2014

The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss

Ryan W. McCreery
Boys Town National Research Hospital, ryan.mccreery@boystown.org

Joshua Alexander
Purdue University

Marc A. Brennan
Boys Town National Research Hospital, marc.brennan@unl.edu

Brenda Hoover
Boys Town National Research Hospital

Judy Kopun
Boys Town National Research Hospital

See next page for additional authors

Authors
Ryan W. McCreery, Joshua Alexander, Marc A. Brennan, Brenda Hoover, Judy Kopun, and Patricia G. Stelmachowicz

Published in final edited form as:
Ear Hear. 2014 ; 35(4): 440–447. doi:10.1097/AUD.0000000000000027.

The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss

Ryan W. McCreery, Ph.D.[1], **Joshua Alexander, Ph.D.**[2], **Marc A. Brennan, Ph.D.**[1], **Brenda Hoover, M.A.**[1], **Judy Kopun, M.A.**[1], and **Patricia G. Stelmachowicz, Ph.D.**[1]

[1]Hearing and Amplification Research Laboratory, Boys Town National Research Hospital, Omaha, NE

[2]Purdue University, West Lafayette, IN

Abstract

Objective—The primary goal of nonlinear frequency compression (NFC) and other frequency lowering strategies is to increase the audibility of high-frequency sounds that are not otherwise audible with conventional hearing-aid processing due to the degree of hearing loss, limited hearing aid bandwidth or a combination of both factors. The aim of the current study was to compare estimates of speech audibility processed by NFC to improvements in speech recognition for a group of children and adults with high-frequency hearing loss.

Design—Monosyllabic word recognition was measured in noise for twenty-four adults and twelve children with mild to severe sensorineural hearing loss. Stimuli were amplified based on each listener's audiogram with conventional processing (CP) with amplitude compression or with NFC and presented under headphones using a software-based hearing aid simulator. A modification of the speech intelligibility index (SII) was used to estimate audibility of information in frequency-lowered bands. The mean improvement in SII was compared to the mean improvement in speech recognition.

Results—All but two listeners experienced improvements in speech recognition with NFC compared to CP, consistent with the small increase in audibility that was estimated using the modification of the SII. Children and adults had similar improvements in speech recognition with NFC.

Conclusion—Word recognition with NFC was higher than CP for children and adults with mild to severe hearing loss. The average improvement in speech recognition with NFC (7%) was consistent with the modified SII, which indicated that listeners experienced an increase in audibility with NFC compared to CP. Further studies are necessary to determine if changes in audibility with NFC are related to speech recognition with NFC for listeners with greater degrees of hearing loss, with a greater variety of compression settings, and using auditory training.

Correspondence: Ryan McCreery, Boys Town National Research Hospital, Omaha, NE 68131, Fax: 402 452 5015, Phone: 402 498 6668, Ryan.McCreery@boystown.org.

INTRODUCTION

Listeners who use conventional hearing-aid amplification often have reduced access to high-frequency sounds due to the degree of hearing impairment, the limited bandwidth of the hearing-aid receiver (<5000–6000 Hz) or a combination of both factors. Relative to wide-bandwidth conditions (> 9000 Hz), limited bandwidth negatively impacts speech recognition (Stelmachowicz et al. 2001) and novel word learning (Pittman 2009) in children with hearing loss, as well as speech recognition (Ricketts et al. 2008) and sound quality ratings (Moore & Tan 2003) in adults. Frequency-lowering signal processing strategies have been implemented in hearing aids to improve the transmission of energy from the high-frequency spectrum (see Simpson 2009, McCreery et al. 2012, and Alexander, 2013 for reviews). By moving acoustic speech cues from high-frequency regions to lower frequencies where aided audibility is better, these strategies can potentially increase the bandwidth that is accessible to the listener.

This paper will focus on one specific form of frequency lowering, nonlinear frequency compression (NFC), in which the input bandwidth above a specified start frequency is compressed into a narrower bandwidth as determined by a specified compression ratio. After processing with NFC, a wider range of frequencies is made audible to the listener, albeit with reduced spectral detail. It is unknown how much the newly available information can be used by individual listeners with hearing loss due to reductions in the distinctiveness of acoustic speech cues caused by the bandwidth compression. As a first step, the current investigation focused on documenting how much improvement in speech intelligibility with NFC can be estimated using a model of audibility that attempts to quantify the amount of speech information that is audible after lowering. The extent to which listeners underperform compared to the model estimates will indicate how much intelligibility is affected by the signal distortion introduced by the frequency compression. The results may lead to the development of a more complex model of speech intelligibility for frequency-lowering amplification in future studies.

The audibility of speech is often quantified using the Speech Intelligibility Index (SII; ANSI S3.5-1997, R2007), which quantifies the proportion of the speech signal that is audible based on the sensation level of speech in a number of discrete frequency bands. Because frequency-lowering changes the distribution of frequency bands in the hearing aid output, one approach to calculating audibility with NFC would be to estimate the sensation level for each frequency band at the lower frequency where it occurs in the output, while retaining individual band importance weights. Such an approach quantifies the amount of speech information that is audible after lowering. This model does not account for how the usability of speech information in audible bands of the input might be affected by frequency lowering. Despite this potential limitation, McCreery et al. (2012) found that nonword recognition for adults with normal hearing for stimuli processed with NFC followed a predictable pattern based on the audible bandwidth. When more of the source bandwidth was made audible by manipulating NFC start frequency and compression ratio, recognition scores were higher. Whereas a predictable pattern based on audibility was observed for listeners with normal hearing, results from listeners with hearing loss are needed to determine if a similar pattern would be observed in the population for which the processing is intended.

Children and adults have had different patterns of speech recognition outcomes with NFC in previous studies. Glista and colleagues (2009) conducted the only previous study that included children and adults. Their results indicated that children were more likely than older adults (50 – 81 years-old) to experience improvements with NFC compared to conventional processing (CP). However, it is difficult to determine whether these findings are due to developmental differences or to the fact that children had 5- to 10-dB more gain than adults per the Desired Sensation Level prescriptive algorithm. Characterizing changes in audibility could help to determine if age-related difference in speech recognition with NFC are related to prescribing more audibility for children or developmental factors that affect speech recognition. Children generally need a wider audible bandwidth to achieve maximum levels of speech understanding (Stelmachowicz et al. 2001; McCreery & Stelmachowicz, 2011), but the extent of this difference depends on the type of stimuli and the task. Assuming that improvements in audible bandwidth are achieved with NFC, data describing the developmental effects of bandwidth on speech recognition may be informative for predicting outcomes.

The primary purpose of the present study was to determine if a modification of the SII that estimated the difference in audibility between CP and NFC would reflect changes in speech recognition in adults and children with mild to severe sensorineural hearing loss. The primary hypothesis is that improvements in audibility with NFC will lead to improvements in speech recognition. Because none of the listeners had experience listening with NFC previously and previous research has suggested that speech recognition with NFC may depend on experience (Wolfe et al. 2010; 2011; Glista et al. 2012), listeners were exposed to NFC by listening to an audio-visual recording of stories processed with NFC. Speech recognition was evaluated prior to and after the story. Based on previous research on the influence of increasing bandwidth without frequency lowering on speech recognition, larger improvements were expected for children than for adults.

METHOD

Participants

Twenty-four adults ages 19–65 years (mean =53.0, SD = 14.0) and 12 children ages 8–16 years old (mean = 12.0, SD=2.4) with mild to severe hearing loss participated in this study. Figure 1 shows the mean (Left - X and Right - O) and range (shaded area) of thresholds for listeners in the current study. At the time of enrollment in the study, ten of the adult listeners had used binaural hearing aids for an average of 14 years (range: 0–29 years) with an average reported daily use time of 11.2 hours (range 6–16). All of the children had used binaural hearing aids for an average of 8 years (range: 4–15 years) with an average reported daily use time of 9.6 hours (range: 5–15 hours) by parental report. None of the listeners in either age group reported prior experience with any frequency lowering technology. While using their personal amplification, the children’s articulation was screened using the Bankson Bernthal Quick Screen of Phonology (BBQSP) and their vocabulary was assessed using the Expressive Vocabulary Test, Form A (EVT-A). The articulation screening was used to determine the presence of significant articulation problems that could complicate scoring, while the expressive language test was used to ensure that all of the children had expressive vocabulary within the normal range for age. All children had BBQSP and EVT-A results within the normal range for age. Listeners were paid $15/hour for their participation. Children were also given a book of their choice.

Stimuli

Three hundred monosyllabic words containing one of 9 fricatives or affricates (/s/, /z/, /f/, /v/, /ʧ/, /ʤ/, /ʃ/, /ʒ/, /θ/) in the initial or final position and six vowel contexts (/a/,/ i/, /I/, /ε/, /u/, /ʌ/) were used to assess speech recognition with conventional processing and NFC. Monosyllabic words were selected from word recognition tests (Phonetically Balanced Kindergarten, Haskins, 1949; Word Intelligibility by Picture Identification, Ross & Lerman, 1971; California Consonant Test, Owens & Schubert, 1977; Computer Assisted Speech Perception Assessment, Boothroyd, 2006) that were intended for children in the lower end of the age range of listeners participating in this study. All of the words were determined to be within the average child lexicon (Storkel & Hoover, 2010). Stimuli were spoken by a 22 year-old female talker. Recorded monosyllabic words were balanced for fricative/affricate content across four 75-word lists, such that each list contained approximately the same number of each fricative and affricate target in initial and final positions. Speech-shaped competing noise was generated by computing a Fast Fourier Transform (FFT) of the female talker, randomizing the phase of the signal at each frequency bin, and then taking the inverse FFT. This process preserved the long-term average spectrum, but eliminated temporal and spectral dips. The words were mixed with the speech-shaped noise at a signal to noise ratio (SNR) of +6 dB. This SNR was chosen because speech recognition data for monosyllabic words from McCreery et al. (2010) suggested that +6 dB SNR would not result in ceiling performance.

Two stories were written to correspond to the pictures in the two children's books "A Boy, a Dog and a Frog" by Mercer Mayer and "One Frog Too Many" by Mercer and Marianna Mayer. Vocabulary was selected that would maximize fricative content. Audio-visual recordings were made of two preschool teachers of children with hearing loss reading the stories. The head and shoulders of the teachers were filmed as they read the stories using a JVC GY-HM710 HD Camera with a Fujinon Zoom Lens TH17X5BRMU. The video resolution of the recording was 1920 × 1080 with a frame rate of 21 frames per second. Video editing was completed using Final Cut Pro X software (Apple, Inc. Cupertino, CA). Audio recordings of the stories were made using a Lectrosonics M175-LS wireless lavaliere microphone connected to a Lectrosonics CR175 receiver at a sampling rate of 48 kHz and 16 bit resolution. The audio recordings were excised and then divided into 55–65 second segments that corresponded to the pages in the books. Photographs taken of the book pages were displayed for 5 seconds before and after each reading. The length of each story was 10 and 12 minutes for "A Boy, a Dog and a Frog" and "One Frog Too Many," respectively.

Instrumentation

Stimulus presentation and audio response recording was performed using custom software on a personal computer with a Lynx Studio Two-B sound card. Sennheiser HD-25-1 II headphones were used for stimulus presentation. A Shure 53 BETA head-worn boom microphone with the standard filter cap was used to record listener responses for later scoring. Videos of the stories were presented via a computer monitor using Max Runtime 5 software (Cycling ’74 Software, San Francisco, CA) to synchronize the processed audio and video recordings.

Stimulus processing and hearing aid simulation

A hearing aid simulator was used to process stimuli in order to provide greater experimental control over hearing-aid signal processing characteristics across listeners and conditions than would be possible using real hearing aids. An 8-channel hearing aid simulator program processed audio files for the monosyllabic words and stories. Two word lists were processed using CP only, which included wide-dynamic range compression (WDRC) based on real-ear aided response targets for Desired Sensation Level v5.0a (DSL; Scollie et al, 2005). Two word lists were processed using NFC followed by WDRC. To determine the WDRC settings, listener’s audiometric thresholds were converted to dB SPL using a transfer function based on the TDH-50 headphones coupled with a KEMAR, which were then entered into the DSL program. The DSL child algorithm was selected for the children, and the DSL-adult algorithm without the binaural correction was selected for the adults. The DSL program generated compression thresholds and ratios for each channel as well as real-ear aided response target levels for 60 dB SPL speech and output limiting levels for each 1/3 octave frequency. Since DSL does not generate a target for 8000 Hz, target sensation levels (SL) for 6000 Hz were used for 8000 Hz.

Presentation level for each listener was estimated with Sennheiser HD-25-1 II headphones and an IEC 711 coupler attached to KEMAR using two iterations of the following steps. First, the sound level of the long-term average speech spectrum based on a calibrated speech signal from a male talker (the “carrot” passage) from the Verifit hearing-aid analyzer was reproduced at 60 dB SPL to the input of the hearing aid simulator. Second, the output of the hearing aid simulator was filtered into 1/3-octave bands (ANSI 2004), and the mean level of each band in dB SPL was computed and compared to the prescribed DSL targets for a 60 dB SPL speech input. Third, 1/3-octave band SPL levels were averaged according to which WDRC channel they fell within and then the average difference between the output and target level was calculated for each channel and used to adjust the gain in that channel. The overall average difference between the output of the HA simulator and DSL targets for CP across all octave frequencies and listeners was 0.686 dB (SD = 1.56). For individual listeners, the average difference between HA simulator output and DSL targets across octave frequencies from 250 Hz – 4000 Hz ranged from −0.99 dB to +1.83 dB. The largest individual negative difference from prescriptive targets at an octave band frequency was −4.83 dB for one listener at 2000 Hz, whereas the largest positive difference was +7.5 dB for one listener at 250 Hz.

Amplification using amplitude compression (WDRC) and frequency-shaping of the stimuli was accomplished for each listener using the MATLAB program reported in McCreery et al. (2013). Specifically, attack and release times were set at 5 and 50 ms, respectively and referenced to the ANSI (2009) standard. Equation 8.1 from Kates (2008) was used to set the gain control circuit.

if $|x(n)| \geq d(n-1)$

$d(n) = \alpha d(n-1) + (1-\alpha)|x(n)|$

else

$d(n) = \beta d(n-1)$

end

where $x(n)$ is the acoustic input signal, $d(n)$ is the local mean level used to generate the gain signal which was applied to $x(n)$ to form the output signal, α is a constant derived from the attack time, and β is a constant derived from the release time. When the signal is increasing in level, the first part of the equation reduces gain, otherwise gain is increased as the signal decreases in level. Minimum gain was limited to 0 dB and maximum gain to 65 dB.

Each pre-recorded digital sound file was filtered into 8 channels. The WDRC output above the kneepoint recommended by DSL and below the output-limiting compression (OLC) kneepoint was determined by the input signal level and the compression ratio. Above the OLC kneepoint, the signal was compressed at a 10:1 ratio with 1 ms attack and 50 ms release times. The maximum output was limited to 105 dB SPL and the signals were scaled for output to the Sennheiser HD-25-1 II headphones. The maximum audible frequency for CP conditions was limited at 5000 Hz for all listeners to simulate a plausible hearing aid bandwidth based on the Phonak Naida SP hearing aid that was the basis for the simulation of NFC. Gain for the last channel, which had a low-frequency crossover at 5000 Hz, was set to 0 dB and a 1024-tap low-pass filter was applied at 5000 Hz that reduced the output by 80 – 100 dB at 5500 Hz.

For the NFC conditions, signals were frequency compressed prior to WDRC using an algorithm based on Simpson et al (2005). Overlapping 128-point Fast Fourier Transforms (FFTs) were computed every 32 samples (1.45 ms) and used to estimate instantaneous frequency. Instantaneous frequencies encompassing an input band of approximately 4500 Hz above the start frequency were synthesized at lower frequencies using phase vocoding; with frequency-reassignment determined using Eq. (1) from Simpson et al. (2005). The 4500 Hz input band was selected to correspond to processing in the Phonak CORE signal processing platform. The nominal start frequency was rounded to the nearest FFT bin, which was a multiple of 172.26 Hz (22050 sampling rate/128 FFT rate). Following NFC, the signals were amplified using the same amplitude compression parameters (gain, compression thresholds and ratios, etc.) as for the CP condition.

NFC settings

Twelve combinations of start frequency and compression ratio were considered for each listener. The combinations were selected from actual values found in the Phonak iPFG v2.0 fitting software assuming a Phonak Naida SP hearing aid and audiograms representing the intended range of participants. Although the hearing-aid simulator allows for any potential combination of start frequency and compression ratio, the available combinations for the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

group of listeners in the current study were selected to provide a more realistic approximation of the combinations of start frequency and compression ratio that are clinically available.

The NFC setting (combination of start frequency and compression ratio) was chosen for each listener using a recommended clinical procedure (Alexander, 2013). The first step in the recommended procedure is to document the highest frequency that can be made audible with CP alone, known as the "maximum audible output frequency" since it determines the output bandwidth available to the listener prior to remapping frequency with NFC. To determine the maximum audible output frequency for each listener, audiometric thresholds in dB SPL were plotted against 1/3-octave band spectral analysis of speech ("carrot" passage) amplified with CP for a 65-dB input level. The maximum audible output frequency was then estimated by graphically by locating the highest frequency where the long term average speech spectrum (root-mean-square) intersected with thresholds. For each NFC setting, the input frequency corresponding to the maximum audible output frequency was computed using the equation from Simpson et al. (2005). The NFC setting that resulted in the widest audible bandwidth was selected. The one exception was when multiple settings produced reasonably close maximum audible output frequencies (operationally defined as within 7%), then the setting with highest start frequency was selected in this group with the intent of preserving formant spacing for vowels.

NFC settings were selected for both ears of each listener. In cases of audiometric asymmetry, the NFC parameters were selected independently for each ear to maximize the audible input bandwidth. All of the children had audibility up to 5000 Hz in both ears for CP alone and so were assigned a start frequency of 3800 Hz and a compression ratio of 2.6:1 based on the procedures described above. Figure 2 plots the frequency input/output function for CP and NFC for the children. The homogeneity of audiograms selected for the study and controlled acoustic settings related to using headphones were likely the source of limited variability in NFC parameters across listeners. For adults, the average frequency that was audible with CP alone was 4630 Hz (range: 2500 – 5000 Hz), and the mean start frequency was 3440 Hz (range: 1500 – 3800) with a mean compression ratio of 2.5:1 (range 2.2:1 – 3.2:1). With this range of start frequencies, input frequencies up to 8300 Hz were audible for all listeners following NFC.

Audibility calculation

The audibility of the long-term average speech spectrum (LTASS) for the female talker for this study was calculated for each listener and each ear using the SII (ANSI S3.5-1997, R2007). The 1/3-octave-band method was used with a weighting function that assumed equal importance across bands (1/18 = 0.0556), since specific importance weights were not available for the stimuli used in the experiment. The calculation assumed a non-reverberant environment. The band levels of the concatenated speech and noise stimuli were measured using a Larson Davis System 824 sound level meter with a Larson Davis AEC 101 IEC 318 headphone coupler. The levels of speech and noise were converted to free-field using the free-field to eardrum transfer function from the SII. Audiometric thresholds were converted from dB HL to dB SPL and then interpolated and extrapolated to correspond with 1/3-

octave-band frequencies. A frequency-specific bandwidth adjustment was used to convert pure tone thresholds to equivalent 1/3-octave-band levels (Pavlovic, 1987). The spectrum levels of speech and noise for each listener's audiogram and condition (CP or NFC) were entered into a program to calculate sensation level (SL) for each 1/3-octave band. For conditions with CP, The SL for each band was multiplied by the importance weight for that band, and the sum of these products for all bands generated the SII for each condition. For conditions with NFC, the SII calculation was the same as for conditions with CP, except that the SL for each frequency band above the start frequency was calculated at the frequency where that band occurred in the output after NFC. Estimates of audibility for frequency-compressed signals are noted as SII-NFC, whereas estimates of audibility from CP alone are noted as SII-CP.

Procedures

All testing was completed in a sound-treated audiometric suite. Audiometric thresholds were tested using TDH-49 earphones. Listeners were seated at a table in front of a computer monitor. The listeners were instructed that they would hear real words and should repeat each word. They were instructed to repeat the words exactly as they heard them. Listeners viewed a slide show of landscapes and pictures on the computer to maintain their attention. Listeners first completed a practice list of ten monosyllabic words with CP that were not part of the lists used for experimental conditions. The words in the first half of the practice list were in quiet and the second half were mixed with noise at +6 dB SNR. After the practice list, stimuli were presented with one CP list of 75 words and one NFC list of 75 words, referred to as the pre-exposure blocks. Following the pre-exposure blocks, all listeners were exposed to the audio-visual presentation of the two stories. Listeners were instructed to carefully attend because they would be asked 4 simple questions after each story. Responses to the questions were recorded. All listeners correctly answered at least 7/8 questions, except one child and one adult that each missed two questions. After the exposure, two more lists of 75 words, one with CP and one with NFC, were presented, referred to as the post-exposure blocks. The processing types (CP or NFC) for pre- and post-exposure blocks were counterbalanced across listeners using a modified Latin square design. The presentation order of the stimuli within each word list was randomized. Whole words and phonemes were scored as correct/incorrect after the experimental session by a single examiner using audio-visual recordings of listeners' responses. Ten percent of the trials were scored by a second examiner who was blinded to the original scoring with a reliability of Cohen's kappa = 0.9789 between the two examiners.

RESULTS

Figure 3 displays word recognition scores for each age group and condition. To address questions related to the effect of processing (CP vs. NFC), exposure (pre- and post-exposure interval), age (adult vs. child) and hearing aid use (yes vs. no) on speech recognition, a four-way mixed analysis of variance was completed. There was a significant main effect of processing [$F(1,34)=27.6$, $p <0.001$, $\eta_p^2 = 0.455$] with NFC (M = 73.60% correct, SD = 11.82) greater than CP (M = 66.49% correct, SD =10.83). The main effect of exposure interval was also significant [$F(1,34)=13.0$, $p =0.001$, $\eta_p^2 = 0.282$] with higher post-exposure word recognition scores (M = 72.35%, SD=10.15) than pre-exposure scores (M = 68.66%, SD = 10.81). None of the other main effects or higher-order interactions reached significance. Specifically, the lack of an age group by processing interaction [F(1,34)=1.3, p =.268, $\eta_p^2 = 0.037$], suggested that the main effect for processing (NFC > CP) was consistent for both children and adults. The lack of a processing by exposure interval interaction [F(1,34)=1.0, p=.328, $\eta_p^2 = 0.029$], suggests that NFC exposure did not preferentially benefit post-exposure word recognition in the NFC condition vs. the CP condition. The lack of a main effect of age indicates that children and adults did not differ in word recognition overall [F(1,34)=0.2, p=.630, $\eta_p^2 = 0.007$] (M = 69.2% children; 70.8% adults). Finally, the lack of a main effect of hearing aid use [F(1,34)=2.9, p=.101, $\eta_p^2 = 0.081$] indicates that hearing aid users and non-users had similar speech recognition across conditions.

Figure 4 displays individual differences between NFC and CP as a function of the difference in audibility between conditions. The improvement in audibility between the SII-NFC and SII-CP was calculated by subtracting the SII-CP from the SII-NFC to represent the increase in potentially usable information with the specific NFC settings and audiogram for each listener. The mean difference in audibility between SII-NFC and SII-CP was 0.1262 (SD =0.031; range =0.09 – 0.22). Examination of the individual data suggested that most of the participants had either improved speech recognition or no change in speech recognition with NFC compared to CP.

DISCUSSION

The purpose of this study was to evaluate the influence of audibility on recognition of speech processed by NFC in children and adults with mild to moderately-severe hearing loss. In support of our hypothesis that increased audibility would support increased speech recognition with NFC, word recognition with NFC was 7% higher than CP. In further support of our hypotheses, this improvement in word recognition was consistent with an index of audibility modified to account for the relocation of frequency bands following NFC. Improvements in audibility were confined to frequencies above 3000 Hz in the current study, which used NFC settings that were selected to maximize the bandwidth of the input signal for listeners with mild to moderate high-frequency hearing loss. Contrary to our hypotheses, children did not differ from adults in overall word recognition and did not preferentially benefit from NFC.

The improvements in word recognition with NFC compared to CP are consistent with some previous studies of children with hearing loss (Wolfe et al. 2010; 2011) and studies of both children and adults with hearing loss (Glista et al. 2009). Adults with hearing loss have shown less consistent improvements in speech recognition in investigations where the start frequency was set at the likely maximum audible frequency so that audibility of the lowered signal was limited (Simpson et al. 2005; 2006) or where frequency lowering parameters were determined by each listener’s sound quality preference (Bohnert et al. 2010). Increased audibility with NFC has been associated with improvements compared with CP in previous research. In studies by Glista et al. (2009) and Wolfe et al. (2010; 2011), audibility of high-frequency sounds was verified either behaviorally, electroacoustically or through a combination of both approaches. The current study verified audibility of high-frequency sounds by estimating the audible bandwidth before and after the application of NLFC. While a number of factors could have contributed to the variability in findings between studies and within samples of listeners from the same studies, if NFC does not improve high-frequency audibility or results in equivocal audibility, improvements in speech recognition would not be anticipated.

In an effort to quantify changes in audibility with NFC in the current study, the SII was adapted to account for the location of each 1/3-octave band in the output of the hearing aid simulator using an importance function with equal weighting across frequency bands. The improvement in audibility for listeners in this study was relatively small, since NFC increased audibility for only the 6300 Hz and 8000 Hz bands for most listeners. However, the amount of improvement in speech recognition with NFC compared to CP is similar to previous studies that have compared speech recognition of children and adults with normal hearing in conditions with and without the 8000 Hz octave band (McCreery & Stelmachowicz, 2011). This suggests that when NFC is limited primarily to frequencies above 3800 Hz, the magnitude of improvement in speech recognition with NFC is similar to what has been observed in previous studies when high-frequency audibility is restored by extending the bandwidth. For example, Hornsby et al. (2011) reported speech recognition of adults with hearing loss under conditions with limited high-frequency audibility to wideband conditions. Hornsby et al. reported either no difference or an improvement in speech recognition for most of their adult listeners with hearing loss when the bandwidth of the signal was extended above 4000 Hz, similar to the magnitude of the average effect observed in the current study with NFC. Several subjects did not experience a significant change in speech recognition with NFC. Two adults in the current investigation showed a significant decrement in performance with NFC, despite an estimated improvement in audibility from the SII-NFC. The two adult listeners who exhibited degradation in word recognition with NFC did not differ from the group in terms of their NFC parameters, age, slope of hearing loss or other factors that might have predicted reduced performance. Some listeners may need greater listening experience with NFC to have improved speech recognition (Glista et al. 2012) than was provided in this cross-sectional study. Clearly, the relationship between audibility and speech recognition for frequency-lowered sounds is complex, and further research is needed to resolve these issues.

Despite the positive results when the SII was adapted to account for frequency lowering, the applicability of these estimates is currently limited and requires further validation. Audibility is a necessary, but not sufficient condition for improving speech recognition in listeners with hearing loss. The audibility-based index used in this study only reflects information that can be transmitted after frequency lowering and does not account for distortion or loss of spectral distinctiveness that may occur with compression. Many factors likely determine how much information individual listeners can extract from the lowered bands, including the frequency regions altered by lowering, the severity of loss in the regions to which information is moved and the frequency resolution of the individual listener with hearing loss. For example, NFC with a high start frequency that alters mostly the broadband frication spectrum might improve fricative recognition, but the same amount of frequency compression at a lower start frequency could alter vowel formant energy and have a negative effect on vowel recognition, although reduced vowel perception has not been consistently documented in the literature. The SII-NFC as applied here assumes that the contribution of each frequency band to speech recognition is the same regardless of whether or not the spectrum has been altered by frequency lowering. Such an assumption may be reasonable for conditions where NFC is limited to the high frequencies (>3800 Hz) as in this study, but may not hold as the start frequency reaches frequency regions where the frequency of vowel formant energy is important for perception. This model of speech intelligibility, which assumes that speech cues made audible with NFC contribute to speech recognition for the listener with hearing loss when spectrally distorted, may not yield accurate predictions of speech recognition across a range of hearing losses or NFC settings.

Additionally, the monosyllabic words used for the speech recognition task in this study were selected to contain phonemes with high-frequency energy, including fricatives and affricates. Phonemes with high-frequency energy were targeted because those sounds may be beyond the bandwidth of CP and potentially would be more audible with NFC. However, the improvements observed with NFC for a stimulus set that has a higher occurrence of fricatives and affricates than usually occur in language may overestimate the magnitude of improvements and generalizability of the model of audibility to more realistic contexts. In a study describing how speech perception with NFC changes over time, Glista and colleagues (2012) reported differences across tasks (detection vs. recognition) and stimuli (phonemes vs. nonwords). Likewise, Wolfe and colleagues (2010) found significant improvements for phonemes and plural words, but not for sentences in noise. Clearly, the type of stimulus can influence speech recognition with NFC. Further investigation into how task and stimulus factors affect speech recognition with NFC will help to guide clinical decisions about how to assess efficacy of frequency lowering.

The current investigation used a software-based hearing-aid simulator to allow for more precise control over signal processing than would be possible using real hearing aids. Although this decision provided greater experimental control over certain hearing-aid parameters, differences between simulations of hearing-aid signal processing under headphones and real hearing aids may impact the generalizability of these findings to clinical populations. The use of headphones also prevented inclusion of listeners with greater degrees of hearing loss because of the maximum output limitations of the headphones. The simulator also capped the maximum audible frequency at 5 kHz with CP for each listener, which may be more or less bandwidth than would be audible in real hearing aids for listeners in this study depending on the degree of high-frequency hearing loss. The decision to impose the same bandwidth for all participants for CP may have resulted in more consistent differences in audibility between CP and NFC than would be observed with wearable hearing aids. Further research should seek to replicate the use of maximum audible frequency for selecting NLFC settings and SII-NFC using real hearing aids and listeners with greater degrees of hearing loss.

Rather than drawing from an unlimited number of combinations of the start frequency and compression ratio, the combinations of parameters used in the study were selected based on the settings that were available in the manufacturer's programming software. It may have been possible to provide greater audibility with NFC by using different combinations of start frequencies and compression ratios than what were available with the manufacturer's programming software. However, such results would not be representative of NFC as implemented in real hearing aids. Additionally, individually manipulating start frequency and compression ratio yields multiple combinations of parameters that result in the same audible bandwidth, and future research should directly evaluate the efficacy NFC settings by independently manipulating start frequency and compression ratio. The improvements in audibility provided by NFC may be less in this study than if start frequency and compression ratio had been individually selected to provide the broadest bandwidth, but this study design was more representative of what could be achieved with NFC as currently implemented in hearing aids.

The adults and children with hearing loss in this investigation had mild to severe high-frequency hearing loss that necessitated a relatively limited number of combinations of start frequency and compression ratio compared to a clinical population. For that reason, frequency lowering was constrained in most cases to frequencies above 3500 Hz. All of the children and all but five of the adults had the same combination of start frequency (3800 Hz) and compression ratio (2.6:1) due to similarities in the maximum audible frequency (5000 Hz) with CP through the hearing aid simulator. Therefore, the extent to which this method of calculating audibility can be extended to listeners with greater degrees of hearing loss cannot be determined from these results. Listeners with greater degrees of hearing loss may require lower start frequencies and higher compression ratios to maximize audibility, which, as noted, would lead to greater distortion of the speech spectrum. Increased spectral distortion could make some high-frequency speech cues difficult to distinguish. For example, results from Glista et al. (2009) suggest that some listeners may experience confusion of /s/ and /ʃ/ sounds with NFC, which could be anticipated with lower start frequencies that distort the relationship and location of the spectral peaks for each of these sounds.

Conclusion

Adults and children with mild to severe high-frequency hearing loss demonstrated improvements in speech recognition with NFC compared to CP. Improvements in audibility were quantified using an adaptation of the SII based on the frequency compression settings for each listener. The average improvement in audibility with NFC was equivalent to the mean improvement in speech recognition. NFC improved audibility of the 6000 and 8000 Hz bands for most of the listeners. The amount of improvement in speech recognition with NFC was similar to that observed in previous studies that extended the bandwidth for frequency bands encompassing 6000 and 8000 Hz. Examination of individual data found that several listeners had similar speech recognition with CP and NFC. Only two adult listeners experienced degradation in speech recognition with NFC with improved audibility. Future studies should seek to extend these approaches to listeners with greater degrees of hearing loss to develop more complex models that could be used to support candidacy decisions for frequency lowering.

Acknowledgments

The authors wish to express their thanks to Prasanna Aryal for computer programming assistance. This research was supported by grants to Dr. Stelmachowicz (R01 DC04300), Dr. McCreery (R03 DC012635), Dr. Brennan (F32

DC012709), Dr. Alexander (RC1 DC010601) and the Post-doctoral training (T32 DC00013) and Human Research Subjects Core (P30-DC004662) grants to Boys Town National Research Hospital.

References

Alexander JM. Individual variability in recognition of frequency-lowered speech. Seminars in Hearing. 2013; 34(2):86–109.

ANSI. ANSI S3.5–1997, American National Standard Methods for Calculation of the Speech Intelligibility Index. New York: American National Standards Institute; 1997.

Bohnert A, Nyffeler M, Keilmann A. Advantages of a non-linear frequency compression algorithm in noise. European Archives of Oto-rhino-laryngology. 2010; 267(7):1045–1053. [PubMed: 20127102]

Boothroyd, A. Computer-aided speech perception assessment (CASPA) 5.0 software manual. San Diego, CA: 2006.

Glista D, Scollie S, Bagatto M, Seewald R, Parsa V, Johnson A. Evaluation of nonlinear frequency compression: clinical outcomes. International journal of audiology. 2009; 48(9):632–644. [PubMed: 19504379]

Glista D, Scollie S, Sulkers J. Perceptual acclimatization post nonlinear frequency compression hearing aid fitting in older children. Journal of Speech, Language and Hearing Research. 2012; 55(6):1765.

Haskins, J. Kindergarten phonetically balanced word lists (PBK). St. Louis: Auditec; 1949.

Hornsby B, Johnson E, Picou E. Effects of degree and configuration of hearing loss on the contribution of high-and low-frequency speech information to bilateral speech understanding. Ear and hearing. 2011; 32(5):543–555. [PubMed: 21336138]

Kates, James M. Digital Hearing Aids. San Diego: Plural Pub.; 2008.

McCreery RW, Brennan MA, Hoover B, Kopun J, Stelmachowicz PG. Maximizing audibility and speech recognition with nonlinear frequency compression by estimating audible bandwidth. Ear and Hearing. 2013; 34(2):e24–e77. [PubMed: 23104144]

McCreery R, Ito R, Spratford M, Lewis D, Hoover B, Stelmachowicz PG. Performance-intensity functions for normal-hearing adults and children using computer-aided speech perception assessment. Ear and hearing. 2010; 31(1):95–101. [PubMed: 19773658]

McCreery R, Stelmachowicz P. Audibility-based predictions of speech recognition for children and adults with normal hearing. Journal of the Acoustical Society of Am. 2011; 130(6):4070–4081.

McCreery RW, Venediktov RA, Coleman JJ, Leech HM. An Evidence-Based Systematic Review of Frequency Lowering in Hearing Aids for School-Age Children With Hearing Loss. American Journal of Audiology. 2012; 21(2):313. [PubMed: 22858615]

Moore BCJ, Tan C-T. Perceived naturalness of spectrally distorted speech and music. The Journal of the Acoustical Society of America. 2003; 114(1):408. [PubMed: 12880052]

Owens E, Schubert ED. Development of the California consonant test. Journal of Speech, Language and Hearing Research. 1977; 20(3):463.

Pavlovic CV. Derivation of primary parameters and procedures for use in speech intelligibility predictions. J Acoust Soc Am. 1987; 82:413–422. [PubMed: 3624646]

Pittman A. Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. Journal of Speech, Language and Hearing Research. 2008 Jun.51:785–797.

Ricketts TA, Dittberner BA, Johnson EE. High frequency bandwidth and sound quality in listeners with normal through moderate hearing loss. J Speech Lang Hear Res. 2008; 51(1):160–172. [PubMed: 18230863]

Ross M, Lerman J. Word Intelligibility by Picture Identification. 1971

Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Laurnagaray D, Beaulac S, et al. The Desired Sensation Level multistage input/output algorithm. Trends in Amplification. 2005; 9(4): 159–197. [PubMed: 16424945]

Simpson A. Frequency-lowering devices for managing high-frequency hearing loss: A review. Trends in amplification. 2009; 13(2):87–106. [PubMed: 19447764]

Simpson A, Hersbach Aa, McDermott HJ. Frequency-compression outcomes in listeners with steeply sloping audiograms. International journal of audiology. 2006; 45(11):619–629. [PubMed: 17118905]

Simpson A, Hersbach Aa, McDermott HJ. Improvements in speech perception with an experimental nonlinear frequency compression hearing device. International Journal of Audiology. 2005; 44(5): 281–292. [PubMed: 16028791]

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. J. Acoust. Soc. Am. 2001; 110:2183–2190. [PubMed: 11681394]

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. Aided perception of /s/ and /z/ by hearing-impaired children. Ear Hear. 2002; 23:316–324. [PubMed: 12195174]

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE, Moeller MP. The importance of high-frequency audibility in the speech and language development of children with hearing loss. Archives of Otolaryngology—Head & Neck Surgery. 2004; 130(5):556. [PubMed: 15148176]

Storkel HL, Hoover JR. An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. Behavior Research Methods. 2010a; 42:497–506. [PubMed: 20479181]

Wolfe J, John A, Schafer E, Nyffeler M, Boretzki M, Caraway T. Evaluation of nonlinear frequency compression for school-age children with moderate to moderately severe hearing loss. Journal of the American Academy of Audiology. 2010; 21(10):618–628. [PubMed: 21376003]

Wolfe J, John A, Schafer E, Nyffeler M, Boretzki M, Caraway T, Hudson M. Long- term effects of nonlinear frequency compression for children with moderate hearing loss. International journal of audiology. 2011; 50(6):396–404. [PubMed: 21599615]

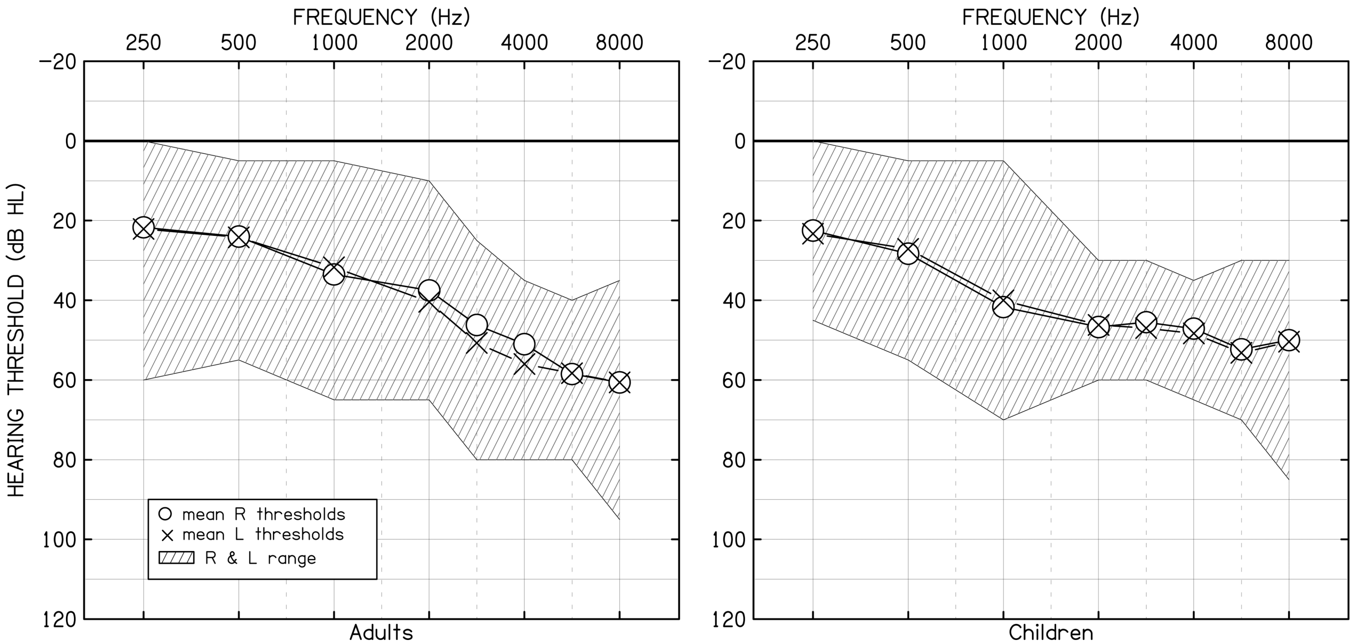

Figure 1.

The mean right ear (circles) and left ear (X) behavioral audiometric thresholds (dB HL) at each test frequency (Hz) for adults (left panel) and children (right panel). The range of thresholds for each age group is plotted as the hatched area.

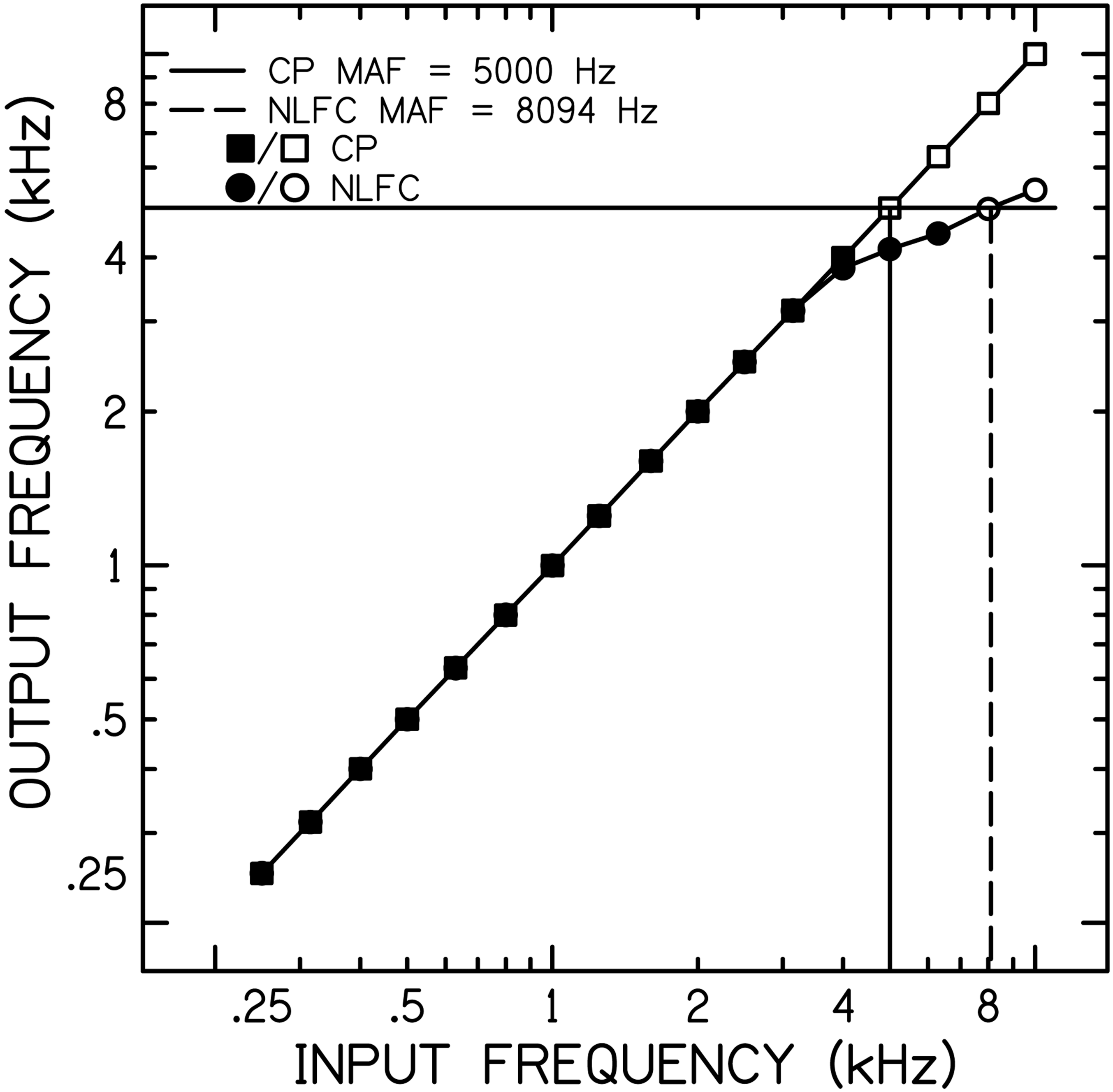

Figure 2.

Example of output frequency (kHz) as a function of input frequency (kHz) for a start frequency of 3800 Hz and a compression ratio (CR) of 2.6:1, which was the most common combination of parameters for listeners in the study. Filled symbols represent frequencies where the signal is audible. Open symbols represent frequencies where the signal is inaudible. Squares represent conventional processing (CP), and circles represent nonlinear frequency compression (NFC). The horizontal solid line represents the maximum audible output frequency. The solid vertical line represents the maximum audible frequency (MAF;

5000 Hz) with CP. The dashed vertical line represents the MAF (8094 Hz) with nonlinear frequency compression (NFC).

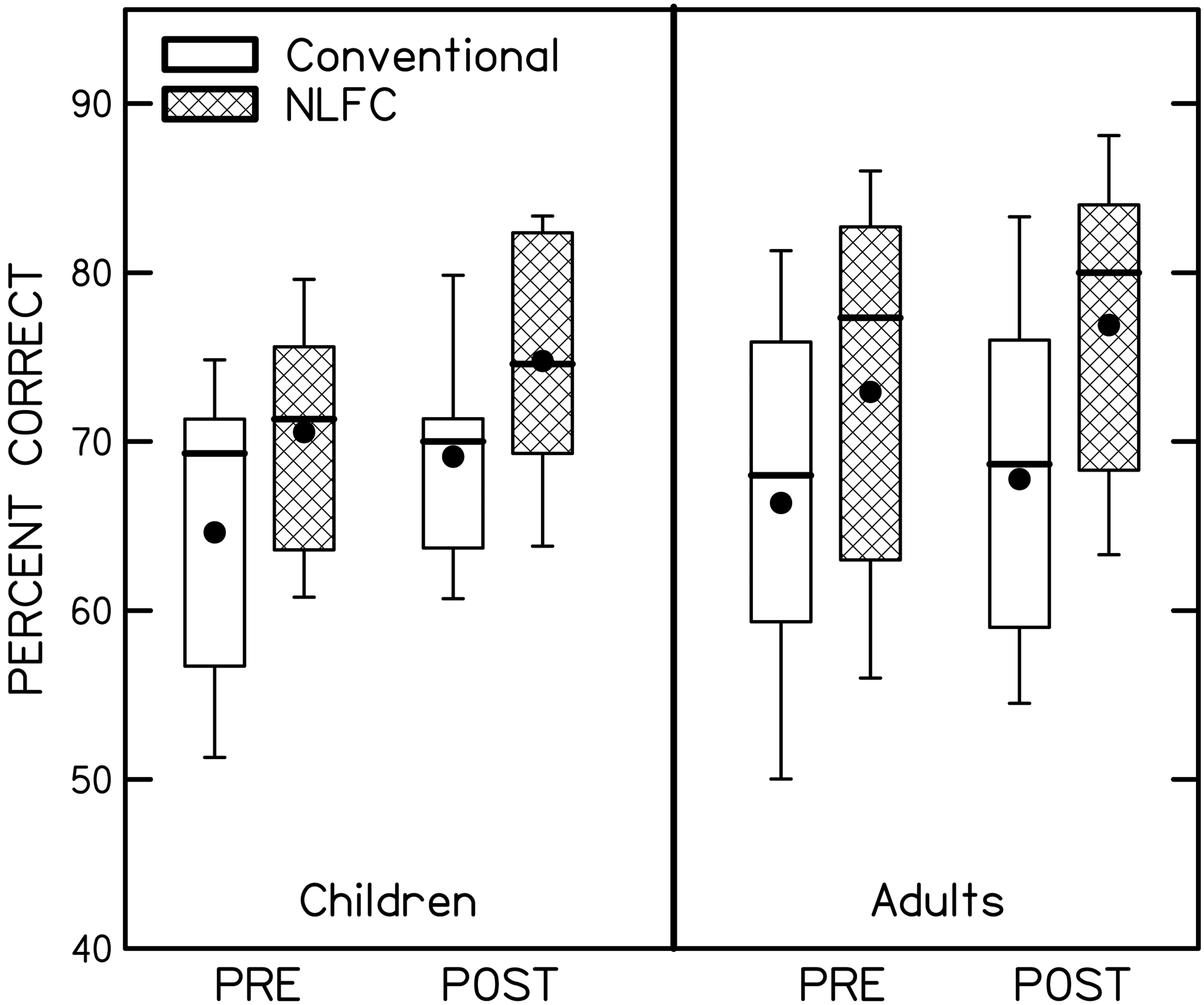

Figure 3.
Percent correct word recognition across conditions of conventional processing and nonlinear frequency compression before (PRE) and after (POST) the story processed with NFC. White box plots are conventional processing (CP), and hatched box plots are nonlinear frequency compression (NFC). Boxes represent the interquartile range (25th – 75th percentile) and whiskers represent the 5th – 95th percentile range. The mean is plotted as a filled circle for each condition, and the median is plotted as a solid horizontal line within each box.

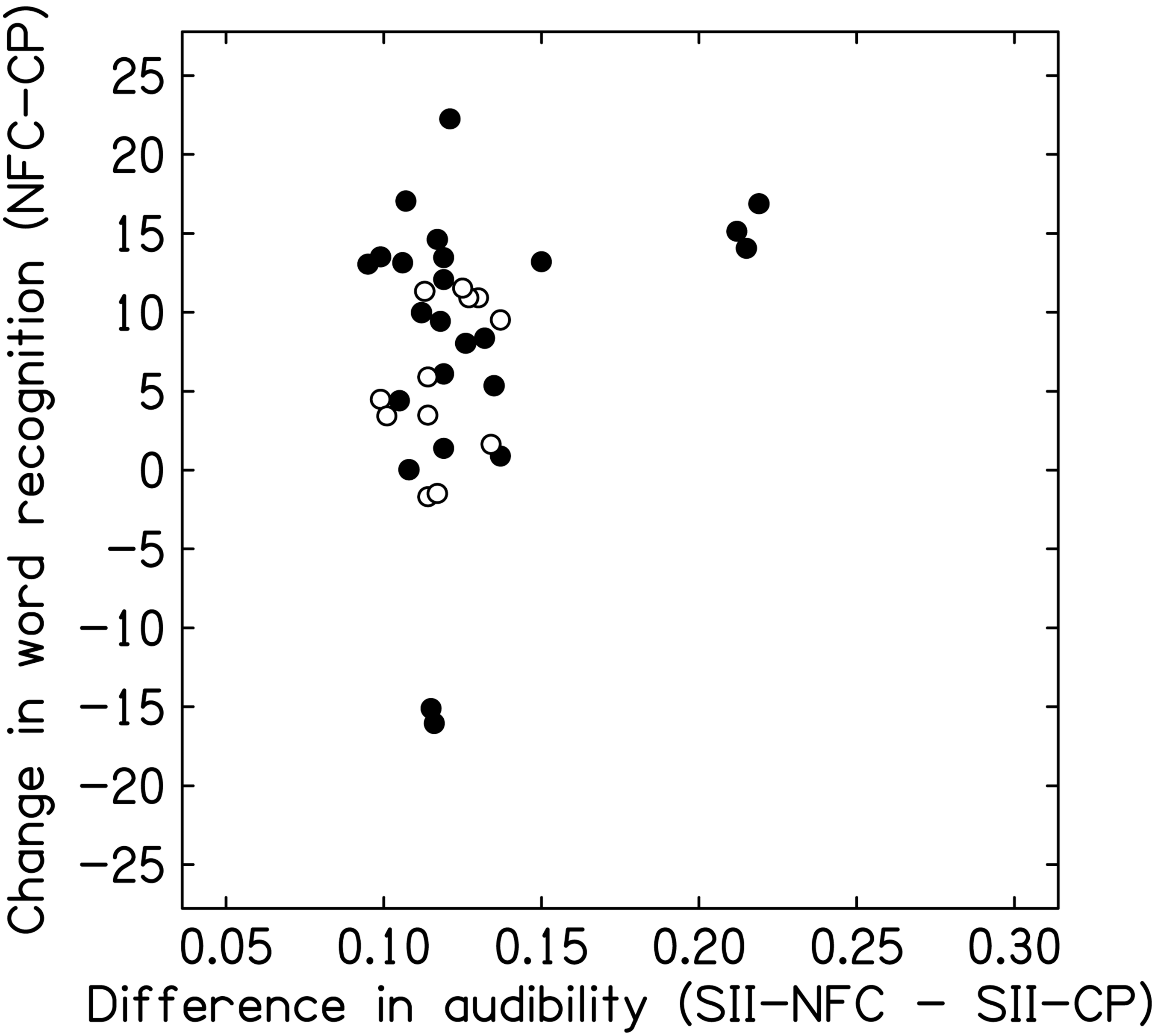

Figure 4.
The difference in word recognition between conventional processing (CP) and nonlinear frequency compression (NFC) as a function of the difference in estimated audibility between the Speech Intelligibility Index for CP (SII-CP) and the SII for NFC (SII-NFC). Open circles represent the difference for adults, and closed circles represent the difference for children.