

Sound Quality Effects of an Adaptive Nonlinear Frequency Compression Processor with Normal-Hearing and Hearing-Impaired Listeners

DOI: 10.3766/jaaa.16179

Danielle Glista*
Marianne Hawkins*
Jonathan M. Vaisberg*
Nazanin Pourmand†
Vijay Parsa*
Susan Scollie*

Abstract

Background: Frequency lowering (FL) technology offers a means of improving audibility of high-frequency sounds. For some listeners, the benefit of such technology can be accompanied by a perceived degradation in sound quality, depending on the strength of the FL setting.

Purpose: The studies presented in this article investigate the effect of a new type of FL signal processing for hearing aids, adaptive nonlinear frequency compression (ANFC), on subjective speech quality.

Research Design: Listener ratings of sound quality were collected for speech stimuli processed with systematically varied fitting parameters.

Study Sample: Study 1 included 40 normal-hearing (NH) adult and child listeners. Study 2 included 11 hearing-impaired (HI) adult and child listeners. HI listeners were fitted with laboratory-worn hearing aids for use during listening tasks.

Intervention: Speech quality ratings were assessed across test conditions consisting of various strengths of static nonlinear frequency compression (NFC) and ANFC speech. Test conditions included those that were fine-tuned on an individual basis per hearing aid fitting and conditions that were modified to intentionally alter the sound quality of the signal.

Data Collection and Analysis: Listeners rated speech quality using the Multi Stimulus test with Hidden Reference and Anchor (MUSHRA) test paradigm. Ratings were analyzed for reliability and to compare results across conditions.

Results: Results show that interrater reliability is high for both studies, indicating that NH and HI listeners from both adult and child age groups can reliably complete the MUSHRA task. Results comparing sound quality ratings across experimental conditions suggest that both the NH and HI listener groups rate the stimuli intended to have poor sound quality (e.g., anchors and the strongest available parameter settings) as having below-average sound quality ratings. A different trend in the results is reported when considering the other experimental conditions across the listener groups in the studies. Speech quality ratings measured with NH listeners improve as the strength of ANFC decreases, with a range of bad to good ratings reported, on average. Speech quality ratings measured with HI listeners are similar and above-average for many of the experimental stimuli, including those with fine-tuned NFC and ANFC parameters.

*National Centre for Audiology, Faculty of Health Sciences, Western University, London, ON, Canada; †Knowles Intelligent Audio, Mountain View, CA

Corresponding author: Danielle Glista, National Centre for Audiology, Faculty of Health Sciences, Western University, 1201 Western Rd., London, ON N6G 1H1, Canada; Email: daglista@nca.uwo.ca

Findings from this study were in part presented at the American Academy of Audiology's annual conference, AudiologyNow! in Phoenix, AZ, April, 2016, the International Hearing Aid Research Conference: IHCON 2016, and at A Sound Foundation through Early Amplification, 7th International Pediatric Conference, 2016.

This study was supported by Phonak AG.

Conclusions: Overall, HI listeners provide similar sound quality ratings when comparing static and adaptive forms of frequency compression, especially when considering the individualized parameter settings. These findings suggest that a range in settings may result in above-average sound quality for adults and children with hearing impairment. Furthermore, the fitter should fine-tune FL parameters for each individual listener, regardless of type of FL technology.

Key Words: adults, children, frequency lowering, hearing aids, nonlinear frequency compression, sound quality

Abbreviations: ANFC = adaptive nonlinear frequency compression; CF = cutoff frequency; CR = compression ratio; FL = frequency lowering; FT = fine-tuned; HI = hearing impaired; ICC = intraclass correlation coefficient; MAOF = maximum audible output frequency; MUSHRA = multiple stimuli and hidden reference and anchor; NFC = nonlinear frequency compression; NH = normal-hearing

INTRODUCTION

Many clinically available hearing aids now include frequency lowering (FL) technology. This type of technology is one of the many digital signal processing options available in modern devices. All types of FL devices share a common goal of improving audibility of high-frequency sounds by shifting them down in the frequency domain. This can help overcome bandwidth/gain limitations imposed by conventional hearing aid devices and/or limitation due to the hair cell damage in hearing-impaired (HI) listeners. However, for some listeners, the benefit of FL can be accompanied by a perceived degradation in sound quality (Souza et al, 2013). The resulting perceived sound quality will depend on a number of variables, including the type of FL processing, strength of the FL settings, hearing level, the type of sound used during sound quality assessment, and highly individual subjective factors such as personal experiences, expectations, and preferences (Parsa et al, 2013; Huber et al, 2014).

FL technologies have evolved over many decades, leading to a variety of terms used to describe different types of FL processing. Current FL types are described using terms such as frequency compression, frequency transposition, frequency composition or multi-layered transposition, and frequency translation; each describing what is suggested to be a unique type of FL algorithm. The effects of each type of processor are difficult to directly compare, given the manufacturers' software constraints around the available fitting parameters when applying FL. Studies comparing FL types suggest large differences in the measured effect on the hearing aid signal; these differences can be observed for FL hearing aids programmed with seemingly similar settings (Alexander, 2013; Scollie et al, 2016). These studies suggest that the numerical FL setting displayed in the fitting software should not be used to determine an optimal setting for a given listener, but instead, this should be determined through the use of real-ear measurement and verification. The number of available clinical FL types continues to evolve, with many hearing aid manufacturers now offering

adaptive FL processors, in addition to static ones. The term "conditional lowering" is also used in the literature to refer to adaptive FL processing; it has been suggested that this type of lowering warrants further research, especially for listeners with extensive high-frequency hearing loss (Salorio-Corbetto et al, 2017). From a sound quality perspective, adaptive FL warrants investigation because the purpose of such processor is sometimes aimed at improving sound quality (Rehmann et al, 2016). This article focuses on the effects of nonlinear frequency compression (NFC) and an adaptive version of it (referred to as ANFC for the purpose of this article) on perceived quality of speech in normal-hearing (NH) and in HI listeners.

NFC selects a band of high-frequency energy on the input side and compresses it into a smaller bandwidth on the output side of the hearing aid response, effectively lowering the respective energy in the frequency domain. For the static form of NFC, this is accomplished by specifying a nominal cutoff frequency (CF) value, above which compression is applied using a specified ratio [referred to as the compression ratio (CR)]. The literature on the effectiveness of NFC has mainly focused on speech perception, with several recent studies reporting NFC benefit for adult and child listeners who are otherwise receiving restricted audibility and/or benefit from conventional hearing aids (Glista et al, 2009; Wolfe et al, 2010; Wolfe et al, 2011; Glista et al, 2012; Alexander, 2013; Hopkins et al, 2014; McCreery et al, 2014; Ellis and Munro, 2015).

This type of lowering can also be accomplished adaptively, using the same principles, but only applying them when the high-frequency components of speech are dominant in the signal. To achieve ANFC, an alternative CF value (which will be referred to as CF2 in this study) can be applied adaptively, and only in the presence of low- and mid-frequency dominated signals. In this case, signal processing is used to analyze the input signal for a difference in the energy distribution between low- to mid-frequency and high-frequency sounds. This prevents the application of FL to low- and mid-frequency signals that may not require lowering to improve audibility. For example, in the case of more

low-frequency content, frequency compression takes place using the higher CF2 value to protect the low-frequency sounds from being compressed (Rehmann et al, 2016). Preliminary studies including adaptive processors have emerged in the literature. A study by Wolfe and colleagues, including children with severe-to-profound high-frequency hearing loss, concluded that the adaptive form of NFC allows for significantly better detection of word-final plurality and the recognition of monosyllabic words in quiet relative to the static form of NFC. Results pertaining to high-frequency phoneme detection and recognition with NFC were reported to be similar to or better than those measured with the static NFC processor. The use of individualized fine-tuning of the adaptive parameters was suggested as an important consideration when optimizing performance with the adaptive form of compression (Wolfe et al, 2017).

Overall, the literature suggests that frequency compression can offer speech perception benefits for some listeners, but that it can also result in a disruption of spectral relationships and formant variations. Perceptually, these disruptions might result in sound quality degradation, particularly at stronger processor settings. For example, vowel formant spectral peaks have been shown to be altered (lowered) by NFC, especially for the higher frequency components (McDermott, 2011). Future investigations are needed to determine if ANFC technology and other forms of FL preserve the spectral relationships and vowel formant structures of speech that are potentially distorted using traditional NFC.

Various studies have focused on the impact of NFC on perceived sound quality. Early studies on the efficacy of NFC reported adverse sound quality affects for some listeners at certain NFC settings (Simpson et al, 2006). Parsa et al (2013) investigated the impact of NFC on perceived sound quality. NH and HI listeners were asked to rate sound quality across different NFC settings. Results showed that both types of listeners were sensitive to the effects of NFC on speech in quiet, with HI listeners being more tolerant to adverse sound quality effects resulting from strong NFC settings, relative to NH listeners. The CF value directly impacted ratings, with poorer sound quality ratings reported for speech when the NFC settings had either higher CR or lower CF frequencies (Parsa et al, 2013). These results are consistent with those reported in a study by Souza et al (2013). In this study, sound quality ratings were greatly affected by a strong amount of NFC. Specifically, low CF values and/or a high CR value adversely affected perceived quality for NH adults, and to a lesser extent, HI adults. Listeners with better high-frequency hearing were more sensitive to the effects of NFC on speech (Souza et al, 2013). The factor of degree of high-frequency hearing loss was highlighted by Salorio-Corbetto et al (2017), who evaluated the sound quality preferences of adult listeners

with extensive high-frequency dead regions in the cochlea, while listening to recordings of different strengths of NFC. They found that most participants had no preference between NFC and conventional processing, at weak or moderate NFC settings. However, at strong NFC settings, more participants preferred the sound quality of conventional processing, and strength of preference tended to increase with processor strength. The pattern of benefit reported by Salorio-Corbetto is consistent with the idea that FL can be viewed in terms of an improved audibility versus increased distortion trade-off that varies by individual listeners (Souza et al, 2013).

Johnson and Light (2015) evaluated the sound quality preferences of adult listeners with severe high-frequency losses, wearing different clinical strengths of NFC compared with conventional hearing aid processing. They found that as the strength of FL was increased, more participants preferred the sound quality of the weaker NFC settings (Johnson and Light, 2015). The authors concluded that modifications to the strength of NFC should be completed in combination with measurements of effective audibility. When studied in adult listeners with mild to moderate hearing losses, NFC had little impact on both speech perception and acceptableness of sound quality (Picou et al, 2015); it is possible that the strength of NFC needed for these listeners, being weaker settings based on the inclusion of listeners with less severe hearing loss, may have contributed to the findings reported. For example, weaker settings using minimal change to either the CF or CR will likely result in minimal impact on perceived speech quality. Furthermore, there will be less room for improvement in speech perception ability for listeners with milder forms of hearing impairment already receiving audibility of high frequencies with conventional technology.

Current FL technologies can be fine-tuned using fitting software specific to the FL algorithm used in the device. With such software, the fitter is able to adjust the strength of FL on a case-by-case basis. Fitting protocols specific to the fitting and fine-tuning of FL technologies are available in the literature. These offer step-by-step guidance for fitting FL clinically; an example of such by Scollie et al (2016) incorporates real ear measurement into the verification process and recommends clinically available stimuli for use in probe microphone measurement systems. These stimuli were tested in a study designed to examine the relationship between the hearing aid verification measures and HI listeners' abilities to detect and discriminate the sounds /s/ and /ʃ/. This study concluded that verified settings which provided a positive sensation level of /s/ and the maximal frequency difference between /s/ and /ʃ/, also resulted in the best speech perception performance (Scollie et al, 2016). However, sound quality outcomes associated

with this fitting protocol have not been reported in the literature.

In summary, current literature suggests that NFC technology offers a means of improving audibility of high-frequency sounds for some listeners. Perceived sound quality of NFC-processed speech is reported to vary across settings, specifically related to the values of CF and CR used; careful consideration of the NFC settings used for each individual fitting is therefore warranted. Research is needed to investigate subjective sound quality associated with ANFC hearing aid fittings. This article will present two studies; the first study discusses NH and the second study discusses HI participants, and is designed to evaluate the following research questions: Is there a difference in the perceived speech quality of static versus adaptive frequency-compressed speech? Do ratings of speech quality differ between adults and children? How do the ratings of speech quality change as a function of ANFC setting strength and with fine-tuning?

STUDY 1

Research Design

Participants

A total of 40 NH listeners participated in this study, including 20 children (mean age = 8.9, range = 6–16 years) and 20 adults (mean age = 21.7, range = 18–30 years). Air conduction thresholds were screened at octave and inter-octave frequencies between 250 and 8000 Hz using insert phones coupled to foam tips to be within 20 dB HL for the adults and 15 dB HL for the children. Tympanometry results were judged to be within normal limits for all participants. This research study was approved by the Western University Health Science Research Ethics Board. Participants were compensated for their time.

Sound Quality Rating Procedure and Experimental Stimuli

The speech quality ratings were completed using custom experimental software using the MULti Stimulus test with Hidden Reference and Anchor (MUSHRA) test paradigm (ITU, 2014). Ratings were completed for a total of eight stimuli. The MUSHRA paradigm includes experimenter selection of a reference stimulus, anchor stimuli, and experimental stimuli. The reference stimulus provides the listener with an example of high-quality speech, such as unprocessed speech. Alternatively, the anchor stimuli are typically processed to provide the listener with examples of poor-quality speech, and are intended to encourage the use of the full rating scale. The remaining experimental stimuli include a

collection of processed stimuli spanning a range in speech quality.

When completing the MUSHRA task, listeners were seated in a double walled sound booth, in front of a loudspeaker (positioned at 0° azimuth) and computer monitor. Stimuli were routed from the PC to external amplifiers and patched into the sound booth to the loudspeaker. The sound field presentation level was 65 dBA. Listeners were instructed to listen to the reference stimulus carefully at the start of each trial. Sound quality ratings were completed using a computer-controlled mouse and response buttons represented on the computer monitor. A separate sound presentation button was provided for the reference, in addition to the inclusion of the reference stimulus as a hidden experimental stimulus. Listeners were then asked to use the mouse to play each stimulus separately and rate the sound quality in comparison with the reference stimulus. Listeners were able to play and rate each of the stimuli multiple times and until they were satisfied with the relative sound quality ratings indicated. Ratings were completed using sliders positioned above each of the eight sound presentation buttons. The sliders were adjustable over a range spanning 1–100, with the following labels appearing at equal intervals: “very bad,” “bad,” “average,” “good,” and “very good.” For example, if the listener clicked at a point on the scale corresponding to “average,” a value to 50 was assigned to the score.

Stimuli included two sets of sentences from the Rainbow Passage; each set was spoken by a female talker. The stimuli were presented in random order over two trials (per sentence). With the exception of the reference stimulus, all stimuli were pre-processed offline. The reference stimulus was equal to the original stimulus (i.e., unprocessed speech). The first anchor was processed using a 2 kHz low-pass filter and a 20 dB slope per octave above the cutoff. The second anchor was processed to include static NFC at the strongest possible setting available clinically (herein referred to as NFC strong). The remaining five experimental stimuli were processed using NFC or ANFC settings chosen based on a simulated hearing aid fitting for an N5 audiogram and in-laboratory electroacoustic measures. These have been labeled as NFC FT, referring to NFC that has been fine-tuned for the N5 audiogram, and ANFC 1 through 4, referring to fine-tuned ANFC settings with varying CF1 and CF2 values. This standard N5 audiogram (Bisgaard et al, 2010) was chosen to be representative of a clinical candidate for FL. In brief, the N5 audiogram represents audiometric thresholds for a severe hearing loss, sloping from 65 dB HL at 250 Hz to 80 dB HL at 6000 Hz.

Experimental settings for both NFC and ANFC were chosen using simulated electroacoustic measures. Specifically, electroacoustic procedures described by Scollie et al (2016), incorporating the recommended /s/ stimulus

in the Verifit2 hearing aid analyzer (Audioscan), were used during the simulated fine-tuning exercises. Briefly, this procedure defines the maximum audible output frequency (MAOF) for the fitter (McCreery et al, 2013) and recommends FL adjustments using a prerecorded calibrated female /s/ stimulus into a recommended MAOF range (Scollie et al, 2016). Research-grade fitting software was used to select final parameter settings based on electroacoustic measurement of the MAOF for /s/. Final settings per condition were measured to fall within the recommended MAOF range of 5300–6406 Hz, indicating that they were set to provide similar amounts of FL for high-frequency stimuli. Numerically, the setting values for CF1 and CF2 differed across experimental conditions, even though the conditions provided a similar amount of lowering, as indicated by the verified frequency location of the /s/ stimulus. The adaptive variable of interest for the purpose for this study was CF2. CF2 was adjusted on the input side of the processor and independent from all other setting parameters. Because of a software limitation, CF1 could not be adjusted independently in the software and instead varied according to the setting selection for all other parameters. The custom software provided by Phonak was used to calculate the output frequency associated with the input frequencies of each parameter, a strategy that has been reported in previous studies of NFC (Salorio-Corbetto et al, 2017). The constant input parameters in “Study 1” included the CF0 and CR. The CF0 parameter defined the slope of the frequency compression function, and was available for adjustment for the purpose of this research study. Minor adjustments were made to the CF0 parameter, on a case-by-case basis, to achieve the desired strength of FL, during verification and fine tuning. The CF0 parameter is not available for adjustment in clinical software. Table 1 reports all parameter values used to create all experimental stimuli for “Study 1.”

Spectrograms of the FL effect for all eight experimental stimuli are shown in Figure 1. Spectrograms are shown to illustrate the effect of the different experimental conditions on the stimuli, including the use of different ANFC parameters. These spectrograms were created using Praat software (version 6.0.29) (Boersma and Weenink, 2018). Specifically, the consonant-vowel pair /is/ was isolated in the word “prism” from the Rainbow Passage sentence and extracted

as a new sound file. They illustrate the broad frequency range of the original speech token, contrasted with significant loss of high-frequency energy in the low-pass filtered condition. Compared with these, the strong NFC condition provides a significant alteration of the original signal and compresses the /s/ frication band to a narrower and lower frequency location. Fine-tuned NFC provides some FL effect because of the moderate settings used and ANFC across settings also provide a moderate effect. Furthermore, the spectrograms illustrate that the ANFC settings from strongest (4) to weakest (1) change in formant structure, with formant locations more similar to the original signal when the ANFC processor is set to provide more adaptation in settings 1 and 2. This illustrates that although the ANFC processor is not a vowel detector per se, it may be possible to select CF2 settings that minimize the impact of frequency compression on the vowel portions of speech. This can also result in less lowering of the /s/ frication band when a highly adaptive (and therefore weaker) setting is created as illustrated in ANFC 1. Overall, we aimed to evaluate whether sound quality ratings would be related to the electroacoustic changes illustrated in Figure 1.

The CF2 settings chosen for the experimental ANFC conditions were selected to provide different amounts of FL for vowels. Specifically, a female utterance of the vowel /i/ was extracted from the Ling-6(HL) test (Glista et al, 2014) and measured in the Verifit2 hearing aid analyzer. This vowel was chosen based on the presence of greater high-frequency content in the upper formant areas, relative to other vowels, which allowed the fitter to observe the effects of CF2 parameter adjustment during verification and fine tuning. Figure 2 includes corresponding spectral displays for these aided measurements across ANFC settings 1–4 and without NFC processing. This visual depiction of the effects of ANFC on the formant structure for /i/ suggests that the upper formants were lowered more for stronger ANFC settings (e.g., ANFC 4) and less when the ANFC CF2 parameter had higher frequency settings (e.g., ANFC 1); lower formant structure remained the same across ANFC CF2 settings. The high-frequency components of the spectra varied in terms of bandwidth and level according to the chosen parameter, in that the overall bandwidth and level increased as the frequency value

Table 1. A List of All Settings Used to Select Both NFC and ANFC Experimental Conditions in “Study 1”

Experimental Condition	Settings		Experimental Condition	Settings			
	CF (Hz)	CR		CF0 (Hz)	CR	CF1 (Hz)	CF2 (Hz)
NFC strong	1440	4:1	ANFC 1	800	1.4:1	3342	5920
NFC FT	2560	2:1	ANFC 2	800	1.4:1	2526	4000
			ANFC 3	800	1.4:1	2076	3040
			ANFC 4	800	1.4:1	1583	2080

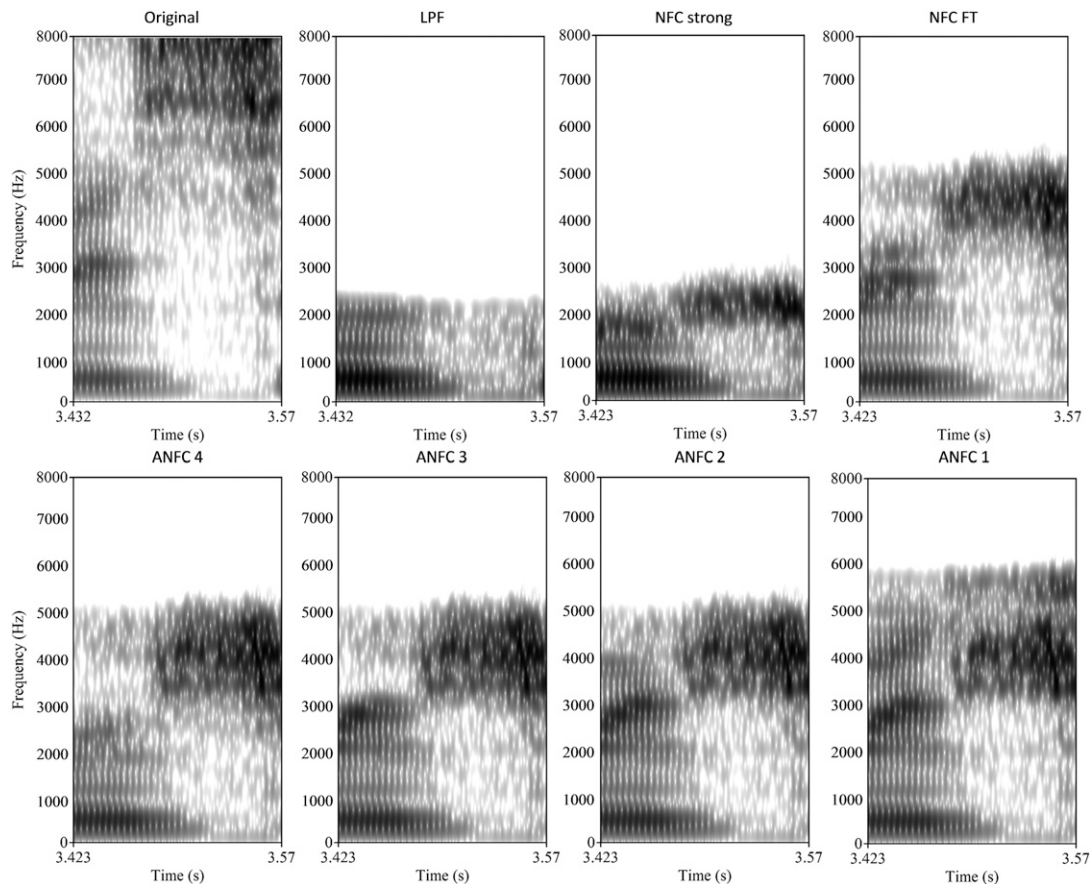


Figure 1. Spectrograms displaying the vowel–consonant pair /is/ for the eight experimental stimuli used in “Study 1.”

of the CF2 was increased. These experimental stimuli were judged to differ in sound quality according to listening checks performed by the experimenters. Specifically, sound quality was judged to decrease as the upper formant of /i/ was lowered.

RESULTS

Analyses for this study were completed using SPSS software (Version 24). Interrater reliability for each participant group (adults and children) was calculated using a one-way random effects intraclass correlation coefficient (ICC) model. The reported ICC for test–retest scores obtained using two Rainbow Passage sentences, collapsed across all experimental stimuli, was 0.93 and 0.89, for the adults and children, respectively.

Figure 3 displays the average sound quality ratings across all eight test conditions for the group mean, and separately for the adults and children. A repeated measures analysis of variance with a Greenhouse-Geisser correction determined that mean sound quality scores differed significantly when comparing the eight test conditions [$F_{(5.12, 194.7)} = 330.3, p < 0.001, \eta_p^2 = 0.897$]; age was not found to be a statistically significant

between-subjects factor. Post hoc tests using a Bonferroni correction (collapsed across participant group) revealed a statistically significant difference between all pairs of stimuli ($p < 0.05$), with the exception of NFC FT and ANFC 3 ($p = 1.000$).

DISCUSSION

The results from the first study indicate that the MUSHRA task can reliably support the evaluation of speech sound quality when used with both adults and children aged 6–30 years with normal hearing. The stimuli used in this study included sentences spoken by a female speaker, processed with various strengths of static and ANFC processing. Overall, NH children and adults were highly sensitive to differences between conditions. Listeners indicated a pattern of reduced sound quality as the strength of ANFC was increased or the frequency range of the adaptive parameter was reduced. Conversely, as the strength of ANFC decreased (or the CF2 values increased in frequency), speech quality ratings increased. Recall that the amount of frequency compression was similar for the /s/ stimulus across these conditions, so it is likely that these sound quality effects were mainly because of

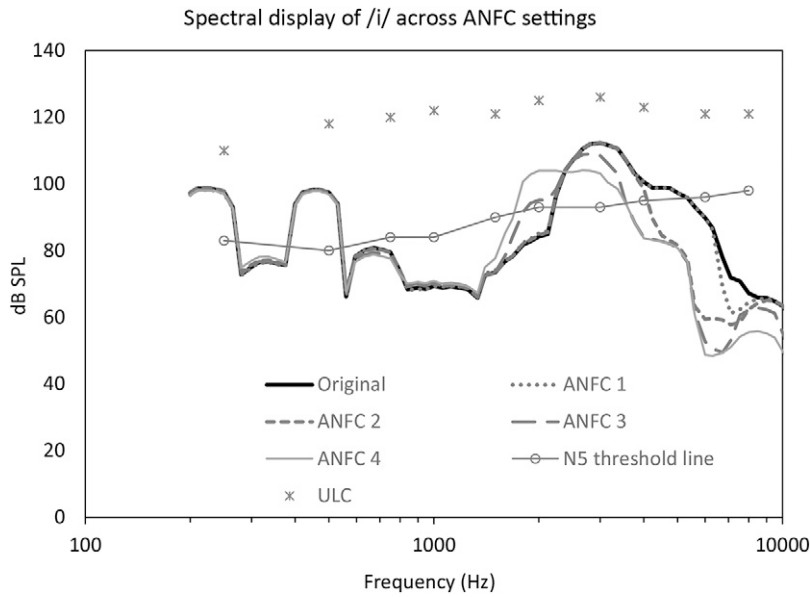


Figure 2. Spectral display of aided measurements for the vowel /i/ completed for the original signal and across all ANFC experimental conditions (ANFC 1–4) for a 65-dB SPL input level. The N5 hearing thresholds are displayed using open-circle markers and the upper-limit-of-comfort using asterisks at octave and inter-octave frequencies.

differences in vowel processing attributable to the adaptive parameter. Listeners judged that the ANFC condition with a moderate CF2 of 3040 Hz to be of “average” sound quality and to be similar to the quality of the fine-tuned static NFC condition. Recall that the same fine-tuning protocol was applied to both the NFC FT and ANFC stimuli. The strongest ANFC setting was judged to have “bad” sound quality and to be of similar sound quality to that produced by a 2 kHz low-pass filter. The poorest rating, on average, was associated with the

strongest form of NFC. This setting was equal to the strongest form of frequency compression available in the hearing aid fitting software, with the CF for NFC set to 1440 Hz to match clinically available settings. The weakest ANFC setting, which provided more of the adaptive effects possible with this processor, was judged to have “good” sound quality, and to be of better sound quality to that produced with the static processor. This indicates that the adaptive version of NFC can improve sound quality of speech and was judged

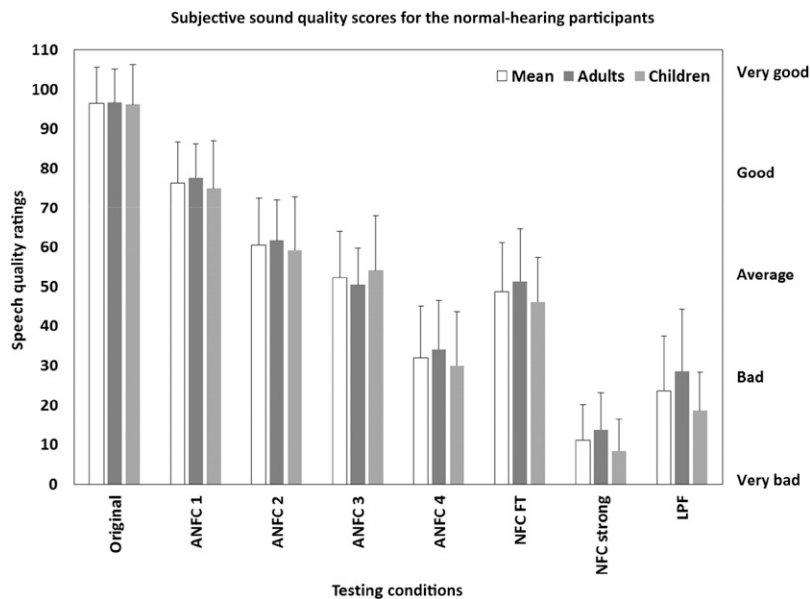


Figure 3. Mean subjective sound quality scores for the NH participants, displayed across eight test conditions for all participants and for the adult and child groups separately. Vertical error bars denote standard deviation values.

to be better than that measured with the static version, for NH listeners. Determining if this finding is present in HI listeners is of considerable clinical interest. The purpose of “Study 2” was to evaluate the sound quality difference between a static NFC and ANFC in a sample of HI listeners.

STUDY 2

Research Design

Participants

A total of 11 HI listeners participated in this study including six children (mean age = 12, range = 9–16 years) and five adults (mean age = 61, range = 24–95 years). Pure-tone audiometry was conducted and real-ear-to-coupler difference values were measured using ER-3A insert earphones (Etymotic) coupled to personal earmolds. All participants presented with sloping, high-frequency sensorineural hearing loss and were judged to be appropriate candidates for NFC. Table 2 lists the air conduction hearing thresholds for both ears and across participants. Participants were arranged according to hearing loss severity, calculated using better-ear high-frequency pure-tone average values at 2000, 3000, 4000, 6000, and 8000 Hz, in order

from 69 to 119 dB HL. Participants were experienced hearing aid users and had at least one full year of previous hearing aid experience, in which they were consistently wearing their hearing aids on a daily basis. All participants had previous exposure to some form of FL processing via participation in other research studies or through their personal hearing aids. Study-worn hearing aids included Naida Q SP or UP devices. Study-worn hearing aids were fitted to either the Desired Sensation Level (DSL) prescription (version 5.0) child or adult prescriptive fitting method depending on age and preference. Refer to Table 3 for participant details pertaining to age, hearing aid model used in this study, and DSL prescription type. This research study was approved by the Western University Health Science Research Ethics Board. Participants were compensated for their time.

Sound Quality Rating Procedure and Experimental Stimuli

The testing procedure used for the HI participants included the use of custom experimental software employing the MUSHRA test paradigm (described in “Study 1”). HI participants completed ratings for a total of nine hearing aid settings chosen to span a range in speech quality. The equipment set-up was similar to that described in “Study 1,” with the exception of the addition

Table 2. Hearing Threshold Values, Displayed in dB HL, Across All Octave and Inter-Octave Frequencies for the Left (L) and Right (R) Ears of Each Participant

Participant	Ear	250	500	750	1000	1500	2000	3000	4000	6000	8000	BEHFPTA
1	R	30	50	55	60	65	70	75	70	65	65	69
	L	25	40	50	55	65	70	70	65	70	70	
2	R	60	65	70	70	75	75	80	80	65	65	73
	L	55	65	70	70	75	75	80	80	65	70	
3	R	30	45	50	65	80	85	85	75	60	65	74
	L	35	40	50	70	80	90	85	75	60	60	
4	R	40	50	60	65	65	65	75	80	80	75	75
	L	40	50	60	65	70	70	75	80	80	80	
5	R	30	30	40	50	50	55	70	85	85	95	78
	L	50	50	70	85	90	90	105	95	NR	NR	
6	R	25	65	100	105	100	95	95	95	85	85	80
	L	25	60	95	95	90	85	85	90	70	70	
7	R	80	70	60	55	55	60	70	80	85	NR	83
	L	80	75	70	65	65	60	80	85	80	NR	
8	R	20	25	30	50	50	60	75	85	100	NR	88
	L	35	40	50	55	60	60	65	80	95	NR	
9	R	25	20	25	25	50	60	90	105	90	NR	93
	L	30	20	25	40	50	70	105	105	100	NR	
10	R	60	55	65	80	105	105	110	NR	NR	NR	115
	L	85	100	105	100	105	110	115	115	NR	NR	
11	R	55	70	95	105	115	115	NR	NR	NR	NR	119
	L	55	75	90	105	115	NR	NR	NR	NR	NR	
Mean		44.1	52.7	63.0	69.8	76.1	79.3	87.7	90.2	88.0	96.4	86.1

Mean threshold values are indicated for the group of listeners. NR = no response, is indicated where a threshold was not measurable within the limits of the audiometer. Participants are arranged in order of least to most severe hearing loss, according to better-ear high-frequency pure-tone average values (BEHFPTA) in dB HL.

Table 3. Participant Factors Including Age at Time of Study, Study-Worn Hearing Aid Model, DSL Prescription Used per Participant and a List of All Fine-Tuned NFC and ANFC Settings Used in “Study 2”

Participant	Age (years)	Hearing Aid Model	Prescription	NFC Fine-Tuned Setting		ANFC Fine-Tuned Setting			
				CF (Hz)	CR	CF0 (Hz)	CR	CF1 (Hz)	CF2 (Hz)
1	13	Naida Q SP	DSL child	4000	2:1	1120	1.3:1	3253	4480
2	11	Naida Q SP	DSL child	4320	2.5:1	1120	1.3:1	3253	4480
3	9	Naida Q SP	DSL child	4160	2.1:1	1120	1.26:1	3460	4640
4	16	Naida Q SP	DSL child	4000	2:1	960	1.24:1	3325	4480
5	10	Naida Q SP	DSL child	4000	2.8:1	800	1.28:1	3244	4800
6	12	Naida Q UP	DSL child	4160	2.1:1	160	1.2:1	2874	5120
7	95	Naida Q SP	DSL adult	1600	1.6:1	160	1.2:1	2798	4960
8	72	Naida Q SP	DSL adult	3200	2.3:1	160	1.2:1	3096	5600
9	87	Naida Q SP	DSL adult	2400	3:1	160	1.3:1	2356	5280
10	25	Naida Q UP	DSL child	1600	4:1	160	1.43:1	1013	2240
11	24	Naida Q UP	DSL child	1440	4:1	320	1.67:1	936	1920

of real-time hearing aid programming. This was accomplished using a direct connection to a Hi-Pro2 hearing instrument fitting interface (GN Otometrics), allowing manipulation of hearing aid settings during the rating task in real-time. The instructions and response format used were the same as that described in “Study 1.”

Stimuli presented to the HI participants included two sets of sentences from the Rainbow Passage; each set was spoken by a female talker. The stimuli were presented in random order over two trials (per sentence). All stimuli were processed via experimental software that, in combination with the manufacturer fitting software, processed the stimuli in real time using the hearing aid processing settings for each FL test condition. All stimuli were therefore heard by the HI participants through individually programmed hearing aids. The reference stimulus condition was equivalent to the original stimulus (i.e., individualized gain per participant but no FL). Two anchor conditions included the use of the strongest available NFC and ANFC parameters; these were labeled as NFC strong and ANFC strong. The parameter values used for NFC strong included a CF value of 1440 Hz and a CR of 4:1. The parameter values used for ANFC strong included a CR of 4:1, CF0 of 160 Hz, CF1 of 285 Hz, and a CF2 value of 1440 Hz.

In addition to the reference and anchor conditions, fine-tuned conditions were created for both the ANFC and NFC processors. This is another exception in study design compared with “Study 1.” The fine-tuned conditions were individualized per participant hearing loss, using electroacoustic procedures described by Scollie et al (2016). This involved defining the MAOF for each participant and using the calibrated /s/ stimulus to adjust the FL parameters into the recommended MAOF range (Scollie et al, 2016). Table 3 displays the NFC and ANFC parameters selected in the experimental fitting software to achieve each participant’s fine-tuned

fitting. The remaining four experimental conditions were created by systematically modifying the CF2 value of each participant’s fine-tuned setting to reflect the same CF2 values as those used in the NH study (all other parameters remained the same as what was decided in the fine-tuned setting). All other parameters were held constant. These are labeled as ANFC 1 through 4 and reflect CF2 frequency values ranging from 2080 through 5920 Hz.

Results

Analyses for the second study were completed using SPSS software (Version 24). Interrater reliability for each HI participant group (adults and children) was calculated using a one-way random effects ICC model. The reported ICC for test–retest scores obtained using two Rainbow Passage sentences, collapsed across all experimental stimuli, was 0.78 and 0.87 for the adults and children, respectively, suggesting reliable ratings within groups, although less reliable compared with NH participants.

Figure 4 displays mean sound quality ratings across nine test conditions and for 11 listeners, separately for the adults and children. A repeated measures analysis of variance with a Greenhouse–Geisser correction determined that mean sound quality scores differed significantly when comparing results obtained for the nine test conditions [$F_{(4,22, 38.0)} = 13.4, p < 0.001, \eta_p^2 = 0.599$]; age was not found to be a statistically significant between-subjects factor. Post hoc analyses using a Bonferroni correction (collapsed across participant group) revealed a statistically significant difference for the following pairwise comparisons ($p < 0.05$): Sound quality ratings were lower for both the ANFC and NFC strong conditions (i.e., the anchor stimuli) when compared with the ANFC FT, NFC FT, ANFC 1, and 3 and to the original condition (i.e., the reference stimulus).

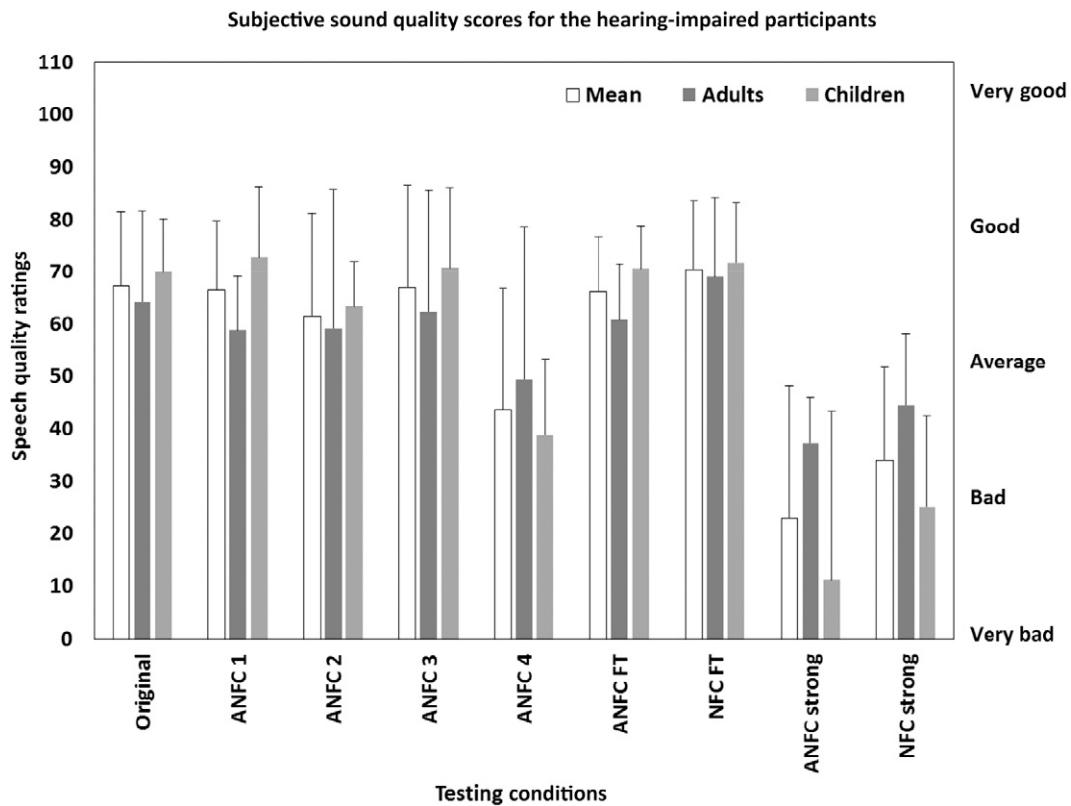


Figure 4. Mean subjective sound quality scores for the HI participants, displayed across nine test conditions for all participants and for the adult and child groups separately. Vertical error bars denote standard deviation values.

All other conditions were not judged to be statistically different.

Discussion

The results from “Study 2” indicate that the MUSHRA task can reliably support the evaluation of speech sound quality when used with both adults and children aged 9–95 years with hearing loss. The hearing losses included in the study presented with better-ear high-frequency pure-tone average values ranging from 69 to 119 dB HL, indicating variability across participant hearing levels within the high-frequency region of interest in this study. Experimental modifications to the adaptive CF2 parameter of the ANFC processor were perceived differently by HI relative to NH listeners. On average, HI children and adults were not as sensitive to differences in the test conditions, specifically to modifications made to the CF2. This is consistent with previous studies of NFC sound quality, in which listeners with hearing loss were less sensitive to sound quality differences between NFC processing conditions compared with their NH counterparts (Parsa et al, 2013). When considering all experimental FL conditions, listener ratings were most consistent (i.e., smaller standard deviation and range values) for the fine-tuned settings (including both

NFC and ANFC). These results may speak to the importance of fine-tuning FL strength on an individual basis.

Recall that the CF2 value for the ANFC 4 condition was set to be lower in frequency than the other experimental ANFC conditions, which may have resulted in more lowering being applied to the vowel portion of the stimulus, for example. The lower frequency setting of the ANFC CF2 resulted in a different amount of lowering being applied to lower frequency stimuli (such as vowels), when compared with higher frequency ones (such as fricative stimuli), thereby limiting the adaptation effect overall. Group-level results suggest that the listeners, on average, could not distinguish the difference in the sound quality of the ANFC 2 condition, compared with the other conditions that were intentionally created to have poor sound quality (i.e., ANFC/NFC strong). Other ANFC settings that had higher CF2 values may therefore have provided more adaptation across phonemes, and resulted in significantly higher sound quality. Overall, there was a range of CF2 settings that resulted in average to good sound quality across HI listeners. Adjustment of the ANFC parameters should be completed on an individual basis, allowing for the adaptive effect of the processor to be applied across a broad frequency range, whenever possible.

Clinical adjustment of FL parameters is largely determined by the fitting method used (i.e., verification measures) and the flexibility of the fitting software. The results presented in this research article pertain to parameter settings available through the use of experimental software and that are not directly comparable with those available through clinical fitting software. Within current clinical fitting software available for this type of FL processing (Rehmann et al, 2016), a subset of all possible parameter settings is made available, rather than all possible combinations. Therefore, adjustment of the adaptive CF parameters, for example, must be adjusted in combination with adjustment to the other FL parameters using a hybrid fitting slider in the manufacturer software. This hybrid strategy differs from the parameterized software used in the present study. For the purpose of this experiment, the investigators chose to use research software that allowed for independent control of the parameters of interest. This may limit the generalizability of our findings to clinical settings.

Summary and Concluding Discussion

Findings suggest that NH listeners are highly sensitive to sound quality effects associated with unprocessed versus NFC-/ANFC-processed speech, to variations in the strength of the setting used for both types of NFC, and to variations of the ANFC parameter designed to preserve the quality of low- and mid-frequency sounds. This was true for both the adult and child participants. HI listeners may be less sensitive to such processing effects, as demonstrated in the results presented in this research article. Previous investigations of sound quality with NFC suggest that listeners with severe hearing loss are less likely to report adverse sound quality effects when compared with listeners with NH (Parsa et al, 2013) or when compared with HI listeners with more residual hearing (Souza et al, 2013; Picou et al, 2015). On average, the HI listeners rated many of the experimental conditions as having average-good sound quality, as compared with the below-average ratings assigned to the strong NFC and ANFC conditions.

The recruitment of participants for FL-related studies is becoming an increasingly challenging task. FL processors are designed to overcome high-frequency audibility limitations, and are often suggested for use with hearing aid candidates who have significant high-frequency hearing loss. However finding candidates who meet this criteria can be challenging as the listeners presenting with this degree of hearing loss often pursue cochlear implantation treatment, thus preventing their inclusion in methodologies such as these. Recruitment for this study aimed to maximize the study sample size and resulted in a total of 11 HI listeners being included in "Study 2." This sample size is similar to other small-sample studies of FL (e.g., Wolfe et al, 2017).

Although the studies in this article included laboratory testing only, it is possible that with acclimatization time and real-world experience with specific FL parameters, the HI listeners may have learned to accept stronger amounts of frequency compression. It is also worth noting that such a study design would require many different lengthy trials to evaluate performance change directly linked to a specific FL parameter selection (Glista et al, 2012). This type of a study would require a different design than what was proposed for the current set of studies.

The results from "Study 1" suggest that ANFC may offer additional sound quality benefit when compared with static NFC. When replicated with HI listeners using a similar study design, a different trend emerged: HI listeners were equally sensitive to ANFC as they were to NFC when fine-tuned using individualized FL parameters and also when compared with many of the other experimental ANFC conditions. The HI listeners were also less sensitive to the introduction of distortions across experimental conditions, when compared with NH listeners, and more likely to judge the experimental stimuli to have similar sound quality. Nonetheless, the HI listeners rated adaptive settings as having higher sound quality for settings that allowed for medium and high levels of adaptation (e.g., ANFC 1 and ANFC 3), compared with a setting that allowed for minimal adaptation (e.g., ANFC 4). This suggests that there may be a wide range of settings that could potentially offer acceptable sound quality, although further validation with a larger sample is warranted. It is worth noting that the sound quality ratings associated with the reference stimulus were much lower for the HI group, when compared with the NH group. This may relate to highly individualized factors such as familiarity of sound, expectations, and intelligibility, in addition to degree and configuration of hearing loss. In particular, the effect of age group was nonsignificant in this study, but age group has emerged as an important variable in other studies with larger samples (Glista et al, 2009; Parsa et al, 2013). This may indicate that our sample size was not large enough to detect group effects between adults and children. Further research is needed to validate individual and group results with other types of FL and in larger samples of listeners. Preliminary research suggests that adaptive frequency compression with listeners with severe-profound loss in the high frequencies can result in similar or improved speech perception ability, when compared with static frequency compression (Glista et al, 2017; Wolfe et al, 2017). The effect of speech perception as it relates to sound quality, across adaptive FL parameters, is something that needs further research. The use of patient-driven listening tasks, such as the paired comparison approach used in the Johnson and Light (2015) study, may help with further exploration of sound quality effect across different FL settings, for a given fitting.

The findings reported in “Study 2” suggest the need to fit and fine-tune FL parameters on an individual basis. In this research article, the fine-tuned fittings were less variable, in comparison with the other experimental conditions, and resulted in above-average sound quality ratings, on average. Overall, the findings suggest that a range in ANFC settings may result in above-average sound quality for adult and child listeners with hearing impairment. Verification methods specific to fitting FL are available for use in clinical hearing instrument fitting equipment. The use of probe microphone measures during the fitting and verification of FL fitting is recommended, as well as the application of individualized FL parameters that maximizes effective audibility and sound quality for the listener and the chosen hearing aid device.

Acknowledgments. The authors would like to thank Charla Levy and Mark Sandor for their contributions to the data collection process and to Julia Rehmann and Michael Boretzki for assistance with the processing of sound files and fitting software.

REFERENCES

- Alexander JM. (2013) Individual variability in recognition of frequency-lowered speech. *Semin Hear* 34(2):86–109.
- Bisgaard N, Vlaming MS, Dahlquist M. (2010) Standard audiograms for the IEC 60118-15 measurement procedure. *Trends Amplif* 14(2):113–120.
- Boersma P, Weenink D. (2018) Praat: doing phonetics by computer [Computer program]. Version 6.0.37. <http://www.praat.org/>. Accessed March 14, 2018.
- Ellis RJ, Munro KJ. (2015) Benefit from, and acclimatization to, frequency compression hearing aids in experienced adult hearing-aid users. *Int J Audiol* 54(1):37–47.
- Glista D, Hawkins M, Bohnert A, Rehmann J, Wolfe J, Scollie S. (2017) The effect of adaptive nonlinear frequency compression on phoneme perception. *Am J Audiol* 26(4):531–542.
- Glista D, Scollie S, Bagatto M, Seewald R, Parsa V, Johnson A. (2009) Evaluation of nonlinear frequency compression: clinical outcomes. *Int J Audiol* 48(9):632–644.
- Glista D, Scollie S, Moodie S, Easwar V; Network of Pediatric Audiologists of Canada. (2014) The Ling 6(HL) test: typical pediatric performance data and clinical use evaluation. *J Am Acad Audiol* 25(10):1008–1021.
- Glista D, Scollie S, Sulkers J. (2012) Perceptual acclimatization post nonlinear frequency compression hearing aid fitting in older children. *J Speech Lang Hear Res* 55(6):1765–1787.
- Hopkins K, Khanom M, Dickinson AM, Munro KJ. (2014) Benefit from non-linear frequency compression hearing aids in a clinical setting: the effects of duration of experience and severity of high-frequency hearing loss. *Int J Audiol* 53(4):219–228.
- Huber R, Parsa V, Scollie S. (2014) Predicting the perceived sound quality of frequency-compressed speech. *PLoS ONE* 9(11):e110260.
- ITU. (2014) *Methods for the Subjective Assessment of Intermediate Quality Level of Audio Systems: Broadcasting Service (Sound)*. Geneva, Switzerland: International Telecommunication Union.
- Johnson EE, Light KC. (2015) A patient-centered, provider-facilitated approach to the refinement of nonlinear frequency compression parameters based on subjective preference ratings of amplified sound quality. *J Am Acad Audiol* 26(8):689–702.
- McCreery RW, Alexander J, Brennan MA, Hoover B, Kopun J, Stelmachowicz PG. (2014) The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss. *Ear Hear* 35(4):440–447.
- McCreery RW, Brennan MA, Hoover B, Kopun J, Stelmachowicz PG. (2013) Maximizing audibility and speech recognition with nonlinear frequency compression by estimating audible bandwidth. *Ear Hear* 34(2):e24–e27.
- McDermott HJ. (2011) A technical comparison of digital frequency-lowering algorithms available in two current hearing aids. *PLoS ONE* 6(7):e22358.
- Parsa V, Scollie S, Glista D, Seelisch A. (2013) Nonlinear frequency compression: effects on sound quality ratings of speech and music. *Trends Amplif* 17(1):54–68.
- Picou EM, Marcrum SC, Ricketts TA. (2015) Evaluation of the effects of nonlinear frequency compression on speech recognition and sound quality for adults with mild to moderate hearing loss. *Int J Audiol* 54(3):162–169.
- Rehmann J, Jha S, Allegro Baumann S. (2016) *SoundRecover2—The Adaptive Frequency Compression Algorithm [Technical Report]*. Stafa, Switzerland: Phonak AG.
- Salorio-Corbetto M, Baer T, Moore BCJ. (2017) Quality ratings of frequency-compressed speech by participants with extensive high-frequency dead regions in the cochlea. *Int J Audiol* 56(2):106–120.
- Scollie S, Glista D, Seto J, Dunn A, Schuett B, Hawkins M, Parsa V. (2016) Fitting frequency-lowering signal processing applying the AAA Pediatric Amplification Guideline: updates and protocols. *J Am Acad Audiol* 27(3):219–236.
- Simpson A, Hersbach AA, McDermott HJ. (2006) Frequency-compression outcomes in listeners with steeply sloping audiograms. *Int J Audiol* 45(11):619–629.
- Souza PE, Arehart KH, Kates JM, Croghan NBH, Gehani N. (2013) Exploring the limits of frequency lowering. *J Speech Lang Hear Res* 56(5):1349–1363.
- Wolfe J, Duke M, Schafer EC, Rehmann J, Jha S, Allegro Baumann S, Jones C. (2017) Preliminary evaluation of a novel non-linear frequency compression scheme for use in children. *Int J Audiol* 56(12):976–988.
- Wolfe J, John A, Schafer E, Nyffeler M, Boretzki M, Caraway T. (2010) Evaluation of non-linear frequency compression for school-age children with moderate to moderately-severe hearing loss. *J Am Acad Audiol* 21(10):618–628.
- Wolfe J, John A, Schafer E, Nyffeler M, Boretzki M, Caraway T, Hudson M. (2011) Long-term effects of non-linear frequency compression for children with moderate hearing loss. *Int J Audiol* 50(6):396–404.