methods, and procedures were the same as described earlier in this article. There were three conditions: NFC-off, NFC, and WB. The order of the conditions was counterbalanced using a Latin squares design. All stimuli were recorded through a Phonak Naída V SP BTE, which has an output frequency response similar to that used in the Naída V UP. One difference is that, depending on the start frequency, NFC in the Naída V SP can lower frequencies up to 10 kHz. The NFC setting used had a start frequency of 4.6 kHz and a compression ratio of 2.8:1.‡ With this setting, the output frequency response rolled off around 5.6 to 5.7 kHz, which means that the highest frequency in the input bandwidth was around 8.6 kHz, the same approximate effective bandwidth as WB and SED.

RESULTS

To assess differences between conditions, a Wilcoxon signed rank test was performed after controlling for possible order effects. Performance with NFC ($M = 0.724$, $SD = 0.102$) was

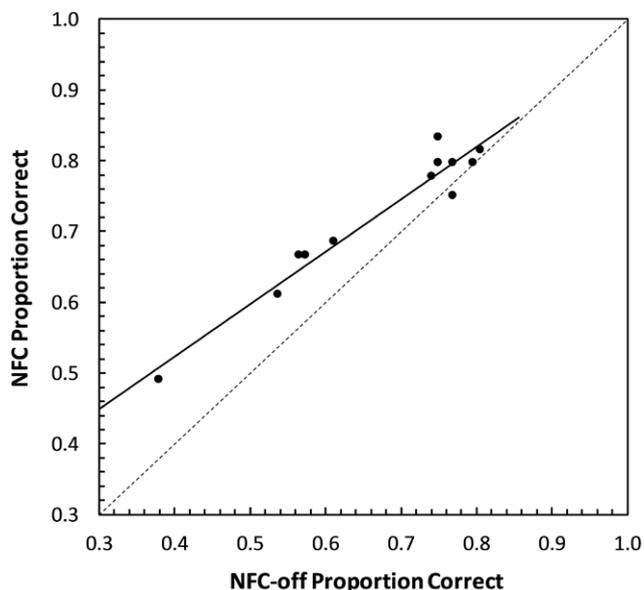


Fig. 11. Scatterplot showing the proportion of correct responses for HI listeners in the follow-up experiment where stimuli were recorded through a Phonak Naída V SP hearing aid with NFC activated and deactivated (NFC-off). The slope of the linear least squares regression fit to the data (solid line) is significantly less than 1.0, indicating that listeners who performed more poorly with NFC-off tended to show the most improvement with NFC. HI indicates hearing impaired; NFC, nonlinear frequency compression.

significantly better than NFC-off ($M = 0.671$, $SD = 0.134$), [$p < 0.01$]. Figure 11 shows a scatterplot of the proportion of correct responses for NFC versus NFC-off. Nearly every data point is above the diagonal, indicating that most listeners performed the same or better with NFC compared with NFC-off. The slope of the linear least squares regression fit to the data (solid line) was significantly less than 1.0 [$t(10) = 12.8$, $p < 0.001$], indicating that listeners who performed more poorly with NFC-off tended to show the most improvement with NFC. Individual differences in audibility would not be expected to contribute to these performance differences because high-frequency gain for NFC was less than for NFC-off (cf. Fig. 2), meaning that improvement was likely due to the content, rather than the audible bandwidth, of the high-frequency spectrum after lowering. As a check, a Pearson correlation between the mean 4 to 6 kHz sensation level of the NFC-off stimuli for each individual and the NFC versus NFC-off difference score revealed no significant relationship ($r = 0.17$, $p > 0.05$).

Performance with WB ($M = 0.717$, $SD = 0.091$) was not significantly different from NFC-off or NFC. The lack of a significant difference between NFC-off and WB seems at odds with the first experiment; however, because the means and standard deviations across the two experiments are comparable, it is possible that the difference would have become significant if all 24 HI listeners had completed the follow-up experiment.

DISCUSSION

The results reported here demonstrate the potential benefit of using some form of frequency compression with a high input bandwidth (SED or NFC) for individuals with mild to moderate high-frequency SNHL, but caution against using FT for this population. It is important to note that neither technology was designed for mild to moderate losses. In fact, the use of

FT for the individuals in this study is inconsistent with the recommended criteria established by the manufacturer (Kuk et al. 2009a). For both NH and HI listeners, FT degraded the perception of /s/ and /z/, resulting in a significant increase in place of articulation confusions with all the voiceless (except /tʃ/) and voiced phonemes, respectively, in the stimulus set. At first, this might seem to make sense because spectral peak location can be a strong cue for place of articulation and because spectral peaks in the 4.2 to 7 kHz region were moved to the 2.1 to 3.5 kHz region. However, the errors one would predict based on spectral peak location alone would be /f/, /ʒ/, or /dʒ/, which have lower spectral peaks in natural speech. Furthermore, as Kent and Read (2002) point out, the nonstrident fricatives (/f/, /v/, /θ/, /ð/) are generally of low intensity and spectrally diffuse in contrast to the strident fricatives (/s/, /z/, /ʃ/, /ʒ/), which can be up to 10 dB higher in intensity, thus making “it unlikely that a strident would be confused with a nonstrident, or vice versa.” This suggests that /s/ and /z/ confusions were more random and cannot be explained simply by lowered spectral peaks. One possible reason could be that the hearing aid with FT only lowered spectral information up to around 7 kHz, thus missing all or some of the spectral peak information for /s/ and /z/. It is interesting to note that the hearing aid with NFC in the first experiment only lowered information up to 6 kHz and an increase in these confusions was not observed, so input bandwidth alone cannot explain the errors observed with FT.

Two factors, the level of the lowered speech and the presence of background noise, might jointly explain the confusions associated with FT. First, the relatively high intensities associated with the production of /s/ and /z/ might not have been maintained after lowering because only part of the high-frequency spectrum (a 1 octave-wide band relative to the target region) was lowered with the FT algorithm. The relative level of the lowered speech is an adjustable parameter in the programming software and ranges from –16 to +14 dB. The default value, 0 dB, was chosen; therefore, the results may have differed if the level had been increased. However, it is also unlikely that the lowered speech was inaudible because expected performance would have been the same as FT-off. The negative influence on /s/ and /z/ identification resulting from the lowered spectral information indicates that it was audible enough to be integrated with the speech stream. In particular, the continuous background noise combined with the energy of the lowered speech might have created a percept consistent with one of the nonstrident fricatives. That is, when energy from /s/ and /z/ was lowered by FT, listeners detected that the spectrum differed from the background noise and thus identified the most probable phoneme that matched the spectrum of the mixed speech and noise (i.e., one with spectrally diffuse energy in the low frequencies). This effect could be similar to the phonemic restoration phenomenon whereby missing speech sounds are perceived when replaced by masking noise (Warren 1970).

The results for FT are in contradiction to findings reported for adults (Kuk et al. 2009b, 2011) and children (Auriemma et al. 2009) with severe to profound SNHL above 2 to 3 kHz, who identified nonsense syllables in a CVCVC context with five different vowels spoken by a female talker. In comparison to the studies by Kuk and colleagues, the obvious limitation of the present study is that listeners did not actually have the experience of wearing the hearing aid with FT; therefore, they did not have an opportunity to acclimate to the novel speech code.

In the the study by Kuk et al. (2009b) study, in addition to daily wear, listeners received 20 to 30 min of daily training with FT for 4 to 6 weeks. In response to training/acclimatization, Kuk et al. (2009b) and Auriemma et al. (2009) reported improvements of about 10% for fricatives compared with baseline performance. Kuk et al. (2011) later demonstrated that training with FT deactivated did not improve performance, suggesting that improvement in earlier studies was due to FT, rather than training. While some improvement in this experiment might be expected if listeners had both training and exposure, some observations are difficult to explain by training effects alone. The substantial decrease in /s/ and /z/ identification with FT (about 30 to 40% for the HI listeners and about 75% for the NH listeners) and the fact that HI listeners received immediate benefit from frequency compression (SED and NFC) are strong indications that other factors need to be considered. In addition, it is interesting to note that Kuk et al. (2009b) found that /s/ and /z/ were among the phonemes that did not show improved performance for FT versus FT-off in noise, even after 2 months of training/acclimatization.

An important difference between this study and that by Kuk et al. (2009b) is severity of hearing loss. Listeners in that study by Kuk et al. had severe to profound SNHL that limited aided audibility in the FT-off condition to 2 to 3 kHz, whereas, listeners in this study had aided audibility out to 5 to 6 kHz. This bandwidth difference can account in part for the difference in baseline performance for fricatives and affricates in the FT-off condition (about 20% and 40% in the study by Kuk et al. versus 58% and 87% in this study). When deficits in performance are great, there is more opportunity for subtle acoustic changes to result in measureable benefit. Furthermore, the differences in hearing loss had consequences for where the lowered information was located with respect to listeners' residual bandwidth. For this study, the highest start frequency for FT was 6 kHz, indicating that the friction noise was lowered to the 2.1 to 3.5 kHz region—somewhere in the middle of the listeners' audible frequency range. It is possible that this interfered with the natural spectral profile that was audible out to 5 to 6 kHz, thus disrupting cues for phoneme identity (e.g., causing a spectral profile for /s/ to sound more like that for /f/). With more severe loss, there is no additional information that is audible above the frequency range of the lowered signal to cause interference. In summary, it is possible that FT can yield benefit for listeners with severe to profound SNHL, but negatively affect perception for listeners with mild to moderate hearing SNHL.¶

In contrast to FT, frequency compression (SED and NFC) performed as well as or better than the control condition with frequency compression deactivated. When the input bandwidth for NFC was limited to 6 kHz in the first experiment, performance was the same as NFC-off. The lack of improvement is not surprising because the effective input bandwidth was essentially the same for the two conditions. The finding that perception was not degraded is reassuring given that the natural high-frequency spectrum from 4 to 6 kHz was altered by NFC. Reasons for this difference in outcomes compared with FT could include the fact that the spectrum below 4 kHz was left unaltered and that when

¶ Using a method of frequency transposition different from the one implemented in Widex hearing aids, Robinson et al. (2009) did not find a benefit for listeners with severe SNHL. Additional research is clearly needed to elucidate who are candidates for the different frequency-lowering technologies.

the spectrum above it was compressed, the cues related to the overall spectral profile and intensity of frication were preserved.

Improvement over NFC-off was observed when the effective input bandwidth was extended out to about 8.6 kHz using natural bandwidth extension (WB in the first experiment), linear frequency remapping (SED), or nonlinear frequency remapping (NFC in the follow-up experiment). The most notable reduction in errors was for /s/ and /z/, a finding consistent with prior research that extended the bandwidth of the high-frequency spectrum without compression (e.g., Stelmachowicz et al. 2001). These results are also consistent with reports examining the efficacy of NFC in wearable devices for adults (Simpson et al. 2005; Glista et al. 2009; Bohnert et al. 2010) and children (Glista et al. 2009; Wolfe et al. 2011) with more severe SNHL.

The findings reported here support the hypothesis that NFC can provide some benefit for fricative identification to individuals with mild to moderate SNHL with little risk of introducing new phonemic confusions. However, it is important to keep in mind that the conditions tested were intentionally designed have significant high-frequency importance because they consisted of low-context speech sounds dominated by high-frequency content, spoken by female talkers embedded in low-frequency, speech-shaped noise. Several questions remain with regard to whether the results will generalize to actual daily use with sentences in a variety of vowel contexts, syllable structures, and talkers in a variety of background settings. The best evidence to date comes from listeners with more severe SNHL. However, as already noted, benefit relative to a no-treatment control may be more difficult to measure with mild to moderate SNHL because performance is not as degraded. Furthermore, even in a controlled laboratory study it is unknown whether NFC will degrade perception of vowels and other consonant sound classes, especially stops, which can be cued by the burst of high-frequency energy after the closure segment. It is likely that vowel perception will be unaffected when high NFC start frequencies are used for listeners with mild to moderate SNHL because vowels are cued primarily by the frequencies of the first two formants, which are generally <3 kHz (Kent & Read 2002). Recent research by Alexander (2012), who used a simulation of NFC to examine the effects of start frequency and compression ratio on consonant and vowel identification in listeners with mild to moderate SNHL, indicated that when the NFC start frequency was ≥ 2.7 kHz, significant improvements in fricative identification could be achieved without decrements in identification for vowels or other consonants. However, when lower start frequencies were combined with relatively high compression ratios, significant decreases in vowel identification were observed. It was hypothesized that this occurred because the relationship between formant frequencies was altered by NFC.

Finally, because a significant amount of the literature on the negative effects of restricted hearing aid bandwidth involves children (e.g., Stelmachowicz et al. 2002; Moeller et al. 2007, 2010), we will speculate on how our findings extend to a pediatric population with similar degrees of SNHL. We might expect to see a greater amount of benefit for fricative perception in children compared with adults simply because they appear to have a greater deficit when listening to speech under identical conditions. Reasons for this include the fact that adults are able to take advantage of linguistic knowledge to help them fill in the information they are missing. Because

children, especially HI children, do not have as rich a linguistic background, they need a more advantageous signal-to-noise ratio when identifying speech in noise (Nittrouer & Boothroyd 1990; Fallon et al. 2002; Hall et al. 2002). Important for this discussion, they also seem to benefit more from an increase in signal bandwidth compared with adults (Kortekaas & Stelmachowicz 2000; Stelmachowicz et al. 2001). It is reasonable, then, to hypothesize that the same difference in benefit will occur when additional information is made available with NFC. Future research is needed to confirm these hypotheses and the extent to which the benefits observed here with adults extend to different listening contexts.

ACKNOWLEDGMENTS

The authors thank Sandy Estee who aided with subject recruitment; Merry Spratford and Brenda Hoover who aided with conducting the experiments; and Dawna Lewis and Ryan McCreery who aided with experimental design.

This work was supported by the National Institutes of Health—National Institute on Deafness and Other Communication Disorders: an R01 (#DC004300-13) awarded to Dr. Patricia Stelmachowicz, a P30 Research Core grant (# DC004662-11) and a T32 training grant (# DC0000-13) awarded to Boys Town National Research Hospital.

The authors declare no other conflict of interest.

Address for correspondence: Joshua M. Alexander, Department of Speech, Language, and Hearing Sciences, Purdue University, Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907, USA. E-mail: alexan14@purdue.edu

Received July 26, 2012; accepted July 30, 2013.

REFERENCES

- Alexander, J. M. (2012). Nonlinear frequency compression: Balancing start frequency and compression ratio. Poster presented at the annual meeting of the American Auditory Society, March 8–10, 2012, Scottsdale, AZ.
- Alexander, J. M. (2013a). Individual variability in recognition of frequency-lowered speech. *Semin Hear*, 34, 86–109.
- Alexander, J. M. (2013b). 20Q: The highs and lows of frequency lowering amplification. *AudiologyOnline*, Article #11772.
- Allen, J. B. (1977). Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE ASSP*, 25, 235–238.
- ANSI (1996). *ANSI S3.6-1996, American National Standards Specification for Audiometers*. New York, NY: American National Standards Institute.
- ANSI (1997). *ANSI S3.5-1997, The Calculation of the Speech Intelligibility Index*. New York, NY: American National Standards Institute.
- Aurimmo, J., Kuk, F., Lau, C., et al. (2009). Effect of linear frequency transposition on speech recognition and production of school-age children. *J Am Acad Audiol*, 20, 289–305.
- Bilger, R. C., & Wang, M. D. (1976). Consonant confusions in patients with sensorineural hearing loss. *J Speech Hear Res*, 19, 718–748.
- Bohnert, A., Nyffeler, M., Keilmann, A. (2010). Advantages of a non-linear frequency compression algorithm in noise. *Eur Arch Otorhinolaryngol*, 267, 1045–1053.
- Braida, L., Durlach, I., Lippman, P., et al. (1978). Hearing aids—A review of past research of linear amplification, amplitude compression and frequency lowering. *ASHA Monog*, 19.
- Byrne, D., Dillon, H., Tran, K., et al. (1994). An international comparison of long-term average speech spectra. *J Acoust Soc Am*, 96, 2108–2120.
- Carney, E., & Schlauch, R. S. (2007). Critical difference table for word recognition testing derived using computer simulation. *J Speech Lang Hear Res*, 50, 1203–1209.
- Ching, T. Y., Dillon, H., Byrne, D. (1998). Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*, 103, 1128–1140.
- Danhauer, J. L., Abdala, C., Johnson, C., et al. (1986). Perceptual features from normal-hearing and hearing-impaired children's errors on the NST. *Ear Hear*, 7, 318–322.

- Dillon, H. (2001). *Hearing Aids*. New York, NY: Thieme.
- Dubno, J. R., & Dirks, D. D. (1982). Evaluation of hearing-impaired listeners using a nonsense-syllable test. I. Test reliability. *J Speech Hear Res*, 25, 135–141.
- Dubno, J. R., Dirks, D. D., Ellison, D. E. (1989). Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. I: The contribution of selected frequency regions. *J Acoust Soc Am*, 85, 347–354.
- Fallon, M., Trehub, S. E., Schneider, B. A. (2002). Children's use of semantic cues in degraded listening environments. *J Acoust Soc Am*, 111(5 Pt 1), 2242–2249.
- Fox, R. A., & Nissen, S. L. (2005). Sex-related acoustic changes in voiceless English fricatives. *J Speech Lang Hear Res*, 48, 753–765.
- Glista, D., Scollie, S., Bagatto, M., et al. (2009). Evaluation of nonlinear frequency compression: Clinical outcomes. *Int J Audiol*, 48, 632–644.
- Hall, J. W., III, Grose, J. H., Buss, E., et al. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear Hear*, 23, 159–165.
- Hedrick, M. (1997). Effect of acoustic cues on labeling fricatives and affricates. *J Speech Lang Hear Res*, 40, 925–938.
- Hedrick, M. S., & Younger, M. S. (2003). Labeling of /s/ and /f/ by listeners with normal and impaired hearing, revisited. *J Speech Lang Hear Res*, 46, 636–648.
- Heinz, J. M., & Stevens, K. N. (1961). On the properties of voiceless fricative consonants. *J Acoust Soc Am*, 33, 589–596.
- Hornsby, B. W., Johnson, E. E., Picou, E. (2011). Effects of degree and configuration of hearing loss on the contribution of high- and low-frequency speech information to bilateral speech understanding. *Ear Hear*, 32, 543–555.
- Horwitz, A. R., Dubno, J. R., Ahlstrom, J. B. (2002). Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. *J Acoust Soc Am*, 111(1 Pt 1), 409–416.
- Kent, R. D., & Read, C. (2002). *Acoustic analysis of speech* (2nd ed.) Canada: Thomson Learning.
- Killion, M. C. (2003). Earmold acoustics. *Semin Hear*, 24, 299–312.
- Kortekaas, R. W., & Stelmachowicz, P. G. (2000). Bandwidth effects on children's perception of the inflectional morpheme /s/: acoustical measurements, auditory detection, and clarity rating. *J Speech Lang Hear Res*, 43, 645–660.
- Kuk, F., & Keenan, D. (2011). Frequency transposition: Training is only half the story. *Hear Rev*, 17(12), 38–46.
- Kuk, F., Keenan, D., Auriemma, J., et al. (2009a). Re-evaluating the efficacy of frequency transposition. *ASHA Lead*, 14(1), 14–17.
- Kuk, F., Keenan, D., Korhonen, P., et al. (2009b). Efficacy of linear frequency transposition on consonant identification in quiet and in noise. *J Am Acad Audiol*, 20, 465–479.
- McDermott, H. J., Dorkos, V. P., Dean, M. R., et al. (1999). Improvements in speech perception with use of the AVR TranSonic frequency-transposing hearing aid. *J Speech Lang Hear Res*, 50, 1194–1202.
- Miller, G., & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am*, 27, 338–352.
- Moeller, M. P., Hoover, B., Putman, C., et al. (2007). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I—Phonetic development. *Ear Hear*, 28, 605–627.
- Moeller, M. P., McCleary, E., Putman, C., et al. (2010). Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. *Ear Hear*, 31, 625–635.
- Moore, B. C., Füllgrabe, C., Stone, M. A. (2010). Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. *J Acoust Soc Am*, 128, 360–371.
- Moore, B. C. J., Stone, M. A., Füllgrabe, C., et al. (2008). Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss. *Ear Hear*, 29, 907–922.
- Mueller, H. G., Alexander, J. M., Scollie, S. (2013). 20Q: Frequency lowering—The whole shebang. *AudiologyOnline*, Article #11913.
- Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *J Phonetics*, 20, 351–382.
- Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *J Acoust Soc Am*, 87, 2705–2715.
- Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *J Acoust Soc Am*, 101, 2253–2266.
- Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *J Speech Hear Res*, 30, 319–329.
- Parent, T. C., Chmiel, R., Jerger, J. (1998). Comparison of performance with frequency transposition hearing aids and conventional hearing aids. *J Am Acad Audiol*, 9, 67–77.
- Pavlovic, C. V. (1987). Derivation of primary parameters and procedures for use in speech intelligibility predictions. *J Acoust Soc Am*, 82, 413–422.
- Pittman, A. L. (2008). Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. *J Speech Lang Hear Res*, 51, 785–797.
- Pittman, A. L., & Stelmachowicz, P. G. (2000). Perception of voiceless fricatives by normal-hearing and hearing-impaired children and adults. *J Speech Lang Hear Res*, 43, 1389–1401.
- Repp, B. H., & Mann, V. A. (1980). Influence of vocalic context on perception of the /s/-/s/ distinction: Spectral factors. *Percept Psychophys*, 28, 213–228.
- Robinson, J. D., Stainsby, T. H., Baer, T., et al. (2009). Evaluation of a frequency transposition algorithm using wearable hearing aids. *Int J Audiol*, 48, 384–393.
- Rudmin, F. (1983). The why and how of hearing /s/. *Volta Review*, 85, 263–269.
- Scollie, S. (2013). 20Q: The Ins and outs of frequency lowering amplification. *AudiologyOnline*, Article #11863.
- Seewald, R. C., Cornelisse, L. E., Ramji, K. V., et al. (1997). *DSL V4.1 for Windows: A Software Implementation of the Desired Sensation Level (DSL[i/o]) Method for Fitting Linear Gain and Wide-dynamic-range Compression Hearing Instruments*. London, Ontario, Canada: University of Western Ontario.
- Simpson, A., Hersbach, A. A., McDermott, H. J. (2005). Improvements in speech perception with an experimental nonlinear frequency compression hearing device. *Int J Audiol*, 44, 281–292.
- Simpson, A., Hersbach, A. A., McDermott, H. J. (2006). Frequency-compression outcomes in listeners with steeply sloping audiograms. *Int J Audiol*, 45, 619–629.
- Stelmachowicz, P. G., Lewis, D. E., Choi, S., et al. (2007). Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear Hear*, 28, 483–494.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., et al. (2001). Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J Acoust Soc Am*, 110, 2183–2190.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., et al. (2002). Aided perception of /s/ and /z/ by hearing-impaired children. *Ear Hear*, 23, 316–324.
- Sullivan, J. A., Allsman, C. S., Nielsen, L. B., et al. (1992). Amplification for listeners with steeply sloping, high-frequency hearing loss. *Ear Hear*, 13, 35–45.
- Thornton, A. R., & Raffin, M. J. (1978). Speech-discrimination scores modeled as a binomial variable. *J Speech Hear Res*, 21, 507–518.
- Turner, C. W., & Robb, M. P. (1987). Audibility and recognition of stop consonants in normal and hearing-impaired subjects. *J Acoust Soc Am*, 81, 1566–1573.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392–393.
- Whalen, D. H. (1981). Effects of vocalic transitions and vowel quality on the English /s/-/ʃ/ boundary. *J Acoust Soc Am*, 69, 275–282.
- Wolfe, J., John, A., Schafer, E., et al. (2011). Long-term effects of nonlinear frequency compression for children with moderate hearing loss. *Int J Audiol*, 50, 396–404.
- Zeng, F. G., & Turner, C. W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *J Speech Hear Res*, 33, 440–449.